Joint Temperature-Lasing Mode Compensation for Time-of-Flight LiDAR Sensors

Anas Alhashimi *, Damiano Varagnolo and Thomas Gustafsson

Control Engineering Group, Department of Computer Science, Electrical and Space Engineering,
Luleå University of Technology, Luleå 97187, Sweden; damiano.varagnolo@ltu.se (D.V.);
thomas.Gustafsson@ltu.se (T.G.)
* Correspondence: anas.alhashimi@ltu.se; Tel.: +46-92-492-167

Abstract: We propose an expectation maximization (EM) strategy for improving the precision of time of flight (ToF) light detection and ranging (LiDAR) scanners. The novel algorithm statistically accounts not only for the bias induced by temperature changes in the laser diode, but also for the multi-modality of the measurement noises that is induced by mode-hopping effects. Instrumental to the proposed EM algorithm, we also describe a general thermal dynamics model that can be learned either from just input-output data or from a combination of simple temperature experiments and information from the laser's datasheet. We test the strategy on a SICK LMS 200 device and improve its average absolute error by a factor of three.

1. Introduction

ToF LiDARs estimate distances by emitting short bursts of laser light and by measuring the time it takes for the reflected photons to arrive back to the device [1]. Despite being based on a very simple principle, they are very much both accurate and precise devices [2]: for example, precisions can reach 10 mm of standard error when the object is 10 m away. Due to these favorable properties, they are commonly used in critical industrial applications where there is the need for high quality measurements.

It is well known that these devices need temperature compensation mechanisms, since changing their temperature leads to changes in the statistics of the returned measurements. The effect of temperature may be huge: experiments by [3] on an amplitude-modulated continuous-wave laser radar pointing at a target six meters away from the sensor showed that measurements at 21 °C and measurements at 45 °C were differing by 40 cm. Since thermal stabilization of a laser scanner may take up to 30 min [4], it is clear that these sensors are affected by a warm-up-induced time drift that must be compensated. Manufacturers of ToF devices thus usually embed opportune algorithms in their products that implement this temperature compensation mechanism.

Unfortunately, temperature is not the only physical factor that deserves compensation: as described in detail in Section 2, lasers can suddenly change their lasing mode. This property, called the mode-hopping effect, has a substantial impact on the measure returned by ToF devices, since changing lasing mode means to change the spectral content of the laser burst, *i.e.*, change its time of flight. Remarkably, to the best of our knowledge, the existing literature does not focus on managing this effect, but rather, considers only temperature compensation mechanisms.

We would like to mention here that there are also other methods with high temperature compensation at the nano-scale measurement, such as [5]. The method reduces the offset, the

temperature characteristic of the main sensing element, the temperature drift and the noise by the switching method.

1.1. Statement of Contributions

We propose an expectation maximization (EM) algorithm that compensates mode-hopping effects by modeling the induced measurement noise as a Gaussian mixture. Thus, from the mathematical perspective, we introduce some latent variables (namely, from which Gaussian the noise comes from) as additional estimands. This EM algorithm is also coupled to a temperature compensation filter that is built on a physics-based linear model for the thermal dynamics of the laser scanner. Summarizing, thus, our contributions are, with respect to the aforementioned literature:

1. A thorough motivation for why it is meaningful to consider mode-hopping effects in laser scanners, arising from a physical description of the lasing mechanism in laser diodes;
2. A thermodynamical model for the thermal dynamics of a whole laser scanner, needed by the proposed strategy to account for temperature effects;
3. A statistical model describing the measurement process that decouples the effects of the mode-hopping from temperature effects;
4. A numerically-efficient EM strategy based on the statistical model above;
5. A validation of the proposed compensation strategy on real devices.

With the validation, we also show that it is possible to improve the absolute error of a SICK LMS 200 device by a factor of three.

1.2. Organization of the Manuscript

Section 2 analyzes the effects of the laser temperature on the measured distance. Section 3 proposes a general model for the thermal dynamics of a pulsed ToF LiDAR. Section 4 presents a general measurement model accounting for both temperature and mode-hopping effects. Section 5 describes how to train the EM algorithm, while Section 6 describes how to use the same algorithm for testing purposes. Section 7 then proposes a likelihood ratio test for calibrating the hyperparameters of the EM algorithm. Section 8 presents some numerical results on commercial devices. Finally, Section 9 draws some conclusions and future research directions.

2. Effects of the Laser Temperature on the Measured Distance

This section lays down interpretations motivating the structure of the novel compensation procedures. We thus here describe the functioning principle of ToF scanners, explain why the measured distance depends on the temperature of the device and motivate why the measurement noise of a LiDAR is intrinsically multi-modal.

Consider then the basic operation of pulsed ToF LiDARs in Figure 1 and in its caption.

The measurement of the distance derives from ideal considerations: if the temporal width of the pulse is null, then the distance d between the sensor and the object should satisfy:

$$d = \frac{c\,\tau}{2} \tag{1}$$

where c is the speed of light and τ is the measured ToF between when the laser pulse is emitted and when it is received. Assume ideally that the laser pulse contains photons with a unique nominal wavelength λ_0. Since:

$$c = \lambda_0 f \tag{2}$$

with f the light frequency and λ_0 the nominal wavelength, and since the light frequency f remains the same through different media, to know c, it is sufficient to know λ_0. Thus, from knowing λ_0, one can compute d, since τ is measured.
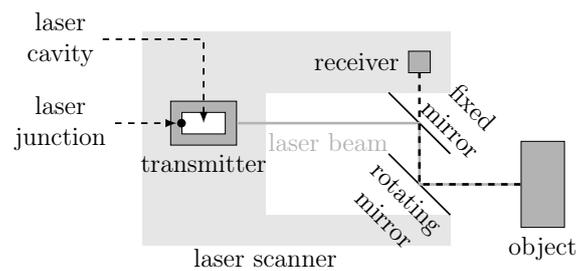
Figure 1. Graphical description of the operating principle for pulsed ToF LiDARs. A pulsed infra-red laser beam is first emitted from the transmitter. The case of the transmitter, in dark gray, encloses a laser junction and a laser cavity. The emitted laser beam is then deflected by a rotating mirror (resulting in a fan-shaped scan pattern) and, finally, reflected from the object surface. The time of flight τ between the transmission and the reception of the laser beam is then used to estimate the distance d between the scanner and the object.

We can already now notice the first effect of the temperature of the device on the measurement: according to its datasheet, the nominal wavelength of the laser diode SPL PL90 from OSRAM [6] is 905 nm at 25 °C, and 907.8 nm at 35 °C. Assuming the target to be at a one-meter distance, this 2.8 nm variation in the wavelength λ then results in a 10.3-ps variation in the ToF τ, *i.e.*, a variation in the measured distance of approximately 3.1 mm.
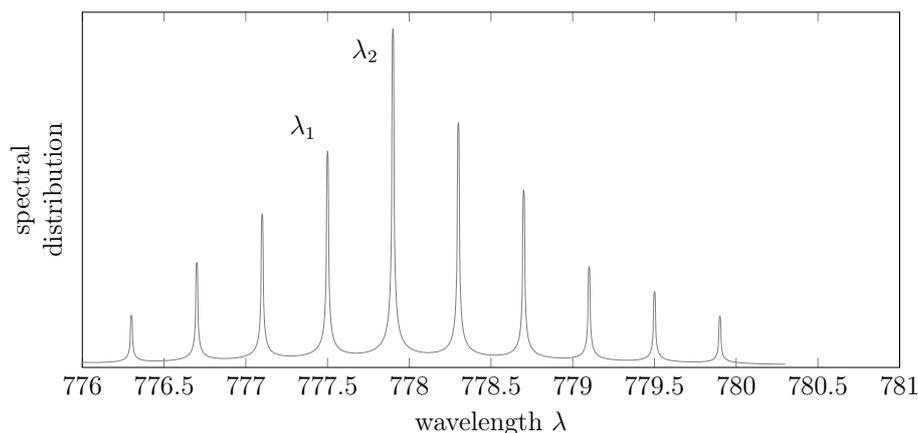

Figure 2. Average spectral distribution for GaAsP lasers at the nominal temperature of 21 °C [7].

The previous consideration is nonetheless simplistic. We can indeed notice another three distinct effects:

1. In general, lasers do not emit at a unique frequency λ. Indeed, the average spectral distribution of the laser pulses follows a "comb"-like density, like the one in Figure 2.

2. Lasers are affected by the so-called mode-hopping effect [8] and, indeed, oscillate between different lasing modes (the teeth of the "comb" of Figure 2), for which two different pulses generated under the same temperature and external conditions may have different λ's (e.g., referring to the same figure, the first pulse may contain only photons with wavelengths λ_1, while the second pulse may contain only photons with wavelengths λ_2). In other words, the actual distribution of one specific pulse may contain only a subset of the teeth of the average spectral distribution. Using naively Equations (1) and (2) to estimate d without being aware of the mode hopping, *i.e.*, assuming a certain λ_0 without actually knowing that the average λ jumps between different lasing modes, reflects thus in a multimodal measurement of d, as clearly shown in Figure 3.

3. The average spectral distribution of the laser pulses is not fixed, but rather depends on the temperature of the transmitter [9]. More precisely, the positions and amplitudes of the modes in Figure 2 depend on both the current flowing through the laser junction and the geometry of the laser cavity, but eventually, these two effects are inter-combined: the current flow produces heat that will modify the geometry of the cavity. Eventually, thus, the temperature affects the position and amplitude of the modes of the average spectral distribution. This temperature effect can be clearly seen in Figure 4: even if the device is nominally already compensated in temperature, one can clearly see two different lasing modes shifting in temperature.
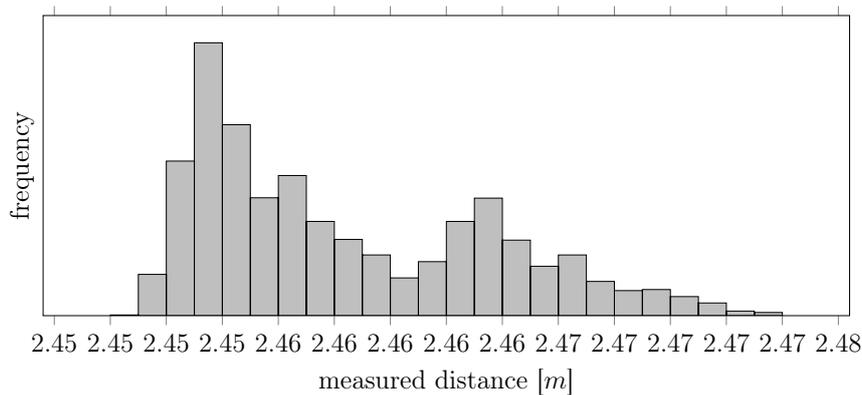

Figure 3. Histogram of consecutive measurements returned by a SICK 200 device in thermodynamical and electrical equilibrium pointing at a fixed object and in a controlled environment.
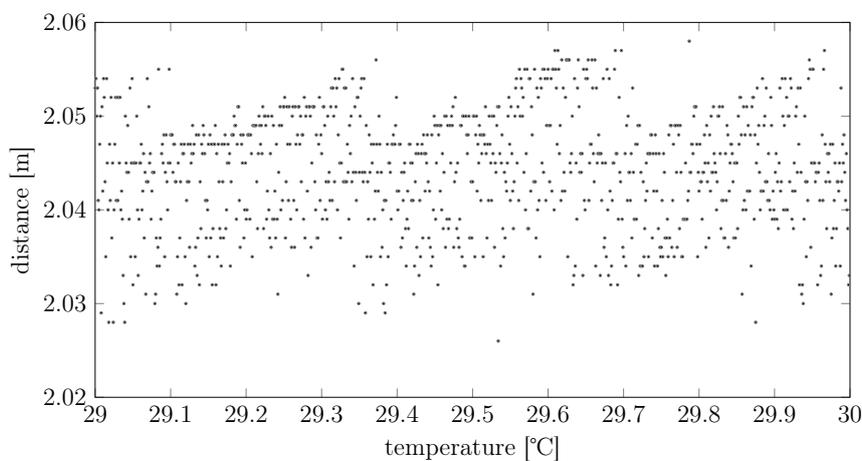

Figure 4. Dependency of the distance measurements on the device temperature for a SICK 200 device pointing to a fixed object in a controlled environment. Despite the true distance and other parameters potentially affecting the measurements being constant in time, the distributions of the measurements are temperature varying. We can notice how the device compensates for the temperature change by adding a temperature-varying negative bias, but that it does not compensate the mode-hopping effect.

To summarize, the actual λ of a laser pulse is in general different from the nominal λ_0 because of two effects: first, the laser may oscillate between different modes; second, these positions of these modes vary with temperature.

3. A General Model for the Thermal Dynamics of a Pulsed ToF LiDAR

Instrumental for compensating the effects of temperature offsets highlighted in the previous section, we here discuss a general thermodynamic model describing the temperature dynamics of

the whole laser scanner that will lead us to being able to represent the temperature changes of the laser cavity.

3.1. Physical Modeling

Just like any power electronic device, LiDARs generate heat that is then exchanged with the environment, so that the temperature of a scanner depends on the temperature of the environment. The main sources of heat inside the LiDAR are thus the laser diode, the motor and the electronic components of the system. The heat generated inside the scanner is then transferred to the surrounding environment through the case. Since our experience indicates that motors and other electronic components induce negligible thermal effects, we consider only the heat produced by the laser diode.

We thus represent the thermal model of a generic scanner as the equivalent electrical circuit shown in Figure 5, interpretable as follows: when the laser is turned on, the heat generated by the laser junction is dissipated in the surrounding environment through first the transmitter case and then, second, through the laser scanner case.
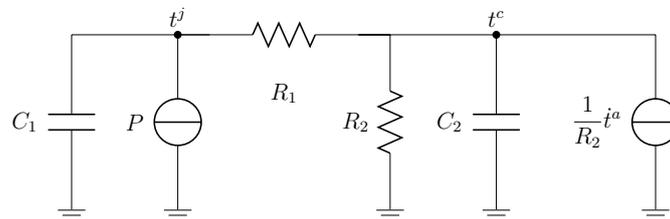

Figure 5. The proposed thermodynamic model for a generic pulsed ToF LiDAR.

Considering the notation:

P	Heat power generated by the junction (equal to zero when the device is off)
t^j	Temperature of the junction
t^c	Temperature of the transmitter case
t^a	Temperature of the external ambient
t^s	Noisy measurement of the temperature of the transmitter case
C_1	Thermal inertia of the transmitter case
C_2	Thermal inertia of the laser scanner case
R_1	Thermal resistance between the transmitter case and the laser scanner case
R_2	Thermal resistance between the laser scanner case and the ambient

It follows that the dynamics for the temperatures of the laser junction and scanner case when the laser is on are thus:

$$\begin{cases} \dot{t}^c = \dfrac{t^j}{R_1 C_2} - t^c \left(\dfrac{1}{R_2 C_2} + \dfrac{1}{R_1 C_2} \right) + \dfrac{1}{C_2 R_2} t^a \\[2ex] \dot{t}^j = \dfrac{P}{C_1} + \dfrac{t^c}{R_1 C_1} - \dfrac{t^j}{R_1 C_1} \end{cases} \tag{3}$$

Discretizing the previous dynamics with a discretization step of T s and letting:

$$\bar{t}_k := \begin{bmatrix} t^c_k \\ t^j_k \end{bmatrix} \qquad \bar{u}_k := \begin{cases} P & \text{if on,} \\ 0 & \text{if off,} \end{cases} \qquad \bar{v}_k := t^a_k - t^a_{k-1} \tag{4}$$

implies the following discrete-time state-space representation of the model:

$$\begin{cases} \bar{t}_{k+1} = A\bar{t}_k + B\bar{u}_k + B'\bar{v}_k \\ t^s_k = C\bar{t}_k + \bar{\mu}_k \end{cases} \tag{5}$$

where t_k^s is a noisy temperature measurement of the case at time k, $\overline{v}_k$ and $\overline{\mu}_k$ are independent process and measurement noises, and where:

$$A = \begin{bmatrix} 1 - \dfrac{T}{R_2 C_2} - \dfrac{T}{R_1 C_2} & \dfrac{T}{R_1 C_2} \\ \dfrac{T}{R_1 C_1} & 1 - \dfrac{T}{R_1 C_1} \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ \dfrac{T}{C_1} \end{bmatrix} \qquad B' = \begin{bmatrix} \dfrac{T}{R_2 C_2} \\ 0 \end{bmatrix} \qquad C = \begin{bmatrix} 1 & 0 \end{bmatrix} \quad (6)$$

Notice that through Equation (5), we introduce t_k^s, *i.e.*, a noisy measurement of the temperature of the transmitter case. This correspond to the practical assumption that perfect knowledge of the actual case temperature t_k^c is in general unavailable, since temperature sensors attached to the case of the scanner will never give noiseless recordings.

3.2. Identifying Model Equation (5)

If one has a dataset of recorded temperatures t_k^j, t_k^c and t_k^a, then one may identify Equation (5) using standard system identification approaches, e.g., a prediction error method (PEM) as in [10]. For the common case where it may be difficult to obtain direct measurements of the quantities, we propose to resort to the following general strategy that uses the datasheet of the laser scanner in conjunction with noisy case temperature measurements t_k^s.

Algorithm 1 Identification of model Equation (5) starting from the datasheet of a laser scanner.

1: From the datasheet of the laser scanner infer:

- its thermal resistance R_1 (directly from the datasheet);
- its thermal capacity C_1 (by estimating the volume of the laser diode from the datasheet and multiplying it for the heat capacity of the material, also indicated in the datasheet);

2: From measuring the weight and material of the case, infer its thermal capacity C_2;

3: Estimate the generated heat power P by measuring the electrical power absorbed by the device and multiplying this quantity by 0.5 (for the estimated efficiency of generic laser diodes [11]);

4: From situations where the scanner is in thermal equilibrium, calculate R_2 by calculating the difference between the measured case temperature t_k^s and the ambient temperature t_k^a divided by the estimated generated heat power P.

3.3. Estimating t^j from t^s

Assume that model Equation (5) has been identified, either from measured data or using Algorithm 1, and observe that the thermal model is observable, reachable and has stable dynamics. Due to their favorable theoretical and numerical properties, we thus devise to estimate t_k^j from noisy measurements of the case temperature t_k^s via Kalman smoothers/filters, as in [12], *i.e.*, to let:

$$\left[\widehat{t}_1^j, \ldots, \widehat{t}_K^j \right] = W \left[t_1^s, \ldots, t_K^s \right] \tag{7}$$

for an opportune (and potentially time varying) matrix W. As before, notice that the filtering starts from noisy measurements of the case temperature t_k^s, rather than from perfect measurements t_k^c.

4. A General Measurement Model Accounting for Temperature and Mode Hopping Effects

We recall that our aim is to model the effect of the temperature of the laser junction t_k^j on the measured distance d and understand how noisy case temperature measurements t_k^s help improving the accuracy on the final estimate of d. To this aim, we propose the following measurement model at the generic time instant k:

$$y_k = d + H\left(t_k^j\right)\boldsymbol{\theta} + (1 - \Delta_k)\,w_k^1 + \Delta_k w_k^2 \tag{8}$$

where:

- y_k is the distance returned by the sensor;
- d is the true distance from the object (assumed deterministic);
- t_k^j is the temperature of the laser cavity at time k;
- The two modes $w_k^1 \sim \mathcal{N}\left(\mu_1, \sigma_1^2\right)$ and $w_k^2 \sim \mathcal{N}\left(\mu_2, \sigma_2^2\right)$ account for a bimodal Gaussian and white additive measurement noise. The Bernoulli random variable (r.v.) $\Delta_k \sim \mathcal{B}\left(\pi\right)$ selects the active mode at time k, so that π reflects the relative importance of the modes. Intuitively, Δ_k represent which lasing mode has been active during measurement y_k. We notice that here, we consider bimodal noises (*i.e.*, only two lasing modes) just for notational simplicity. It is nonetheless immediate to generalize the subsequent findings for an *M*-modal case;
- $H\left(\cdot\right)\boldsymbol{\theta}$ is a non-linear transformation of the temperature of the laser junction t_k^j in a measurement bias. In the following Examples 1 and 2, we show how different $H\left(\cdot\right)$'s and $\boldsymbol{\theta}$'s express different maps from the junction temperature t_k^j to the measurement bias.

Notice that, induced by our experience, we let the measurement noise modes w_k^1 and w_k^2 be independent from the laser junction temperature t_k^j. Our experiments indeed indicated that the moments of the noises are not affected by changing temperatures (at least for a range between 0 °C and 40 °C). Notice moreover that model Equation (8) is linear in $\boldsymbol{\theta}$; this restrictive assumption is nonetheless essential for building distance estimation algorithms that are numerically fast.

Example 1 (Polynomial model)**.** *Let:*

$$H_k = H\left(t_k^j\right) := \left[t_k^j, \left(t_k^j\right)^2, \ldots, \left(t_k^j\right)^N\right] \tag{9}$$

and $\boldsymbol{\theta} := [\theta_1, \ldots, \theta_N]$. *Then, the generic model Equation* (8) *specializes into:*

$$y_k = d + \sum_{n=1}^{N}\left(t_k^j\right)^n + (1 - \Delta_k)\,w_k^1 + \Delta_k w_k^2 \tag{10}$$

i.e., a measurement model where the temperature plays the role of an N-th order polynomial bias (see Figure 6).
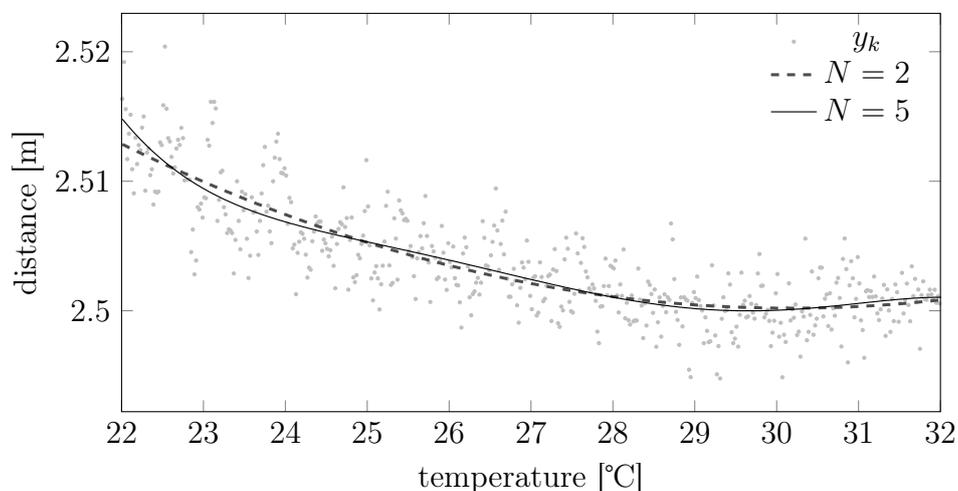

Figure 6. Potential temperature-bias dependencies for the polynomial model of Example 1.

Example 2 (Fourier expansion model). *Let:*

$$H_k = H\left(t_k^j\right) := \left[\cos\left(2\pi f_0 t_k^j\right),\ \sin\left(2\pi f_0 t_k^j\right),\ \ldots,\ \cos\left(2N\pi f_0 t_k^j\right),\ \sin\left(2N\pi f_0 t_k^j\right)\right] \quad (11)$$

and $\boldsymbol{\theta} := \left[\theta_1', \theta_1'', \ldots, \theta_N', \theta_N''\right]$, *where the fundamental frequency* f_0 *is assumed to be known. Then, the generic model Equation* (8) *specializes into:*

$$y_k = d + \sum_{n=1}^{N}\left(\theta_n' \cos\left(2n\pi f_0 t_k^j\right) + \theta_n'' \sin\left(2n\pi f_0 t_k^j\right)\right) + (1 - \Delta_k)\, w_k^1 + \Delta_k w_k^2 \quad (12)$$

i.e., a measurement model where the temperature plays the role of a bias that is periodic with frequency f_0 *(see Figure* 7*).*
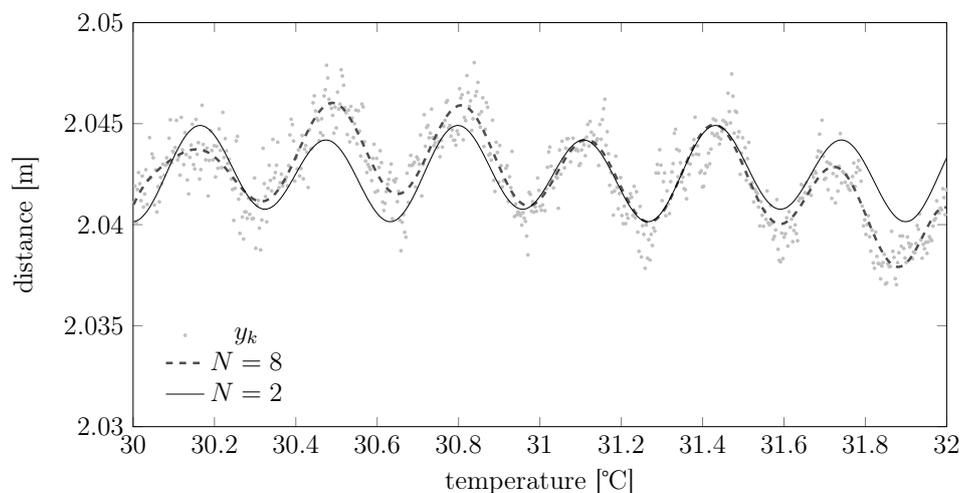

Figure 7. Potential temperature-bias dependencies for the Fourier expansion model of Example 2.

We then use model Equation (8) for practical purposes following the classical scheme:

- Design the model, *i.e.*, decide the structure for $H\left(\cdot\right)$ (e.g., between Example 1, Example 2 or also different ones depending on the need) starting from data collected in a controlled environment (see Section 7);
- Train the model, *i.e.*, estimate $\boldsymbol{\theta}$ and the statistics of w_k^1, w_k^2 and Δ_k from data collected in a controlled environment (see Section 5);
- Test the model, *i.e.*, use the previous estimated quantities during the normal operation of the laser scanner, so as to improve the estimation of d from data collected in a non-controlled environment (see Section 6).

The following sections are dedicated in detail to how to implement the previous three points. Notice that Section 7, on the design of the model, is presented after Sections 5 and 6 for notational convenience.

5. Training Model Equation (8)

In this section, we devise a numerical algorithm for learning $\boldsymbol{\theta}$ (*i.e.*, the coefficients multiplying $H(\cdot)$), μ_1, μ_2, σ_1^2, σ_2^2, π (*i.e.*, the statistics of the noises w_k^1, w_k^2 and of the random variables Δ_k) and the values of the mode selection variables Δ_k starting from a dataset containing:

- y_k for $k \in \{1, \ldots, K\}$;
- d, *i.e.*, the real distance;

- t_k^s, *i.e.*, measurements of the temperature of the laser scanner case (to be transformed into estimates of t_k^j through Equation (7)).

Let then:

$$
\begin{aligned}
\boldsymbol{y} &:= \begin{bmatrix} y_1 & \cdots & y_K \end{bmatrix}^T \\
\boldsymbol{t}^s &:= \begin{bmatrix} t_1^s & \cdots & t_K^s \end{bmatrix}^T \\
\boldsymbol{t}^j &:= \begin{bmatrix} t_1^j & \cdots & t_K^j \end{bmatrix}^T \\
\boldsymbol{\Delta} &:= \begin{bmatrix} \Delta_1 & \cdots & \Delta_K \end{bmatrix}^T \\
\overline{\boldsymbol{\theta}} &:= \begin{bmatrix} \boldsymbol{\theta}^T & \mu_1 & \mu_2 & \sigma_1^2 & \sigma_2^2 & \pi \end{bmatrix}^T.
\end{aligned}
\tag{13}
$$

Given the frequentist assumptions on the unknowns, we would like to perform maximum likelihood (ML) estimation for $\overline{\boldsymbol{\theta}}$, *i.e.*, seek for:

$$
\arg\max_{\overline{\boldsymbol{\theta}}, \boldsymbol{\Delta}} \ \mathbb{P}\left[\boldsymbol{y}, \boldsymbol{t}^j \, ; \, d, \overline{\boldsymbol{\theta}}, \boldsymbol{\Delta} \right]
\tag{14}
$$

Since the laser junction temperature $\boldsymbol{t}^j$ is unavailable, Equation (14) cannot be solved. We thus resort to solving the approximated problem:

$$
\left(\widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\Delta}} \right) := \arg\max_{\overline{\boldsymbol{\theta}}, \boldsymbol{\Delta}} \ \mathbb{P}\left[\boldsymbol{y}, \widehat{\boldsymbol{t}^j} \, ; \, d, \overline{\boldsymbol{\theta}}, \boldsymbol{\Delta} \right]
\tag{15}
$$

where $\widehat{\boldsymbol{t}^j}$ is the estimate of $\boldsymbol{t}^j$ given $\boldsymbol{t}^s$ as in Equation (7). We also notice that, instead of considering the joint density of $\boldsymbol{y}$ and $\widehat{\boldsymbol{t}^j}$, it is sufficient to consider the conditional density of $\boldsymbol{y}$ given $\widehat{\boldsymbol{t}^j}$:

Proposition 1.

$$
\left(\widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\Delta}} \right) = \arg\max_{\overline{\boldsymbol{\theta}}, \boldsymbol{\Delta}} \ \mathbb{P}\left[\boldsymbol{y} \ \middle| \ \widehat{\boldsymbol{t}^j} \, ; \, d, \overline{\boldsymbol{\theta}}, \boldsymbol{\Delta} \right]
\tag{16}
$$

It is important to notice that the ML problem in Equation (16) contains the latent variables $\boldsymbol{\Delta}$; to estimate them, we thus resort to a tailored EM approach [13]. To this aim, define then the auxiliary variables:

$$
\begin{aligned}
\Sigma_1 &:= \operatorname{diag}\left(1 - \widehat{\Delta}_k \right) & \widetilde{y}_k &:= y_k - d \\
\Sigma_2 &:= \operatorname{diag}\left(\widehat{\Delta}_k \right) & \widetilde{\boldsymbol{y}} &:= \boldsymbol{y} - d\mathbb{1}
\end{aligned}
\qquad
H := \begin{bmatrix} H_1 \\ \vdots \\ H_K \end{bmatrix} = \begin{bmatrix} H\left(\widehat{t}_1^j \right) \\ \vdots \\ H\left(\widehat{t}_K^j \right) \end{bmatrix}
\tag{17}
$$

with $\mathbb{1}$ being a vector of K ones. The computation of $\left(\widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\Delta}} \right)$ is performed through the iteration up to convergence (stopping criteria for EM algorithms are usually based on relative or absolute changes in the parameter estimate or in the value of the log likelihood; see, e.g., [14]; in our implementations, we used the absolute changes in the parameter estimate $\widehat{\pi} < \epsilon$; where ϵ is a small number 1×10^{-5} in the implementation) of the following two steps:

E-step

$$
\begin{aligned}
\delta_k &= (1 - \widehat{\pi}) \, \mathcal{N}\left(\widetilde{y}_k - H_k\widehat{\boldsymbol{\theta}} - \widehat{\mu}_1, \widehat{\sigma}_1^2 \right) + \widehat{\pi} \, \mathcal{N}\left(\widetilde{y}_k - H_k\widehat{\boldsymbol{\theta}} - \widehat{\mu}_2, \widehat{\sigma}_2^2 \right), \qquad k = 1, \ldots, K \\
\widehat{\Delta}_k &= \frac{\widehat{\pi}\mathcal{N}\left(\widetilde{y}_k - H_k\widehat{\boldsymbol{\theta}} - \widehat{\mu}_2, \widehat{\sigma}_2^2 \right)}{\delta_k}, \qquad k = 1, \ldots, K
\end{aligned}
\tag{18}
$$

M-step

$$
\begin{aligned}
C &= \widehat{\sigma}_1^2\,\Sigma_1 + \widehat{\sigma}_2^2\,\Sigma_2 \\
\widehat{\boldsymbol{\theta}} &= \left(H^T C^{-1} H\right)^{-1} H^T C^{-1} \widetilde{\boldsymbol{y}} \\
\widehat{\mu}_i &= \frac{\mathbb{1}^T \Sigma_i}{\mathbb{1}^T \Sigma_i \mathbb{1}} \left(\widetilde{\boldsymbol{y}} - H\widehat{\boldsymbol{\theta}}\right) & i = 1,2 \\
\widehat{\sigma}_i^2 &= \frac{\left(\widetilde{\boldsymbol{y}} - H\widehat{\boldsymbol{\theta}}\right)^T \Sigma_i \left(\widetilde{\boldsymbol{y}} - H\widehat{\boldsymbol{\theta}}\right)}{\mathbb{1}^T \Sigma_i \mathbb{1}} & i = 1,2 \\
\widehat{\pi} &= \frac{1}{K} \sum_k \widehat{\Delta}_k
\end{aligned}
\tag{19}
$$

In our numerical experiments, we empirically found it convenient to use the following initial conditions:

$$
\widehat{\boldsymbol{\theta}} = \boldsymbol{0}, \qquad \widehat{\mu}_1 = 0.1, \qquad \widehat{\mu}_2 = -0.1, \qquad \widehat{\sigma}_1^2 = 0.1, \qquad \widehat{\sigma}_2^2 = 0.1, \qquad \widehat{\pi} = 0.5
\tag{20}
$$

As for the convergence of the EM to the true ML estimate, we notice that EM algorithms are not in general ensured to have convergence properties. A sufficient condition for convergence is in [15], where the authors show that EM algorithms are convergent if the maximizer of the M-step is unique (a condition that is almost always satisfied in practice). In our case, the M-step maximizer is unique as long as in the update for $\widehat{\boldsymbol{\theta}}$ in Equation (19), the matrix $H^T C^{-1} H$ admits the inverse. In general, e.g., in both the polynomial case of Example 1 and in the Fourier expansion case of Example 2, this translates into the need for at least N samples associated with N different laser junction temperatures t_k^j.

6. Testing Model Equation (8)

We now devise a numerical algorithm for estimating d and the values of the lasing mode selection variables Δ_k starting from the model trained in Section 5 (*i.e.*, an estimated vector $\widehat{\overline{\boldsymbol{\theta}}} := \left[\widehat{\boldsymbol{\theta}}^T,\ \widehat{\mu}_1,\ \widehat{\mu}_2,\ \widehat{\sigma}_1^2,\ \widehat{\sigma}_2^2,\ \widehat{\pi}\right]^T$ and the statistics of w_k^1, w_k^2 and Δ_k) and a set of measurements y_k and t_k^s for $k = 1, \ldots, K$.

Assuming once again to transform the temperature sensor measurements $\boldsymbol{t}^s$ into estimated laser junction temperatures $\widehat{\boldsymbol{t}^j} = W\boldsymbol{t}^s$ through Equation (7), the problem of estimating d and the Δ_ks can be cast as:

$$
\left(\widehat{d}, \widehat{\boldsymbol{\Delta}}\right) = \arg\max_{d, \boldsymbol{\Delta}}\ \mathbb{P}\left[\boldsymbol{y}\ \middle|\ \widehat{\boldsymbol{t}^j}\ ;\ d, \widehat{\overline{\boldsymbol{\theta}}}, \boldsymbol{\Delta}\right]
\tag{21}
$$

As before, we compute this ML estimate through an EM approach:

E-step

$$
\begin{aligned}
\delta_k &= (1 - \widehat{\pi})\,\mathcal{N}\left(y_k - \widehat{d} - H_k\widehat{\boldsymbol{\theta}} - \widehat{\mu}_1, \widehat{\sigma}_1^2\right) + \widehat{\pi}\mathcal{N}\left(y_k - \widehat{d} - H_k\widehat{\boldsymbol{\theta}} - \widehat{\mu}_2, \widehat{\sigma}_2^2\right), \quad k = 1, \ldots, K \\
\widehat{\Delta}_k &= \frac{\widehat{\pi}\mathcal{N}\left(y_k - \widehat{d} - H_k\widehat{\boldsymbol{\theta}} - \widehat{\mu}_2, \widehat{\sigma}_2^2\right)}{\delta_k}, \quad k = 1, \ldots, K
\end{aligned}
\tag{22}
$$

M-step

$$
\begin{aligned}
C &= \Sigma_1 \widehat{\sigma}_1^2 + \Sigma_2 \widehat{\sigma}_2^2 \\
\widehat{d} &= \left(\mathbb{1}^T C^{-1} \mathbb{1}\right)^{-1} \mathbb{1}^T C^{-1} \left(\boldsymbol{y} - H\widehat{\boldsymbol{\theta}}\right)
\end{aligned}
\tag{23}
$$

The same values of μ_1, μ_2, σ_1^2 and σ_2^2 at the end of the training step will be used in the testing step. There is no need to recompute them again. In our experiments, we found it beneficial to start from the initial conditions $\widehat{d} = \dfrac{\mathbb{1}^T \boldsymbol{y}}{K}$ and $\widehat{\Delta}_k = 0.5$. The convergence properties of this EM procedure are then very similar to the EM in Section 5.

7. Designing Model Equation (8)

This section is divided into two parts:

1. Section 7.1, suggesting some hints for designing different structures for $H(\cdot)$ (e.g., choosing the order for Example 1, for Example 2 or also designing different functional structures depending on the collected information);

2. Section 7.2, suggesting a numerical algorithm for discriminating between different competing structures for H starting from data collected in a controlled environment.

7.1. Designing $H(\cdot)$

The proposed EM algorithms have the numerically-favorable property of having both the E and the M steps solvable in closed form. It is important to notice that this is induced by the fact that model Equation (8) is linear in $\boldsymbol{\theta}$, *i.e.*,

$$H\left(t_k^j\right)\boldsymbol{\theta} = \left[H_1\left(t_k^j\right),\ldots,H_N\left(t_k^j\right)\right]\begin{bmatrix}\theta_1\\\vdots\\\theta_N\end{bmatrix} = \sum_{n=1}^{N}H_n\left(t_k^j\right)\theta_n \tag{24}$$

Thus, with this structure, the designer can model the effect of the temperature t_k^j on the measurement y_k as the sum of N independent effects, each one represented as an opportune generic function of t_k^j (the weight of which is actually assumed unknown before the training phase).

As shown in Figures 6 and 7, the structures proposed in Examples 1 and 2 have quite general generalization capabilities. Nonetheless, the designer can tailor H so that it resembles other structures; our suggestion is to start from raw measured data spanning different temperatures, check visually how the macroscopic temperature trend behaves and then decompose this trend as the sum of different functions that will become the various $H_n(\cdot)$ in model Equation (8).

7.2. Determining the Best $H(\cdot)$ among Different Competing Potential Structures

The process described in Section 7.1 may lead to different competing structures for $H(\cdot)$. In other words, the designer may propose different structures $H^{(1)}$, $H^{(2)}$, *etc.*, and would like to choose the "best" $H^{(i)}$ given a dataset containing $\boldsymbol{y}$, $\boldsymbol{t}^s$ and the true distance d.

We propose to use the classical approach of discriminating the various $H^{(i)}$'s considering their goodness of fit, *i.e.*, to use generalized likelihood ratios (GLRs) [16], for which we first estimate the best estimates $\widehat{\overline{\boldsymbol{\theta}}}^{(i)}$ and $\widehat{\boldsymbol{\Delta}}^{(i)}$ for each $H^{(i)}$ given the dataset and then select the best hypothesis considering their resulting log-likelihoods. More formally, the suggested procedure is as in Algorithm 2.

Algorithm 2 Selection of the best $H^{(i)}$.

1: **for** $i = 1, 2, \ldots$ **do**
2:     Compute $\widehat{\overline{\boldsymbol{\theta}}}^{(i)}$ and $\widehat{\boldsymbol{\Delta}}^{(i)}$ as in Section 5;
3:     Compute

$$\begin{aligned}\ell^{(i)} := \quad &\log(\widehat{\sigma}_1^2)\mathbb{1}^T\Sigma_1\mathbb{1} + \log(\widehat{\sigma}_2^2)\mathbb{1}^T\Sigma_2\mathbb{1}\\ &+\frac{1}{2\widehat{\sigma}_1^2}\left(\boldsymbol{y} - d\mathbb{1} - H^{(i)}\widehat{\boldsymbol{\theta}}^{(i)} - \widehat{\mu}_1^{(i)}\mathbb{1}\right)^T\Sigma_1^{(i)}\left(\boldsymbol{y} - d\mathbb{1} - H^{(i)}\widehat{\boldsymbol{\theta}}^{(i)} - \widehat{\mu}_1^{(i)}\mathbb{1}\right)\\ &+\frac{1}{2\widehat{\sigma}_2^2}\left(\boldsymbol{y} - d\mathbb{1} - H^{(i)}\widehat{\boldsymbol{\theta}}^{(i)} - \widehat{\mu}_2^{(i)}\mathbb{1}\right)^T\Sigma_2^{(i)}\left(\boldsymbol{y} - d\mathbb{1} - H^{(i)}\widehat{\boldsymbol{\theta}}^{(i)} - \widehat{\mu}_2^{(i)}\mathbb{1}\right)\end{aligned} \tag{25}$$

    where $\Sigma_1^{(i)} := \text{diag}\left(1 - \widehat{\Delta}_k^{(i)}\right)$ and $\Sigma_2^{(i)} := \text{diag}\left(\widehat{\Delta}_k^{(i)}\right)$;
4: **end for**
5: Select that $H^{(i)}$ that corresponds to the maximal $\ell^{(i)}$.

---

8. Experiments

We now validate the proposed thermal model and estimation strategy on real data, aiming to show their effectiveness. We thus consider the simple scenario where a target is fixed in front of a sensor, under constant light and electrical conditions, so that the actual distance between the sensor and the target d is fixed. We consider a SICK LMS 200 LiDAR, one of the most widely-used LiDARs in industry and robotics applications.

8.1. Training and Validation of the Thermal Model Equation (5)

To train and validate the thermal model Equation (5), we conducted the experiment summarized in Figure 8: in a thermally-controlled room at 22°, we performed several on-off cycles of the device and measured the corresponding case temperature t_k^s. We then used the data represented as "training data" in the figure, the data-sheet of the SICK LMS 200 LiDAR and Algorithm 1 to train model Equation (5). After that, we drove the trained model with the $\bar{u}_k$ in the "test data" as inputs and obtained a predicted temperature $\widehat{y}_k^s$. The goodness of fit of the predicted temperatures, computed as:

$$100\left(1 - \frac{\sum_k \left(\widehat{y}_k^s - y_k^s\right)^2}{\sum_k \left(y_k^s\right)^2}\right) \tag{26}$$

is then 92.79%. This indicates a very good fit, *i.e.*, a good approximation capability of our proposed thermal model (and associated learning algorithm).
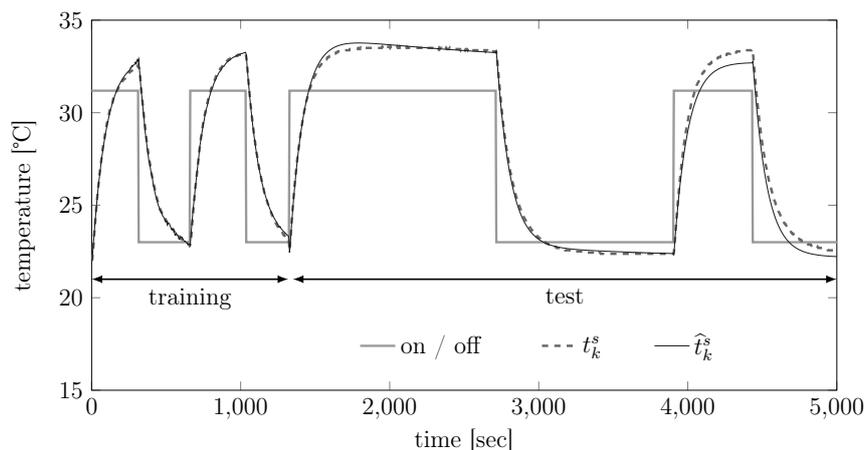
Figure 8. Validation of the thermal model Equation (5) on a SICK 200 LiDAR.

8.2. Selection of the Optimal $H(\cdot)$ for the SICK 200 LiDAR

We used Algorithm 2 on real data from a SICK LMS 200 to discriminate between different $H^{(i)}$'s in Examples 1 and 2, *i.e.*, the hypotheses:

$$\mathbb{H}_i : \boldsymbol{y} \sim (1-\Delta)\mathcal{N}\left(d + H^{(i)}\boldsymbol{\theta} + \mu_1, \sigma_1^2\right) + \Delta\mathcal{N}\left(d + H^{(i)}\boldsymbol{\theta} + \mu_2, \sigma_2^2\right) \tag{27}$$

against the null hypothesis:

$$\mathbb{H}_0 : \mathbf{y} \sim \mathcal{N}\left(d, \sigma^2\right) \tag{28}$$

so that the log likelihood ratio test between two hypotheses can be defined as:

$$\Lambda_i(\boldsymbol{y}) = \frac{\ell\left(\widehat{d}, \widehat{\sigma}^2 | \boldsymbol{y}\right)}{\ell\left(\widehat{d}, \widehat{\sigma}_1^2, \widehat{\sigma}_2^2, \widehat{\mu}_1, \widehat{\mu}_2, \widehat{\boldsymbol{\theta}}, \widehat{\Delta} | \boldsymbol{y}, H^{(i)}\right)} \tag{29}$$

where we explicitly mention the dependence of the log likelihood ℓ from the various parameters of the considered model.

Figures 9 and 10 show the likelihood ratios obtained for different candidates for polynomial and Fourier expansions, respectively. For the polynomial case, we obtained the highest likelihood for the polynomial order $N = 16$, while for the Fourier case, we obtained the optimal value for $N = 180$. We motivate this latter order, much higher than that of the polynomial case, to be due to the periodic nature of the Fourier model.
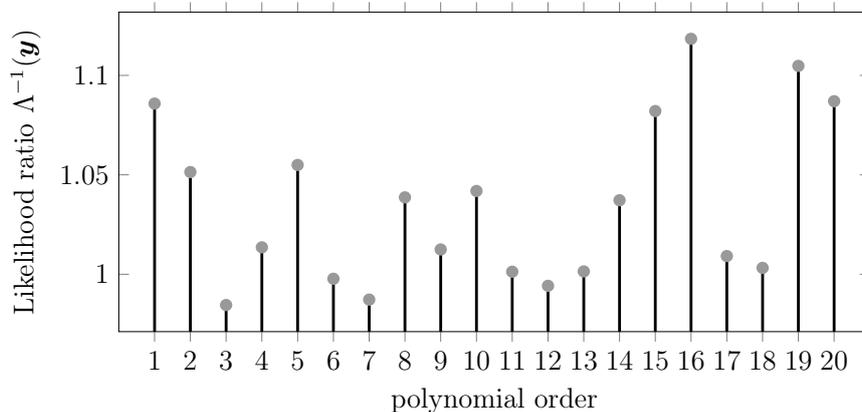
Figure 9. Generalized likelihood ratio (GLR) for Example 10 (polynomial model). The vertical axis values represent the likelihood ratio, and the horizontal axis represents the polynomial order used for generating the $H^{(i)}$ matrix.
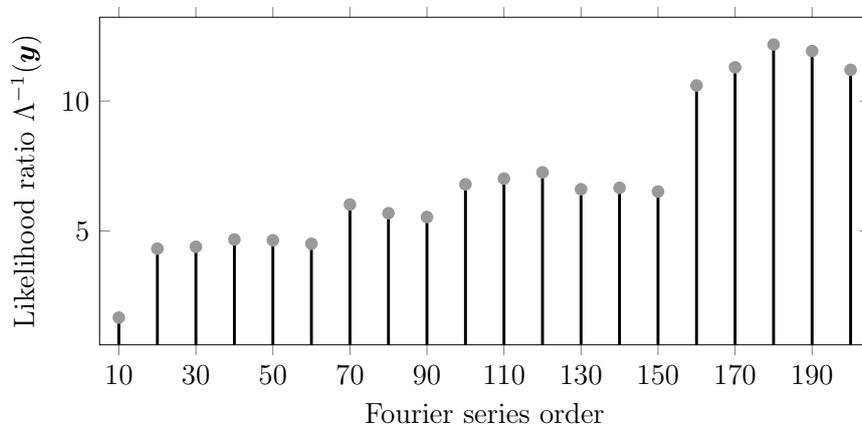
Figure 10. GLR for Example 12 (Fourier model) with different orders. The vertical axis values represent the likelihood ratio and the horizontal axis represents the Fourier series order used for generating the $H^{(i)}$ matrix.

8.3. Assessment of the Performance Improvement for the SICK 200 LiDAR

Assume considering the best model order as in Section 8.2 and applying the main EM estimation algorithm for testing purposes. Figure 11 then shows an example of the outcome of this test on real data: first, the figure plots for each raw measurement the associated lasing modes detected by the EM testing procedure. The same plot shows also the bias induced by the temperature, *i.e.*, $H\widehat{\boldsymbol{\theta}}$.

Once the raw measurements have been associated with the corresponding lasing modes, it is then possible to subtract from them the corresponding noise bias $\widehat{\mu}_1$ or $\widehat{\mu}_2$; in other words, it is possible to remove a certain lasing mode-dependent bias from each of the various measurements. This leads to a new distribution of the measurements, as shown in Figure 12. This figure shows the cumulative distribution of the measurements that we plotted in Figure 11 before and after subtracting the biases

induced by the lasing modes. It is clearly visible that after this removal, the cumulative distribution becomes sharper, indicating that the novel dataset has a smaller variance.
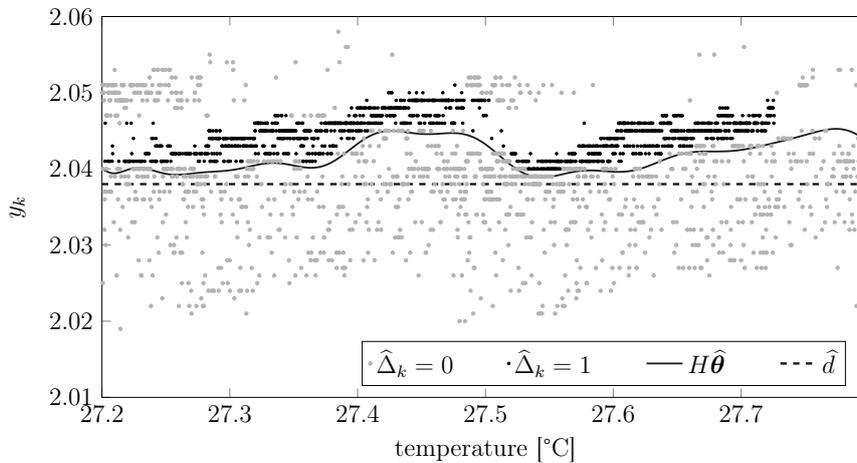

Figure 11. Application of the EM testing algorithm on real data, where the raw measurements are plotted *versus* the temperature of the device. The black dots are the raw measurements that have been associated with lasing Mode 1, while the gray dots are the measurements associated with lasing Mode 2. The solid line denotes the temperature-induced bias on the measured distance.
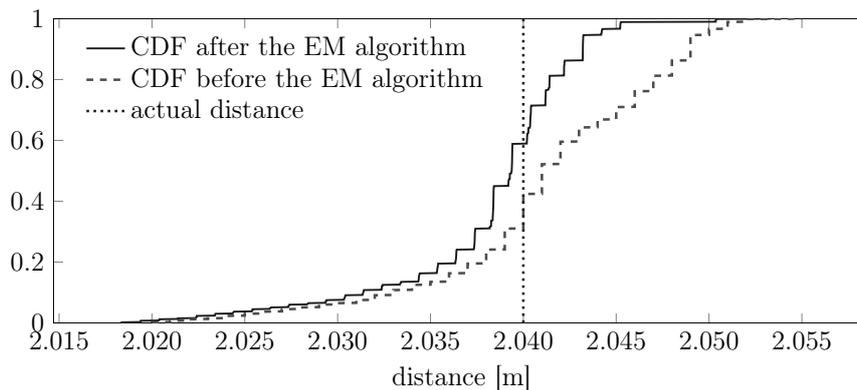
Figure 12. Effects of compensating some raw data with the biases corresponding to the lasing mode of each sample.

8.4. Convergence Properties for the EM Algorithms

We now report some numerical results for the convergence properties of the EM training and test algorithms. More precisely, Figure 13 shows the evolution of the estimates during the training phase, while Figure 14 shows the same evolution for the testing phase. We also plot in Figure 15 the computational burden of performing a fixed number of EM test steps for different measurement vector lengths and show empirically how the computational efforts for testing some points are quadratic with the number of samples. Importantly, this computational effort does not depend heavily on the order of the model.

We show in Figure 16 how the precision and accuracy of the estimate $\widehat{d}$ changes with the number of samples in the test set. More precisely, we show the following cumulative distribution: given the set of 20,000 raw measurements, extract every subset of 200, 400 or 800 consecutive measurements and, for each of these subsets, compute an estimate $\widehat{d}$. The result is intuitive: estimating from more measurements leads to a better estimator (*i.e.*, a cumulative distribution that steepens in correspondence of the true distance).
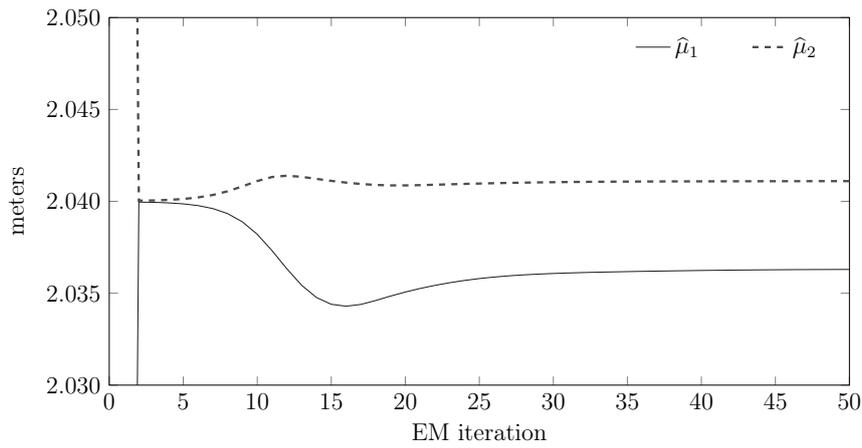
Figure 13. Convergence of the estimated means to their final estimates during the training phase for a training set containing 4000 raw measurements.
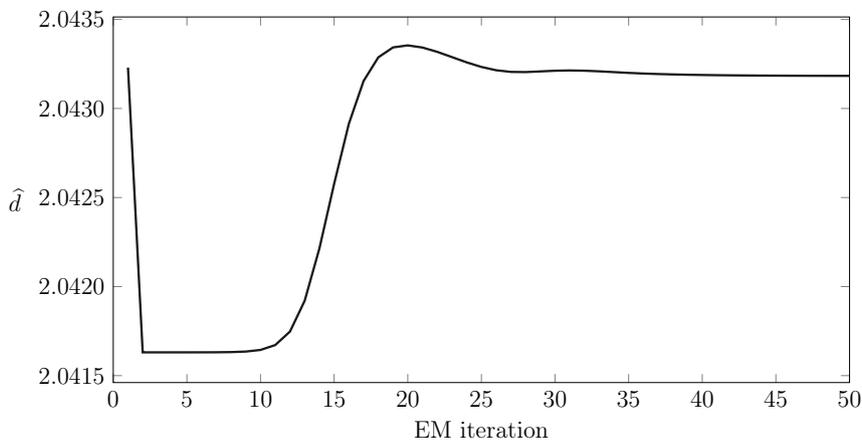

Figure 14. Convergence of the estimated distance to its final estimate during the test phase for a test set containing 4000 raw measurements.
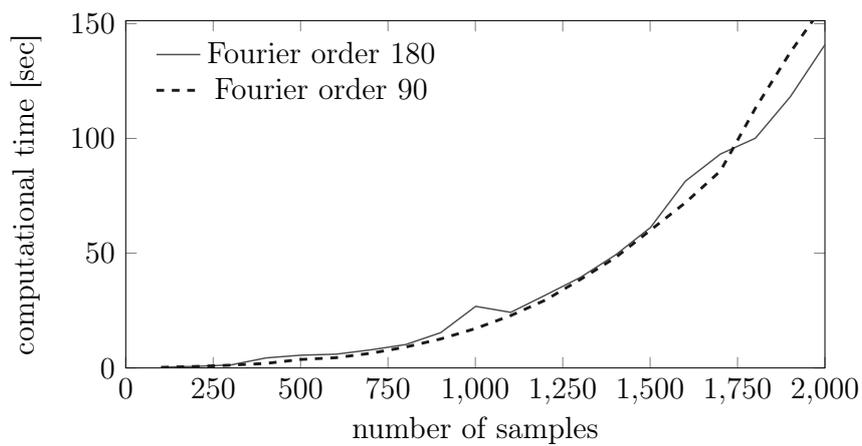

Figure 15. Dependence of the computational complexity for the EM testing phase for two different Fourier order models as a function of the number of samples in the test set.
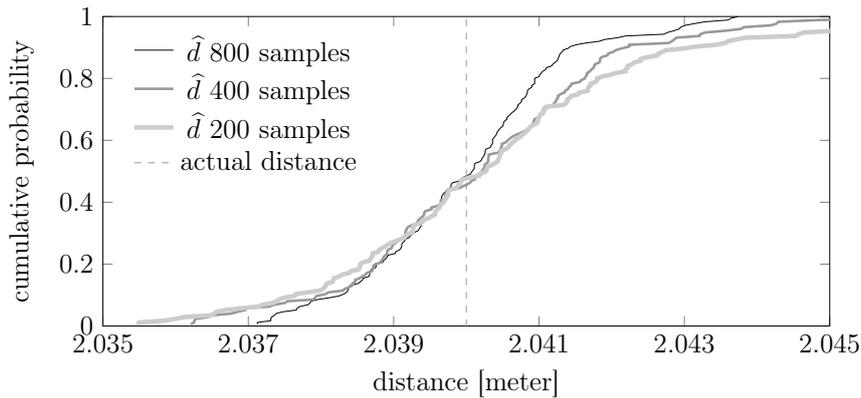
Figure 16. Cumulative distributions of the estimate $\widehat{d}$ for different testing sample sizes. The device temperature ranged in these experiments between 24.8 °C and 28.8 °C.

We show in Figure 17 the realizations of the errors committed by three different estimators: $\widehat{d}$ (our estimator), $\overline{y}$ (the sample average of a window of measurements) and $\overline{y - H\widehat{\theta}}$ (the sample average of a window of measurements corrected by the temperature compensation term $H\widehat{\theta}$). We notice that $\widehat{d}$ commits the smallest errors almost everywhere. To integrate the information, we also plot the temperature of the device in the lower plot. This helps with recognizing that $\widehat{d}$ has better performance, especially where the effect of the temperature is higher.
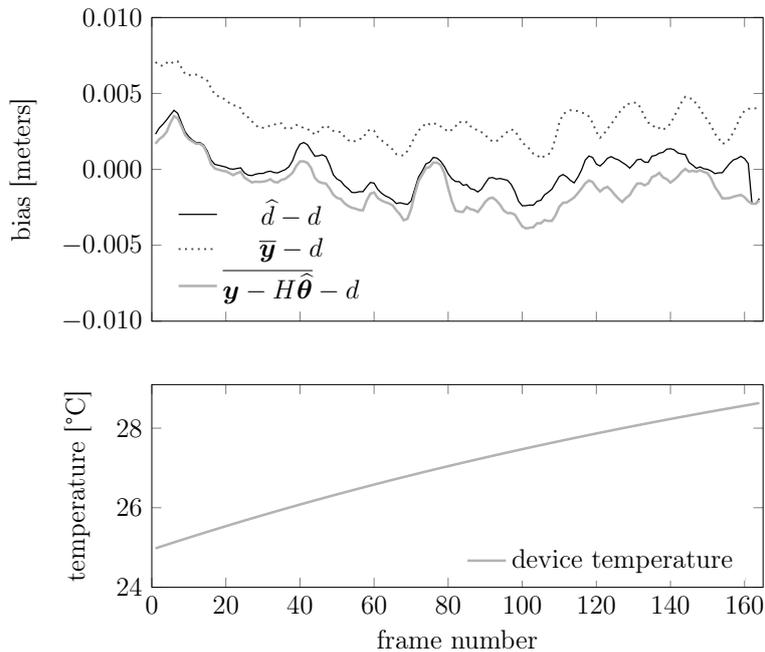

Figure 17. Realizations of the errors committed by the three estimators $\widehat{d}$, $\overline{y}$ and $\overline{y - H\theta}$. In this experiment, the temperature of the device ranged between 24.8 °C and 28.8 °C; the sample size is 1000 points.

We also plot in Figures 18 and 19 how the statistical moments of the empirical error and of the empirical absolute error of these estimators behave with the test sample size. Each point in the plot represents the mean or the variance of an error (or absolute error) sequence similar to those shown in Figure 17. We see that the performance of $\widehat{d}$ improves monotonically, while for the other estimators, this does not happen. Similar results happen also in the plots for the variances, where the proposed estimator outperforms the other two.
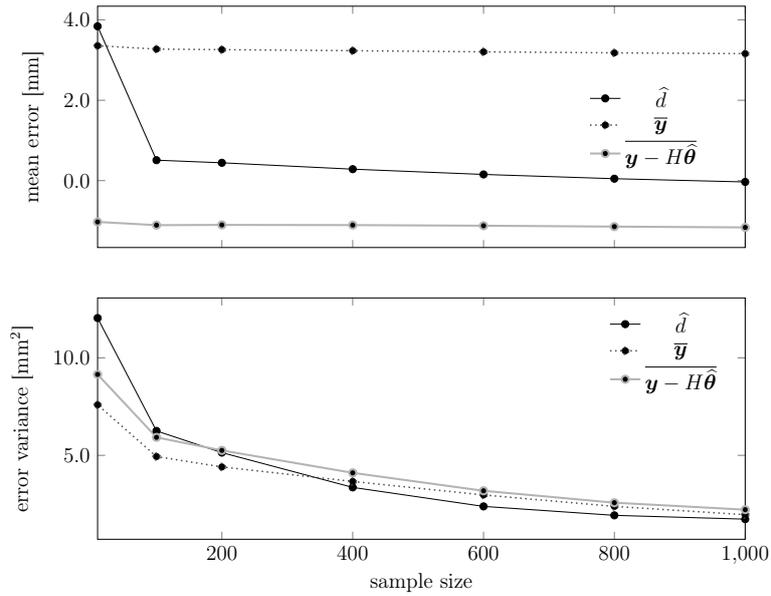
Figure 18. Dependency of the statistical moments of the various estimators on the test sample size. The upper plot shows the mean error, while the lower plot shows the variance of the error. In these experiments, the temperature of the device ranged between 24.8 °C and 28.8 °C.
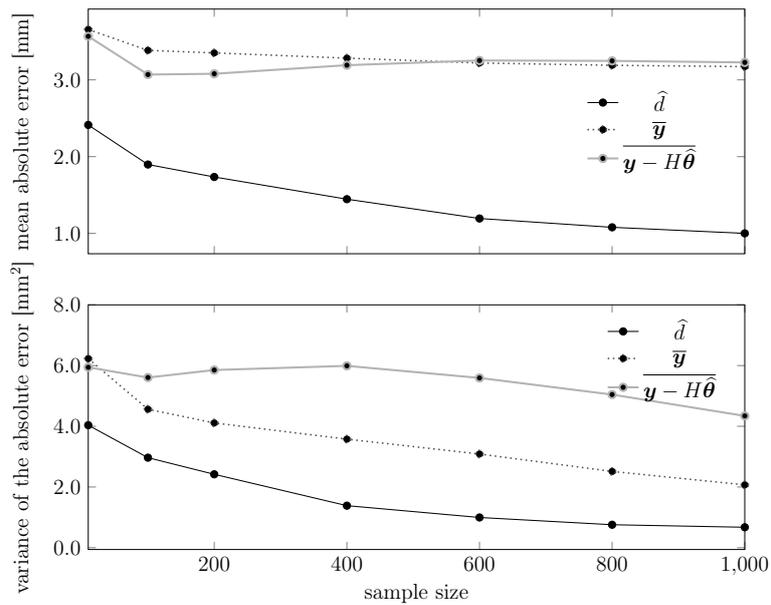

Figure 19. Dependency of the statistical moments of the absolute error of the various estimators on the test sample size. The upper plot shows the mean absolute error, while the lower plot shows the variance of the absolute error. In these experiments, the temperature of the device ranged between 24.8 °C and 28.8 °C.

9. Conclusions

Physical considerations on the mode-hopping effect lead to the consideration that the measurement noise of time of flight (ToF) laser scanners is intrinsically multi-modal. In its turn, this implies that estimating the actual distance between scanners and the surrounding objects should be performed using latent variable-based statistical models, where the latent variables correspond to the lasing modes of the laser. Since no literature seems to account for this multi-modality, we aimed at closing this gap.

We thus proposed an expectation maximization (EM) algorithm on top of a model that captures biases on the measured distance induced by temperature changes, plus mode-hopping effects through an opportune Gaussian mixture on the measurement noise. Importantly, thanks to a separable model, the EM iterations can be performed analytically. The computational advantages are clear: non-separable models may indeed need to perform the EM iterations numerically, and this would lead to a computational burden hindering the usability of mode-hopping correction procedures in on-line settings.

The proposed strategy incorporates an accurate model of the temperature dynamics of the laser diode. Moreover, to account for the fact that it may be difficult to collect temperature data on the laser, we proposed a strategy for the identification of the parameters of this model that exploit the datasheet of the laser device and some very simple experiments.

Overall, the proposed temperature compensation strategy led to diminishing the spread of the distribution of real measurements from a SICK LMS 200 around the true value *d*, as shown in Figure 12, with corresponding decays of the variance of the absolute error from 2.0 mm² to 0.68 mm² and shift of the expected absolute error from 3 mm to 1 mm, as shown in Figure 19.

We eventually notice that the proposed strategy is still in its infancy: indeed, we considered a frequentist case for which the actual distance *d* is a deterministic and fixed quantity, aiming at showing that it is possible to improve the overall precision of a laser scanner through accounting for mode-hopping effects. Nonetheless, in real scenarios, *d* will vary; we thus devise future efforts focusing on strategies for which *d* is a stochastic process that varies according to some given *a priori* distribution.

Author Contributions: The authors contributed equally to the development of the paper.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix

Proof of Proposition 1. By the definition of conditional probability,

$$
\begin{aligned}
\mathbb{P}\left[y, \widehat{t^j} \; ; \; d, \overline{\theta}, \Delta\right] &= \mathbb{P}\left[y \mid \widehat{t^j} \; ; \; d, \overline{\theta}, \Delta\right] \mathbb{P}\left[\widehat{t^j} \; ; \; d, \overline{\theta}, \Delta\right] \\
&= \mathbb{P}\left[y \mid \widehat{t^j} \; ; \; d, \overline{\theta}, \Delta\right] \mathbb{P}\left[\widehat{t^j}\right]
\end{aligned}
\tag{30}
$$

since $\widehat{t^j}$ does not statistically depend on the (deterministic) parameters d, $\overline{\theta}$ and Δ. □

Proof of the structure of the EM training algorithm. The probability density of y given $\widehat{t^j}$ (and thus H) is:

$$
\begin{aligned}
p\left(y \mid \widehat{t^j} \; ; \; d, \overline{\theta}, \Delta\right) = \\
(1 - \pi)\mathcal{N}\left(y - d\mathbb{1} - H\theta - \mu_1\mathbb{1}, \sigma_1^2\right) \\
+ \pi\mathcal{N}\left(y - d\mathbb{1} - H\theta - \mu_2\mathbb{1}, \sigma_2^2\right)
\end{aligned}
\tag{31}
$$

Knowing the latent variables Δ leads thus to the log-likelihood:

$$
\begin{aligned}
\ell\left(\overline{\theta} \; ; \; y, \widehat{t^j}, d, \Delta\right) = \\
(y - d\mathbb{1} - H\theta - \mu_1\mathbb{1})^T \Sigma_1 (y - d\mathbb{1} - H\theta - \mu_1\mathbb{1}) \\
+ (y - d\mathbb{1} - H\theta - \mu_2\mathbb{1})^T \Sigma_2 (y - d\mathbb{1} - H\theta - \mu_2\mathbb{1})
\end{aligned}
\tag{32}
$$

Calculating the zero of the score with respect to μ_i and then calculating the zero of the score with respect to θ leads to the update rules for both $\widehat{\mu}_i$ and $\widehat{\theta}_i$.

As for the update rule for $\widehat{\sigma}_i^2$, we notice that it derives immediately from the fact that the score with respect to σ_i is:

$$-\frac{K}{\sigma_i^2} + \frac{1}{\sigma_i^4}\left((\boldsymbol{y} - d\mathbb{1} - H\boldsymbol{\theta} + \mu_1\mathbb{1})^T \frac{\Sigma_1}{K}(\boldsymbol{y} - d\mathbb{1} - H\boldsymbol{\theta} + \mu_1\mathbb{1})\right.$$

$$\left. + (\boldsymbol{y} - d\mathbb{1} - H\boldsymbol{\theta} - \mu_2\mathbb{1})^T \frac{\Sigma_2}{K}(\boldsymbol{y} - d\mathbb{1} - H\boldsymbol{\theta} - \mu_2\mathbb{1})\right) \tag{33}$$

□

Proof of the structure of the EM testing algorithm. The update rules come directly from the structure of the log-likelihoods in the proof of the structure of the EM training algorithm. □

References

1. One Photon Per Pixel Produces 3-D LiDAR Image, 2013. Available online: www.photonics.com (accessed on 8 December 2015).
2. Pulse Ranging Technology (PRT)—Highly Accurate Positioning over Large Distances. Available online: www.pepperl-fuchs.us (accessed on 8 December 2015).
3. Hebert, M.; Krotkov, E. 3-D measurements from imaging laser radars: How good are they? In Proceedings of the IEEE/RSJ International Workshop on Intelligent Robots and Systems' 91. Intelligence for Mechanical Systems (IROS'91), Osaka, Japan, 3–5 November 1991; pp. 359–364.
4. Reina, A.; Gonzales, J. Characterization of a radial laser scanner for mobile robot navigation. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, (IROS'97), Grenoble, France, 6 Septembre 1997; pp. 579–585.
5. Matko, V. Next generation AT-cut quartz crystal sensing devices. *Sensors* **2011**, *11*, 4474–4482.
6. Semiconductor, O.O. Pulsed Laser Diode in Plastic Package 25 W Peak Power SPL PL90. In *Data sheet, OSRAM Opto Semiconductor*; OSRAM Opto Semiconductors GmbH: Regensburg, Germany, 2014; pp.244–248.
7. Hawkes, J.; Latimer, I. *Lasers: Theory and Practice*; Prentice Hall: Maylands Ave, UK, 1995.
8. Paschotta, R. *Article on "Mode Hopping" in the Encyclopedia of Laser Physics and Technology*; Wiley-VCH: Weinheim, Germany, 2008.
9. Tell, B.; Brown-Goebeler, K.; Leibenguth, R.; Baez, F.; Lee, Y. Temperature dependence of GaAs-AlGaAs vertical cavity surface emitting lasers. *Appl. Phys. Lett.* **1992**, *60*, 683–685.
10. Ljung, L. *System Identification: Theory for the User*; Prentice-Hall Information and System Sciences Series, Prentice Hall PTR: Englewood Cliff, NJ, USA, 1999.
11. Paschotta, R. *Article on "laser Diodes" in the Encyclopedia of Laser Physics And Technology*; Wiley-VCH: Weinheim, Germany, 2008.
12. Anderson, B.D.; Moore, J.B. Optimal Filtering; In *Courier Corporationz*; Dover Publications, Inc.: Mineola, NY, USA, 2012.
13. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Ser. B Methodol.* **1977**, *39*, 1–38.
14. McLachlan, G.; Peel, D. *Finite Mixture Models*; John Wiley & Sons: NY, USA, 2004.
15. Vaida, F. Parameter convergence for EM and MM algorithms. *Stat. Sin.* **2005**, *15*, 831–835.
16. Basseville, M.; Nikiforov, I.V.; *Detection of Abrupt Changes: Theory and Application*; Prentice Hall Englewood Cliffs: Englewood Cliff, NJ, USA, 1993.