Generalized Method of Wavelet Moments for Inertial Navigation Filter Design

calibration procedure. Regarding this, we apply the GMWM on error signals issued from MEMS-based inertial sensors by building and estimating composite stochastic processes for which classical methods cannot be used. In a first stage, we validate the resulting models using AV and PSD analyses and then, in a second stage, we study the impact of the resulting stochastic models design in terms of positioning accuracy using an emulated scenario with statically observed error signatures. We demonstrate that the GMWM-based calibration framework enables to estimate complex stochastic models in terms of the resulting navigation accuracy that are relevant for the observed structure of errors.

YANNICK STEBLER
École Polytechnique Fédérale de Lausanne
Switzerland

STÉPHANE GUERRIER
University of Illinois at Urbana-Champaign
USA

JAN SKALOUD
École Polytechnique Fédérale de Lausanne
Switzerland

MARIA-PIA VICTORIA-FESER
University of Geneva
Switzerland

The integration of observations issued from a satellite-based system (GNSS) with an inertial navigation system (INS) is usually performed through a Bayesian filter such as the extended Kalman filter (EKF). The task of designing the navigation EKF is strongly related to the inertial sensor error modeling problem. Accelerometers and gyroscopes may be corrupted by random errors of complex spectral structure. Consequently, identifying correct error-state parameters in the INS/GNSS EKF becomes difficult when several stochastic processes are superposed. In such situations, classical approaches like the Allan variance (AV) or power spectral density (PSD) analysis fail due to the difficulty of separating the error processes in the spectral domain. For this purpose, we propose applying a recently developed estimator based on the generalized method of wavelet moments (GMWM), which was proven to be consistent and asymptotically normally distributed. The GMWM estimator matches theoretical and sample-based wavelet variances (WVs), and can be computed using the method of indirect inference. This article mainly focuses on the implementation aspects related to the GMWM, and its integration within a general navigation filter

NOMENCLATURE

Conventions

$H(\cdot)$	Refers to a function
$\{h_l\}$	Refers to a sequence of values indexed by integer l
$\{h_l^\star\}$	Refers to a simulated sequence indexed by integer l
h_l	Refers to the lth value of a sequence
$\mathbf{h}$	Vector containing $\{h_l\}$
Y_k	Refers to a random variable indexed by integer k
y_k	Refers to a realization of Y_k indexed by integer k

Notations

f	Frequency of a sinusoid
$\{h_l\}$	Discrete wavelet transform filter
$\{\tilde{h}_l\}$	Maximal overlap discrete wavelet transform filter
$F(\boldsymbol{\theta})$	Data generating model parameterized by $\boldsymbol{\theta}$
k	Unitless index
N	Sample size
$\mathcal{N}(\cdot, \cdot)$	Gaussian distribution function
$\mathbb{R}^p$	Space of real-valued p dimensional vectors
$S_Y(\cdot)$	Power spectral density function of $\{Y_k\}$
$\{Y_k\}$	Time series or portion of a stochastic process
$\nu(\tau_j)$	Wavelet variance at scale τ_j
$\hat{\nu}(\tau_j)$	Unbiased maximal overlap discrete wavelet transform estimator of wavelet variance
$\boldsymbol{\nu}$	Theoretical wavelet variance vector containing $\{\nu(\tau_j)\}$
$\boldsymbol{\nu}(\boldsymbol{\theta})$	Wavelet variance implied by $\boldsymbol{\theta}$ assuming that $F(\boldsymbol{\theta})$ corresponds to the true data generating process
$\hat{\boldsymbol{\nu}}$	Estimated wavelet variance based on maximal overlap discrete wavelet transform estimator
$\rho_{XY}(\tau)$	Cross-covariance sequence between $\{X_k\}$ and $\{Y_k\}$ at lag τ
$\sigma_{\bar{Y}}^2(\tau)$	True Allan variance of signal $\{Y_k\}$ at scale τ_j
$\hat{\sigma}_{\bar{y}}^2(\tau)$	Sample-based (estimated) Allan variance of signal
τ_j	Dyadic wavelet/Allan variance scale ($\tau_j = 2^{j-1}$)
$\boldsymbol{\theta}$	Parameter vector such that $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^p$

I. INTRODUCTION

The estimation of a vehicle's position, velocity, and attitude in space in a precise and reliable way, at any time, with and without reception of satellite signals, is an important issue for many applications. To this end, navigation by means of Global Navigation Satellite Systems (GNSS) is now a standard approach for performing localization in outdoor environment. Conditions where satellite signals are partially or completely unavailable severely degrade the performance of such systems. Moreover, GNSS sensor bandwidth (typically below 10 Hz) may be too low for some applications and the tracking of satellites is difficult to maintain in vibrating conditions. Also, no information on attitude (e.g. sensor orientation) is provided by GNSS. A well-accepted and largely proven approach for improving navigation in such situations is to integrate GNSS with inertial sensors. A conventional strapdown inertial measurement unit (IMU) is composed of a triad of usually orthogonally mounted accelerometers and gyroscopes observing specific force and angular rate or change, respectively. After initialization, these signals are integrated with respect to time to yield velocity, position, and attitude at a high rate (typically, higher than 50 Hz). This procedure is the core of a strapdown inertial navigation system (INS). The combination of inertial navigation with GNSS is usually performed through Bayesian techniques among which the most popular is the extended Kalman filter (EKF). During periods of poor GNSS signal quality or the total absence of its reception, inertial navigation operates in coasting mode, i.e., the navigation states are determined independently from GNSS data. In such case, the overall navigation performance becomes strongly dependent on the errors corrupting inertial signals. These errors are integrated in the INS and their impact consequently grows with time. Correct error modeling and estimation of their systematic part is thus very important for improving, as well as correctly predicting, its quality.

Inertial sensors are corrupted by errors (e.g. scale factors, biases, drifts) of deterministic and stochastic nature. A large part of the deterministic errors (e.g. axes misalignment, temperature effects) is compensated through physical models during calibration procedures. These error types and their calibration have been investigated in many works (see e.g. [1–3]) and are not treated in this study. On the other hand, the stochastic errors contain components that have random behavior (e.g. dynamics-dependent errors) or are too complicated to model deterministically (e.g. environmental changes, internal sensor noise). These types of errors are modeled using stochastic processes in the augmented part of the navigation filter (for details on state space augmentation, see e.g. [4]). The questions of which stochastic processes to use for best describing the random part of the inertial sensors behavior and the determination of the process parameters are tasks of the challenging navigation filter
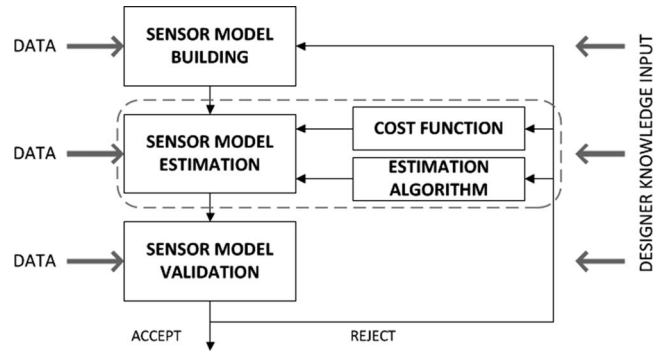

Fig. 1.   Filter design loop. In this work, the main focus is set on elements enclosed by dashed line (adapted from [7]).

design. The problem of estimating the stochastic process model (from noisy data) is the main objective of this paper, and is referred to herein as the stochastic modeling problem.

The rest of the paper is organized as follows. In Section II, we precisely state the problem handled in this article by providing a general description of the navigation filter design problem and the major drawbacks of the conventional methods. Our work is based on the generalized method of wavelet moments (GMWM, [5]) estimator, for which we provide the theoretical background in Section III. Important implementation issues related to the GMWM-based calibration procedure are discussed in Section IV. An example of calibration of a MEMS-based IMU is done in Section V. Section VI further analyses the impact of the resulting EKF design with the previously estimated model parameters on the final navigation performance. Section VII provides conclusions and perspectives.

II. PROBLEM STATEMENT

Consider the sequence $\{y_k : k = 1, \ldots, N\}$ representing the observed one-dimensional noncompensated error signal of an accelerometer or a gyroscope. This sequence is assumed to be a realization of a univariate Gaussian time series $\{Y_k : k \in \mathbb{Z}\}$ that is stationary[1] and to which the conditional distribution $F(Y_k|Y_{k-1}, Y_{k-2}, \ldots, \boldsymbol{\theta})$ with parameters $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subseteq \mathbb{R}^p$ is associated. Therefore,

$$Y_k \sim F(Y_k|Y_{k-1}, Y_{k-2}, \ldots, \boldsymbol{\theta}). \tag{1}$$

Navigation filter design is an iterative process that can be represented as a loop (see Fig. 1) aiming to answer the following three questions which may have significant impact on navigation performance.

1) Which processes should be considered for building $F(\cdot)$ that best describe the underlying dynamics of the observed sensor error sequence $\{y_k\}$? This question refers

---

[1]$\{Y_k\}$ can also be nonstationary but with stationary backward differences of order d. The first-order backward difference of Y_k is $Y_k^{(1)} = Y_k - Y_{k-1}$ and the backward difference of order d is $Y_k^{(d)} = Y_k^{(d-1)} - Y_{k-1}^{(d-1)}$.

TABLE I
PSD and Time Representation of Commonly Used Stochastic Processes

Process	Abbr.	θ	Power Spectral Density	Differential Equation Discrete State-Space Model	Note
Gaussian White Noise	WN	σ_{WN}^2	$S_X(f) = \sigma_{WN}^2$	$\dot{X}(t) = W(t)$	
				$X_k = W_k$	$W_k \overset{iid}{\sim} N\left(0, \sigma_{WN}^2\right)$
Random Walk	RW	σ_{RW}^2	$S_X(f) = \frac{\sigma_{RW}^2}{(2\pi f)^2}$	$\dot{X}(t) = W(t)$	
				$X_{k+1} = X_k + W_k$	$W_k \overset{iid}{\sim} N\left(0, \sigma_{RW}^2\right)$
Gauss-Markov	GM	β, σ_{GM}^2	$S_X(f) = \frac{2\sigma_{GM}^2\beta}{(2\pi f)^2+\beta^2}$	$\dot{X}(t) = -\beta X(t) + W(t)$	
				$X_{k+1} = e^{-\beta\Delta t}X_k + W_k$	$W_k \overset{iid}{\sim} N\left(0, \sigma_{GM}^2(1-e^{2\beta\Delta t})\right)$
Bias Instability	BI	T_{BI}, σ_{BI}^2	$S_X(f) = \begin{cases} \left(\frac{\sigma_{BI}^2}{2\pi}\frac{1}{f}\right) & \text{if } f \le f_0; \\ 0 & \text{if } f > f_0. \end{cases}$	$-$	
				$X_{k+1} = \begin{cases} W_{k+1} & \text{if } \mod(t_{k+1}, T_{BI})=0; \\ X_k & \text{otherwise.} \end{cases}$	$W_k \overset{iid}{\sim} N\left(0, \sigma_{BI}^2\right)$
Quantization Noise	QN	Q	$S_X(f) = 4\Delta t\, Q^2 \sin^2\left(\frac{\pi f}{\Delta t}\right), f < \frac{\Delta t}{2}$	$-$	
				$X_k = \sqrt{Q}\dot{U}_k\delta t$	$U_k = \sqrt{12}U_k^*, U_k^* \sim U(0, 1)$
Random Ramp	RR	c_{RR}	$S_X(f) = \frac{c_{RR}}{(2\pi f)^3}$	$\dot{X}_1(t) = X_2(t)$	$c_{RR} = X_2$
				$\dot{X}_2(t) = 0$	
				$X_{1,k+1} = X_{1,k} + X_{2,k}\cdot\Delta t$	
				$X_{2,k+1} = X_{2,k}$	

to the model building problem, which aims to find a plausible model to be estimated. In the filter design context, the tradeoff between the accuracy of the model with respect to the application, as well as the increased computational load involved by the augmented states, must be taken into account by the designer at this stage.

2) How can the parameters θ of the resulting model $F(\theta)$ be accurately estimated? This question tackles the model estimation problem, which becomes nontrivial when multiple processes are mixed. At this stage, a model estimation algorithm together with a cost function must be chosen in order to select a model from the set and estimate its parameters.

3) How can the model be validated? The model validation step is a very challenging task in inertial navigation for several reasons. First, only one realization of the trajectory is generally observed, making any statistical analysis difficult or even impossible. Second, influence of environmental conditions (e.g. temperature changes or vibrations) may largely affect the behavior of the sensors and thus change the underlying error dynamics. Finally, an inadequate set of observations $\{y_k\}$ and/or too many states in $P(\cdot)$ may result in a problem referred to as the observability problem. Although the system dynamics and the observation models should reflect the real navigation situation by including a representative number of states, the interrelationships within the dynamic model as well as the external observations may affect the observability of these states [6]. In our study, we validate models at observation level by comparing the noise structure between observed signals and synthetic signals generated under $F(\hat{\theta})$ (this can be seen as the cost function of Fig. 1), and at state level by analysing navigation

performance with respect to a reference when inertial navigation is operating in coasting mode.

Within the scope of our work, the problem of modeling and estimation focuses on the stochastic error components affecting accelerometers and gyroscopes. Further, we restrict the set of possible models $F(\cdot)$ to composite models that we define as linear combinations of independent stochastic processes which are usually used within the navigation filters: Gaussian white noise (WN), random walk (RW), random ramp (RR), quantization noise (QN), and especially first-order Gauss-Markov (GM) processes for approximating a bias instability process, and more generally, flicker noise. All these processes are defined in Table I. As already mentioned, issues like computational load or observability must be taken into account when designing an EKF. We believe that this set of model structure is fairly general to be used with a large class of sensors. In particular, combinations of multiple GM processes are of interest since they can approximate many random processes [8–10].

Although GM processes are very commonly used in inertial sensor models, the estimation of their parameters $\theta = \{\beta, \sigma_{GM}^2\}$ is a nontrivial task if the error structure is mixed with other types of noise. Such an estimation may be attempted by computing and interpreting power spectral density (PSD) or Allan variance (AV) plots [11]. The latter is probably the most commonly used method for model identification and sensor calibration, together with other metrics such as the Hadamard or the total variance [12–15]. The AV is a well-established method for identifying stochastic processes affecting the output of a sensor [3]. It can also be used to estimate the parameters

of some model that is believed to describe the stochastic processes of interest. Although this method was originally intended for studying the stability of oscillators, it has been successfully applied to problems dealing with different types of sensors, among which stands the modeling of inertial sensor errors [17–22]. The AV at scale τ (denoted as $\sigma_{\bar{Y}}^2(\tau)$) is defined as

$$\sigma_{\bar{Y}}^2(\tau) = \frac{1}{2}\mathbb{E}\left[\left(\bar{Y}_k(\tau) - \bar{Y}_{k-\tau}(\tau)\right)^2\right] \qquad (2)$$

where $\bar{Y}_k(\tau)$ is the sample average of τ consecutive observations, i.e., $\bar{Y}_k(\tau) = \frac{1}{\tau}\sum_{j=0}^{\tau-1} Y_{k-j}$. The AV can be expressed in the frequency domain through the relationship between $\sigma_{\bar{Y}}^2(\tau)$ and the PSD $S_Y(f)$ of the intrinsic processes [20] which links the parameter vector $\boldsymbol{\theta}$ to $\sigma_{\bar{Y}}^2(\tau)$. This relationship is due to the known form of the PSD function characterizing different noise processes and enables to express $\boldsymbol{\theta}$ as a function of $\sigma_{\bar{Y}}^2(\tau)$ (a detailed discussion on how to express this link can be found in [19]). In general, only the processes listed in Table I excepting GM are considered with the AV. These processes correspond to linear regions in a "$\hat{\sigma}_{\bar{y}}(\tau)$ vs. τ" log-log plot. Therefore, $\boldsymbol{\theta}$ is usually estimated by performing linear regressions with slope β of (visually) identified linear regions in such plots.[2] Unfortunately, this methodology suffers from severe drawbacks.

1) The AV method works reasonably well only for processes that are clearly identifiable and separable in the spectral domain and not subject to spectral ambiguity [23]. However, it does not allow to directly read out the parameters of a GM process as large values of β make this process similar to WN, while small values of β approximate an RW.

2) Inference about the estimated parameters is in many cases impossible. Indeed, the system parameters are indirectly estimated through functions of coefficients estimated by linear regression (say $\widehat{\boldsymbol{\beta}}$). The standard solution for deriving the (asymptotic) distribution of $\hat{\boldsymbol{\theta}}$ from the distribution $\widehat{\boldsymbol{\beta}}$ is achieved through a first-order approximation. In statistics, this approach is called the delta method (see [24] for details). In order to apply this method, it is required that the function $\boldsymbol{\beta} \mapsto \boldsymbol{\theta}$ is one-to-one and that $\widehat{\boldsymbol{\beta}}$ is a consistent estimator. Unfortunately, this is generally not the case here. Consequently, deriving the (asymptotic) distribution of $\hat{\boldsymbol{\theta}}$ is not possible in general.

3) The conventional AV methodology is limited to models composed of processes characterized by linear regions in a "$\hat{\sigma}_{\bar{y}}(\tau)$ v.s. τ" log-log plot and therefore this approach is far from being general.

4) In [25] we showed that the AV approach leads in general to inconsistent estimators for signals comprising at least two processes.

As an alternative to the AV approach, [26] proposed estimating $\hat{\boldsymbol{\theta}}$ by maximizing the log-likelihood of the state space model associated to $F(\cdot)$ using the EM algorithm [27]. This approach is more general than the AV approach and works very well with relatively simple models. Unfortunately, when the model complexity increases, this methodology becomes numerically challenging as it becomes very sensitive to the initial approximation of parameters and the convergence to global minimum is not guaranteed. In practice, composite models comprising GM processes are estimated through ad-hoc tuning, by using available sensor specifications, or by experience [28]. Recent investigations like [22] have studied this type of model and proposed a methodology in which GM processes are used to overbound the sensor error. But the success of this methodology was quite limited in our experience. For all these reasons, we propose employing the GMWM framework for such purpose. GMWM has been developed for the estimation of composite stochastic processes and is excepted from the previously mentioned inconveniences. The theoretical basis of this estimation approach was first introduced in [5] and its practical use validated and demonstrated by the use of simulation studies in [29]. This method relies on wavelet variances (WVs) which can be seen as a generalization of the AV.

III. THE GENERALIZED METHOD OF WAVELET MOMENTS

A. The Wavelet Variance

As pointed out by [30], the WV can be interpreted as the variance of a process after it has been subject to an approximate bandpass filter. The WV can be built using wavelet coefficients issued from a modified discrete wavelet transform (DWT) [31, 32] called the maximal overlap DWT (MODWT) [32, 33]. The wavelet coefficients are built using wavelet filters $\{\tilde{h}_{j,l} : j = 1, \cdots, J\}$ which for $j = 1$ and for the MODWT satisfy

$$\sum_{l=0}^{L_1-1} \tilde{h}_{1,l} = 0, \sum_{l=0}^{L_1-1} \tilde{h}_{1,l}^2 = \frac{1}{2} \text{ and } \sum_{l=-\infty}^{\infty} \tilde{h}_{1,l}\tilde{h}_{1,l+2m} = 0 \qquad (3)$$

where $\tilde{h}_{1,l} = 0$ for $l < 0$, $l \geq L_1$, L_1 is the length of $\tilde{h}_{1,l}$, m is a nonzero integer. Considering the transfer function of $\tilde{h}_{1,l}$ as

$$\tilde{H}_1(f) = \sum_{l=0}^{L_1-1} \tilde{h}_{1,l} e^{-i2\pi f l}, \qquad (4)$$

the jth level wavelet filters $\{\tilde{h}_{j,l}\}$ of length $L_j = (2^j - 1)(L_1 - 1) + 1$ can be obtained by computing the inverse discrete Fourier transform of

$$\tilde{H}_j(f) = \tilde{H}_1(2^{j-1}f)\prod_{l=0}^{j-2} e^{i2\pi 2^l f(L_1-1)}\tilde{H}_1(\tfrac{1}{2} - 2^l f). \qquad (5)$$

The MODWT filter is actually a rescaled version of the DWT filter $h_{j,l}$, i.e., $\tilde{h}_{j,l} = h_{j,l}/2^{j/2}$. Filtering an infinite

---

[2]Note that the PSD function can be used with the same methodology.

sequence $\{Y_k; k \in \mathbb{Z}\}$ using the filters $\{\tilde{h}_{j,k}\}$ yields the MODWT wavelet coefficients

$$\bar{W}_{j,k} = \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} Y_{k-l}, \ k \in \mathbb{Z}. \quad (6)$$

We define the WV at dyadic scales $\tau_j = 2^{j-1}$, as the variances of the $\{\bar{W}_{j,k}\}$ sequences, i.e.,

$$\nu(\tau_j) = \text{var}\left[\bar{W}_{j,k}\right]. \quad (7)$$

Note that the WVs are assumed not to depend on time. The condition for this property to hold is that the integration order d for the series $\{y_k\}$ to be stationary is such that $d \leq L_{1/2}$ and $\{\tilde{h}_{j,l}\}$ is based on a Daubechies wavelet filter (see [34] and [32, ch. 8]). This is due to the fact that Daubechies wavelet filters of width L_1 contain an embedded backward difference filter of order $L_{1/2}$. In such a case, the series of wavelet coefficients $\{\bar{W}_{j,k}\}$ is stationary with PSD $S_{W_j}(f) = |\tilde{H}_j(f)|^2 S_{F_\theta}(f)$, where $|\cdot|$ is denoting the modulus, and $F_\theta = F(\theta)$. This means that the variance of wavelet coefficient series is equal to the integral of its PSD [35], i.e.,

$$\nu(\tau_j) = \int_{-1/2}^{1/2} S_{W_j}(f) df = \int_{-1/2}^{1/2} |\tilde{H}_j(f)|^2 S_{F_\theta}(f) df. \quad (8)$$

Hence, there is an implicit link between the WV and the parameters of the data generating model $F(\theta)$. We exploit this connection when defining an estimator for θ, namely by matching a sample estimate of $\nu(\tau_j)$ together with the model-based expression of the WV given by (8). For WVs based on Haar wavelet filters (see (18) below) and for the processes considered in this work, the integral in (8) can be solved (see [5] for their expression), based on the results of [36]. WVs for other models can be computed using the same methodology.[3]

For a finite (observed) process $\{y_k : k = 1, \dots, N\}$, the MODWT WV estimator given by

$$\hat{\nu}(\tau_j) = \frac{1}{M_j} \sum_{k=L_j}^{N} W_{j,k}^2 \quad (9)$$

where $W_{j,k} = \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} y_{k-l}, k \in (L_j; N)$ and $M_j = N - L_j + 1$, is a consistent estimator for $\nu(\tau_j)$. With this respect, [35] also shows that under suitable conditions, $\sqrt{M_j} \left(\hat{\nu}(\tau_j) - \nu(\tau_j)\right)$ is asymptotically normal with mean 0 and variance

$$S_{W_j}(0) = 2 \int_{-1/2}^{1/2} S_{W_j}^2(f) df$$
$$= 2 \int_{-1/2}^{1/2} |\bar{H}_j(f)|^4 S_{F_\theta}^2(f) df. \quad (10)$$

Equation (10) can be estimated by means of

$$\hat{S}_{W_j}(0) = \sum_{\tau=-M_j}^{M_j} \left[\frac{1}{M_j} \sum_{k=L_j}^{N} W_{j,k} W_{j,k+|\tau|}\right]^2 \quad (11)$$

and the asymptotic properties of (11) are given in [32, p. 312]. These results were extended to the multivariate case in [5] where we demonstrated that under some regularity conditions, the asymptotic distribution of $\hat{\nu}$ is given by

$$\sqrt{N} \left(\hat{\nu} - \mathbb{E}\{\boldsymbol{\nu}\}\right) \overset{\mathcal{D}}{\underset{N\to\infty}{\mapsto}} N\left(0, \mathbf{V}_{\hat{\nu}}\right) \quad (12)$$

where

$$\boldsymbol{\nu} = \left[\nu(\tau_j)\right]_{j=1,\dots,J} \quad \text{and} \quad \hat{\boldsymbol{\nu}} = \left[\hat{\nu}(\tau_j)\right]_{j=1,\dots,J}. \quad (13)$$

The matrix $\mathbf{V}_{\hat{\nu}}$ has size $(J \times J)$ and is given by

$$\mathbf{V}_{\hat{\nu}} = \begin{bmatrix} \sigma_{1,1}^2 & \cdots & \sigma_{1,J}^2 \\ \vdots & \ddots & \vdots \\ \sigma_{J,1}^2 & \cdots & \sigma_{J,J}^2 \end{bmatrix}. \quad (14)$$

The elements of $\mathbf{V}_{\hat{\nu}}$ can be obtained through

$$\sigma_{ml}^2 = 2\pi S_{W_m W_l}(0), \quad \text{for } m, l = 1, \dots, J \quad (15)$$

where

$$S_{W_m W_l}(f) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} \rho_{W_m W_l}(\tau) e^{-if\tau} \quad (16)$$

and $\rho_{W_m W_l}(\tau) = \text{cov}\left[W_{m,k}, W_{l,k+\tau}\right]$ is the cross-covariance function. The estimation of σ_{ml}^2 is in general not straightforward. In [5], we show that under the assumption of a Gaussian process for $\{y_k\}$, a suitable estimator is given by

$$\hat{\sigma}_{ml}^2 = \frac{1}{2} \sum_{\tau=-M(T_{ml})}^{M(T_{ml})} \left[\frac{1}{M(T_{ml})} \sum_{k\in T_{ml}} W_{m,k} W_{l,k+\tau}\right]^2$$
$$+ \frac{1}{2} \sum_{\tau=-M(T_{ml})}^{M(T_{ml})} \left[\frac{1}{M(T_{ml})} \sum_{k\in T_{ml}} W_{m,k-\tau} W_{l,k}\right]^2 \quad (17)$$

where T_{ml} is the smallest set of time indices containing both the indices in T_m and T_l [see (9)], and $M(T_{ml})$ their number. Alternatively, when the process is not Gaussian or when the sample size is very large as is the case with the dataset analysed in [29] for which the computation of (17) was infeasible, one can use a parametric bootstrap to estimate $\text{cov}[\hat{\nu}(\tau_m), \hat{\nu}(\tau_l)]$. In such a scheme Q samples of size N are simulated from $F(\hat{\theta})$ on which Q sequences, $\{\hat{\nu}_q(\tau_m)\}$ and $\{\hat{\nu}_q(\tau_l)\}$ for $q = 1, \dots, Q$ are computed, and σ_{ml}^2 are estimated by their empirical covariance.

A particular choice for the wavelet filter is given by the Haar wavelet filter which first DWT filter ($j = 1$) is

$$\{h_{1,0} = 1/\sqrt{2}, \ h_{1,1} = -1/\sqrt{2}\} \quad (18)$$

with length $L_1 = 2$. If the process is stationary with backward differences of order $d > 1$ one can use other

---

[3]However, as shown later in Section III-C, the possibility of expressing these error models analyticaly in the space of WVs is not a necessary prerequisite for applying the new estimator.

wavelet filters such as Daubechies wavelet filters [34]. When the WV is evaluated with Haar wavelet filters, it is actually equal to half the AV.

B. GMWM Estimator

We saw in (8) that the variance of a wavelet coefficient sequence is equal to the integral of its PSD. Therefore, there exists a mapping

$$\theta \mapsto \nu(\theta). \qquad (19)$$

In some sense, the GMWM approach inverses this mapping by trying to approximate the point $\theta(\hat{\nu})$ where $\hat{\nu}$ is the WV estimated from the observed sample using MODWT estimator. In other words, the GMWM aims to find the value of θ implied by the empirical WV $\hat{\nu}$. For that, some discrepancy measure between the empirical WV and the ones implied by the model F_θ has to be minimised, as is done e.g. with the generalised method of moments [2].

Therefore, we propose to follow [1] and estimate the model's parameters using an estimator which combines on the one hand the WV and on the other hand the generalized least-squares (GLS) principle, using the relationship given in (8). More precisely, we propose to find $\hat{\theta}$ such that the WV implied by the model, say $\phi(\theta)$, matches the empirical WV, say $\hat{\phi}$, and solves the following GLS optimization problem:

$$\hat{\theta} = \underset{\theta \in \Theta}{\operatorname{argmin}} \ \left(\hat{\phi} - \phi(\theta)\right)^T \Omega \left(\hat{\phi} - \phi(\theta)\right) \qquad (20)$$

in which Ω, a positive definite weighting matrix,[4] is chosen in a suitable manner (see below). Equation (20) defines the GMWM estimator. The vector $\phi(\cdot) = [\phi_j(\cdot)]_{j=1,\ldots,J}$ is a binding function between θ and ν such that $\phi(\theta) = \nu$, and $\hat{\phi} = \hat{\nu}$ and $\phi(\hat{\theta})$ are two estimators.

As already mentioned, we proved in [5] that $\hat{\theta}$ consistent for composite processes composed of the sum of one or more but only one of each WN, QN, RW, RR and $k < \infty$ AR(1) (i.e. GM) processes. Hence, $\hat{\theta}$ has the following distribution:

$$\sqrt{N}\left(\hat{\theta} - \theta\right) \underset{N \to \infty}{\overset{\mathcal{D}}{\mapsto}} \mathcal{N}\left(0, V_{\hat{\theta}}\right) \qquad (21)$$

where

$$V_{\hat{\theta}} = B V_{\hat{\phi}} B^T \qquad (22)$$

and

$$B = \left(D^T \Omega D\right)^{-1} D^T \Omega. \qquad (23)$$

The matrix D is given by

$$D = \frac{\partial \phi(\theta)}{\partial \theta^T} \qquad (24)$$

___
[4] Ω has to be positive definite in order to ensure the convexity of (20).

and $V_{\hat{\phi}} = V_{\hat{\nu}}$. When $\Omega = I$, then

$$V_{\hat{\theta}} = \left(D^T D\right)^{-1} D^T V_{\hat{\phi}} D \left(D^T D\right)^{-1}. \qquad (25)$$

The most efficient estimator is obtained by choosing $\Omega = V_{\hat{\phi}}^{-1}$, leading then to

$$V_{\hat{\theta}} = \left(D^T V_{\hat{\phi}}^{-1} D\right)^{-1}. \qquad (26)$$

In practice, the matrix D is computed at $\hat{\theta}$.

Obviously, the number of scales J should be $J \geq p$, with p the dimension of θ, but at the same time, a too large J introduces variability in the estimator. In [37], we propose a method for selecting the scales J that minimize the determinant of $V_{\hat{\theta}}$ and propose a method to remove the finite sample bias of $\hat{\theta}$ using a simulation-based approach. Note that when $J > p$, i.e., the number of WV is greater than the dimension of the parameter vector θ, the goodness-of-fit of the model $F(\theta)$ to the data can be assessed by testing the hypotheses $H_0 : \mathbb{E}\left[\hat{\phi} - \phi(\theta)\right] = 0$, $H_1 : \mathbb{E}\left[\hat{\phi} - \phi(\theta)\right] \neq 0$ using the χ^2-test statistic

$$N \left(\hat{\phi} - \phi(\hat{\theta})\right)^T V_{\hat{\phi}}^{-1} \left(\hat{\phi} - \phi(\hat{\theta})\right) \underset{N \to \infty}{\overset{\mathcal{D}}{\mapsto}} \chi^2_{J-p} \qquad (27)$$

under H_0 (see [38]) and provided that $p < J < \infty$.

Alternatively, a model selection criteria was developed specifically for the GMWM in [39] (see also [40, sect. 3.5.2]). The connection between model selection and navigation performance are left for further research.

C. From the Generalized Method of Moments to Indirect Inference

The analytical expressions of $\nu(\theta)$ used in (20) using the Haar wavelet filter defined in (18) can be computed for several well-known models such as the ones listed in Table I, as well as AR($\cdot$), sums of AR($\cdot$), ARMA ($\cdot$,$\cdot$) and others [5] using the general results of [36] on the AV. In addition, the analytical WV of sums of independent processes corresponds to the sum of the WV of individual processes within the model. Indeed, for composite processes, i.e., $Y_k = \sum_m (Y_m)_k$, (8) can be expanded to

$$\nu(\tau_j) = \int_{-1/2}^{1/2} |\tilde{H}_j(f)|^2 \left(\sum_m S_{Y_m}(f)\right) df = \sum_m \nu_m(\tau_j) \qquad (28)$$

with $S_{Y_m}(\cdot)$ the PSD and $\nu_m(\tau_j)$ the WV at scale τ_j of $\{(Y_m)_k\}$. Therefore, when an analytical expression for $\nu(\theta)$ is available, the estimator defined in (20) can be seen as a generalized method of moments estimator [38].

As a possible extension of the GMWM when analytical expressions for $\phi(\theta)$ in (20) are too complicated to compute, one can resort to simulations to compute $\phi(\theta)$ and hence place the GMWM in the framework of indirect inference [41–43]. Basically, given a sample of observations $\{y_k : k = 1, \ldots, N\}$ and a hypothetical model $F(\theta)$, we can define $\hat{\phi}_j$ as the WV $\hat{\nu}(\tau_j)$ estimated

from the sample using (9), and $\hat{\phi}_j^\star(\boldsymbol{\theta})$ as the WV estimate $\hat{\nu}_\star(\tau_j)$ computed on a simulated series

$$\left\{ y_k^\star(\boldsymbol{\theta}) : k = 1, \ldots, R \cdot N \right\}, R \geq 1 \qquad (29)$$

from $F(\boldsymbol{\theta})$. Then $\widehat{\boldsymbol{\phi}} = [\hat{\phi}_j]_{j=1,\cdots,J}$ and $\widehat{\boldsymbol{\phi}^\star}(\boldsymbol{\theta}) = [\hat{\phi}_j^\star(\boldsymbol{\theta})]_{j=1,\cdots,J}$ are used in (20) to obtain an estimate $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$, which properties are described in e.g. [43]. In particular, for R sufficiently large, $\mathbf{V}_{\hat{\boldsymbol{\theta}}} \approx \mathbf{B}\mathbf{V}_{\hat{\boldsymbol{\phi}}}\mathbf{B}^T$. In that case, $\mathbf{B}$ can be computed numerically.

D. Inference on PSD vs. Wavelet Variance

Note that one could wonder why we use WVs instead of making inference directly on the PSD for which a similar estimation procedure could be used. We believe that the wavelet-based approach is more suitable for the following reasons that we also mentioned in [5].

1) The computation of empirical WV is more straightforward than nonparametric PSD. For example, the periodogram is an inconsistent estimator of $S_Y(\cdot)$ and can be badly biased even for large samples sizes (because of the frequency leakage effects). Therefore, more sophisticated PSD estimators and/or smoothing techniques such as prewhitening or tapering will be employed with a hope to approach the consistency provided by the GMWM estimator.

2) Inference on the PSD would make the optimization of a least-squares-type measure (between the empirical and model-based PSD) more difficult to solve when the PSD has large variability over very narrow frequency bands. As shown in [32], the wavelet coefficient at scale τ_j is associated with frequencies in the interval $[1/2^{j+1}, 1/2^j]$ and (8) can be approximated by

$$\nu^2(\tau_j) \approx 2 \int_{1/2^{j+1}}^{1/2^j} S_Y(f) df. \qquad (30)$$

This means that the WV summarizes the information in the PSD using just one value per octave frequency band. This property is particularly useful when the PSD is relatively featureless within each octave band. In the case of the widely used pure power law processes ($S_Y(f) \propto |f|^\alpha$) for example, from (30) one gets $\nu(\tau_j) \propto \tau_j^{-\alpha-1}$, meaning that no information is lost when using the "PSD summary" given by the WV. According to [32], as the width L of the wavelet filter $\{\tilde{h}_l\}$ used to form $\{\tilde{h}_{j,l}\}$ increases, the approximation expressed in (30) improves because $\{\tilde{h}_{j,l}\}$ then becomes a better approximation to an ideal bandpass filter. Therefore, by making the assumption that L is chosen such that (30) is a reasonable approximation, it is possible to estimate $S_Y(\cdot)$ using a function $\bar{S}_Y(\cdot)$ that is piecewise constant over octave bands $\left[\frac{1}{2^{j+1}\Delta t}, \frac{1}{2^j\Delta t}\right]$ for $j = 1, \ldots, J$. Then, when $\frac{1}{2^{j+1}\Delta t} < f \leq \frac{1}{2^j\Delta t}$, we assume
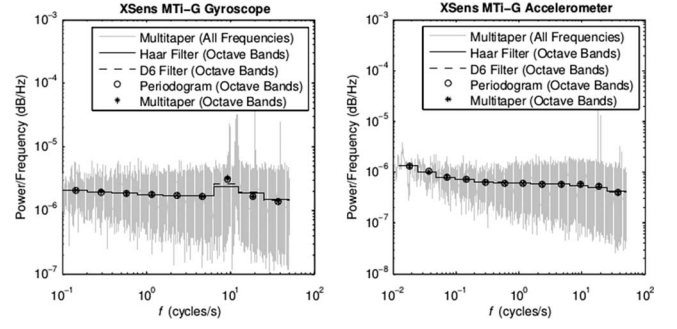
$$\bar{S}_Y(f) = C_j \qquad (31)$$

Fig. 2. Comparison of octave band PSD estimates for the XSens MTi-G gyroscope (left panel) and accelerometer (right panel) error signal based on the periodogram, a multitaper PSD estimate and Haar and D(6) WV estimates.

where C_j is a constant defined such that [32]

$$\int_{1/(2^{j+1}\Delta t)}^{1/(2^j\Delta t)} S_Y(f) df = \int_{1/(2^{j+1}\Delta t)}^{1/(2^j\Delta t)} \bar{S}_Y(f) df = \frac{C_j}{2^{j+1}\Delta t}. \qquad (32)$$

From (30), we have

$$\nu(\tau_j) \approx \frac{C_j}{2^j \Delta t} \qquad (33)$$

and hence we can use $\hat{C}_j = 2^j \hat{\nu}(\tau_j) \Delta t$ to estimate PSD levels. The panels of Fig. 2 show estimated PSD levels $\hat{C}_j$ plotted as a constant line over each octave band for an XSens MTi-G MEMS-based gyroscope and accelerometer error signal (we study these sensors later in Section V), respectively. The full line corresponds to the Haar-based estimate, while the dashed line represents the estimate using a higher order Daubechies wavelet filter (D6). In this case, there is good agreement between the two PSD estimates over all frequencies, meaning that the use of the Haar filter is sufficient. The well-known frequency leakage phenomenon was also mentioned in the previous point. Therefore, we also plotted PSD estimates based upon the leakage-prone periodograph (circles) and a multitaper estimator (asterisks), which should be relatively free of leakage (for more details on periodogram and multitaper PSD estimators, see [32]). In order to obtain the periodogram and multitaper PSD estimates, we averaged the PSD estimates obtained at each Fourier frequency $f_k = \frac{k}{N\Delta t}$, $k = 1, 2, \cdots$ (gray line corresponding to the multitaper PSD estimate) over octave frequency bands, meaning that each of the octave band averaged estimates is plotted versus the average of the Fourier frequencies associated with the estimates. In this case, the Haar-based PSD estimate does not suffer from frequency leakage, which should be visible at the low power portion of the estimated PSD (at high frequencies).

3) The PSD of two important processes in sensor error models, namely the RR (drift) and the RW, cannot be distinguished (both have slope of –2 in a log-log representation of the PSD).

4) The MODWT on which the WV computation is based requires a number of multiplications of order $N \log_2 N$, which is the same order as the widely used fast

Fourier transform algorithm. For this reason, the employment of WV does not increase the computational burden.

IV. IMPLEMENTATION

A. Optimizer Initialization

The core of the GMWM estimation is in solving the optimization problem defined by (20). Sensitivity to initial values at which the optimizer starts is therefore a classical issue. We shortly describe the implemented strategy to overcome numerical issues related to convergence. The GMWM minimization problem can be geometrically interpreted as a minimization of the distance between the $\widehat{\boldsymbol{\phi}}$ and $\boldsymbol{\phi}(\hat{\boldsymbol{\theta}})$ curves, considering weights $\boldsymbol{\Omega}$. However in cases where $\widehat{\boldsymbol{\phi}}$ has strong variations over all scales τ, the optimizer rapidly converges and matches well the left part of the curve $\boldsymbol{\phi}$. This is due to the weighting matrix $\boldsymbol{\Omega}$, which puts more importance on WV at small scales. A simple geometrical reformulation of (20)

$$\widehat{\boldsymbol{\theta}}^{(0)} = \underset{\boldsymbol{\theta} \in \boldsymbol{\Theta}}{\operatorname{argmin}} \left(1 - \frac{\boldsymbol{\phi}(\boldsymbol{\theta})}{\widehat{\boldsymbol{\phi}}}\right)^T \left(1 - \frac{\boldsymbol{\phi}(\boldsymbol{\theta})}{\widehat{\boldsymbol{\phi}}}\right) \qquad (34)$$

"flattens" the curves $\boldsymbol{\phi}(\cdot)$ and $\widehat{\boldsymbol{\phi}}$ and significantly decreases the convergence time. Note that during initialization, $\boldsymbol{\Omega}$ can safely be omitted. Using such strategy, only a few iterations are necessary to solve the original problem of (20).

B. Note on the Estimation of the Wavelet Variance Covariance

Similarly to [29], we present a small simulation study to demonstrate how "far" the matrix $\mathbf{V}_{\hat{v}}$ is in reality from $\sigma_\epsilon^2 \mathbf{I}$ with σ_ϵ^2 the uncertainty in the linear regression performed in the log-log "$\sigma_{\bar{y}}(\tau)$ vs. τ" plot which is a common assumption in the standard AV methodology. Indeed, B sequences of WN processes, say $\{w_k^{(i)} : k = 1, \dots, 1000, \ i = 1, \dots, B\}$ with unit variance were generated. The WV of each process, denoted as $\hat{v}_{(i)}$, was estimated. Then, the empirical covariances (and correlations) of $\hat{v}$ were computed using

$$\widehat{\boldsymbol{\Sigma}}_B = \frac{1}{B-1} \sum_{i=1}^{B} \left(\hat{\boldsymbol{v}}_{(i)} - \bar{\boldsymbol{v}}\right) \left(\hat{\boldsymbol{v}}_{(i)} - \bar{\boldsymbol{v}}\right)^T \qquad (35)$$

where $\bar{\boldsymbol{v}}$ is the mean of the $\hat{\boldsymbol{v}}_{(i)}$. For large B, the matrix $\widehat{\boldsymbol{\Sigma}}_B$ is a fairly good approximation of $\mathbf{V}_{\hat{v}}$ and corresponds to the (parametric) bootstrap estimator of $[\hat{\boldsymbol{v}}]$. Fig. 3 shows the variances of $\hat{\boldsymbol{v}}$ (i.e., the diagonal of $\mathbf{V}_{\hat{v}}$) estimated using the asymptotic formula version defined in (17) and the bootstrap estimator $\widehat{\boldsymbol{\Sigma}}_B$ of (35). The variance (i.e., $\hat{\sigma}_\epsilon^2$) implied by the model $\sigma_\epsilon^2 \mathbf{I}$ is also depicted. As expected, the bootstrap and asymptotic estimators are very close and practically demonstrate, at least in this example, the validity of (17). In addition, these two estimates are very far from $\hat{\sigma}_\epsilon^2$ which illustrates the fact that the standard regression approach used in the AV methodology is unsuitable in terms of efficiency.
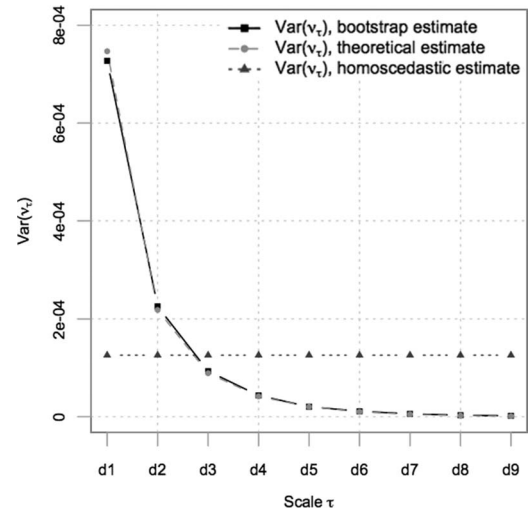

Fig. 3. Estimated variances of $\hat{v}$ with bootstrap (with $B = 1000$) and asymptotic (theoretical) approaches as well as variances implied by homoscedastic model (adapted from [29]).
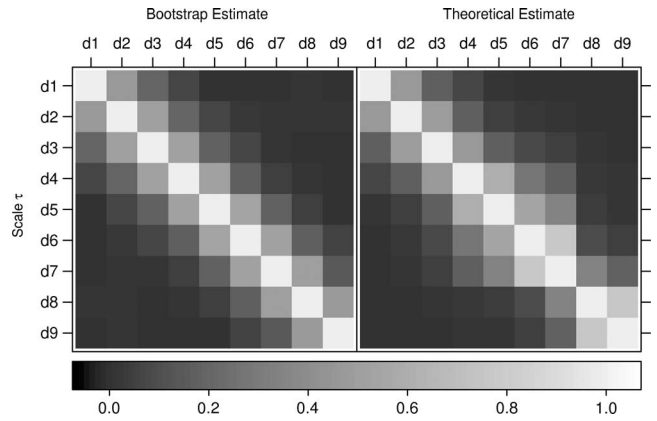

Fig. 4. Comparison between WV correlations estimated using bootstrap approach with $B = 1000$, on the left part, and theoretical correlations computed using (17), on the right part (adapted from [29]).

Fig. 4 presents the correlations estimated with the bootstrap and the asymptotic estimators, which are also very close. Note also that the WVs are highly correlated between neighboring scales. This confirms the inadequacy of the implicit model $\sigma_\epsilon^2 \mathbf{I}$.

V. STUDY OF STOCHASTIC ERRORS IN MEMS-BASED INERTIAL SENSORS

In [29], we first demonstrated (mainly through simulations) that the GMWM estimator is able to handle complex error models for which the AV-based estimation failed and the EM algorithm did not converge. In this section, we further investigate its applicability in the context of the design of an INS mainly based on low-cost MEMS sensors that are widely spread in many aerospace as well as terrestrial applications.

A. Error Signal Construction

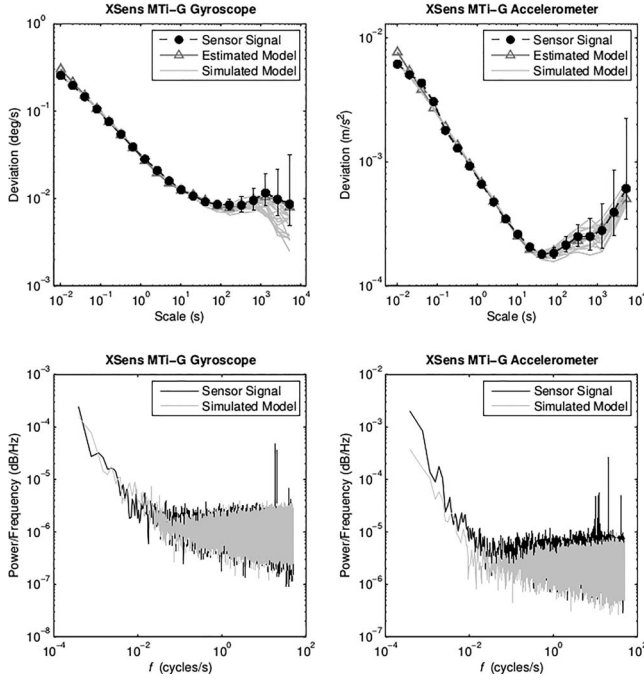Static data were collected from a MEMS-based IMU (XSens MTi-G) sampling at 100 Hz during 4 hr. The

Fig. 5. Result of the XSens MTi-G gyroscope and accelerometer error modeling using (36) with $M \in (2, 3)$, respectively. Upper panels depict WV of X-axis gyroscope (left panel) and accelerometer (right panel) error signal (black circles) with associated 95% confidence interval, WV resulting from estimated model (gray triangles), and WV of 20 synthetic signals issued from estimated model (gray lines). Black lines in lower panels show PSD of gyroscope (left panel) and accelerometer (right panel) error signals, and gray lines correspond to PSD of estimated models.

TABLE II
Result of the Gyroscope and Accelerometer Error Model Estimation by Means of the GMWM using (36) with $M \in (2, 3)$, respectively

θ	Gyroscope		Accelerometers	
	Value	Units	Value	Units
$\hat{\sigma}_{WN}$	0.032	$deg/s/\sqrt{Hz}$	85.5	$\mu g\sqrt{Hz}$
$\hat{\beta}_1$	0.024	Hz	25.8	Hz
$\hat{\sigma}_{GM,1}$	$7.63 \cdot 10^{-4}$	$deg/s/\sqrt{Hz}$	143	$\mu g/\sqrt{Hz}$
$\hat{\beta}_2$	$9.03 \cdot 10^{-4}$	Hz	0.04	Hz
$\hat{\sigma}_{GM,2}$	$8.64 \cdot 10^{-4}$	$deg/s/\sqrt{Hz}$‘	103	$\mu g/\sqrt{Hz}$
$\hat{\beta}_3$	–	–	$2.32 \cdot 10^{-4}$	Hz
$\hat{\sigma}_{GM,3}$	–	–	900	$\mu g/\sqrt{Hz}$

where M was determined empirically by going several times through the filter design loop (see Fig. 1). In each loop, the value of M was changed and the final model was accepted as a compromise between the model complexity and the matching quality (the goodness-of-fit). In this case, a value of $M = 2$ was chosen for the gyroscope, and $M = 3$ for the accelerometer model, respectively. Hence, we have the following parameters to estimate:

$$\boldsymbol{\theta} = \left\{ \sigma_{WN}^2, \beta_m, \sigma_{GM,m}^2 \right\}_{m=1,\ldots,M} \tag{37}$$

from the observed gyroscope and accelerometer error signal $\{y_k\}$.

C. Model Estimation

The estimated parameter values for the gyroscope and accelerometer error models are listed in Table II. The quality of the fit can be judged with the gray line with triangle markers in the upper left panel of Fig. 5. Despite the slight deviation from the first WV point of the signal with respect to the model, the estimated model still corresponds well to the observed signal. Note that despite the difficulty to compare models, the angular RW provided by the manufacturer is 0.05 deg/s/$\sqrt{Hz}$ [44]. For the accelerometer, the velocity RW PSD level indicated by the manufacturer is 0.002 m/s^2/$\sqrt{Hz}$ [44], which is slightly higher than our estimated $\hat{\sigma}_{WN}$ value of $8.55 \cdot 10^{-4}$ m/s^2/$\sqrt{Hz}$. Similarly to the gyroscope, the short-term noise structure is quite complex and none of the considered models in this work could easily approximate this.

D. Model Validation

In Fig. 5, the PSD of one realization issued from $F(\hat{\boldsymbol{\theta}})$ is depicted as a gray line superposed to the signal PSD (black line) for both, the MTi-G gyroscope (lower left panel) and the MTi-G accelerometer (lower right panel). The quality of the matching shows that the spectral noise structure generated by both models correctly reproduces the sensors' noise shapes.

Until now, we assumed a unique model for all three sensors composing an IMU. To verify this assumption, we modeled separately each of the three gyroscopes and accelerometers composing the MTi-G unit. Fig. 6 depicts
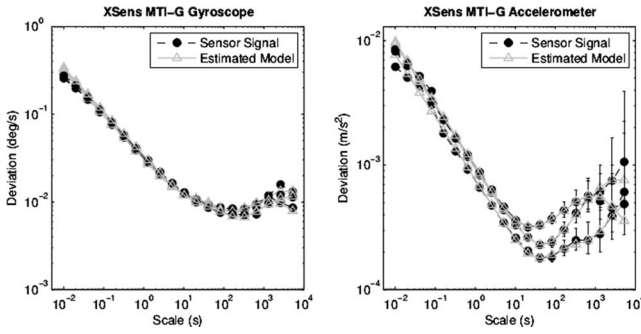
device operated under constant temperature and stationary conditions. After mean removal, the observed sequence $\{y_k : k = 1, \ldots, N\}$ with $N = 2'057'424'$ samples at hand contain measurement errors driven by stochastic processes issued from an unknown model $F(\boldsymbol{\theta})$. First, the GMWM estimator is applied on accelerometer and gyroscope error signals. Second, the obtained models are validated by simulating twenty realizations under $\hat{\boldsymbol{\theta}}$, computing their Haar WV, and comparing them to the Haar WV of the signal under study. Note that exactly the same conclusion could be drawn when comparing the AV curves because of the linear relationship between the Haar WV and the AV. For each error signal under study, we also mention the result obtained by the commonly used AV calibration procedure in order to demonstrate its limits.

B. Model Building

Computation of the Haar WV (or AV) curve on one (X-axis) gyroscope error signal (see black curve in the upper left panel of Fig. 5) and one (X-axis) accelerometer error signal (see upper right panel of Fig. 5) reveals the presence of correlated and uncorrelated noise. For both signals, we use the following composite process $F(\boldsymbol{\theta})$:

$$Y_k = \sum_{m=1}^{M} (Y_{GM,m})_k + W_k, \quad \text{for } k \in \mathbb{Z}. \tag{36}$$

Fig. 6. Result of GMWM modeling for XSens MTi-G accelerometer and gyroscope error signals using (36) with M ∈ (2, 3) for gyroscopes and accelerometers, respectively. For each IMU axis, WV sequences estimated on original error signals are drawn as black circles (together with 95% confidence intervals). WV sequences resulting from estimated models are drawn as gray lines with triangles. Note that confidence intervals associated to model-based WV are not drawn for readability reasons.

the three Haar WV curves corresponding to the error signals of the gyroscopes (left panel) and the accelerometers (right panel) from the MTi-G as black lines. In the same figure, the gray lines represent the Haar WV issued from the GMWM estimation. Since the error structure is similar among each sensor type, the use of one single model per sensor type in the filter design is relevant in this case. Nevertheless, its parameters should be estimated separately.

VI. INFLUENCE ON THE EKF SOLUTION

A. Testing Models

As already mentioned in Section II, validating and comparing different models on the state level basis is extremely tricky. Issues like observability, nonmodeled parasite signals generated from the dynamics, or variations of the environmental conditions may affect the error signal structure as well as the output of the EKF. Therefore, a statement telling that one model is in general more accurate than another must be considered with caution. In this work, we chose to construct models on signals sampled in static conditions and to validate them using tools such as WV (or AV) and PSD on the observation level. This means that the built model is valid for the conditions at hand during calibration.

A navigation filter designer may be interested in analysing the effect of a constructed model on a trajectory, independently from any unmodeled effects that were not experienced during the calibration phase. In this way, the model can be tested under the chosen dynamics and typical operation duration. Although the GMWM estimator opens the door to the estimation of complex composite models such as sums of GM processes, their parameters are not necessarily observable under all dynamical conditions. Moreover, the duration of the trajectory may be too short to estimate some processes with large correlation time. These two reasons may justify

choosing a simplified version of the model, i.e., removing some of the GM processes, or lumping them together.

Simulation provides a mean for performing such an analysis. In this respect, we use the following procedure.

1) An accurate navigation solution for a given trajectory is computed using signals from high-grade sensors (typically tactical/navigation-grade IMUs, L1/L2 carrier-phase differential GNSS positioning).

2) The real error signals acquired by the sensor under static conditions are added to synthetic inertial signals emulated along the reference trajectory (note that the synthetic signal matches the chosen reference perfectly; hence it contains no errors at all).

3) Artificial outages in GNSS position/velocity observations are added to the dataset, which is subsequently processed by INS/GNSS integration implementing a closed-loop EKF.

4) The quality of the model is judged by analysing the actual navigation error as well as the EKF-predicted accuracy during inertial coasting mode (i.e., periods with no external aiding).

B. Example: Low-Cost MEMS-Based IMU Operating on a Flying Platform

In this example, we assume the XSens MTi-G IMU, for which a model was constructed in Section V, to operate on a light flying platform such as an unmanned air vehicle. The target application is mapping via remote sensing sensors which must be accurately georeferenced using the integrated inertial and differential phase and code GNSS observations. Typical mission durations are between 5 to 30 min. We are interested in studying the behavior of the EKF operating in the aimed context with the given set of navigation sensors used for georeferencing.

The left panel of Fig. 7 shows an extract of a trajectory issued from a helicopter flight performing airborne laser scanning [45]. There, the laser data were georeferenced using a trajectory obtained by integrating observations from a tactical-grade Litton LN-200 IMU and a Javad Legacy L1/L2 GNSS receiver (rover). The centimeter-level accurate GNSS solution was obtained by carrier-phase differential postprocessing of the rover observation with a base GNSS receiver (Topcon Hiper Pro). Finally, the smoothed integrated navigation solution provided the trajectory that will serve as a reference as well as a base for emulation. We then emulated synthetic specific force and angular rate observations along this trajectory, and corrupted them with the real error signal observed in static conditions (see Section V).

We are now interested in comparing subversions of the constructed model expressed by (36).

Model 1: $Y_k = W_k$, $k \in \mathbb{Z}$. This model is computationally the most efficient but does not account for any correlated errors.

Model 2: $Y_k = \sum_{m=1}^{M}(Y_{GM,m})_k + W_k$, $k \in \mathbb{Z}$, $m = 1$, $\cdots$, M with $M = 1$ for the gyroscopes, and $M = 2$ for the accelerometers. This model accounts for correlated errors
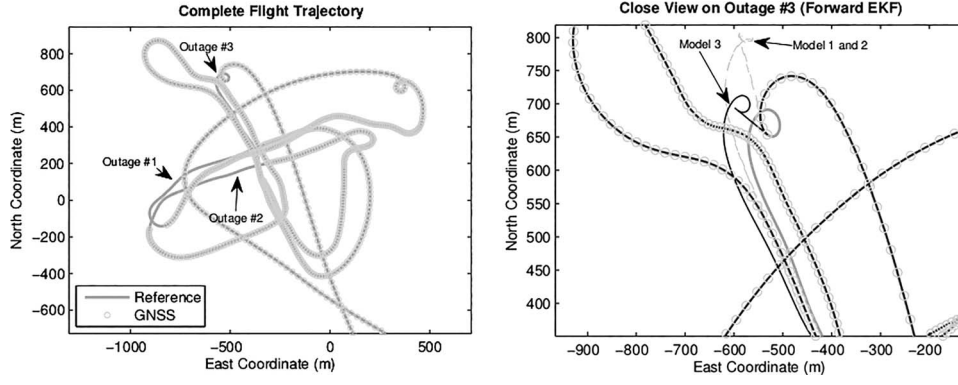
Fig. 7. Emulated trajectory issued from ALS flight. Left panel shows reference trajectory with three artificial GNSS-free periods. Right panel provides close view on third GNSS-free period, together with estimated trajectory using model 3.
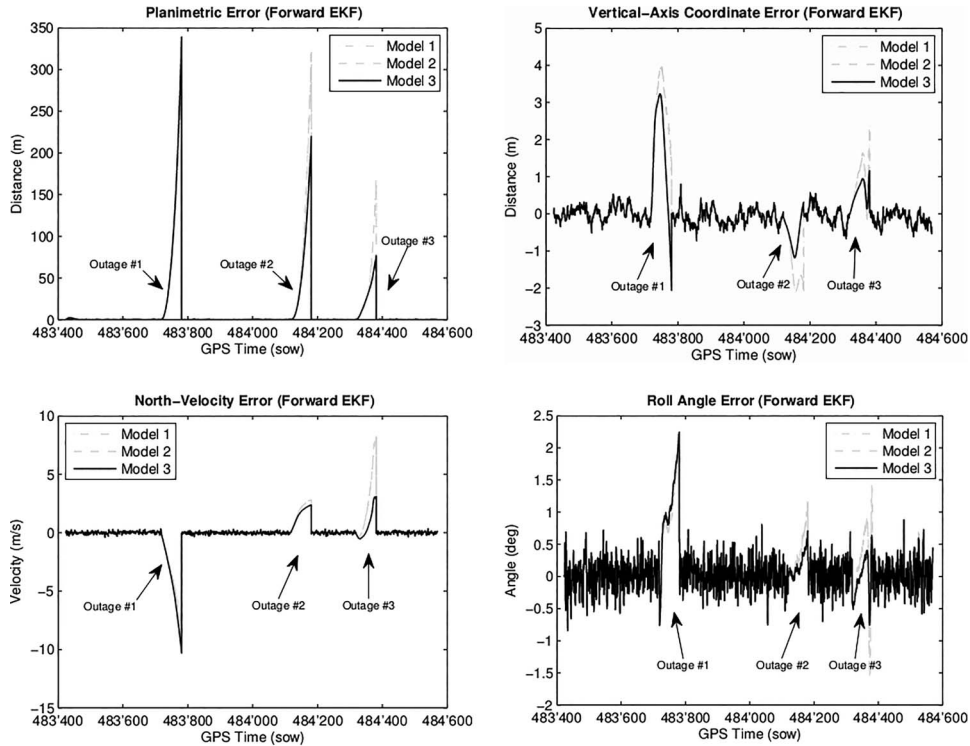

Fig. 8. Navigation performance achieved by XSens MTi-G device operated on emulated trajectory with three GNSS-free periods. Upper panels show planimetric (left) and altimetric (right) positioning errors, while lower panels depict North-axis velocity error (left) and roll angle error (right) for three tested models.

with the correlation time of ~ 40 s for the gyroscopes, and ~ 25 s for the accelerometers.

Model 3: This is a complete model, i.e., with $M = 2$ for the gyroscopes and $M = 3$ for the accelerometers. This model accounts for errors with short and long correlation times, but is computationally more demanding.

Note that the parameters are estimated independently by GMWM for each model structure. Hence, to some extent, the model suboptimality may get absorbed due to the overfitting, e.g., the variance of WN in model 1 is $\hat{\sigma}_{WN} = 0.052 \ deg/s/\sqrt{\text{Hz}}$ for gyroscopes and $91.6 \ \mu g/\sqrt{\text{Hz}}$, respectively, which is larger than in model 3.

The different models with their parameters as estimated by the GMWM become the augmented part of the EKF state vector. We introduce three artificial GNSS-free periods, each of 60 s duration. The first outage occurs approximately 7 min after mission start, the second after 11 min, and the third after 15 min. The error growth during each outage can be visualized in Fig. 8. Although not clearly visible in the figure, a general view reveals that model 2 performs only slightly better than model 1. All models perform similarly during the first outage, meaning that some of the model parameters could not yet be correctly estimated at that time. However, after being longer in mission, model 3 clearly outperforms the two
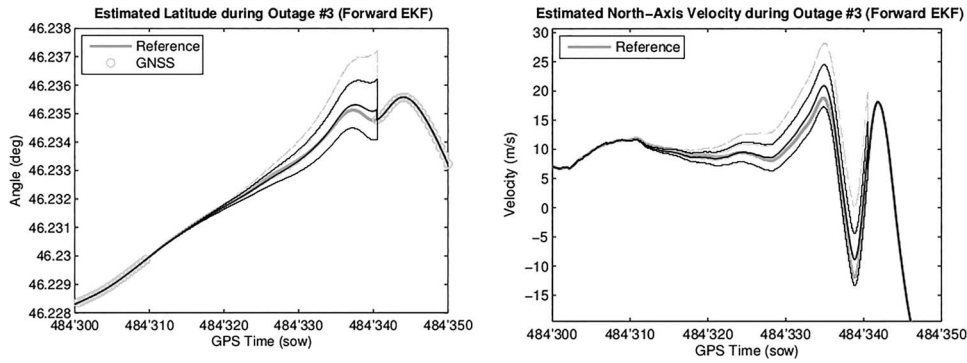
Fig. 9. Estimated latitude (left panel) and North-axis velocity (right panel) using three subversions of constructed model, together with estimated 3-σ confidence levels.

others. Regarding planimetric positioning error (upper left panel), the improvement of model 3 compared with model 1 (and 2) is more than 30% for the second outage, and raises to around 60% at the third outage. The right panel of Fig. 7 shows a close view on the trajectory during this third GNSS-free period. With respect to altimetric error, the upper right panel of Fig. 8 indicates that model 3 always performed significantly better than models 1 and 2. The lower left and right panels show the North-axis velocity (East and vertical velocity have similar behavior) and the roll angle errors (pitch and heading angles have similar behavior), respectively. The same conclusion as for the planimetry can also be drawn for the accuracy of velocity and attitude.

Besides navigation, the committed errors in the EKF-estimated covariance matrix $\mathbf{P}_k, ; k \in \mathbb{Z}$ need to be evaluated for each model. Fig. 9 depicts the true latitude error and North-axis velocity error, considering that similar conclusions can be made from the other state variables. The estimated state variables during the third outage period are depicted as thick black (model 3) and light gray (models 1 and 2) lines, surrounded by their estimated 3-σ error bounds (thiner lines with same color) extracted from the diagonal of $\mathbf{P}_k$. The true state is drawn as the thickest gray line. From both panels it can be seen that the estimated precision encompasses the true error only for the complete model, i.e., model 3, while it is clearly underestimated with the two other models.

At this stage, the filter can be designed with the following recommendations.

1) The first GM process, i.e. $(Y_{GM,1})_k$, can safely be neglected, since it brings no significant improvement over model 1.

2) When model 1 or model 2 is chosen, the suppression of the model(s) accounting for long-term correlated errors, i.e., $(Y_{GM,1})_k$ and/or $(Y_{GM,2})_k$, may be compensated by augmenting the WN level in W_k. This will improve the correspondence between the estimated filter precision with the real error.

As interpreting the comparison of models on the basis of single realization will be taken with caution, we have

TABLE III
Mean Planimetric Error per Outage Evaluated from Different Emulation Scenarios for the 3 Tested Models

	Mean Error at the End of GNSS-Free Period (m)		
Period No.	Model 1	Model 2	Model 3
1	325	319	316
2	322	315	294
3	318	226	141

collected an additional 4 hr of static data that together with the former observations total an 8 hr-long error signal. While the duration of the helicopter flight is about 20 min, we have at our disposition a total of $8 \times 3 = 24$ nonoverlapping error realizations to be evaluated per model and outage, respectively. The mean values of these errors are presented in Table III. It can be concluded from this table that while the difference between models appears not very significant, model 3 outperforms models 1 and 2 about 28% at the second outage and about 52% at the third, respectively. As no systematic effects are estimated in model 1 (WN) to correct the sensor prior to strapdown navigation this model provides the largest navigation error. It should be noted, however, that the error signature may likely change under real vehicle dynamics. Therefore, the provided emulation will be considered not only as a final model for the particular sensor, but rather as a path that can be taken in the construction of models where the sensor is already calibrated for temperature effects but its true remaining model is unknown albeit its structure is complex and its effect is large. The inconvenience in such model construction under the vehicle dynamics is that the signal of reference (i.e., an IMU of higher accuracy) needs to be present for relatively long periods. This is left for future investigations.

VII. CONCLUSION

In this research, we proposed and tested a new methodology for estimating the parameters when modeling inertial sensor errors. Its importance in terms of final navigation accuracy has been proved using real data

sets. The limitations of the commonly adopted method for modeling stochastic errors, i.e., the AV (or PSD) analysis, have been highlighted and the use of a new (consistent) estimator, the GMWM, proposed in this respect. The implemented new calibration framework enabled analysing long error sequences and rapidly (in a few seconds) estimating very complex models such as composite stochastic processes. Moreover, if required, confidence intervals can be computed for each estimated parameter using this approach. The obtained models were then validated at observation levels by comparing the real data Haar WV (or AV) and PSD estimates with their counterparts computed on simulated signals issued from the estimated model. Finally, the impact on the accuracy of the EKF solution implementing the estimated sensor models was evaluated by means of emulated specific forces and angular rates corrupted by the real error signals. This enabled eliminating dynamics-dependent effects, which may not have been present during the calibration phase (e.g. platform vibrations). The performed emulation study is therefore not conclusive generally, nevertheless, provides certain indications, e.g., for applications related to sensor orientation where dynamics are rather limited and the IMU is isolated from platform vibration together with the sensor head of a camera and/or a scanner, e.g. [46].

Of course, as stated above, the encountered dynamics might also affect the sensor error behavior, and by extension, the final filter accuracy. To that end, error signals must be properly constructed under dynamical conditions and precisely studied with respect to dynamics (e.g. angular and linear acceleration and jerk). If some dependency can be observed, individual sensor error models should account for dynamics. However, the construction of such error signals as well as the process of building models for dynamics is far from being trivial and therefore left to future research.

REFERENCES

[1] Fong, W., Ong, S., and Nee, A.
    Methods for in-field user calibration of an inertial measurement unit without external equipment.
    Measurement Science and Technology, **19**, 085202 (2008). DOI 10.1088/0957-0233/19/8/085202.

[2] Titterton, D., and Weston, J.
    Strapdown Inertial Navigation Technology.
    New York: Peter Peregrinus Ltd., 1997.

[3] Berman, Z.
    Efficient error model construction.
    In *Position Location and Navigation Symposium* (PLANS 2012), 2012, pp. 837–848.

[4] Gelb, A.
    Applied Optimal Estimation.
    Cambridge, MA: MIT Press, 1974.

[5] Guerrier, S., Skaloud, J., Stebler, Y., and Victoria-Feser, M. P.
    Wavelet-variance-based estimation for composite stochastic processes.
    Journal of the American Statistical Association, **108** (2013).

[6] Jeleli, C.
    Inertial Navigation Systems with Geodetic Applications.
    Boston: Walter de Gruyter, 2000. DOI 3110159031.

[7] Nelson, A.
    Nonlinear estimation and modeling of noisy time-series by dual Kalman filtering methods.
    Ph.D. thesis, Oregon Graduate Institute of Science and Technology, Hillsboro, OR, 2000.

[8] Brown, R., and Hwang, P.
    Introduction to Random Signals and Applied Kalman Filtering, vol. 3 (2nd ed.).
    New York: Wiley, 1997.

[9] Bryson, A.
    Applied Linear Optimal Control: Examples and Algorithms.
    New York: Cambridge University Press, 2002.

[10] Xing, Z.
    Over-bounding integrated INS/GNSS output errors.
    Ph.D. thesis, University of Minnesota, Minneapolis, MN, 2010.

[11] Allan, D.
    Statistics of atomic frequency standards.
    Proceedings of the IEEE, **54** (1966), 221–230.

[12] Howe, D., Beard, R., Greenhall, C., Vernotte, F., Riley, W., and Peppler, T.
    Enhancements to GPS operations and clock evaluations using a total Hadamard deviation.
    IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, **52**, 8 (2005), 1253–1261.

[13] Vernotte, F., Lantz, E., and Gagnepain, J.
    Oscillator noise analysis: Multivariance measurement.
    IEEE Transactions on Instrumentation and Measurement, **42**, 2 (1993), 342–350.

[14] Vernotte, F., Lantz, E., Groslambert, J., and Gagnepain, J.
    A new multi-variance method for the oscillator noise analysis.
    In *Proceedings of the Frequency Control Symposium*, 1999.

[15] Vernotte, F., and Zalamansky, G.
    A Bayesian method for oscillator stability analysis.
    IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, **46**, 6 (1999), 1545–1550.

[16] *IEEE Standard Specification Format Guide and Test Procedure for Linear, Single-Axis, Non-gyroscopic Accelerometers*.
    IEEE Standard 1293-1998, 1998.

[17] Guerrier, S.
    Integration of skew-redundant MEMS-IMU with GPS for improved navigation performance.
    Master thesis, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, 2008.

[18] Guerrier, S.
    Improving accuracy with multiple sensors: Study of redundant MEMS-IMU/GPS configurations.
    In *Proceedings of the 22nd International Technical Meeting of The Satellite Division of the Institute of Navigation* (ION GNSS 2009), Savannnah, GA, 2009, pp. 3114–3121.

[19] Hou, H.
    Modeling inertial sensors errors using Allan variance.
    Master's thesis, Geomatics Engineering, University of Calgary, AB, Canada, 2004.

[20] El-Sheimy, N., Hou, H., and Niu, X.
    Analysis and modeling of inertial sensors using Allan variance.
    IEEE Transactions on Instrumentation and Measurement, **57**, 1 (2008), 140–149.

[21] Strus, J., Kirkpatrick, M., and Sinko, J.
    Development of a high accuracy pointing system for maneuvering platform.
    In *Proceedings of the 20th International Technical Meeting of The Satellite Division of the Institute of Navigation* (ION GNSS 2007), Fort Worth, TX, 2007.

[22] Xing, Z., and Gebre-Egziabher, D.
Modeling and bounding low cost inertial sensor errors.
In *Position, Location and Navigation Symposium*, 2008, pp.
1122–1132. DOI 10.1109/PLANS.2008.4569999.

[23] Greenhall, C.
Spectral ambiguity of Allan variance.
IEEE Transactions on Instrumentation and Measurement, **47**,
3 (1998), 623–627.

[24] Shao, J.
Mathematical Statistics (Springer Texts in Statistics).
New York: Springer, 2003. Available:
http://books.google.ch/books?id = cyqTPotl7QcC.

[25] Guerrier, S., Stebler, Y., Skaloud, J., and Victoria-Feser, M. P.
Limits of the Allan variance and optimal tuning of wavelet
variance based estimators.
Submitted working paper, 2013. Available:
http://www.hec.unige.ch/guerrier/allan.

[26] Stebler, Y., Guerrier, S., Skaloud, J., and Victoria-Feser, M.
Constrained expectation-maximization algorithm for
stochastic inertial error modeling: Study of feasibility.
Measurement Science and Technology, **22**, 085204, 2011.
DOI 10.1088/0957-0233/22/8/085204.

[27] Dempster, A., Laird, N., and Rubing, D.
Maximum likelihood from incomplete data via the EM
algorithm.
*Journal of the Royal Statistical Society. Series B
(Methodological)*, **39**, 1 (1977), 1–38.

[28] Waegli, A., and Skaloud, J.
Assessment of GPS/MEMS-IMU integration performance in
ski racing.
In *Proceedings of ENC-GNSS 2007* (TimeNav'07), Geneva,
Switzerland, 2007.

[29] Stebler, Y., Guerrier, S., Skaloud, J., and Victoria-Feser, M.
A framework for inertial sensor calibration using complex
stochastic error models.
In Position, Location and Navigation Symposium, 2012,
Session A5.

[30] Percival, D., and Guttorp, P.
Long-memory processes, the Allan variance and wavelets.
Wavelets in Geophysics, **4**, (1994), 325–344.

[31] Mallat, S.
A Wavelet Tour of Signal Processing.
New York: Academic Press, 1999.

[32] Percival, D. and Walden, A.
Wavelet Methods for Time Series Analysis (Cambridge Series
in Statistical and Probabilistic Mathematics).
New York: Cambridge University Press, 2000.

[33] Greenhall, C.
Recipes for degrees of freedom of frequency stability
estimators.
IEEE Transactions on Instrumentation and Measurements, **40**
(1991), 994–999.

[34] Daubechies, I.
Ten Lectures on Wavelets, vol. 61.
Philadelphia, PA: SIAM, 1992.

[35] Serroukh, A., Walden, A., and Percival, D.
Statistical properties and uses of the wavelet variance
estimator for the scale analysis of time series.
Journal of the American Statistical Association, **95**, 449
(2000), 184–196.

[36] Zhang, N.
Allan variance of time series models for measurement data.
Metrologia, **45**, 5 (2008), 549.

[37] Guerrier, S., Skaloud, J., Stebler, Y., and Victoria-Feser, M.
Wavelet-variance-based estimation for composite stochastic
processes.
*Journal of the American Statistical Association (Theory and
Methods)*, **108**, 503 (2013). Available:
http://www.hec.unige.ch/guerrier/GMWM.

[38] Hansen, L.
Large sample properties of generalized method of moments
estimators.
Econometrica: Journal of the Econometric Society, **50**, 4
(1982), 1029–1054.

[39] Guerrier, S., Molinari, R., Mabillard, R., Skaloud, J., and
Victoria-Feser, M. P.
An algorithm for automatic inertial sensors calibration.
In *Proceedings of the ION GNSS*, 2013, to be published.

[40] Guerrier, S.
Two essays in statistics: A prediction divergence criterion for
model selection & wavelet variance based estimation of latent
time series models.
Ph.D. thesis, University of Geneva, Geneva, Switzerland,
2013.

[41] Smith, A.
Estimating nonlinear time-series models using simulated
vector autoregressions.
Journal of Applied Econometrics, **8**, S1 (1993), S63–S84.

[42] Gallant, A., and Tauchen, G.
Which moments to match?
Econometric Theory, **12**, 4 (1996), 657–681.

[43] Gourieroux, C., Monfort, A., and Renault, E.
Indirect inference.
Journal of Applied Econometrics, **8** (1993), 85–118. Available:
http://ideas.repec.org/a/jae/japmet/v8y1993isps85-118.html.

[44] XSens Technologies B. V.
MTi-G user manual and technical documentation.
Technical Report, 2009.

[45] Schaer, P., Skaloud, J., Stebler, Y., Tom, P., and Stengele, R.
Airborne LiDAR in-flight accuracy estimation.
GPS World, **20**, 8 (2009), 37–41.

[46] Skaloud, J., Vallet, J., Keller, K., Veyssiere, G., and Koelbl, O.
An eye for landscapes - Rapid aerial mapping with handheld
sensors.
GPS World, **17**, 5 (May 2006), 26–32.