2018 European Control Conference (ECC)
June 12-15, 2018. Limassol, Cyprus

Bayesian strategies for calibrating heteroskedastic static sensors with unknown model structures

Anas Alhashimi, Simone Del Favero, Damiano Varagnolo, Thomas Gustafsson, Gianluigi Pillonetto

***Abstract*— This paper investigates the problem of calibrating sensors affected by (i) heteroskedastic measurement noise and (ii) a polynomial bias, describing a systematic distortion of the measured quantity. First, a set of increasingly complex statistical models for the measurement process was proposed. Then, for each model the authors design a Bayesian parameters estimation method handling heteroskedasticity and capable to exploit prior information about the model parameters. The Bayesian problem is solved using MCMC methods and reconstructing the unknown parameters posterior in sampled form. The authors then test the proposed techniques on a practically relevant case study, the calibration of Light Detection and Ranging (Lidar) sensor, and evaluate the different proposed procedures using both artificial and field data.**

I. Introduction

Assume that a measurement system is affected by a nonlinear bias and heteroskedastic measurement noise, so that both the bias and the variance of the measurement noise depend on the state of the system. A practically relevant case study considered in this paper is the Lidar sensor, that evaluates the distance to a target by illuminating that target with a pulsed laser light, and measuring the reflected pulses on CCD camera (for detailed description about the Lidar construction and operation see [8], [11]). An example of dataset is depicted in Figure 1. The measurements are taken by a wheeled robot mounting a triangulation Lidar, moving on a flat surface away from a fixed obstacle, nominally following a straight path, see Figure 2. Figure 1 shows clearly that the measurements collected are corrupted by a heteroskedastic noise.

Assume that, as in the particular case of Figure 1, the user has information on the process (i.e., on the actual position of the robot) but no model of the sensor from a data-sheet or other sources. The general problem is then to characterize the bias and variance of the sensor starting from the available training set (see Section II for a more formal statement).

This research deals with the sensor calibration problem, i.e. with the problem to estimate numerically the parameters describing the error noise, assuming a suitable model structure has been selected. Notice that this paper discuss the calibration of the distance measuring mechanism not the rotating mechanism of the sensor.

This paper follows the following approach: starting from a generic training set like the one in Figure 1, the paper: *1)* lists a set of different plausible models of the heteroskedasticity of generic sensors; *2)* derives for each model structure a tailored model-parameter estimation algorithm.

Literature review

Estimating model parameters with presence of heteroskedasticity is a challenging task and usually Ordinary Least Squares (OLS) can be used, the OLS solution to the parameter estimation problem is shown to be unbiased but also an inefficient estimator especially at large heteroskedasticity [1], however, the OLS gives biased estimate of the variance and incorrect confidence intervals and statistical inference [2].

The noise structure in the heteroskedastic systems can either be dependent on the states or independent on the states. For the case where the noise structure is independent on the states, Gibbs sampler proposed for inference in a linear

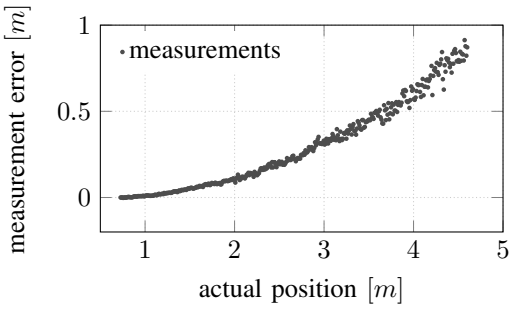

Fig. 1. Example of distance measurements collected by a triangulation Lidar . The measurements are visibly corrupted by a heteroskedastic noise.

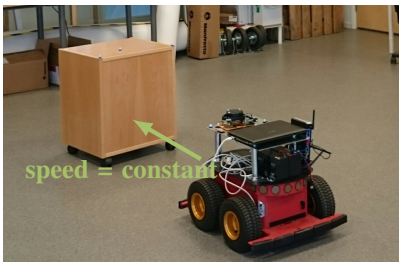

Fig. 2. Photo of the experimental setup used for collecting the dataset. The photo shows the robot, the Lidar and the wooden target.

A. Alhashimi, D. Varagnolo and T. Gustafsson are with the Control Engineering Group, Department of Computer Science, Electrical and Space Engineering, Luleå University of Technology, Luleå 97187, Sweden. Emails: { anas.alhashimi | damiano.varagnolo | thomas.gustafsson } @ ltu.se

The research leading to these results was partially supported by Norrbottens Forskningsråd and University of Baghdad.

heteroskedastic Bayesian model by [3], the measurements were assumed to be independent and identically distributed with Student-t distribution, the author also showed the equivalence in posteriors between the Student-t measurements case and an appropriate mixture of Gaussian measurements. However, this research will focus on the case where the noise structure depends on the states, more specifically when the noise structure is a function with unknown parameters of the states. This kind of models have been studied in econometrics literature, for example [4] proposed a two-step estimation procedure for models where the i^{th} disturbance variance σ_i is of the form $\sigma_i = \sigma^2 x_i^{\lambda}$ where x_i is the state, σ^2 and λ are unknown model parameters. This procedure has been examined in details for more general variance models similar to $\sigma_i = e^{x_i \lambda}$ and compared with the iterative Maximum Likelihood (ML) by [5]. The analysis of [5] shows that the two-step estimation procedure is inconsistent in the heteroskedasticity parameters and then proposed a modified two-step estimation procedure which is consistent. Bayesian estimators for heteroskedastic systems proposed by [6] for variance models similar to $\sigma_i = \sigma^2 x_i^{-\lambda}$. [7] extended the Bayesian approach of [6] to multidimensional and estimated the parameters through opportune Markov chain Monte Carlo (MCMC) sampler then compared the performance with the previous estimators (the two-step and iterative ML), it has been shown that Bayesian estimators give better performance in the Root-Mean-Square Error (RMSE) and the interquartile range sense. The previously cited papers [4], [5], [6], [7] assume models useful for econometrics, however in this research, the authors will present Bayesian analysis for parameter estimation in generic non-linear and heteroskedastic systems, in which, the non-linearity is modeled through a polynomial of opportune order. An example of systems with heteroskedastic noise and non-linear bias is the triangulation Lidar used with robotic applications.

The calibration of the intrinsic parameters of triangulation Lidars have been studied in [8], [9], [10], [11], [12], [13]. More specifically, the technology introduced in [8] performed an early-assessment of the potential for the triangulation Lidars technology. These devices are affected by nonlinearities, as discussed in [9], and by the color of the target, as discussed in [10]. A first calibration procedure that builds on a statistical model of the sensor was proposed in [9] where the model was assumed to be homoskedastic. Heteroskedastic model for the sensor were proposed by [11] with a more general calibration procedure based on Weighted Least Squares (WLS) for parameter estimation and Akaike Information Criterion (AIC) for model selection. This model was extended in [12] to include the beam angle, also a calibration procedure suitable for targets with flat surfaces was proposed. In addition, there have been proposed calibration procedures that do not require independent sources of groundtruth information (unknown states), as in [13] where an approximated Expectation Maximization (EM) procedure is used for joint parameter and states estimation.

Notice that the calibration procedure proposed in [13] has been the inspirational work for the current manuscript. The problem of [13] is nonetheless that it maximizes an approximated likelihood function. Here, instead, this research aims at:

i) maximize the exact likelihood function rather than an approximated one;
ii) using our prior information about the process under the Bayesian framework;
iii) make a preliminary study for the more generic case (and of greater practical importance) that the authors would like to solve in the future, in which, the states are unknown;
iv) generalize the procedure so that it can go beyond the triangulation Lidar case and be applied to more general frameworks.

Statement of contributions

The authors consider the problem of estimating static measurement models starting from training set of inputs-outputs pairs (see Section II for more details).

Our contributions are thus:

i) propose a set of different statistical models for the measurement process that capture the most common statistical behaviors of static (i.e., not subject to hysteresis) sensors;
ii) design a set of different parameters estimation techniques that solve the sensor calibration problem described above;
iii) implement the various estimation algorithms above on some real case scenarios and assess their performance using fit indexes and computational complexity perspectives.

Structure of the manuscript

The next Section II presents the problem formulation. Then Sections III describes the details of the parameter estimation methods. The obtained Numerical results for both simulations and real datasets will be presented in Section IV. And finally the conclusion in Section V.

II. Problem statement

Assume that an unspecified sensor transforms the state $x_i \in \mathbb{R}$ of an unspecified system into a corresponding measurement y_i following the nonlinear static model

$$y_i = f_{\text{mean}}(x_i) + f_{\text{noise}}(x_i)\,\nu_i \tag{1}$$

where the term $f_{\text{mean}}(x_i)$ models the bias as a static nonlinear function depending on the state of the system, the static function $f_{\text{noise}}(\cdot)$ captures the heteroskedasticity of the measurement system and $\nu_i \sim \mathcal{N}(0,1)$ and independent and identically distributed (iid).

i	time / measurement index ($i \in \{1, \dots, M\}$)
$\mathcal{T}$	training dataset
$\mathcal{V}$	verification dataset
M	number of measurements in training dataset
L	number of measurements in verification dataset
x_i	value of the state at time i
y_i	sensor measurement at time i
k	index of the MCMC run
k_{max}	maximal number of MCMC iterations
k_{min}	burn-in MCMC iterations
$f_{\text{mean}}(\cdot)$	(static) measurement bias (see (1))
$\boldsymbol{\alpha}$	parameters of the polynomial defining $f_{\text{mean}}(\cdot)$ (see (3))
N	model order for the bias term f_{mean}
$\boldsymbol{\mu}_{\boldsymbol{\alpha}}$, $\Sigma_{\boldsymbol{\alpha}}$, $\boldsymbol{\tau}_{\boldsymbol{\alpha}}$	parameters defining the prior for $\boldsymbol{\alpha}$ (see (5))
$\boldsymbol{G}_{\boldsymbol{x}}$	Vandermonde matrix associated to the set of states $\boldsymbol{x}$ (see (4))
$f_{\text{noise}}(\cdot)$	(static) heteroskedastic standard deviation of the measurement noise (see (6))
ρ	power factor defining the actual variance of the noise term f_{noise} (see (6))
σ_ν	nominal standard deviation of the measurement noise (see (6))
$\tau_\nu = \sigma_\nu^{-2}$	nominal precision of the measurement noise
a_ν, b_ν	parameters defining the prior for σ_ν (see (7))
ν_i	normalized measurement noise at time i
Σ_ν	covariance of the measurement noise vector (see (2))
γ	acceptance probability
β, β'	proposal variances

TABLE I

SUMMARY AND MEANING OF THE MOST COMMON SYMBOLS.

Assume that, during an ad-hoc sensor calibration experiment, M measurements have been collected from (1) at perfectly known values of the state x_i. The vector of measurements can then be written compactly as

$$\boldsymbol{y} = \boldsymbol{f}_{\text{mean}}(\boldsymbol{x}) + \boldsymbol{f}_{\text{noise}}(\boldsymbol{x}) \odot \boldsymbol{\nu}$$

where $\boldsymbol{y} := [y_1, \dots, y_M]^T$, $\boldsymbol{x} := [x_1, \dots, x_M]^T$, $\boldsymbol{\nu} := [\nu_1, \dots, \nu_M]^T$. $\boldsymbol{f}_{\text{mean}}(\boldsymbol{x}) \in \mathbb{R}^M$ and $\boldsymbol{f}_{\text{noise}}(\boldsymbol{x}) \in \mathbb{R}^M$ is defined similarly (noticing that it can be specialized in three different ways conforming with (6)), and $\odot$ indicates the Hadamard product. The measurement noise covariance matrix is therefore

$$\Sigma_\nu := \text{diag}\left(f_{\text{noise}}(x_1)^2, \dots, f_{\text{noise}}(x_M)^2\right). \quad (2)$$

The authors assume also that the user steers the unspecified system opportunely in a way that is informative for calibrating the sensor. For instance, in the case of Figure 1, the sensor is attached to a wheeled robot moving with nominally constant speed on a flat surface towards a fixed target, so to "explore" all the potential statistical dependencies of y_i on x_i.

For what concerns $f_{\text{mean}}(x_i)$ the authors assume that it can be captured through a polynomial of opportune order N, i.e.,

$$f_{\text{mean}}(x_i) = \underbrace{\begin{bmatrix}1 & x_i & x_i^2 & \dots & x_i^N\end{bmatrix}}_{=: G_{x_i}} \underbrace{\begin{bmatrix}\alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N\end{bmatrix}}_{=: \boldsymbol{\alpha}} = G_{x_i}\boldsymbol{\alpha} \quad (3)$$

for a suitable value of N known and fixed and hence as a consequence $\boldsymbol{f}_{\text{mean}}(\boldsymbol{x})$ will be

$$\boldsymbol{f}_{\text{mean}}(\boldsymbol{x}) = \underbrace{\begin{bmatrix}1 & x_1^1 & \cdots & x_1^N \\ \vdots & & \vdots & \\ 1 & x_M^1 & \cdots & x_M^N\end{bmatrix}}_{=: G_{\boldsymbol{x}}} \underbrace{\begin{bmatrix}\alpha_0 \\ \vdots \\ \alpha_N\end{bmatrix}}_{=: \boldsymbol{\alpha}} = G_{\boldsymbol{x}}\boldsymbol{\alpha}, \quad (4)$$

This paper describes the unknown parameter vector $\boldsymbol{\alpha}$ in (3) as random variable with a known Gaussian prior $\boldsymbol{\alpha} \sim \mathcal{N}(\mu_{\boldsymbol{\alpha}}, \Sigma_{\boldsymbol{\alpha}})$, where

$$\mu_{\boldsymbol{\alpha}} := \begin{bmatrix}0 & 1 & 0 & \cdots & 0\end{bmatrix}^T \quad \Sigma_{\boldsymbol{\alpha}} := \text{diag}\left(\boldsymbol{\tau}_{\boldsymbol{\alpha}}^{-2}\right) \quad (5)$$

and with the vector of precisions $\boldsymbol{\tau}_{\boldsymbol{\alpha}}$ known. Notice that this particular $\mu_{\boldsymbol{\alpha}}$ captures the fact that, a priori, a sensor should follow an expected ideal behavior, i.e., should be s.t. $y_i = x_i$.

The static map $f_{\text{noise}}(\cdot)$ captures the heteroskedasticity of the measurement system.

and the following three cases will be considered:

$$\begin{aligned}&\text{Case I:} && f_{\text{noise}}(x_i) &&= \sigma_\nu \\ &\text{Case II:} && f_{\text{noise}}(x_i) &&= \sigma_\nu x_i^\rho \\ &\text{Case III:} && f_{\text{noise}}(x_i) &&= \sigma_\nu f_{\text{mean}}(x_i)^\rho\end{aligned} \quad (6)$$

Case I describe homoskedastic sensors, Case II sensors whose heteroskedasticity depends on the actual state of the system and Case III sensors whose heteroskedasticity depends on the expected measured state of the system.

σ_ν, called the *nominal standard deviation* of the measurement noise, is also a random variable with prior

$$\sigma_\nu^{-2} = \tau_\nu \sim \text{Gamma}(a_\nu, b_\nu) \quad (7)$$

with known shape and scale parameters a_ν and b_ν (this manuscript will also call $\tau_\nu := \sigma_\nu^{-2}$ the *nominal precision* of the measurement noise). $\rho \in \mathbb{R}_+$ is an unknown parameter to be estimated from the data, having a truncated Gaussian prior and constitutes an other degree of freedom allowing for further modeling flexibility.

The problem of *sensor calibration* consists of obtaining estimates of the functions $f_{\text{mean}}(\cdot)$ and $f_{\text{noise}}(\cdot)$ starting from a training set. Thus the authors aim to estimate the parameters $\boldsymbol{\alpha}$, σ_ν^{-2} and ρ when present.

III. OVERVIEW OF THE METHODOLOGY

The authors propose three different algorithms for the three models of $f_{\text{noise}}(\cdot)$ considered. Each algorithm reconstructs the posterior distribution of the unknown parameters with MCMC techniques. In the Case I the target posterior density is $p(\boldsymbol{\alpha}, \tau_\nu | \boldsymbol{x}, \boldsymbol{y})$ while in Cases II and III this is $p(\boldsymbol{\alpha}, \rho, \tau_\nu | \boldsymbol{x}, \boldsymbol{y})$. The Maximum A Posteriori (MAP) estimators of the unknown model parameters with MCMC techniques is then computed. It is important to remark that N is assumed known at this stage.

1) Case I, $f_{noise}(x_i) = \sigma_\nu$:

Since x is known and N fixed, $G_{\boldsymbol{x}}$ is a fixed matrix and our problem reduces to the classical problem of computing the MAP estimate of $\boldsymbol{\alpha}$ from a homoskedastic noisy measurements of a linear transformation of the unknown parameters,

with the additional complexity that also the model noise variance is unknown. This problem can be solved using a Gibbs sampler, given that all the priors and likelihoods are conjugate, resulting in Algorithm 1.

Algorithm 1 MCMC for the case $f_{\text{noise}}(x_i) = \sigma_\nu$

initialization:

$$\boldsymbol{\alpha}^{(0)} = \mu_{\boldsymbol{\alpha}} \qquad \tau_\nu^{(0)} \sim \text{Gamma}\,(a_\nu, b_\nu)$$

for $k = 0, 1, \ldots k_{max}$ **do**

update τ_ν and $\boldsymbol{\alpha}$ using the Gibbs sampler:

$$\begin{aligned}\boldsymbol{\alpha}^{(k+1)} &\sim p\left(\boldsymbol{\alpha}^{(k)} \middle| \boldsymbol{x}, \tau_\nu^{(k)}\right)\\ \tau_\nu^{(k+1)} &\sim p\left(\tau_\nu^{(k)} \middle| \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\alpha}^{(k+1)}\right)\end{aligned}$$

where

$$p\left(\boldsymbol{\alpha}^{(k)} \middle| \boldsymbol{x}, \boldsymbol{y}, \tau_\nu^{(k)}\right) \propto \mathcal{N}\left(B^{(k)}A^{(k)}, B^{(k)}\right)$$

$$A^{(k)} = \tau_\nu^{(k)} G_{\boldsymbol{x}}^T \boldsymbol{y} - \Sigma_{\boldsymbol{\alpha}}^{-1}\mu_{\boldsymbol{\alpha}}$$

$$B^{(k)} = \left(\tau_\nu^{(k)} G_{\boldsymbol{x}}^T G_{\boldsymbol{x}} + \Sigma_{\boldsymbol{\alpha}}^{-1}\right)^{-1}$$

$$p\left(\tau_\nu^{(k)} \middle| \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\alpha}^{(k+1)}\right) \propto$$

$$\text{Gamma}\left(a_\nu + \frac{M}{2}, \left(\frac{1}{b_\nu} + \frac{1}{2}C^{(k+1)^T}C^{(k+1)}\right)^{-1}\right)$$

$$C^{(k+1)} = \left(\boldsymbol{y} - G_{\boldsymbol{x}}\boldsymbol{\alpha}^{(k+1)}\right)$$

2) Case II, $f_{noise}(x_i) = \sigma_\nu x_i^\rho$:
in this case ρ also has to be estimated from the data. Σ_ν becomes:

$$\Sigma_\nu = \sigma_\nu^2 \text{diag}\left(x_1^{2\rho}, \ldots, x_M^{2\rho}\right)$$

and the MAP estimator is

$$\arg \max_{\boldsymbol{\alpha}\in\mathbb{R}^N, \sigma_\nu^2\in\mathbb{R}_+, \rho\in\mathbb{R}_+} p\left(\boldsymbol{\alpha}, \sigma_\nu^2, \rho \,|\boldsymbol{x}, \boldsymbol{y}\right)$$

It is possible to reconstruct the posterior in sampled form using an MCMC approach. To do so, it has to be noted that $p\left(\rho \middle| \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\alpha}^{(k+1)}, \tau_\nu^{(k+1)}\right)$ is not known, hence the authors suggest to resort to Single Component Metropolis-Hasting scheme, resulting in Algorithm 2.

Remark 1 *The authors notice that the hyper–parameter β determining the proposal variance of the scheme requires manual tuning. As suggested in the literature, it is beneficial to tune β so that the acceptance ratio of the sampler lies around 44% (see [14] and Table 1 in [15] for more details).*

3) Case III, $f_{noise}(x_i) = \sigma_\nu f_{mean}\,(x_i)^\rho$:
in this case Σ_ν can be rewritten as

$$\Sigma_\nu = \sigma_\nu^2 \text{diag}\Big(f_{\text{mean}}(x_1)^{2\rho}, \ldots, f_{\text{mean}}(x_M)^{2\rho}\Big)$$

As before, the posterior $p\left(\rho \middle| \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\alpha}^{(k+1)}, \tau_\nu^{(k+1)}\right)$ is not known but in this case also the posterior $p\left(\boldsymbol{\alpha} \middle| \boldsymbol{x}, \boldsymbol{y}, \tau_\nu^{(k+1)}, \rho^k\right)$ is not known and have to be computed resorting to a acceptance/rejection mechanism, resulting in algorithm 3.

Notice also that the hyperparameters β and β' requires manual tuning so that the acceptance ratio of the sampler lies between 20% and 31% depending on the length of parameter vector to estimate (see Table 1 in [15] for more details).

Algorithm 2 MCMC for the case $f_{\text{noise}}(x_i) = \sigma_\nu x_i^\rho$

initialization:

$$\begin{aligned}\boldsymbol{\alpha}^{(0)} &= \mu_{\boldsymbol{\alpha}}\\ \rho^{(0)} &= 0\\ \tau_\nu^{(0)} &\sim \text{Gamma}\,(a_\nu, b_\nu)\end{aligned}$$

for $k = 0, 1, \ldots$ (up to convergence or up to a maximum number of iterations k_{max}) **do**

update τ_ν and $\boldsymbol{\alpha}$ using the Gibbs sampler:

$$\begin{aligned}\boldsymbol{\alpha}^{(k+1)} &\sim p\left(\boldsymbol{\alpha}^{(k)} \middle| \boldsymbol{x}, \tau_\nu^{(k)}, \rho^{(k)}\right)\\ \tau_\nu^{(k+1)} &\sim p\left(\tau_\nu^{(k)} \middle| \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\alpha}^{(k+1)}, \rho^{(k)}\right)\end{aligned}$$

generate new proposal:

$$\rho^{(k+1)} \sim \mathcal{N}\left(\rho^{(k)}, \beta\right)$$

calculate the acceptance probability:

$$\gamma = \min\left[1, \frac{p\left(\boldsymbol{y} \middle| \boldsymbol{x}, \rho^{(k+1)}, \boldsymbol{\alpha}^{(k+1)}, \tau_\nu^{(k+1)}\right)}{p\left(\boldsymbol{y} \middle| \boldsymbol{x}, \rho^{(k)}, \boldsymbol{\alpha}^{(k+1)}, \tau_\nu^{(k+1)}\right)} P\right]$$

where

$$P = \frac{p\left(\rho^{(k+1)}\right)}{p\left(\rho^{(k)}\right)}$$

accept the proposal if $\gamma > \mathcal{U}\,[0, 1]$ and $0 \geq \rho \geq 10$.

where

$$p\left(\boldsymbol{\alpha}^{(k)} \middle| \boldsymbol{x}, \boldsymbol{y}, \tau_\nu^{(k)}, \rho^{(k)}\right) \propto \mathcal{N}\left(B^{(k)}A^{(k)}, B^{(k)}\right)$$

$$A^{(k)} = \tau_\nu^{(k)} G_{\boldsymbol{x}}^T D^{(k)} \boldsymbol{y} - \Sigma_{\boldsymbol{\alpha}}^{-1}\mu_{\boldsymbol{\alpha}}$$

$$B^{(k)} = \left(\tau_\nu^{(k)} G_{\boldsymbol{x}}^T D^{(k)} G_{\boldsymbol{x}} + \Sigma_{\boldsymbol{\alpha}}^{-1}\right)^{-1}$$

$$D^{(k)} = \text{diag}\left(x_1^{-2\rho^{(k)}}, \ldots, x_M^{-2\rho^{(k)}}\right)$$

$$p\left(\tau_\nu^{(k)} \middle| \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\alpha}^{(k+1)}, \rho^{(k)}\right) \propto$$

$$\text{Gamma}\left(a_\nu + \frac{M}{2}, \left(\frac{1}{b_\nu} + \frac{1}{2}C^{(k+1)^T} D^{(k)} C^{(k+1)}\right)^{-1}\right)$$

$$C^{(k+1)} = \left(\boldsymbol{y} - G_{\boldsymbol{x}}\boldsymbol{\alpha}^{(k+1)}\right)$$

$$p\left(\rho^{(k)}\right) \propto \mathcal{N}\,(0, 1)$$

IV. Numerical Results

The authors evaluate the different proposed procedures using both simulated and field data.

A. Simulated Dataset

The simulated data were used to plot a typical posterior for the case $f_{\text{noise}}(x_i) = \sigma_\nu f_{\text{mean}}\,(x_i)^\rho$, were the actual model order is $N = 3$. More precisely, the ρ–τ_ν space was plotted in Figure 3 using dataset of both 50 and 900 samples, the $\boldsymbol{\alpha}$–τ_ν space in Figure 4 left plots, and the $\boldsymbol{\alpha}$–ρ space in Figure 4 on the right plots to show that for each pair of parameters the posterior is in this case unimodal.

Evaluation of the parameter estimation procedures: Figures 5 and 6 show the estimated densities of the various

Algorithm 3 MCMC for the case $f_{\text{noise}}(x_i) = \sigma_\nu f_{\text{mean}}(x_i)^\rho$

1: initialization:

$$\begin{aligned} \boldsymbol{\alpha}^{(0)} &= \mu_{\boldsymbol{\alpha}} \\ \rho^{(0)} &= 0 \\ \tau_\nu^{(0)} &\sim \text{Gamma}\,(a_\nu, b_\nu) \end{aligned}$$

2: **for** $k = 0, 1, \ldots$ (up to convergence or up to a maximum number of iterations k_{max}) **do**

3: update τ_ν using the Gibbs sampler

$$\tau_\nu^{(k+1)} \sim p\left(\tau_\nu^{(k)} \middle| \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\alpha}^{(k)}\right)$$

4: generate new proposals:

$$\boldsymbol{\alpha}^{(k+1)} \sim \mathcal{N}\left(\boldsymbol{\alpha}^{(k)}, \beta\right)$$

$$\rho^{(k+1)} \sim \mathcal{N}\left(\rho^{(k)}, \beta'\right)$$

5: calculate the acceptance probability:

$$\gamma = \min\left[1, \frac{p\left(\boldsymbol{y} \middle| \boldsymbol{x}, \rho^{(k+1)}, \boldsymbol{\alpha}^{(k+1)}, \tau_\nu^{(k+1)}\right)}{p\left(\boldsymbol{y} \middle| \boldsymbol{x}, \rho^{(k)}, \boldsymbol{\alpha}^{(k)}, \tau_\nu^{(k+1)}\right)} P\right]$$

where

$$P = \frac{p\left(\rho^{(k+1)}\right) p\left(\boldsymbol{\alpha}^{(k+1)}\right)}{p\left(\rho^{(k)}\right) p\left(\boldsymbol{\alpha}^{(k)}\right)}$$

6: accept the proposal if $\gamma > \mathcal{U}[0,1]$ and $0 \geq \rho \geq 10$.

where

$$p\left(\boldsymbol{\alpha}^{(k)}\right) \sim \mathcal{N}\left(\mu_{\boldsymbol{\alpha}}, \Sigma_{\boldsymbol{\alpha}}\right)$$

$$p\left(\tau_\nu^{(k)} \middle| \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\alpha}^{(k)}, \rho^{(k)}\right) \propto \text{Gamma}\left(a_\nu + \frac{M}{2}, \left(\frac{1}{b_\nu} + \frac{1}{2} C^{(k)^T} E^{(k)} C^{(k)}\right)^{-1}\right)$$

$$E^{(k)} = \text{diag}\left(f_{\text{mean}}(x_1)^{-2\rho^{(k)}}, \ldots, f_{\text{mean}}(x_M)^{-2\rho^{(k)}}\right)$$

$$p\left(\rho^{(k)}\right) \propto \mathcal{N}(0,1)$$

model parameters for the three different proposed models. Figure 5 clearly shows that the posterior densities for α's in the heteroskedastic models are greatly peaked.

B. Real Dataset

Description of the experiments: The authors consider datasets where the Lidar sensor is mounted on top of robot, and the robot moves with constant speed (0.05 m/s) towards a fixed wooden target starting at a distance of 4m and ending at a distance of 0.5m. The experiment setup is shown in Figure 2. The low cost triangulation Lidar used in the experiments is form Neato[1], similar to the one described in [8] but with infrared laser. The groundtruth distances x_is were collected using a Vicon motion capture system.

Evaluation of the parameter estimation procedures: Figures 7 and 8 show the estimated densities of the various model parameters for the three different proposed models (assuming model order $N = 3$).

[1]A robotic vacuum cleaner manufacturer, https://www.neatorobotics.com/

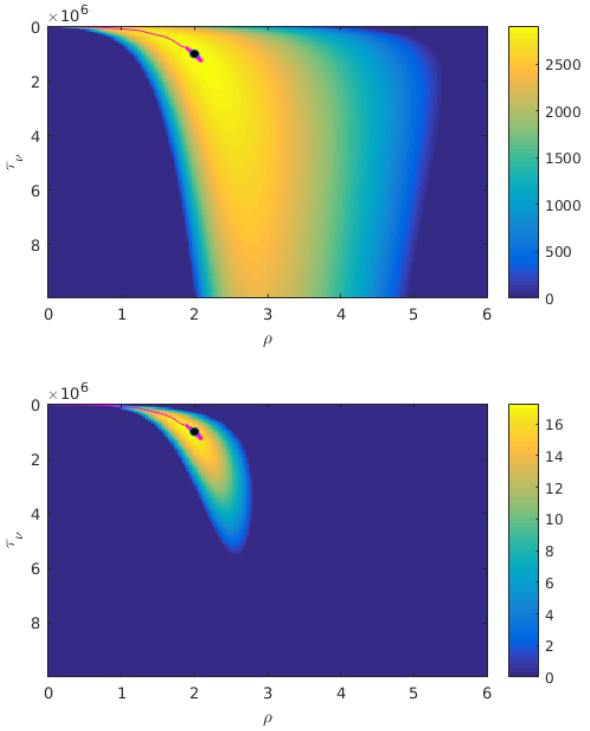

Fig. 3. Density color map of the posterior $p(\rho, \tau_\nu | \boldsymbol{x}, \boldsymbol{y})$ for the case $f_{\text{noise}}(x_i) = \sigma_\nu f_{\text{mean}}(x_i)^\rho$. The black point represents the actual values of the parameters and the solid line in magenta color shows the MCMC convergence path in ρ–τ_ν space. The upper plot produced using 900 samples while the lower produced using only 50 samples (plots produced using artificial dataset).

N	σ_ν	$\sigma_\nu x_i^\rho$	$\sigma_\nu f_{\text{mean}}(x_i)^\rho$
1	1397.59261	50.14214	3220.53529
2	3.15795	0.27043	0.02243
3	0.49000	0.00507	**0.00185**
4	0.48642	0.00404	**0.00088**
5	0.48714	0.00220	**0.00092**
6	0.48675	0.00229	0.01049
7	0.48754	0.00285	0.45820

TABLE II

THE PREDICTION MEAN SQUARED ERROR (MSE) SCORE FOR THE DIFFERENT COMPETING MODELS.

performance evaluation: To evaluate the estimation procedure on real dataset the authors apply the prediction MSE on another dataset (verification Dataset). The total empirical MSE between the predicted output $\widetilde{y}_i$ and the actual output recorded in the verification dataset is defined as

$$\text{MSE} := \sum_{i=1}^{L} \frac{f_{\text{noise}}(x_i)}{\widehat{\sigma_\nu^2}} \left\| \widetilde{y}_i^{(s)} - y_i \right\|^2,$$

where L is the number of measurements in the vitrification dataset, (s) indicates the model index in the models set. The authors select a set of proposed models consisting of model orders from 1 till 7 for each one of the three proposed models in Sections III-.1, III-.2 and III-.3. The prediction MSE scores for the estimated models is presented in Table II. The empirical MSE suggests the model structure

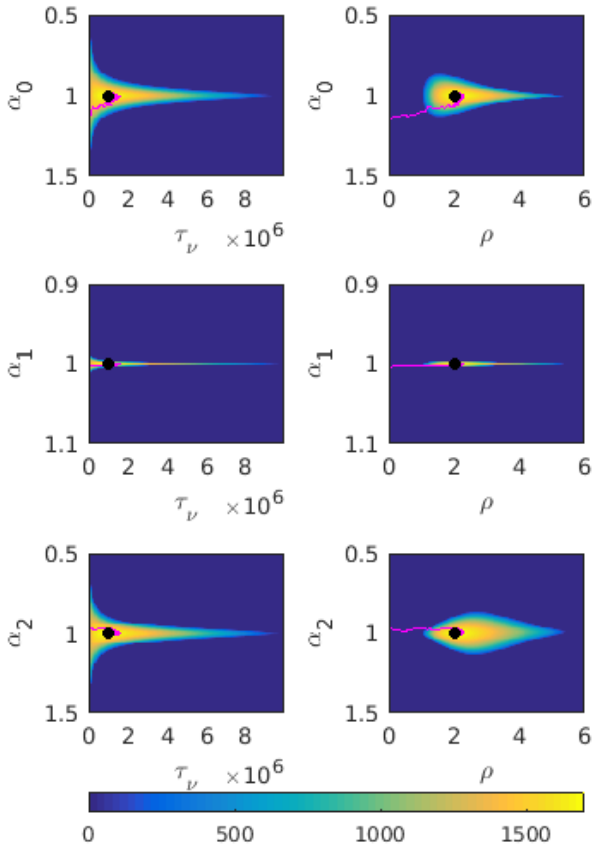

Fig. 4. On the left, the density color map of the posterior $p(\alpha_j, \tau_\nu \,|\boldsymbol{x}, \boldsymbol{y})$ for the case $f_{\text{noise}}(x_i) = \sigma_\nu f_{\text{mean}}(x_i)^\rho$. While on the right the density color map of the posterior $p(\alpha_j, \rho \,|\boldsymbol{x}, \boldsymbol{y})$ for the same case. The black points represent the actual values of the parameters and the solid lines in magenta color show the MCMC convergence path (plots produced using artificial dataset).

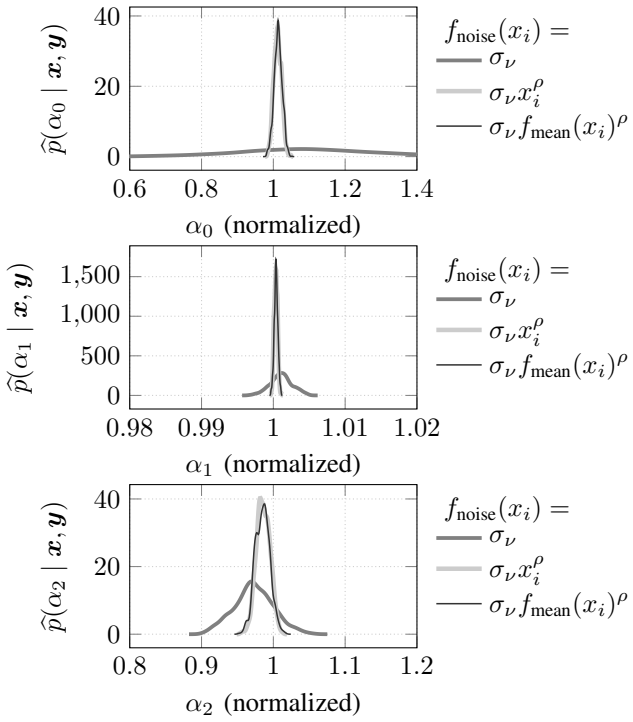

Fig. 5. The estimated posterior densities for the parameters α_0, α_1 and α_2. The densities in both heteroskedastic models are more peaked compared with the homoskedastic model (plots produced using artificial dataset).

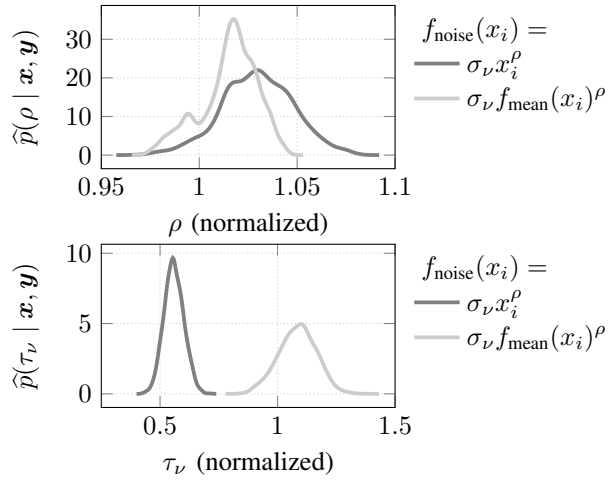

Fig. 6. The estimated posterior density for the parameter ρ in the top plot and for τ_ν in the bottom plot, the values are normalized with the actual parameter values (plots produced using artificial dataset).

with $f_{\text{noise}} = \sigma_\nu f_{\text{mean}}(x_i)^\rho$ and model order $N = 4$ as the one with minimum prediction MSE. The three models with smallest scores are designated with bold fonts in the Table.

The computation time during the parameter estimation process for the models in Table II is presented in Figure 9. The time is calculated using Matlab with standard laptop computer running Ubuntu, on Intel quad core i7-2640 CPU @2.8GHz processor.

V. Conclusions and Future Work

Motivated by the Lidar sensor calibration problem, this paper proposed a Bayesian method for sensor calibration. The method simultaneously estimates (i) an heteroskedastic measurement noise and (ii) a polynomial bias, describing a systematic distortion of the measured quantity. Such a Bayesian formulation allows to exploit prior information on the sensor model parameters and parameter estimation is performed using MCMC techniques.

To take advantage of the identified models in a practical scenario, two further steps have to be developed.
First, there is the need of a mechanism for automatic selection of the most effective model among the three proposed ones. Such a mechanism could be based on cross-validation, i.e. performing the analysis of Section IV-B on a suitable validation set and selecting the model with lower MSE. Alternatively, model selection could be performed on the basis of criteria accounting also for model complexity (e.g. Bayesian Information Criterion (BIC)).
Then, the measurement-model parameters estimated on a

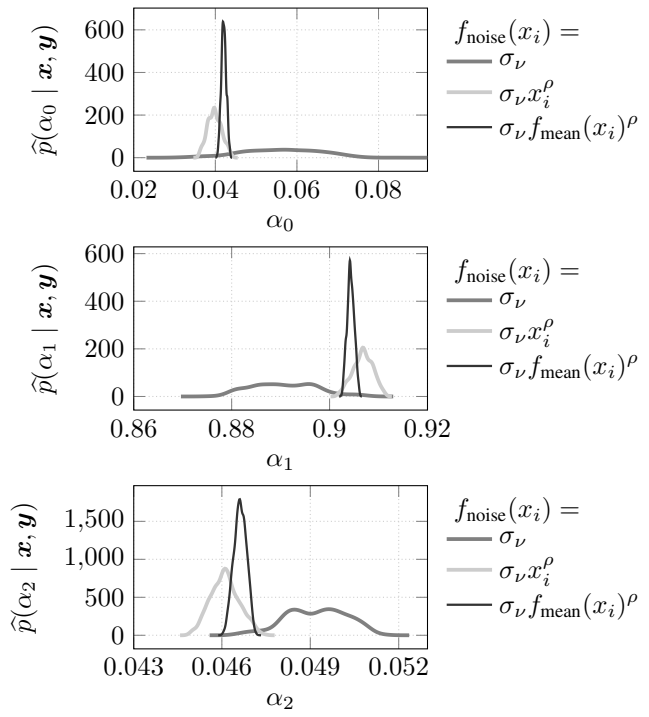

Fig. 7. The estimated posterior densities for the parameters α_0, α_1 and α_2 (plots produced using real dataset).

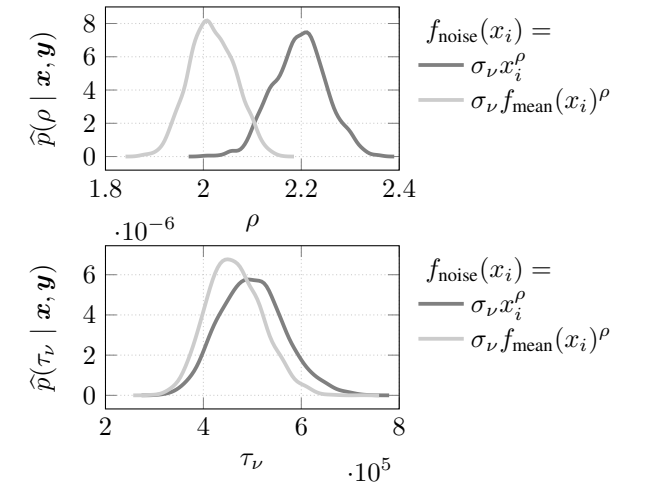

Fig. 8. The estimated posterior density for the parameter ρ in the top plot and for τ_ν in the bottom plot (plots produced using real dataset).

training set should be used to reconstruct the underlying state x associated to a new measurement y collected with the sensor. This could be done in a Bayesian framework reconstructing

$$p\left(x \,|y, \widehat{\boldsymbol{\alpha}}, \widehat{\tau_\nu}, \widehat{\rho}\right)$$

in sampled form using MCMC approaches.
These two further steps are under investigation and will be object of future works.

The proposed calibration parameter estimation could be compared with those obtained through different approaches,

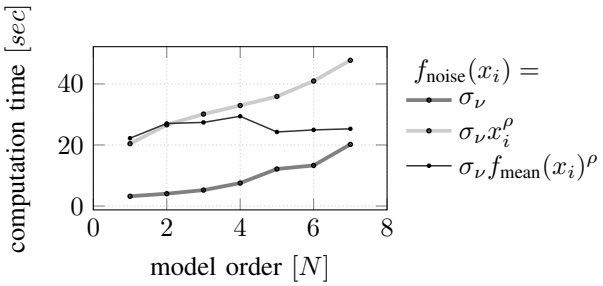

Fig. 9. The MCMC computation time for the various proposed models in Table II.

such as explicit modeling [16].

REFERENCES

[1] G. E. Box and W. J. Hill, "Correcting inhomogeneity of variance with power transformation weighting," *Technometrics*, vol. 16, no. 3, pp. 385–389, 1974.

[2] H. White, "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity," *Econometrica: Journal of the Econometric Society*, pp. 817–838, 1980.

[3] J. Geweke, "Bayesian treatment of the independent student-t linear model," *Journal of applied econometrics*, vol. 8, no. S1, 1993.

[4] R. E. Park, "Estimation with heteroscedastic error terms," *Econometrica (pre-1986)*, vol. 34, no. 4, p. 888, 1966.

[5] A. C. Harvey, "Estimating regression models with multiplicative heteroscedasticity," *Econometrica: Journal of the Econometric Society*, pp. 461–465, 1976.

[6] W. J. Boscardin and A. Gelman, "Bayesian computation for parametric models of heteroscedasticity in the linear model," *University of California, Berkeley*, 1994.

[7] H. Tanizaki and X. Zhang, "Posterior analysis of the multiplicative heteroscedasticity model," 2001.

[8] K. Konolige, J. Augenbraun, N. Donaldson, C. Fiebig, and P. Shah, "A low-cost laser distance sensor," in *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*. IEEE, 2008, pp. 3002–3008.

[9] J. Lima, J. Gonçalves, and P. J. Costa, "Modeling of a low cost laser scanner sensor," in *CONTROLO'2014–Proceedings of the 11th Portuguese Conference on Automatic Control*. Springer, 2015, pp. 697–705.

[10] D. Campos, J. Santos, J. Gonçalves, and P. Costa, "Modeling and simulation of a hacked neato xv-11 laser scanner," in *Robot 2015: Second Iberian Robotics Conference*. Springer, 2016, pp. 425–436.

[11] A. Alhashimi, D. Varagnolo, and T. Gustafsson, "Statistical modeling and calibration of triangulation lidars," in *ICINCO*, 2016.

[12] A. Alhashimi, G. Pierobon, D. Varagnolo, and T. Gustafsson, *Modeling and Calibrating Triangulation Lidars for Indoor Applications*. Cham: Springer International Publishing, 2018, pp. 342–366.

[13] A. Alhashimi, D. Varagnolo, and T. Gustafsson, "Calibrating distance sensors for terrestrial applications without groundtruth information," *IEEE Sensors Journal*, vol. 17, no. 12, pp. 3698–3709, 2017.

[14] G. O. Roberts, A. Gelman, W. R. Gilks, *et al.*, "Weak convergence and optimal scaling of random walk metropolis algorithms," *The annals of applied probability*, vol. 7, no. 1, pp. 110–120, 1997.

[15] A. Gelman, G. O. Roberts, W. R. Gilks, *et al.*, "Efficient metropolis jumping rules," 1996.

[16] M. Sheehan, A. Harrison, and P. Newman, "Self-calibration for a 3d laser," *The International Journal of Robotics Research*, 2012. [Online]. Available: http://ijr.sagepub.com/content/early/2011/12/21/0278364911429475