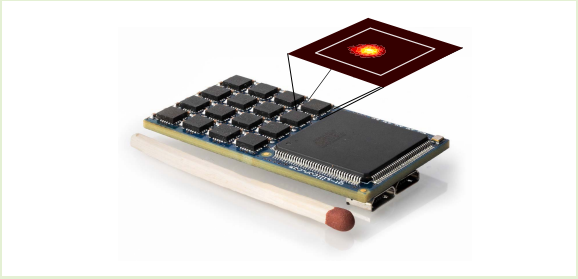
Quantifying the Uncertainty of the Relative Geometry in Inertial Sensors Arrays

Håkan Carlsson, *Student Member, IEEE*, Isaac Skog, *Senior Member, IEEE*, Thomas B. Schön, *Senior Member, IEEE*, and Joakim Jaldén, *Senior Member, IEEE*

***Abstract*—We present an algorithm to estimate and quantify the uncertainty of the accelerometers' relative geometry in an inertial sensor array. We formulate the calibration problem as a Bayesian estimation problem and propose an algorithm that samples the accelerometer positions' posterior distribution using Markov chain Monte Carlo. By identifying linear substructures of the measurement model, the unknown linear motion parameters are analytically marginalized, and the remaining non-linear motion parameters are numerically marginalized. The numerical marginalization occurs in a low dimensional space where the gyroscopes give information about the motion. This combination of information from gyroscopes and analytical marginalization allows the user to make no assumptions of the motion before the calibration. It thus enables the user to estimate the accelerometer positions' relative geometry by simply exposing the array to arbitrary twisting motion. We show that the calibration algorithm gives good results on both simulated and experimental data, despite sampling a high dimensional space.**



***Index Terms*— Accelerometers, gyroscopes, calibration, inertial sensors, self-calibration, sensor arrays, Bayesian estimation, Markov chain Monte Carlo, pseudo-marginal metropolis hastings, Rao-Blackwellization.**

I. Introduction

THE rapid development of the Micro-Electro-Mechanical Systems (MEMS) manufacturing technology has facilitated the construction of small and low-cost sensor arrays of multiple Inertial Measurement Units (IMU) [1], [2]. These so-called inertial sensor arrays have several advantages over a single IMU. The sensor redundancy enables measurement noise reduction through averaging and sensor fault detection [3]. Moreover, measurements from multiple spatially separated accelerometers provide information about both translational acceleration and rotational information. The additional rotational information can increase both the angular velocity measurement accuracy and the angular velocity measurement range of the gyroscopes [4], which typically is a trade-off for gyroscopes. This increase in accuracy and range is useful in high dynamic applications, such as ballistic platform guidance [5] and human motion analysis [6], [7]. Additionally, the angular acceleration can be estimated from a single time sample, omitting the need for time differentiation. Since differentiation amplifies noise and introduces a delay, this is particularly suitable for real-time motion control [8]. Thanks to the versatility of the inertial sensor array, the device has found many different application areas in motion analysis and tracking, such as gyro-free inertial navigation [9], [10], pedestrian tracking [11], biomechanics [6], crash testing [12], [13], and experimental structural dynamics [14]. Reviews of applications and methods concerning inertial sensor arrays can be found in [15], [16].

Reliable motion estimation from the inertial sensor array requires small measurement errors [17]. The measurement errors can principally be grouped into stochastic and deterministic errors. The stochastic measurement errors in the inertial sensor array are typically addressed through sensor redundancy. The deterministic errors are systematic errors that corrupt the mapping between the measured physical quantity and the sensor output. First-order systematic errors include offset, scale factor, misalignment, and mounting errors. These errors are individual to each sensor and will

Manuscript received March 26, 2021; revised May 13, 2021; accepted June 15, 2021. Date of publication June 17, 2021; date of current version August 31, 2021. This work was supported in part by the Wallenberg Autonomous Systems and Software Program (WASP) and in part by the Swedish Foundation for Strategic Research (SSF) through the Project ASSEMBLE. The associate editor coordinating the review of this article and approving it for publication was Prof. Chao Tan. *(Corresponding author: Håkan Carlsson.)*

Håkan Carlsson and Joakim Jaldén are with the Department of Intelligent Systems, KTH Royal Institute of Technology, 114 28 Stockholm, Sweden (e-mail: hakcar@kth.se; jalden@kth.se).

Isaac Skog is with the Department of Electrical Engineering, Linköping University (LIU), 581 83 Linköping, Sweden (e-mail: isaac.skog@liu.se).

Thomas B. Schön is with the Department of Information Technology, Uppsala University, 752 36 Uppsala, Sweden (e-mail: thomas.schon@it.uu.se).

Digital Object Identifier 10.1109/JSEN.2021.3090273

degrade the accuracy of the estimated motion if they are not compensated for. Calibration of offset, scale factor errors, and misalignment errors for accelerometer arrays have been addressed before [18]–[21], where the accelerometer output under stationary conditions is compared with the gravity vector, which has a known magnitude. However, estimation of accelerometer position errors due to erroneous mounting requires the array to undergo rotational motion. An external rig can generate such motion [21]–[23], or the motion can be arbitrary and estimated using a separate sensor system [19]. However, these techniques require external systems that may prohibitively increase the inertial sensor array's cost [2]. The alternative to using dedicated external equipment is to jointly estimate the unknown motion and the calibration parameters. Such joint estimation has been done before using maximum likelihood methods for accelerometer position errors [24]–[26], accelerometer offset, scale factor, and misalignment errors [27] and magnetometers [28]. The drawback of the maximum likelihood method is, however, the difficulty to directly quantify the estimation uncertainty given specific measurements. For biomedical applications [29], it is important to have an uncertainty quantification of the calibration parameters to assess the calibration quality. Uncertainty quantification of estimates has previously been addressed using Kalman filtering [30]. However, the Kalman filter requires a dynamic model for the motion. For arbitrary motion, it may be difficult to specify and tune a dynamic model representing the motion.

The problem of estimating and quantifying the uncertainty of the accelerometer position errors in an inertial sensor array from nonconsecutive measurements of unknown motion is previously not addressed in the literature. To address the problem of uncertainty estimation associated with the maximum likelihood method and to avoid having to specify a motion model, we propose a Bayesian estimator that uses both accelerometer and gyroscope measurements and a non-informative motion prior, to estimate the accelerometer position errors. Bayesian estimators address the uncertainty quantification by providing a posterior distribution of the parameters given a set of measurements. The *main contribution* is a numerical sampling method targeting the posterior distribution of the accelerometer positions, in which the unknown motion is efficiently marginalized *without* assuming a dynamic model. Since the measurement model is non-linear, the estimator marginalizes the motion numerically. However, by explicitly taking advantage of the signal model's linear substructures, the marginalization can partly be done analytically. This increases the efficiency of the numerical marginalization, which is then performed in a lower-dimensional subspace that coincides with the information gained by the gyroscopes.

II. Problem Formulation

This section presents the calibration problem of estimating and quantifying the accelerometer position uncertainty. In Section II-A, we formulate the calibration problem as a Bayesian estimation problem, where the solution is the posterior distribution. In Section II-B we present the signal model of the inertial sensor array, which is used in approximating the posterior distribution.

A. Posterior Distribution

A joint probability density function $p(y_{1:T})$ describes the statistical properties of a measurement sequence $y_{1:T} \triangleq \{y_t\}_{t=1}^{T}$ from the sensors. In each time sample t, all the sensors on the inertial sensor array instantaneously take a measurement, and all these measurements are concatenated into an array measurement y_t. Each measurement y_t depends on a set of calibration parameters θ, independent of the sample index t, and a set of motion parameters η_t, specific to array sample t. The calibration problem is to estimate the parameters θ given a set of measurements $y_{1:T}$ while also considering the unknown motion sequence $\eta_{1:T} \triangleq \{\eta_t\}_{t=1}^{T}$. The motion parameters are not of direct interest, but since the sensor measurements depend on the motion, they have to be accounted for. We pose the calibration problem as a Bayesian estimation problem. The solution to this problem is to compute the posterior distribution of the calibration parameters given a measurement sequence, that is,

$$p(\theta|y_{1:T}) = \frac{p(y_{1:T}|\theta)\,p(\theta)}{p(y_{1:T})}. \tag{1}$$

The posterior distribution $p(\theta|y_{1:T})$ encapsulates all available information about the calibration parameters θ provided by the measurements $y_{1:T}$. Further, $p(y_{1:T}|\theta)$ is the likelihood function of the measurements, $p(\theta)$ is the prior distribution of the calibration parameters and

$$p(y_{1:T}) = \int p(y_{1:T}|\theta)\,p(\theta)\,d\theta \tag{2}$$

is the marginalized likelihood.

Since the inertial sensor measurements also depend on the motion, the likelihood function $p(y_{1:T}|\theta)$ is not directly available. To account for the motion parameters $\eta_{1:T}$ when computing the posterior distribution of the calibration parameters, the motion parameters have to be marginalized out as

$$p(y_{1:T}|\theta) = \int p(y_{1:T}, \eta_{1:T}|\theta)\,d\eta_{1:T}. \tag{3}$$

The integral in (3) requires an assumption of the motion parameters, specified as a prior distribution $p(\eta_{1:T})$. Even though motion has a time dependence according to Newton's law of motion, we do not model this time dependence to simplify the estimation problem. This assumption implies that $\eta_{1:T}$ can be considered independent over time, and the prior distribution $p(\eta_{1:T})$ can be factored as $p(\eta_{1:T}) = \prod_{t=1}^{T} p(\eta_t)$. Moreover, we assume that for each time sample t the prior distribution $p(\eta_t)$ is non-informative. It turns out that interpreting a non-informative prior distribution as a Gaussian distribution with a variance that tends to infinity yield favorable theoretical simplifications. Given these assumptions on the motion, this paper presents a method to efficiently sample the posterior distribution $p(\theta|y_{1:T})$ while marginalizing the motion parameters in (3).

B. Signal Model

A signal model for an inertial sensor array consisting of multiple accelerometer and gyroscope triads can be derived

from the kinematics of rigid body motion [31]. A rotating rigid body yields different point-wise linear accelerations depending on the point location. Thus, each accelerometer triad geometrically dispersed and fixed on a rigid body measures different linear accelerations, which can be decomposed into the translational, centrifugal, and Euler accelerations. Moreover, the angular velocity is equal for all points on the rotating rigid body, and consequently, all the geometrically dispersed gyroscope triads measure the same angular velocity. Altogether, assuming that the inertial sensor array is a rigid body, the measurement output of sensor triad k of type i at time sample t is [4]

$$y^{(\mathrm{a})}_{k,t} = \underbrace{\omega_t \times (\omega_t \times r_k)}_{\text{Centrifugal}} + \underbrace{\dot{\omega}_t \times r_k}_{\text{Euler}} + s_t + e^{(\mathrm{a})}_{k,t}, \tag{4a}$$

$$y^{(\mathrm{g})}_{k,t} = \omega_t + e^{(\mathrm{g})}_{k,t}, \tag{4b}$$

where the superscript $i \in \{\mathrm{a}, \mathrm{g}\}$ denotes accelerometer and gyroscope, respectively. In (4), ω_t, $\dot{\omega}_t$, and s_t, denote the inertial sensor array angular velocity, angular acceleration, and translational acceleration, respectively. The translational acceleration of the inertial sensor array s_t is the acceleration of the origin of the inertial sensor array's coordinate system. All the parameters are defined in this coordinate system. Moreover, r_k denote the accelerometer triad position,[1] and $e^{(i)}_{k,t}$ denotes the measurement noise. The measurement noise of the accelerometer triad k is assumed to be uncorrelated with the measurement noise of gyroscope triad k, that is $e^{(\mathrm{g})}_{k,t}$ is uncorrelated with $e^{(\mathrm{a})}_{k,t}$. Furthermore, for each $i \in \{\mathrm{a}, \mathrm{g}\}$, $e^{(i)}_{k,t}$ is assumed to be uncorrelated over time, uncorrelated over triads, and zero-mean Gaussian distributed with covariance $Q^{(i)}_k$, that is, $e^{(i)}_{k,t} \sim \mathcal{N}(0, Q^{(i)}_k)$. In the signal model (4), we assume that the sensor triads' scale factor-, misalignment-, orientation-, and offset-errors are calibrated, and the measurements are compensated for these errors. Calibration of these errors for accelerometer triads can be performed during stationary conditions (see, e.g. [18]). Calibration of these errors for gyroscope triads, bar the offset error, are more complicated since a rotational reference motion is required. We discuss how to mitigate this problem in Section IV-B.

Concatenating the measurements for A accelerometer triads and G gyroscope triads yields the array signal model [4]

$$y^{(\mathrm{a})}_t = h(\omega_t, \theta) + H(\theta)\varphi_t + e^{(\mathrm{a})}_t, \tag{5a}$$

$$y^{(\mathrm{g})}_t = 1_G \otimes \omega_t + e^{(\mathrm{g})}_t, \tag{5b}$$

where

$$y^{(i)}_t \triangleq \begin{bmatrix} y^{(i)}_{1,t} \\ \vdots \\ y^{(i)}_{K,t} \end{bmatrix}, \quad h(\omega_t, \theta) \triangleq \begin{bmatrix} \Omega^2(\omega_t) r_1 \\ \vdots \\ \Omega^2(\omega_t) r_A \end{bmatrix}, \quad \varphi_t \triangleq \begin{bmatrix} \dot{\omega}_t \\ s_t \end{bmatrix},$$

$$e^{(i)}_t \triangleq \begin{bmatrix} e^{(i)}_{1,t} \\ \vdots \\ e^{(i)}_{K,t} \end{bmatrix}, \quad H(\theta) \triangleq \begin{bmatrix} -\Omega(r_1) & I_3 \\ \vdots & \vdots \\ -\Omega(r_A) & I_3 \end{bmatrix}.$$

[1]Although three sensors inside a sensor triad cannot strictly speaking be located at the same point in space, they are assumed in this work to be sufficiently close to be considered colocated.

Here $\theta \triangleq \{r_k\}_{k=2}^{A}$ defines the considered calibration parameters for the inertial sensor array,[2] $K \in \{A, G\}$, 1_N is a column vector of size N with all entries equal to 1, I_M is the identity matrix of dimension M, and $\Omega(a)b = a \times b$ is the skew-symmetric matrix form of the cross-product. The motion of the inertial sensor array is described by the parameters $\eta_t \triangleq \{\omega_t, \varphi_t\}$, hereafter referred to as the motion parameters of the inertial sensor array. For each $i \in \{\mathrm{a}, \mathrm{b}\}$, the joint measurement noise $e^{(i)}_t$ is Gaussian distributed with a mean equal to zero and with the covariance $Q^{(i)} \triangleq \bigoplus_{k=1}^{K} Q^{(i)}_k$. Inter-triad cross-talk effects could be accounted for by also including off-diagonal elements in the covariance matrix $Q^{(i)}$. Denoting $y_t \triangleq \{y^{(\mathrm{a})}_t, y^{(\mathrm{g})}_t\}$ as the collection of measurements from all the sensors in the array at one time instant, the measurement model in (5) defines the joint probability distribution $p(y_{1:T})$, from which the posterior distribution $p(\theta|y_{1:T})$ is to be computed.

III. Bayesian Estimator

To solve the calibration problem, we estimate the posterior distribution of the calibration parameters, that is $p(\theta|y_{1:T})$ in (1). This section introduces a sampler for the posterior distribution, which is outlined in Algorithm 1. In Section III-A, the overall sampler is derived and presented. A key aspect of this sampler is how to numerically marginalize the motion parameters, which is discussed in Section III-B. Finally, Section III-C discusses how to practically implement the sampler.

A. Sampling the Posterior Distribution

In general, closed-form computation of the posterior distribution of θ in (1) is not possible. In particular, since the measurement model in (5) is non-linear in the angular velocity ω_t, the integral over $\eta_{1:T}$ in (3) is analytically intractable. This integral is required for the computation of the posterior distribution of the calibration parameters $p(\theta|y_{1:T})$, which is thus also analytically intractable.

Consequently, the computation of the posterior distribution $p(\theta|y_{1:T})$ has to be approached with numerical methods. One way to sample the posterior distribution is to use the Markov Chain Monte Carlo (MCMC) technique. Here, a Markov chain is constructed such that the collection of steps in the chain constitutes samples of the posterior distribution. Such a Markov chain can be constructed by modifying the transition probability of some suitably chosen proposal random walk.

An MCMC method that addresses both the calculation of the marginalized likelihood $p(y_{1:T})$ in (2) and the intractable integral (3) is the Pseudo-marginal Metropolis Hastings (PmMH) method [32], [33]. The PmMH method is an extension of the Metropolis Hastings (MH) method [34]. Assuming that $p(\theta|y_{1:T})$ can be evaluated point-wise up to a proportionality constant, a MH step in the Markov chain consists of drawing a proposed state $\theta' \sim q(\cdot|\theta^{(m-1)})$ from a proposal distribution

[2]The accelerometer position r_1 is without loss of generality, defined as the origin of the coordinate system in which the parameters are defined. Hence, it does not need to be estimated. The translational acceleration s_t thus coincides with the acceleration at r_1.

q dependent on the current state $\theta^{(m-1)}$, and accepting the proposal with probability

$$\rho_{\mathrm{MH}}(\theta', \theta^{(m-1)}) = \min\left(1, \frac{p(\theta'|y_{1:T})}{p(\theta^{(m-1)}|y_{1:T})}\right). \tag{6}$$

That is, $\theta^{(m)} = \theta'$ with probability $\rho_{\mathrm{MH}}(\theta', \theta^{(m-1)})$ and $\theta^{(m)} = \theta^{(m-1)}$ with probability $1 - \rho_{\mathrm{MH}}(\theta', \theta^{(m-1)})$. We assume q to be symmetric, specifically $q(a|b) = q(b|a)$, and thus the ratio $q(\theta^{(m-1)}|\theta')/q(\theta'|\theta^{(m-1)})$ accounting for the proposal distribution in (6) is canceled [34]. Using Bayes' rule the acceptance probability in (6) can be reformulated as

$$\rho_{\mathrm{MH}}(\theta', \theta^{(m-1)}) = \min\left(1, \frac{p(y_{1:T}|\theta')p(\theta')}{p(y_{1:T}|\theta^{(m-1)})p(\theta^{(m-1)})}\right), \tag{7}$$

where we note that the marginalized likelihood $p(y_{1:T})$ in (1) cancels. Instead of evaluating the full posterior distribution in (1), the computational problem is now to point-wise evaluate the measurements' likelihood given the calibration parameters $p(y_{1:T}|\theta)$. When the Markov chain in $\theta^{(m)}$ has reached its stationary distribution, the samples $\theta^{(m)}$ can be shown to be drawn from the posterior distribution $p(\theta|y_{1:T})$ [34].

However, the measurements' likelihood given the calibration parameters $p(y_{1:T}|\theta)$ is, as earlier noted, not directly available in our application. Due to the motion dependence in (5), only the measurements' likelihood given the calibration parameters and the motion parameters is available, that is $p(y_{1:T}|\eta_{1:T}, \theta)$. Thus, to evaluate $p(y_{1:T}|\theta)$ point-wise, the motion parameters has to be marginalized out as in (3). Since (5) is non-linear in ω_t, the marginalization in (3) is not analytically tractable. This intractable marginalization is addressed by using the PmMH method, where $p(y_{1:T}|\theta)$ in (7) is replaced with an estimate $\hat{p}(y_{1:T}|\theta)$, that is,

$$\rho_{\mathrm{PmMH}}(\theta', \theta^{(m-1)}) = \min\left(1, \frac{\hat{p}(y_{1:T}|\theta')p(\theta')}{\hat{p}(y_{1:T}|\theta^{(m-1)})p(\theta^{(m-1)})}\right). \tag{8}$$

A sufficient condition for the resulting Markov chain with transition probability defined by $\rho_{\mathrm{PmMH}}(\theta', \theta^{(m-1)})$ to produce samples from $p(\theta|y_{1:T})$ is that the estimator $\hat{p}(y_{1:T}|\theta)$ is *unbiased* up to a proportionality constant, that is, $\mathbb{E}[\hat{p}(y_{1:T}|\theta)] = Cp(y_{1:T}|\theta)$ for some constant C that does not depend on θ [35]. This weaker form of unbiasedness is sufficient, since the constant C cancels out in the acceptance probability (8).

A naive estimator $\hat{p}(y_{1:T}|\theta)$ can be derived by first conditioning on the arbitrary motion parameters $\eta_{1:T}$ in (3) according to

$$p(y_{1:T}|\theta) = \int p(y_{1:T}|\eta_{1:T}, \theta)p(\eta_{1:T})d\eta_{1:T}. \tag{9}$$

This integral can be interpreted as the expectation of the function $p(y_{1:T}|\eta_{1:T}, \theta)$ over the motion parameter $\eta_{1:T}$ distributed according to the prior distribution $p(\eta_{1:T}) = p(\eta_{1:T}|\theta)$. Here it is also recognized that the motion parameters do not depend on the calibration parameters. In principle, the integral in (9) can then be numerically estimated using Monte Carlo integration [36] as

$$\hat{p}^N(y_{1:T}|\theta) = \frac{1}{N}\sum_{n=1}^{N} p(y_{1:T}|\eta_{1:T}^{(n)}, \theta), \tag{10a}$$

$$\eta_{1:T}^{(n)} \sim p(\eta_{1:T}), \tag{10b}$$

where N is the number of independent samples. The superscript N is added to the estimator $\hat{p}^N(y_{1:T}|\theta)$ to highlight the dependence on the number of samples used. Note that $\hat{p}^N(y_{1:T}|\theta)$ is an unbiased estimator of $p(y_{1:T}|\theta)$ for any N, and that the estimator variance is inversely proportional to N.

However, the naive estimator $\hat{p}^N(y_{1:T}|\theta)$ is practically infeasible. Since the motion of the inertial sensors array is assumed to be arbitrary and unknown during the calibration, there is no prior knowledge of $\eta_{1:T}$, and the prior distribution $p(\eta_{1:T})$ has to be assumed to be non-informative. That is, the distribution is effectively flat. A non-informative prior is infeasible to combine with Monte Carlo sampling in (10) since the samples $\eta_{1:T}$ then have to cover an infinite space. This problem is further aggravated by the high dimension of $\eta_{1:T}$, which is $9T$. We thus need to develop an estimator with more favorable properties.

B. Marginalization of Motion Parameters

A remedy to the infeasibility of the naive estimator in (10) can be found by introducing the following key assumptions regarding the signal model in (5). We note that the gyroscope measurements $y_{1:T}^{(\mathrm{g})}$ and the motion parameters $\eta_{1:T}$ are both independent of the calibration parameters θ, that is, the accelerometer positions. This independence means that the gyroscope measurements can be conditioned upon in (9) to yield

$$p(y_{1:T}^{(\mathrm{a})}, y_{1:T}^{(\mathrm{g})}|\theta) = \int p(y_{1:T}^{(\mathrm{a})}|y_{1:T}^{(\mathrm{g})}, \eta_{1:T}, \theta)p(y_{1:T}^{(\mathrm{g})}, \eta_{1:T})d\eta_{1:T}. \tag{11}$$

Further, the accelerometer measurements $y_{1:T}^{(\mathrm{a})}$ are conditionally independent of the gyroscope measurements given the angular velocity $\omega_{1:T}$, that is,

$$p(y_{1:T}^{(\mathrm{a})}|y_{1:T}^{(\mathrm{g})}, \eta_{1:T}, \theta) = p(y_{1:T}^{(\mathrm{a})}|\eta_{1:T}, \theta). \tag{12}$$

Assuming that the measurements are independent over time, the measurements' likelihood given the calibration parameters θ can be factorized as

$$p(y_{1:T}|\theta) = \prod_{t=1}^{T} p(y_t|\theta). \tag{13}$$

Subsequently, the joint distribution of the gyroscope measurements and the motion parameters can be factorized as

$$p(y_t^{(\mathrm{g})}, \eta_t) = p(y_t^{(\mathrm{g})}, \omega_t)p(\varphi_t), \tag{14}$$

where we have used the fact that the gyroscope measurements do no depend on φ_t, introduced in (5a). Although the angular acceleration $\dot{\omega}$ is the time-derivative of ω, we assume no prior knowledge of the relationship between time samples and model these as independent random variables.

Using (14) we can make two simplifications. Firstly, the joint distribution of the gyroscope measurements and the angular velocity $p(y_t^{(\mathrm{g})}, \omega_t)$ can be written as

$$p(y_t^{(\mathrm{g})}, \omega_t) = p(\omega_t|y_t^{(\mathrm{g})})p(y_t^{(\mathrm{g})}). \tag{15}$$

The term $p(y_t^{(\mathrm{g})})$ is a constant that cancels in the acceptance probability (8). Further, the posterior distribution of ω_t given

$y_t^{(\mathrm{g})}$ can be computed in closed-form given the linearity of the gyroscope measurement model in (5b).

Secondly, using the time independence in (13), the motion independence in (14), and noting that the signal model in (5a) has a separation in the non-linear parameter ω_t and the linear parameters φ_t, the integral in (11) can be separated into a product of terms of the form

$$p(y_t^{(\mathrm{a})}, y_t^{(\mathrm{g})}|\theta) = \int p(y_t^{(\mathrm{a})}|\omega_t, \theta)\, p(\omega_t|y_t^{(\mathrm{g})}) d\omega_t, \quad (16a)$$

$$p(y_t^{(\mathrm{a})}|\omega_t, \theta) = \int p(y_t^{(\mathrm{a})}|\varphi_t, \omega_t, \theta)\, p(\varphi_t) d\varphi_t. \quad (16b)$$

If both φ_t and $e_t^{(\mathrm{a})}$ are assumed to be Gaussian distributed, the marginalization of the linear motion parameters φ_t in (16b) can be done analytically, since $y_t^{(\mathrm{a})}$ depends linearly on φ_t. Such analytical marginalization reduces the variance of the likelihood estimator $\hat{p}^N(y_t|\theta)$ in (10) and is known as Rao-Blackwellization in the Monte Carlo literature [36], [37]. Therefore, we assume that φ_t is Gaussian distributed, that is

$$p(\varphi_t) = \mathcal{N}(\varphi_t;\, f_t, \beta F_t), \quad (17)$$

with some mean value f_t and covariance F_t, where $\beta > 0$ is a positive scaling factor. The solution to the marginalization in (16b) is then given by

$$p(y_t^{(\mathrm{a})}|\omega_t, \theta) = \mathcal{N}(y_t^{(\mathrm{a})};\, m_t, M_t), \quad (18a)$$

$$m_t \triangleq h(\omega_t, \theta) + H(\theta) f_t, \quad (18b)$$

$$M_t \triangleq Q^{(\mathrm{a})} + \beta H(\theta) F_t H^\top(\theta). \quad (18c)$$

To investigate the case of a non-informative prior on the motion parameters φ_t, we let β tend towards infinity, and study the distribution

$$\mathcal{N}(y_t^{(\mathrm{a})};\, m_t, M_t) = \frac{\exp\left(-\frac{1}{2}\|y_t^{(\mathrm{a})} - m_t\|^2_{M_t^{-1}}\right)}{\sqrt{|2\pi M_t|}}, \quad (19)$$

where $\|a\|_P$ is the weighted Euclidean norm [38], defined as $\|a\|_P^2 \triangleq a^\top P a$ for a positive definite matrix P. By using the Schur complement [39] the inverse of the covariance matrix M_t is

$$M_t^{-1} = \Lambda - \Lambda H(\theta)\left(\frac{1}{\beta}F_t^{-1} + H^\top(\theta)\Lambda H(\theta)\right)^{-1} H^\top(\theta)\Lambda, \quad (20)$$

where $\Lambda \triangleq (Q^{(\mathrm{a})})^{-1}$, and the determinant of M_t is

$$|M_t| = \left|\frac{1}{\beta}F_t^{-1} + H^\top(\theta)\Lambda H(\theta)\right| |\beta F_t||Q^{(\mathrm{a})}|. \quad (21)$$

In the limit of $\beta \to \infty$, the inverse of M_t becomes $M_t^{-1} \to S$, where

$$S \triangleq \Lambda - \Lambda H(\theta)(H^\top(\theta)\Lambda H(\theta))^{-1} H^\top(\theta)\Lambda. \quad (22)$$

Although the determinant of $|M_t|$ diverge with $\beta \to \infty$, we can use the fact that constants cancel in the acceptance probability (8) and replace $|M_t|$ by

$$\frac{|M_t|}{|\beta F_t||Q^{(\mathrm{a})}|} \to \left|H^\top(\theta)\Lambda H(\theta)\right|. \quad (23)$$

The null space of the information matrix S is the range space of $H(\theta)$, that is, $SH(\theta) = 0$ for all θ. Hence, in the limit, the quadratic form in the exponent of the distribution $\mathcal{N}(y_t; m_t, M_t)$ becomes

$$\|y_t^{(\mathrm{a})} - h(\omega_t, \theta) - H(\theta) f_t\|^2_{M_t^{-1}} \to \|y_t^{(\mathrm{a})} - h(\omega_t, \theta)\|_S^2. \quad (24)$$

Thus, the terms remaining in the accelerometer measurements' likelihood $p(y_t^{(\mathrm{a})}|\omega_t, \theta)$ are terms that do not cancel in the acceptance ratio in (8), and that depends on the motion η_t and the calibration parameters θ, that is,

$$p(y_t^{(\mathrm{a})}|\omega_t, \theta) \propto \frac{\exp\left(-\frac{1}{2}\|y_t^{(\mathrm{a})} - h(\omega_t, \theta)\|_S^2\right)}{|H^\top(\theta)\Lambda H(\theta)|}. \quad (25)$$

Note that (25) becomes independent of the prior values of f_t and F_t when $\beta \to \infty$, which is desirable as we have no particular reason to prefer one particular assumption for the motion over any other.

The naive and practically infeasible likelihood estimator in (10) can thus be replaced with the following Rao-Blackwellized estimator

$$\hat{p}^N(y_{1:T}^{(\mathrm{a})}|\theta) = \prod_{t=1}^{T}\left(\frac{1}{N}\sum_{n=1}^{N}\frac{\exp\left(-\frac{1}{2}\|y_t^{(\mathrm{a})} - h(\omega_t^{(n)}, \theta)\|_S^2\right)}{|H^\top(\theta)\Lambda H(\theta)|}\right), \quad (26a)$$

$$\omega_t^{(n)} \sim p(\omega_t|y_t^{(\mathrm{g})}). \quad (26b)$$

Here, the numerical integration is only performed over ω_t compared to η_t in (10b). The subspace over ω_t has a lower dimension than that of η_t and is orthogonal to the matrix $H(\theta)$. Moreover, samples in this space are drawn from the distribution $p(\omega_t|y_t^{(\mathrm{g})})$, which has a concentrated probability mass compared to the non-informative distribution $p(\eta_t)$ in (10b), and thus the Monte Carlo samples are more efficiently localized. The remaining motion parameters φ_t, which enter the array signal model in (5a) linearly, are analytically marginalized using a wide Gaussian distribution that, in the limit, do not depend on the prior values of φ_t. This method of separating the computational problem into linear and non-linear parts is similar to the likelihood function's concentration in maximum likelihood estimation [4] and separable least squares [40]. The complete procedure for sampling of the posterior distribution $p(\theta|y_{1:T})$ is summarized in Algorithm 1.

C. Implementation Details and Tuning

The PmMH sampler detailed in Algorithm 1 has certain implementation and tuning aspects that need to be considered. For example, when computing the probabilities in the acceptance ratio in (8), the logarithm of the probabilities should be used to avoid numerical underflow. However, when evaluating $\hat{p}^N(y_{1:T}|\theta)$ in (26) the summation of N exponential terms may still cause numerical underflow. The logarithm of a sum of exponential terms can efficiently and accurately be computed using the log-sum-exp function [41], see Appendix A. The summation over T time samples, which are assumed independent, can be evaluated in parallel and do not pose any numerical challenges.

Algorithm 1 Resulting PmMH Sampler

Given $\theta^{(0)}$
Compute $\alpha^{(0)} \leftarrow \hat{p}^N(y_{1:T}^{(a)}|\theta^{(0)})$ using (26a)
for $m = 1, \ldots, M$ **do**
  Sample $\theta' \sim q(\cdot|\theta^{(m-1)})$ using (27)
  for $t = 1, \ldots, T$ and $n = 1, \ldots, N$ **do**
    Sample $\omega_t^{(n)} \sim p(\omega_t^{(n)}|y_t^{(g)})$
  end for
  Compute $\alpha' \leftarrow \hat{p}^N(y_{1:T}^{(a)}|\theta')$ using (26a)
  Compute $\rho \leftarrow \min\left(1, \frac{\alpha' p(\theta')}{\alpha^{(m-1)} p(\theta^{(m-1)})}\right)$
  Sample $u \sim \mathcal{U}(0, 1)$
  if $\rho > u$ **then**
    Set $\theta^{(m)}, \alpha^{(m)} \leftarrow \theta', \alpha'$
  else
    Set $\theta^{(m)}, \alpha^{(m)} \leftarrow \theta^{(m-1)}, \alpha^{(m-1)}$
  end if
end for

Furthermore, a proposal distribution q for θ needs to be chosen. Under mild technical assumptions, the stationary distribution of $\theta^{(m)}$ produced by Algorithm 1 will be the same regardless of the choice of proposal distribution [34]. However, the convergence speed and mixing ability (defined below) may be affected. In this work, we use

$$q(\theta'|\theta) = \mathcal{N}(\theta'; \theta, \sigma_q^2 I_{3(A-1)}), \tag{27}$$

where σ_q is referred to as the step length of the Gaussian random walk and is a hyperparameter that has to be tuned.

The samples produced by the Markov chain are by construction correlated. Intuitively, correlated samples contain less information about the posterior distribution than independent samples. The correlation affects how the empirical average

$$\frac{1}{M}\sum_{m=1}^{M} f(\theta^{(m)}) \tag{28}$$

converges to $\mathbb{E}[f(\theta)]$ for a function f. The degree of correlation is also known as mixing [34], which is a measure of the efficiency of MCMC algorithms. The MH algorithm's mixing efficiency using a Gaussian random walk as the proposal is affected by the step length σ_q. If σ_q is too large, most of the proposed steps will be rejected since they will be in regions of low posterior probability, and the correlation among $\theta^{(m)}$ will be high. On the other hand, if σ_q is too small, most proposed states will be accepted, resulting in a poor support coverage of the posterior distribution. Hence, there is a trade-off when selecting σ_q. The optimal step length for the MH algorithm using a Gaussian random walk as the proposal distribution is obtained when the acceptance probability is approximately 23% [42], which can be empirically estimated by running the chain with different values of σ_q.

Furthermore, the amount of correlation in a Markov chain produced by the PmMH algorithm also depends on the variance of the likelihood estimator in (26). If the variance is too high, an accepted state $\theta^{(m)}$ may yield a high value for $\hat{p}^N(y_{1:T}|\theta^{(m)})$ by chance. Subsequently, the proposed states θ' are then likely to result in $\hat{p}^N(y_{1:T}|\theta') \ll \hat{p}^N(y_{1:T}|\theta^{(m)})$, implying that the acceptance probability in (8) becomes small, that is,

$$\min\left(1, \frac{\hat{p}(y_{1:T}|\theta')p(\theta')}{\hat{p}(y_{1:T}|\theta^{(m)})p(\theta^{(m)})}\right) \approx 0. \tag{29}$$

Thus, the chain is likely to get stuck in the state $\theta^{(m)}$ for many iterations, increasing the correlation. The variance of $\hat{p}^N(y_{1:T}|\theta)$ can be decreased by increasing the number of Monte Carlo samples N. However, the high correlation of $\theta^{(m)}$ given a small N could be compensated by taking more and smaller MCMC steps, that is, increase M and decrease σ_q. Hence, there is a trade-off in the proposal step length σ_q, the number of MCMC steps M, and the number of Monte Carlo samples N. Finding an optimal choice for the hyperparameter N is non-trivial since the variance of $\hat{p}^N(y_{1:T}|\theta)$ depends on θ. The MCMC literature [43], [44] provides guidelines for setting N and σ_q under idealized assumptions. For example, [43] recommends to tune N so that $\mathrm{Var}[\log \hat{p}^N(y_{1:T}^{(a)}|\theta)] \approx 3.3$ while simultaneously tuning σ_q so that the acceptance probability is 7%. Here, we pragmatically choose to first determine N by calculating the variance of (26) for different N evaluated at the prior mean value of θ, and then tune σ_q. Since (26) can be evaluated in parallel over T time samples, it is sensible to increase N instead of M. Concretely, we set N to obtain $\mathrm{Var}[\log \hat{p}^N(y_{1:T}^{(a)}|\theta)] \approx 0.1$ and then set σ_q so that the acceptance probability is between 7% and 23%.

Finally, the Markov chain needs to start at an initial value $\theta^{(0)}$. The initial value $\theta^{(0)}$ may be set to a value where the posterior distribution $p(\theta|y_{1:T})$ has a low probability. The law of large numbers guarantees the convergence of the empirical average in (28) to the expected value as $M \to \infty$, independent of the starting value $\theta^{(0)}$. However, the empirical average in (28) can only be evaluated using a finite number of samples. If too many samples are from low probability regions, (28) will be a poor approximation to the expected value. The practical solution to this problem is to discard samples of the chain before the chain has reached high probability regions. This is known as the burn-in period. The burn-in period is a complicated matter since it depends on the MCMC algorithm and the target distribution [45]. Here we take the pragmatic approach of running two parallel chains, starting from different initial values $\theta^{(0)}$, and then assess by cross-referencing when the two chains have both reached a region of high probability.

IV. Simulations and Experiments

The performance of the proposed sampler for the posterior distribution, summarized in Algorithm 1, was evaluated by simulations and experiments on a real inertial sensor array. The simulations demonstrate the performance obtained under idealized assumptions, and real-world experiments demonstrate the practical applicability of the considered method.

A. Simulations

We evaluated the performance of the Bayesian estimator for the accelerometer positions using simulated data. The inertial sensor array considered in the simulations is shown in Fig. 1,

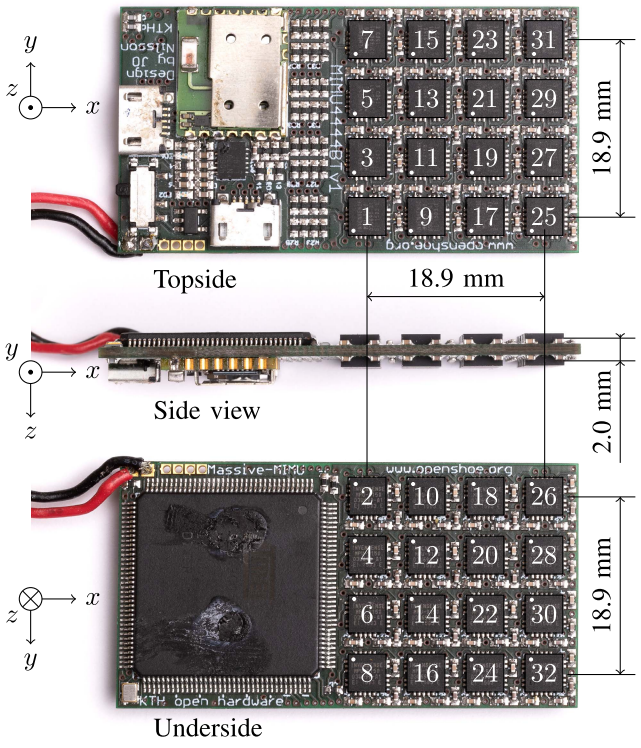


Fig. 1. Layout and dimensions of the inertial sensor array used in the experiments and simulations, along with the numbering of the IMU triads used in tables and figures. The dimensions denote the distance between the center positions of the IMU chip packages. The white dots on each IMU package are the orientation markers, indicating the mounting directions of the IMU packages [46]. The IMU packages on the underside, even numbers, are rotated 180° around the axis $y = x$ relative to the IMU packages on the topside, odd numbers. Each IMU package measures 4 mm × 4 mm × 1 mm [46].

which is an inertial sensor array with 32 IMUs attached to a Printed Circuit Board (PCB). MCMC methods generally do not scale well with high dimensional problems, and testing the method with 32 IMUs would confirm the practical feasibility of the proposed method. Since IMU 1's position, that is r_1, defines the origin of the parameter coordinate system, the dimension of the calibration parameters is $3 \times 31 = 93$. We assume that the measurement errors of the accelerometers and the gyroscopes to be uncorrelated and to have a standard deviation of $\sigma^{(a)} = 0.04$ m/s^2 and $\sigma^{(g)} = 0.06$ deg/s, respectively. These parameters were selected to reflect the performance of typical MEMS-based IMUs [46]. Further, we assume idealistically that the sensors have an infinite dynamic range, specifically, they never saturate. To simulate field conditions, we generate a random motion sequence $\eta_{1:T}$ so that it emulates the dynamics of twisting the array with a hand. Specifically, ω, $\dot{\omega}$, and s are assumed to be zero-mean normally distributed with a standard deviation of 1000 °/s, 35 000 °/s^2, and 20 m/s^2, respectively [26]. We set the true values of the accelerometer positions equal to the center of the IMU packages shown in Fig. 1. The prior distribution of the accelerometer positions $p(\theta)$ is set to be a Gaussian with a standard deviation of $\sigma_r = 3$ mm. Even though the size of the IMU packages is 4 mm × 4 mm × 1 mm, we set this large value

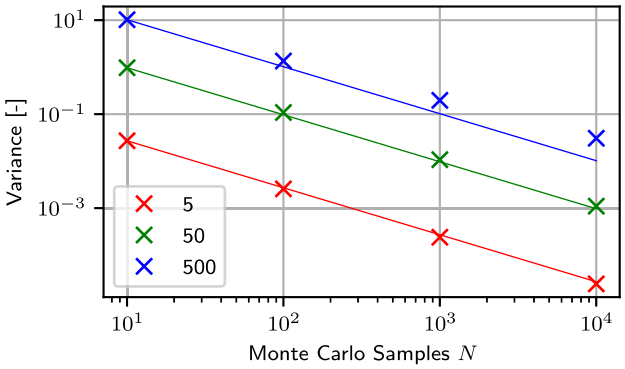
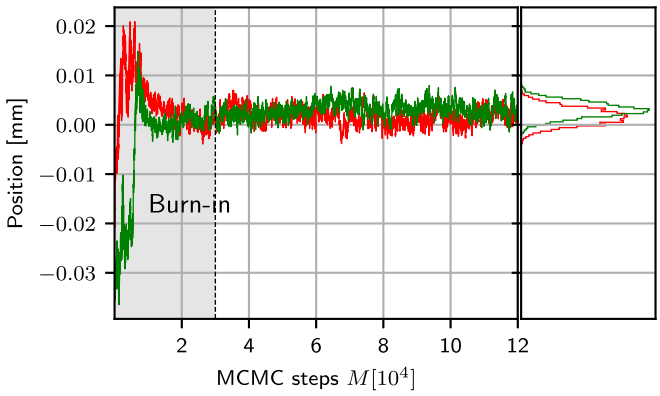
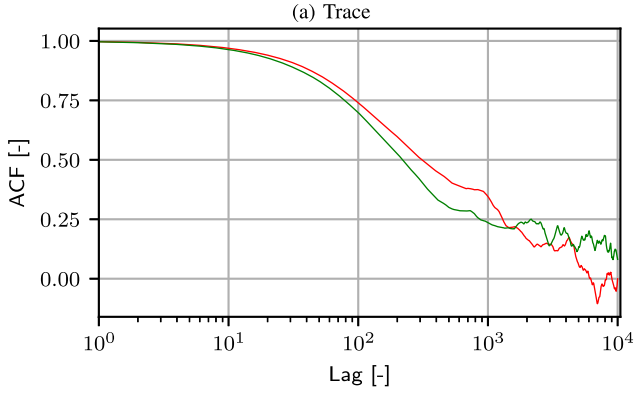


Fig. 2. The variance of $\log \hat{p}^N(y_{1:T}^{(a)}|\theta)$ in (26a) computed with 1000 realizations for different number of Monte Carlo samples N and different number of time samples T. The solid lines correspond to $\sim 1/N$.



(a) Trace



(b) Estimated autocorrelation function

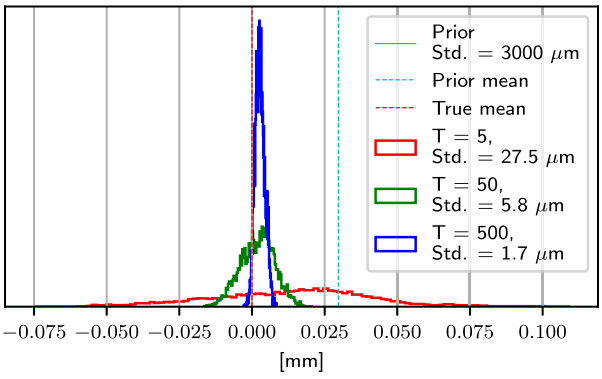
Fig. 3. Trace (a) and estimated autocorrelation (b) of two Markov chains for the x-component of r_2 when $T = 500$. The histogram in (a) and the estimated autocorrelation in (b) uses only samples after the burn-in period.

for the standard deviation to see how robust the method is to large initial uncertainties in the accelerometer positions. The prior mean is equal to a realization of a Normal distribution centered on the true positions and with a standard deviation of 0.1 mm. It should be stressed that the proposed sampler only requires point-wise evaluation of the prior distribution. An alternative prior distribution for the accelerometer positions could be a uniform distribution with bounds equal to the edges of the IMU package.

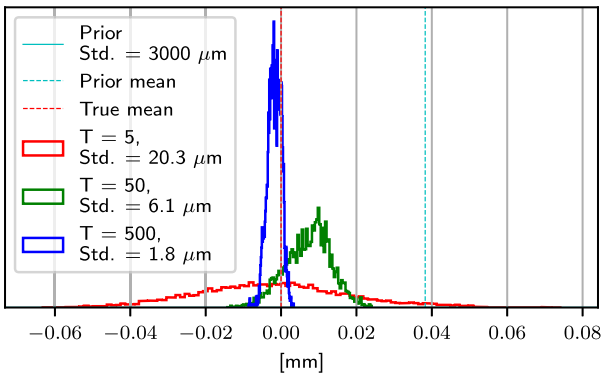
We evaluated Algorithm 1 for three different lengths of measurement sequences $y_{1:T}$, that is, $T \in \{5, 50, 500\}$. For each realization of the measurement sequence $y_{1:T}$, we tuned

TABLE I
VALUES FOR THE TUNING PARAMETERS

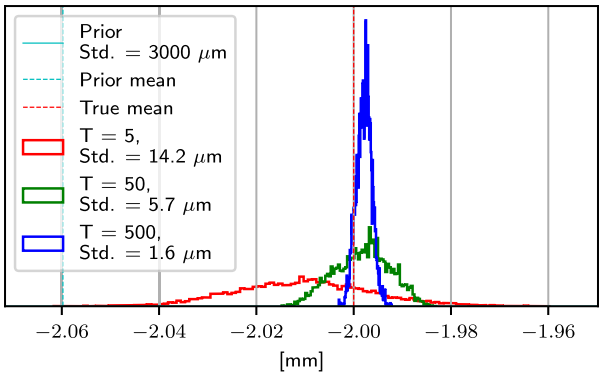
Time samples T	Monte Carlo samples N	Step length σ_q [μm]	$\mathrm{Var}[\log \hat{p}^N(y^{(\mathrm{a})}_{1:T}\|\theta)]$	ρ
5	10	3.16	0.03	23%
50	100	1.41	0.11	9%
500	1000	0.44	0.20	9%



(a) x axis



(b) y axis



(c) z axis

Fig. 4. Posterior distributions of IMU 2's position, in the *x*-axis (a), *y*-axis (b), and *z*-axis (c) directions. The estimated posterior distribution concentrates around the true value and the variance shrinks with an increasing number of time samples *T*.

the number of Monte Carlo samples N for the likelihood estimator in (26) and the step length σ_q for the Gaussian random walk according to Section III-C. Fig. 2 depicts the variance of (26) evaluated at the prior mean of θ. For a fixed T, we observe that the variance of $\log \hat{p}^N(y^{(\mathrm{a})}_{1:T}|\theta)$ is approximately inversely proportional to N. Importantly, to achieve the same level of variance for $\log \hat{p}^N(y^{(\mathrm{a})}_{1:T}|\theta)$ when T increases requires the number of Monte Carlo samples N to increase since (26) is a sum over independent samples. Thus, Algorithm 1 has a limitation in how many time samples can be used in the calibration. Moreover, the values for the chosen tuning parameters are reported in Table I.

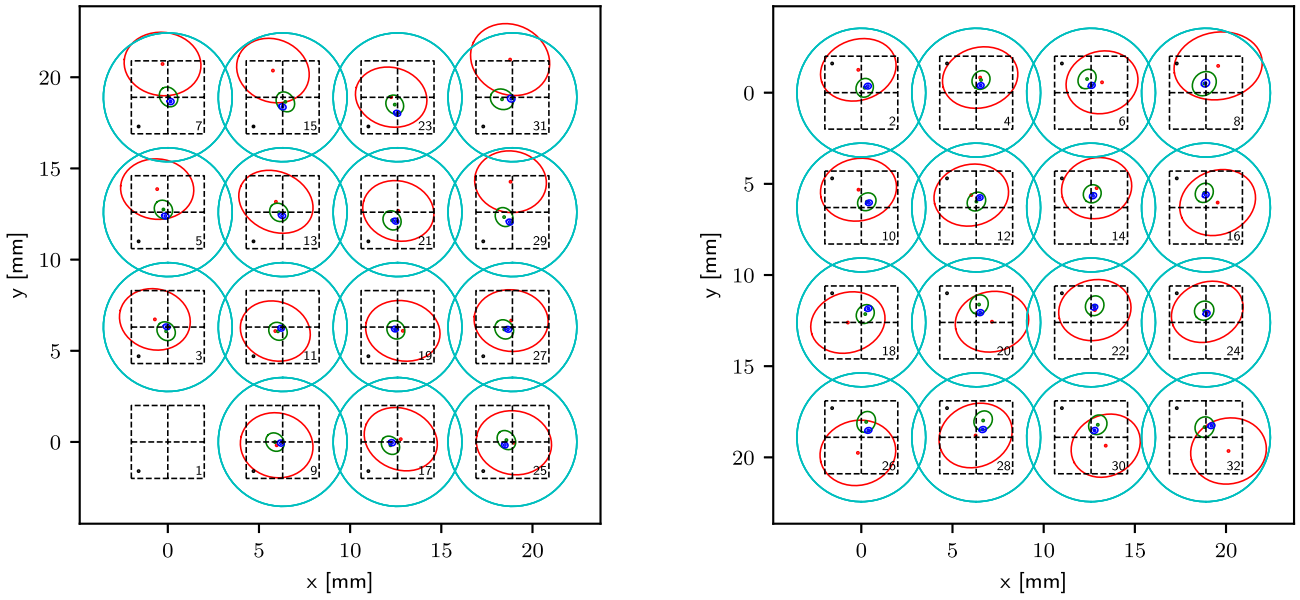
Having fixed the number of Monte Carlo samples N for each T, Fig. 3 shows the Markov chain's trace, burn-in period, and autocorrelation function for $\theta^{(m)}$. Two chains starting from two different $\theta^{(0)}$ are seen to converge to the same empirical distribution. The posterior distribution for IMU 2's position is shown in Fig. 4. For all evaluated time samples, the variance of the posterior distribution is less than the variance of the prior distribution. As the number of time samples T increases, the mean of the posterior distribution moves closer to the true value, and the variance of the posterior distribution decreases. We can thus conclude the practical feasibility of our proposed MCMC algorithm and it's robustness to large initial uncertainties in the accelerometer positions.

B. Practical Considerations

Having established the PmMH sampler's feasibility through simulations, we will now consider the sampler's performance on experimental data. However, before estimating the accelerometers relative geometry, we calibrate the accelerometer and gyroscope triads according to a first-order model during stationary conditions [18]. Specifically, we estimate the sensors' measurement variance, offset, scale factor, misalignment, and orientations using a maximum likelihood estimator. However, the gyroscopes' scale factor-, misalignment-, orientation- errors cannot be estimated during stationary conditions, since these parameters require a rotational motion. The gyroscope scale factor error will increase the angular velocity estimation error for large angular velocities. For high dynamic motions, the posterior distribution of the angular velocity given the gyroscope measurements, $p(\omega_t|y^{(\mathrm{g})}_t)$, may yield a distribution that is not sufficiently centered around the true angular velocity. This inaccurate localization of the distribution becomes a problem when numerically marginalizing the angular velocity in (26), since the drawn samples are from low probability regions. We mitigate this problem by replacing the estimated covariance during stationary conditions with the sample covariance during a single time sample. Specifically, we set $p(\omega_t|y^{(\mathrm{g})}_t) = \mathcal{N}(\bar{y}^{(\mathrm{g})}_t, \bar{Q}^{(\mathrm{g})}_t)$ where

$$\bar{y}^{(\mathrm{g})}_t = \frac{1}{G}\sum_{k=1}^{G} y^{(\mathrm{g})}_{k,t}, \tag{30a}$$

$$\bar{Q}^{(\mathrm{g})}_t = \frac{1}{G-1}\sum_{k=1}^{G}(y^{(\mathrm{g})}_{k,t} - \bar{y}^{(\mathrm{g})}_t)(y^{(\mathrm{g})}_{k,t} - \bar{y}^{(\mathrm{g})}_t)^\top. \tag{30b}$$



(a) Topside (b) Underside

Fig. 5. Estimated 50% confidence ellipses (lines) and means (dots) of the posterior distribution of the accelerometer positions of the inertial sensor array for the IMUs on the topside (a) and the underside (b). The cyan lines define the prior distribution. The red, green, and blue lines are for 5, 50, and 500 time samples, respectively. The IMU in the lower-left corner on the topside defines the reference position and is thus not estimated. The dashed lines define the borders of the IMU packages and also divide the package up into four quadrants. The black dots denote the orientation markers of the IMU packages. The IMU numbering corresponds to that of Fig. 1.

Moreover, the assumption that accelerometer sensors follow a first-order model may not be a sufficiently accurate model for the accelerometer measurements during high dynamic motions. This problem may worsen if the parameters of the first-order model are estimated during stationary conditions. For large signal amplitudes, error sources such as sensor non-linearities may well be the dominating error, which the first-order approximation does not account for. Since the measurement of each accelerometer triad depends on its the position, we unfortunately cannot replace the estimated accelerometer measurement covariance during stationary conditions with the sample covariance, as could be done for the gyroscope triads. We choose a practical approach to this problem by setting the standard deviation for the accelerometers as the expected error due to its sensor non-linearities; a quantity typically found in the sensor data-sheets. For accelerometers in MPU-9150 [46], the expected error due to non-linearities is around 0.5% of the signal amplitude. A conservative value for the signal amplitude is the point of saturation, which for said accelerometers is around 157 $\mathrm{m/s^2}$. Hence, we set the accelerometer standard deviation to $\sigma^{(\mathrm{a})} = 0.79\ \mathrm{m/s^2}$. Scaling the accelerometer measurement covariance widens the posterior distribution of the estimated parameters.

C. Experiments

We evaluated the performance of the Bayesian estimator using real experiments performed on the array in Fig. 1. We collected measurements from the array while it was exposed to hand twisting motion. During one time sample we assume that the inertial sensor array samples all the IMUs simultaneously and instantaneously. The details about the hardware and the sampling process can be found in [47]. From this collection of measurements, three subsets of $T \in \{5, 50, 500\}$ were constructed. The three subsets did not share any time samples. The time samples were selected when there was a nonzero rotation of the array, since a non-rotating motion would not provide any information of the accelerometer positions. Guidelines of how to select informative samples can be found in [25], [26]. Fig. 5 depicts the posterior distribution of the accelerometer triad positions. For Bayesian estimators, the minimum mean square error estimate is the conditional mean of the posterior distribution. We note that the position estimates for underside accelerometer triads ($k = 2, 4, \ldots, 32$) are consistently located in the upper-right quadrant inside its IMU package, while topside accelerometer triads ($k = 1, 3, \ldots, 31$) are mainly located at the center. Underside IMU packages are rotated 180° around the axis $y = x$ relative to topside IMU packages. This relative positioning implies that the estimator places the accelerometer triad sensors in the upper-right quadrant of each package with half the shift from the center of what is observed in Fig. 5b, when viewed from above and when the package is oriented such that the orientation marker is in the upper-left quadrant. Since the origin of the coordinate system is on the topside (accelerometer triad 1's position), position estimates for topside accelerometer triads are not shifted, and positions estimates for underside accelerometer triads are doubly shifted. This result is consistent with the results in [26], where we used a maximum likelihood estimator to estimate the accelerometer triad positions in Fig. 1. Even if the three sensors in an accelerometer triad cannot be perfectly colocated inside the package, the results suggest that the accelerometer triad positions are consistently estimated. Similar results were observed over repeated experimental trails.

TABLE II
COMPARISON OF THE ESTIMATED STANDARD DEVIATIONS FOR DIFFERENT NUMBER OF TIME SAMPLES USED IN THE CALIBRATION

μm	$T = 5$	$T = 50$	$T = 500$
Simulations	1430.7	403.0	145.5
Experiments	1450.7	439.4	144.9

Moreover, we observe that the posterior distribution covariance shrinks for an increasing number of time samples used in the calibration, as we observed for the simulations in Section IV-A. We assess the uncertainty quantification by inspecting the confidence ellipses. Assuming perfectly mounted IMU packages on the PCB, the accelerometers position uncertainty would only stem from the internal placement of the accelerometer sensor axes inside the package. As mentioned, the accelerometer triad position estimates for topside IMUs are unshifted with respect to the origin of the coordinate system. Thus, assuming that each IMU package has the same systematic position error inside the package, the expected accelerometer triad positions are at the center of each package for IMUs on the topside. Viewing each accelerometer triad position estimate as an independent trail, the confidence ellipses at a level of $(1-1/31)\cdot 100\%$ of the posterior distribution should then include the center of the IMU package in 97% of the trails. From the posterior distribution computed with 500 time samples, the number of 97% confidence ellipses covering the center of the packages is $10/15$. Even though the sample size is small, this might suggest that the uncertainty estimation is somewhat underestimated. This underestimation could be attributed to the model mismatch described in Section IV-B and/or the idealistic assumption of perfect IMU mounting on the PCB.

A comparison of the estimated standard deviations between simulations and the experiments is shown in Table II. The simulations were conducted using similar motion dynamics as in the experiments and the accelerometer covariance was also set to the same values as in the experiments. We observe that the estimated standard deviation of the accelerometer triad positions from the simulations are first, similar to the experimental values, and second, approximately 100 times larger than the simulations in Fig. 4 conducted with a smaller covariance. These results are in line with [24], which also saw a similar discrepancy between simulations under idealized conditions and experiments. The resulting widening of the posterior distribution in the experiments can thus be attributed to the enlargement of the accelerometer covariance. The gap between the results of the idealized model and the experimental results could be ascribed to the inaccuracy of the first-order model assumption of the sensors.

V. SUMMARY AND CONCLUSION

We have developed an algorithm to estimate and quantify the uncertainty of the relative geometry of accelerometer triads in inertial sensor arrays. The calibration problem is posed as a Bayesian estimation problem, and we compute the posterior distribution of the accelerometer positions using MCMC. The problem's unknown motion parameters are numerically marginalized using a Rao-Blackwellized estimator that samples the non-linear motion parameters using information from the gyroscopes and analytically integrates out linear motion parameters. Thanks to this effective use of the model structure, the proposed MCMC method is shown to efficiently sample the posterior distribution in a relatively high dimensional space (93 dimensions), both on simulated and experimental data. The performance of the numerical marginalization naturally depends on the number of Monte Carlo samples used. However, increasing the number of time samples in the calibration necessitates an increase in the number of Monte Carlo samples, which increases the computational complexity. The computational time for 5, 50, and 500 time samples were 30 minutes, 1 hour, and 8 hours, respectively, using 40 Intel Xeon 2.2 GHz processors. Thus, the proposed calibration method may be more suitable for short measurement sequences and offline calibration. However, the proposed estimator assumes no time dependence between measurements, which allows the user to omit time samples before the calibration. Omitted time samples could include outliers or time samples when the inertial sensor array is stationary, which contain little information on the accelerometer positions. Moreover, we conducted experiments where the inertial sensor array was exposed to hand twisting motion. In the experiments, we assumed that the individual MEMS sensors on the array to follow a first-order model. The experimental results indicate that the first-order model may be an inadequate model for high dynamic applications, where error sources such as sensor non-linearities may be the dominating error. In our experiments, this could be mitigated by assuming a conservative covariance for the accelerometers, and future research could extend the sensor model to a second-order model. Despite this conservative assumption, the experimental results are consistent with the results given by a maximum likelihood estimator for a similar problem [26] and in line with similar studies [24], indicating the validity of the array signal model and the estimator.

APPENDIX A
LOG-SUM-EXP COMPUTATION

The logarithm of a sum of exponential terms, $\log \sum e^{x_i}$ may over- and underflow using floating-point arithmetic. Such sums can be effectively computed using the following trick to compute the log-sum-exp function [41]. First, for $a, b \in \mathbb{R}$ we have

$$\log(e^a + e^b) = \begin{cases} a + \log(1 + e^{b-a}), & a > b, \\ b + \log(1 + e^{a-b}), & b > a, \end{cases} \tag{31}$$

which is equivalent to

$$\log(e^a + e^b) = \max(a, b) + \log(1 + e^{-|a-b|}). \tag{32}$$

Then $\log \sum e^{x_i}$ can accurately be calculated recursively using (32).

REFERENCES

[1] D. K. Shaeffer, "MEMS inertial sensors: A tutorial overview," *IEEE Commun. Mag.*, vol. 51, no. 4, pp. 100–109, Apr. 2013.

[2] H. Martin, P. Groves, M. Newman, and R. Faragher, "A new approach to better low-cost MEMS IMU performance using sensor arrays," in *Proc. 26th Int. Tech. Meeting Satell. Division Inst. Navigat. (ION GNSS)*, Nashville, TN, USA, Sep. 2013, pp. 2125–2142.

[3] S. Bras, P. Rosa, C. Silvestre, and P. Oliveira, "Fault detection and isolation in inertial measurement units based on bounding sets," *IEEE Trans. Autom. Control*, vol. 60, no. 7, pp. 1933–1938, Jul. 2015.

[4] I. Skog, J.-O. Nilsson, P. Händel, and A. Nehorai, "Inertial sensor arrays, maximum likelihood, and Cramér–Rao bound," *IEEE Trans. Signal Process.*, vol. 64, no. 16, pp. 4218–4227, Aug. 2016.

[5] K. Pamadi, E. Ohlmeyer, and T. Pepitone, "Assessment of a GPS guided spinning projectile using an accelerometer-only IMU," in *Proc. AIAA Guid., Navigat., Control Conf. Exhib.*, Providence, RI, USA, Aug. 2004, pp. 1–13.

[6] D. B. Camarillo, P. B. Shull, J. Mattson, R. Shultz, and D. Garza, "An instrumented mouthguard for measuring linear and angular head impact kinematics in American football," *Ann. Biomed. Eng.*, vol. 41, no. 9, pp. 1939–1949, Apr. 2013.

[7] A. Waegli, J. Skaloud, S. Guerrier, M. E. Parés, and I. Colomina, "Noise reduction and estimation in multiple micro-electro-mechanical inertial systems," *Meas. Sci. Technol.*, vol. 21, no. 6, Apr. 2010, Art. no. 065201.

[8] S. J. Ovaska and S. Valiviita, "Angular acceleration measurement: A review," *IEEE Trans. Instrum. Meas.*, vol. 47, no. 5, pp. 1211–1217, Oct. 1998.

[9] M. Pachter, T. C. Welker, and R. E. Huffman, "Gyro-free INS theory," *Navigation*, vol. 60, no. 2, pp. 85–96, Jun. 2013.

[10] I. Klein, "Analytic error assessment of gyro-free INS," *J. Appl. Geodesy*, vol. 9, no. 1, pp. 49–62, Jan. 2015.

[11] I. Skog, J.-O. Nilsson, and P. Händel, "Pedestrian tracking using an IMU array," in *Proc. IEEE Int. Conf. Electron., Comput. Commun. Technol. (CONECCT)*, Bangalore, India, Jan. 2014, pp. 1–4.

[12] Y.-S. Kang, K. Moorhouse, and J. H. Bolte, "Measurement of six degrees of freedom head kinematics in impact conditions employing six accelerometers and three angular rate sensors (6aω configuration)," *J. Biomech. Eng.*, vol. 133, no. 11, Nov. 2011, Art. no. 111007.

[13] W. R. Bussone, J. Olberding, and M. Prange, "Six-degree-of-freedom accelerations: Linear arrays compared with angular rate sensors in impact events," *SAE Int. J. Transp. Saf.*, vol. 5, no. 2, pp. 194–207, Mar. 2017.

[14] D. Tcherniak and M. Schwaab, "On a method for finding position and orientation of accelerometers from their signals," *Mech. Syst. Signal Process.*, vol. 140, Jun. 2020, Art. no. 106662.

[15] J.-O. Nilsson and I. Skog, "Inertial sensor arrays—A literature review," in *Proc. Eur. Navigat. Conf. (ENC)*, Helsinki, Finland, May 2016, pp. 1–10.

[16] M. W. Givens and C. Coopmans, "A survey of inertial sensor fusion: Applications in sUAS navigation and data collection," in *Proc. Int. Conf. Unmanned Aircr. Syst. (ICUAS)*, Atlanta, GA, USA, Jun. 2019, pp. 1054–1060.

[17] H. Martin, P. Groves, and M. Newman, "The limits of in-run calibration of MEMS inertial sensors and sensor arrays," *Navigation*, vol. 63, no. 2, pp. 127–143, Jun. 2016.

[18] J.-O. Nilsson, I. Skog, and P. Händel, "Aligning the forces—Eliminating the misalignments in IMU arrays," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 10, pp. 2498–2500, Oct. 2014.

[19] D. Dubé and P. Cardou, "The calibration of an array of accelerometers," *Trans. Can. Soc. Mech. Eng.*, vol. 35, no. 2, pp. 251–267, Jun. 2011.

[20] K. Parsa, T. A. Lasky, and B. Ravani, "Design and implementation of a mechatronic, all-accelerometer inertial measurement unit," *IEEE/ASME Trans. Mechatronics*, vol. 12, no. 6, pp. 640–650, Dec. 2007.

[21] P. Cappa, F. Patane, and S. Rossi, "Two calibration procedures for a gyroscope-free inertial measurement system based on a double-pendulum apparatus," *Meas. Sci. Technol.*, vol. 19, no. 5, p. 55204, Apr. 2008.

[22] R. Zhang, F. Hoflinger, and L. M. Reind, "Calibration of an IMU using 3-D rotation platform," *IEEE Sensors J.*, vol. 14, no. 6, pp. 1778–1787, Jun. 2014.

[23] T. Nieminen, J. Kangas, S. Suuriniemi, and L. Kettunen, "An enhanced multi-position calibration method for consumer-grade inertial measurement units applied and tested," *Meas. Sci. Technol.*, vol. 21, no. 10, Aug. 2010, Art. no. 105204.

[24] P. Schopp, H. Graf, W. Burgard, and Y. Manoli, "Self-calibration of accelerometer arrays," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 8, pp. 1913–1925, Aug. 2016.

[25] H. Carlsson, I. Skog, and J. Jaldén, "On-the-fly geometric calibration of inertial sensor arrays," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Sep. 2017, pp. 1–6.

[26] H. Carlsson, I. Skog, and J. Jalden, "Self-calibration of inertial sensor arrays," *IEEE Sensors J.*, vol. 21, no. 6, pp. 8451–8463, Mar. 2021.

[27] F. Olsson, M. Kok, K. Halvorsen, and T. B. Schön, "Accelerometer calibration using sensor fusion with a gyroscope," in *Proc. IEEE Stat. Signal Process. Workshop (SSP)*, Palma de Mallorca, Spain, Jun. 2016, pp. 1–5.

[28] M. Kok and T. B. Schön, "Magnetometer calibration using inertial sensors," *IEEE Sensors J.*, vol. 16, no. 14, pp. 5679–5689, Jul. 2016.

[29] F. Olsson, M. Kok, T. Seel, and K. Halvorsen, "Robust plug-and-play joint axis estimation using inertial sensors," *Sensors*, vol. 20, no. 12, p. 3534, Jun. 2020.

[30] J. Wahlström, I. Skog, P. Händel, and A. Nehorai, "IMU-based smartphone-to-vehicle positioning," *IEEE Trans. Intell. Vehicles*, vol. 1, no. 2, pp. 139–147, Jun. 2016.

[31] L. D. Landau and E. M. Lifshitz, *Mechanics*, 3rd ed. Oxford, U.K.: Butterworth-Heinemann, 1976.

[32] M. A. Beaumont, "Estimation of population growth or decline in genetically monitored populations," *Genetics*, vol. 164, no. 3, pp. 1139–1160, Jul. 2003.

[33] C. Andrieu and G. O. Roberts, "The pseudo-marginal approach for efficient Monte Carlo computations," *Ann. Statist.*, vol. 37, no. 2, pp. 697–725, Apr. 2009.

[34] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, 2nd ed. New York, NY, USA: Springer, 2004.

[35] J. Dahlin and T. B. Schön, "Getting started with particle metropolis-hastings for inference in nonlinear dynamical models," *J. Stat. Softw.*, vol. 88, no. 2, pp. 1–41, 2019.

[36] J. S. Liu, *Monte Carlo Strategies in Scientific Computing*. New York, NY, USA: Springer, 2004.

[37] T. Schön, F. Gustafsson, and P.-J. Nordlund, "Marginalized particle filters for mixed linear/nonlinear state-space models," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2279–2289, Jul. 2005.

[38] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications*, 2nd ed. New York, NY, USA: Springer, 2003.

[39] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. New York, NY, USA: Cambridge Univ. Press, 2012.

[40] M. R. Osborne, "Separable least squares, variable projection, and the Gauss-Newton algorithm," *Electron. Trans. Numer. Anal.*, vol. 28, no. 1, pp. 1–15, Feb. 2007.

[41] P. Blanchard, D. J. Higham, and N. J. Higham, "Accurately computing the log-sum-exp and softmax functions," *IMA J. Numer. Anal.*, pp. 1–20, Aug. 2020.

[42] G. O. Roberts and J. S. Rosenthal, "Optimal scaling for various Metropolis-Hastings algorithms," *Stat. Sci.*, vol. 16, no. 4, pp. 351–367, Nov. 2001.

[43] C. Sherlock, A. H. Thiery, G. O. Roberts, and J. S. Rosenthal, "On the efficiency of pseudo-marginal random walk metropolis algorithms," *Ann. Statist.*, vol. 43, no. 1, pp. 238–275, Feb. 2015.

[44] A. Doucet, M. K. Pitt, G. Deligiannidis, and R. Kohn, "Efficient implementation of Markov chain Monte Carlo when using an unbiased likelihood estimator," *Biometrika*, vol. 102, no. 2, pp. 295–313, Mar. 2015.

[45] S. P. Brooks and G. O. Roberts, "Convergence assessment techniques for Markov chain Monte Carlo," *Stat. Comput.*, vol. 8, no. 4, pp. 319–335, Dec. 1998.

[46] *MPU-9150 Product Specification, Rev. 4.3*, InvenSense, San Jose, CA, USA, Sep. 2013.

[47] I. Skog, J.-O. Nilsson, and P. Händel, "An open-source multi inertial measurement unit (MIMU) platform," in *Proc. Int. Symp. Inertial Sensors Syst. (ISISS)*, Laguna Beach, CA, USA, Feb. 2014, pp. 1–4.

Håkan Carlsson (Student Member, IEEE) received the M.Sc. degree in engineering physics from the Chalmers University of Technology, Gothenburg, Sweden, in 2013. He is currently pursuing the Ph.D. degree in signal processing with the Division of Information Science and Engineering, KTH Royal Institute of Technology, Stockholm, Sweden. His main research interests include sensor fusion and sensor calibration.

Isaac Skog (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering from the KTH Royal Institute of Technology, Stockholm, Sweden, in 2003 and 2005, respectively, and the Ph.D. degree in signal processing, in 2010, with a thesis on low-cost navigation systems. In 2009, he spent five months with the Mobile Multi-Sensor System Research Team, University of Calgary, Canada, as a Visiting Scholar. In 2011, he spent four months with the Indian Institute of Science (IISc), Bengaluru, India, as a Visiting Scholar. From 2010 to 2017, he was a Researcher with the KTH Royal Institute of Technology. He is currently an Assistant Professor with Linköping University, Linköping, Sweden, and a Senior Researcher with the Swedish Defence Research Agency (FOI), Stockholm. He was a recipient of the Best Survey Paper Award by the IEEE Intelligent Transportation Systems Society in 2013.

Thomas B. Schön (Senior Member, IEEE) received the B.Sc. degree in business administration and economics, the M.Sc. degree in applied physics and electrical engineering, and the Ph.D. degree in automatic control from Linköping University, Sweden, in January 2001, September 2001, and February 2006, respectively. He has held a visiting positions with the University of Cambridge, U.K., the University of Newcastle, Australia, and Universidad Técnica Federico Santa María, Valparaíso, Chile. In 2018, he was elected to the Royal Swedish Academy of Engineering Sciences (IVA) and the Royal Society of Sciences in Uppsala. He is the Beijer Professor of Artificial Intelligence with the Department of Information Technology, Uppsala University. He has a broad interest in developing new algorithms and mathematical models capable of learning and acting based on data. He is a Fellow of the ELLIS society. He received the Tage Erlander Prize for Natural Sciences and Technology by the Royal Swedish Academy of Sciences (KVA) in 2017 and the Arnberg Prize by the Royal Swedish Academy of Sciences (KVA) in 2016. He was awarded the Automatica Best Paper Prize in 2014. He received the Best Ph.D. Thesis Award by The European Association for Signal Processing (EURASIP) in 2013 and the Best Teacher Award at the Institute of Technology, Linköping University, in 2009.

Joakim Jaldén (Senior Member, IEEE) received the M.Sc. and Ph.D. degrees in electrical engineering from the KTH Royal Institute of Technology, Stockholm, Sweden, in 2002 and 2007, respectively.

From July 2007 to June 2009, he held a postdoctoral research position with the Vienna University of Technology, Vienna, Austria. From September 2000 to May 2002, he also studied with Stanford University, CA, USA. From August to September, 2008, he worked as a Visiting Researcher with ETH, Zürich, Switzerland. In July 2009, he returned to KTH, where he is currently a Professor of Signal Processing. His recent work includes work on signal processing for biomedical data analysis, including single-cell tracking for time-lapse microscopy and inverse diffusion for immunoassays. He was an Associate Editor of the IEEE COMMUNICATIONS LETTERS from 2009 to 2011 and the IEEE TRANSACTIONS IN SIGNAL PROCESSING from 2012 to 2016. He has been a member of the IEEE Signal Processing for Communications and Networking Technical Committee (SPCOM-TC) since 2013, where he serves as the Chair (2019–2020). Since 2016, he has been responsible for the five-year B.Sc. and M.Sc. degree program in electrical engineering at KTH. For his work on MIMO communications, he has been awarded the IEEE Signal Processing Society's 2006 Young Author Best Paper Award and the Best Student Conference Paper Award at IEEE ICASSP 2007. He was also a recipient of the Ingvar Carlsson Career Award issued in 2009 by the Swedish Foundation for Strategic Research.