Dynamic and relative calibration of temperature sensors in the case of uncertain parameters

ISSN 0253-2778
CODEN CKHPDZ

中国科学技术大学学报

JOURNAL OF UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA

2012 5 第 42 卷 /Vol.42

中国科学技术大学学报

第四十二卷

第五期

二〇一二年五月

中文科技综合类核心期刊
中国科技核心学术期刊

JOURNAL OF UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA

Vol. 42 No. 5(Serial No. 229) May 2012

CONTENTS

中国科学技术大学学报

第 42 卷　第 5 期(总第 229 期)　2012 年 5 月

目　　次

[期刊基本参数]　CN34-1054/N * 1965 * m * A4 * 88 * zh/en * P * ￥30.00 * 1200 * 11 * 2012-05

责任编辑(组稿编辑)　李玉红　　**英文编辑**　崔海建　Federick Firstbrook

第42卷第5期 中国科学技术大学学报 Vol. 42, No. 5
2012年5月 JOURNAL OF UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA May 2012

Article ID: 0253-2778(2012)05-0345-12

Dynamic and relative calibration of temperature sensors in the case of uncertain parameters

Canuto Enrico, Ospina José A, Acuna-Bravo Wilber

(*Politecnico di Torino, Dipartimento di Automatica e Informatica, Corso Duca degli Abruzzi* 24, 10129 *Torino, Italy*)

Abstract: The paper addresses sensor bias calibration when only differential bias is of interest as in relative calibration of different sensors. Temperature sensors are considered. Relative calibration is of interest in control problems where precise thermal gradients must be generated, or uniform and stable temperatures must be guaranteed. Static (or steady-state) and dynamic calibration are compared both theoretically and experimentally. Dynamic calibration has the advantage of employing any available measurement, but requires that a suitable dynamic model of the calibration equipment be available. A simple equipment with three sensors is considered in the paper, but the results can be extended to more complex ones. As a further advantage, calibration can be performed on any apparatus with the constraint (used in the paper) that the dynamic model is of the same order as the sensor size. Identifiability conditions are proved. Calibration is obtained through a nonlinear weighted least squares problem, which is solved in an iterative way. Convergence, consistency and asymptotic efficiency are proved and verified with Monte Carlo simulations.

Key words: relative calibration; system identification; weighted least squares; thermal sensors; identifiability

CLC number: TH811 **Document code**: A doi: 10.3969/j.issn.0253-2778.2012.05.001

Citation: Canuto Enrico, Ospina José A, Acuna-Bravo Wilber. Dynamic and relative calibration of temperature sensors in the case of uncertain parameters[J]. Journal of University of Science and Technology of China, 2012, 42(5): 345-356.

具有不确定参数的温度传感器的动态相对校准

Canuto Enrico, Ospina José A, Acuna-Bravo Wilber

(都灵理工大学自动化与电子信息科学系,意大利都灵 10129)

摘要:存在动态偏差情况下的传感器校准,与多个传感器的相对校准类似.实现相对校准对许多控制问题都很重要,比如需要产生精确温度梯度的问题、保持均匀与稳定温度场的问题.本文以温度传感器为例,对静态(稳态)校准与动态校准从理论和实验两个角度进行了对比.动态校准的一个优势是其适用于任何测量数据,但其需要建立一个测量仪器的动态模型.尽管本文只研究了一个由三个传感器构成的简单设备,但其结果也适用于更复杂的情况.这正是因为动态校准的另一个显著优势,即其能在任何设备上进行校准,校准的一个

Received: 2012-04-13; **Revised**: 2012-05-18

Biography: Canuto Enrico (corresponding author), male, born in 1945, Professor. E-mail: enrico.canuto@polito.it

The earlier version has been accepted for 2012 *Chinese Control Conference.*

前提是所用的动态模型与传感器具有相同的阶次.本文证明了可辨识性(可校准性)条件,并通过以迭代方式求解一个非线性加权最小二乘问题,实现了校准.对解的收敛性、一致性和渐近效率,不但给予了理论证明,且通过蒙特卡罗模拟实验进行了验证.

关键词:相对校准;系统辨识;加权最小二乘;温度传感器

0 Introduction

The paper addresses sensor bias calibration when only differential bias is of interest as in the relative calibration of different sensors. Temperature sensors are considered. Relative calibration is distinguished from calibration by comparison[1] where a sensor is calibrated with respect to a reference. Here none of the sensors need to be a reference, but instead their differential uncertainty is estimated, and no effort is done of relating uncertainty to either sensor. As such, calibration by comparison is a specific case of relative calibration, since the estimated uncertainty is entirely related to non reference sensors. The principles of dynamic calibration are outlined and employed in the paper, marking a difference from the usual calibration by comparison, as the latter is performed under steady state conditions and only aims to identify the sensor response uncertainty. Dynamic calibration, to be performed under any sensor condition, calls for system (nonlinear) identification. Among the several methods available in the Refs. [2-5] covariance inversion is justified and compared to other methods. A brief outline and discussion of the different concepts that have been just expressed, follows.

(Ⅰ) Relative calibration

Calibration by comparison is a method for estimating the response curve of sensors. In the case of thermometers, two sensors (one to be calibrated and the other to be used as a reference) are brought to the same temperature by placing them inside a thermal well. The reference sensor measures the varying temperature of the well. Moving the well along a pre-specified scale of temperatures, a table is obtained, relating the sensor output to be calibrated to the well temperature. Each pair of tabulated values is accompanied by the uncertainty of the temperature value, which is sometimes called "tolerance" (see Ref. [1] for terminology). The tabulated uncertainty of the temperature values is obtained by combining all uncertainty sources affecting the calibration[6].

It is common to find applications where the absolute temperature of some objects is not of particular interest; on the contrary, the uncertainty of the difference between two or more temperatures at two or more points is of interest. This occurs in applications where temperature must be uniform and stable, such as in Refs. [7] and [8], or when precise thermal gradients must be generated as in Ref. [9]. In fact, the results shown here are motivated by the design of a temperature regulator for an optical cavity which must generate precise thermal gradients to create appropriate temperature profiles along the cavity[10]. In these applications, calibration by comparison can be employed even in the absence of a high quality reference sensor. The analysis of the reduction of the differential uncertainty after calibration by comparison is reported in Ref. [11]. To mark the difference between calibration with respect to a reference (calibration by comparison) and calibration of the differential uncertainty, the term relative calibration is adopted.

(Ⅱ) Dynamic relative calibration

Relative calibration[11] calls for a thermodynamic equipment in which the sensors being calibrated are placed inside a calibration well and are taken to the same temperature, thus reducing as much as possible any gradient among them. Design errors, such as inappropriate or asymmetrical ambient insulation, lead to calibration errors. Sensors are moved to the same temperature by a symmetrical heating system and by waiting enough time for the assembly to reach

steady-state conditions. However gradient reduction cannot be obtained when calibrating large sensor networks and regression with respect to sensor location can be a solution assuming steady state[12].

In this paper an extension is considered. Instead of assuming that sensors are at the same temperature and that the assembly has reached a steady-state condition, a thermodynamic model of the equipment is considered. Based on this model, significant parameters of the calibration equipment can be estimated, improving the estimation of the sensor bias as well. To this end, the estimation algorithm includes the dynamical behavior of the assembly, which renders it unnecessary to wait for the steady-state. Here dynamic and relative calibration is applied to a pair of sensors, but it can be extended to an arbitrary set of sensors.

(Ⅲ) Identification method and results

The measurement equations for applying regression to dynamic models are affected by colored noise and correlation between the past variables entering the regression matrix. As a result, the parameter estimate loses consistency, and alternative methods like prediction error or instrumental variables methods are employed to recover consistency[2]. The regression variables are filtered to create a prediction error which is white and therefore uncorrelated with the past variables[2]. Error whitening being the goal, an alternative method is to directly whiten the error by inverting the covariance matrix, which may become awkward in the case where the error covariance depends on unknown parameters as in the cases treated here and in Ref. [2]. It is shown that covariance inversion becomes effective if the error components affected by unknown parameters can be made sufficiently negligible, without impairing identification performance.

Consider the first-order multivariate ARMAX model[2]

$$\boldsymbol{y}(i+1) = -\boldsymbol{A}\boldsymbol{y}(i) + \boldsymbol{B}\boldsymbol{u}(i) + \boldsymbol{e}(i+1) - \boldsymbol{C}\boldsymbol{e}(i) \tag{1}$$

where $\boldsymbol{y}$ is the output vector, $\boldsymbol{u}$ is a known input vector, $\boldsymbol{e}$ is an unknown vector, and $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ are matrices to be identified. Differences from Eq. (1) can be appreciated by writing the calibration equation to be investigated with notations similar to Eq. (1) (other notations will be used throughout), namely

$$\begin{aligned}(\boldsymbol{I} - \boldsymbol{H}(\mathrm{A}))\,\boldsymbol{y}(i+1) - \boldsymbol{y}(i) = \\ \boldsymbol{A}(\boldsymbol{y}(i) - \boldsymbol{s}) + \boldsymbol{B}(\boldsymbol{u}(i) - \boldsymbol{s}_u) + \\ \boldsymbol{e}(i+1) - \boldsymbol{e}(i) - \boldsymbol{A}\boldsymbol{e}(i) - \boldsymbol{B}\boldsymbol{e}_u(i)\end{aligned} \tag{2}$$

The regression is nonlinear (actually bilinear in the parameters) because of the unknown biases $\boldsymbol{s}$ and $\boldsymbol{s}_u$ affecting output and input vectors and because of the matrix $\boldsymbol{H}(\boldsymbol{A})$. The unknown matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ enter the error component. If $|\boldsymbol{A}| \ll 1$, $|\boldsymbol{B}| \ll 1$, $|\boldsymbol{H}| \ll 1$, and the variance of $\boldsymbol{B}\boldsymbol{e}_u$ is of the same order of $\boldsymbol{e}$, then the error in Eq. (2) can be approximated to $\boldsymbol{e}(i+1) - \boldsymbol{e}(i)$, which justifies the covariance inversion.

Monte Carlo runs confirm that the differential bias estimate is unbiased also for a finite number of observations unlike the static case. The overall parameter estimate is proven to be consistent and asymptotically efficient.

1 Dynamic model

1.1 State and measurement equations

The thermodynamic apparatus in Fig. 1 consists of two bodies of capacitance C_1 and C_2 (unit: J/K) where the sensors to be calibrated are immersed. The sensors measure the mean temperatures (unit: K) θ_1 and θ_2 of the bodies. Each body is supplied with thermal power by separate actuators, whose commands u_1 and u_2 represent the fraction of the peak power $P_{k,\max}$, $k=1,2$. The bodies are thermally linked (g (unit: W/K) in Fig. 1) to facilitate thermal equilibrium. The ambient surrounding the bodies is a metal chamber which can be thermally regulated, but not in this

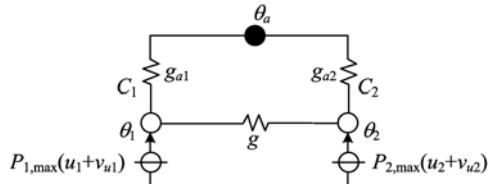

Fig. 1　Lumped-parameter model of the thermal apparatus

treatment. The ambient temperature θ_a is measured by a third sensor. The thermal conductances body-to-chamber ($g_{a1} = g_a + \Delta g_a$, $g_{a2} = g_a - \Delta g_a$ (unit: W/K) in Fig. 1) are much lower than g and nominally equal to g_a.

The pairs (θ_1, θ_2) and (u_1, u_2) are denoted with the vectors $\boldsymbol{\theta}$ and $\boldsymbol{u}$, respectively. Using the total capacitance $C_{12} = C_1 + C_2$ as a scale factor, the continuous-time dynamic model of the assembly in Fig. 1 can be written as

$$\dot{\boldsymbol{\theta}}(t) = \boldsymbol{A}\boldsymbol{\theta}(t) + \boldsymbol{B}_a\theta_a(t) + \boldsymbol{B}_u(\boldsymbol{u} + \boldsymbol{v}_u)(t),\ \boldsymbol{\theta}(0) = \boldsymbol{\theta}_0 \tag{3}$$

where

$$\boldsymbol{A} = \boldsymbol{\Gamma}^{-1}\boldsymbol{G},\ \boldsymbol{B}_a = \boldsymbol{\Gamma}^{-1}\boldsymbol{G}_a,\ \boldsymbol{B}_u = \boldsymbol{\Gamma}^{-1}\boldsymbol{P} \tag{4}$$

and

$$\left.\begin{aligned} \boldsymbol{\Gamma} &= C_{12}\begin{bmatrix}\gamma_1 & 0\\ 0 & 1-\gamma_1\end{bmatrix},\ \gamma_1 = \frac{C_1}{C_2}\\ \boldsymbol{G} &= \begin{bmatrix}-g-g_{a1} & g\\ g & -g-g_{a2}\end{bmatrix}\\ \boldsymbol{P} &= \begin{bmatrix}P_{1,\max} & 0\\ 0 & P_{2,\max}\end{bmatrix},\ \boldsymbol{G}_a = \begin{bmatrix}g_{a1}\\ g_{a2}\end{bmatrix}\end{aligned}\right\} \tag{5}$$

Given a time unit T and the discrete time i the sampled-data version of Eq. (3) becomes

$$\begin{aligned}&(\boldsymbol{I} - \boldsymbol{H})(\boldsymbol{\theta}(i+1) - \boldsymbol{\theta}(i)) =\\ &\quad \boldsymbol{F}\boldsymbol{\theta}(i) + \boldsymbol{F}_a\boldsymbol{\theta}_a(i) + \boldsymbol{F}_u(\boldsymbol{u} + \boldsymbol{v}_u)(i)\end{aligned} \tag{6}$$

where the correction matrix $\boldsymbol{H}$ holds

$$\boldsymbol{H} = \boldsymbol{I} - \boldsymbol{F}(e^{\boldsymbol{F}} - \boldsymbol{I})^{-1} - \boldsymbol{I} = \boldsymbol{F}/2(\boldsymbol{I} + \boldsymbol{F}/3 + \cdots) \tag{7}$$

Entries of the matrices in Eq. (6) are

$$\left.\begin{aligned} \boldsymbol{F} &= \boldsymbol{A}\boldsymbol{T} = \begin{bmatrix}-q_1 - q_{a1} & q_1\\ q_2 & -q_2 - q_{a2}\end{bmatrix}\\ \boldsymbol{F}_a &= \boldsymbol{B}_a\boldsymbol{T} = \begin{bmatrix}q_{a1}\\ q_{a2}\end{bmatrix}\\ \boldsymbol{F}_u &= \boldsymbol{B}_u\boldsymbol{T} = \begin{bmatrix}q_{u1} & 0\\ 0 & q_{u2}\end{bmatrix}\end{aligned}\right\} \tag{8}$$

Parameters in Eq. (8), namely

$$\left.\begin{aligned} q_1 &= gT/(C_{12}\gamma_1)\\ q_2 &= gT/(C_{12}(1-\gamma_1))\\ q_{a1} &= g_{a1}T/(C_{12}\gamma_1)\\ q_{a2} &= g_{a2}T/(C_{12}(1-\gamma_1))\\ q_{u1} &= P_{1,\max}T/(C_{12}\gamma_1)\\ q_{u2} &= P_{2,\max}T/(C_{12}(1-\gamma_1))\end{aligned}\right\} \tag{9}$$

are collected in the vector

$$\boldsymbol{q}_p^{\mathrm{T}} = [q_1 \quad q_2 \quad q_{a1} \quad q_{a2} \quad q_1 \quad q_2] \tag{10}$$

The measurement equations are

$$\left.\begin{aligned} \boldsymbol{y}(i) &= \boldsymbol{\theta}(i) + \boldsymbol{s} + \boldsymbol{v}(i) + \boldsymbol{e}(i)\\ \boldsymbol{y}_u(i) &= \boldsymbol{u}(i)\\ y_a(i) &= \theta_a(i) + s_a + v_a(i)\end{aligned}\right\} \tag{11}$$

where all the variables in the first two equations are bi-dimensional vectors. In Eq. (11) $\boldsymbol{s}$ is the bias vector of sensors 1 and 2, s_a is the bias of the ambient sensor, $\boldsymbol{v}$ and v_a are zero-mean white noises, and $\boldsymbol{e}$ is the model error encoding the neglected dynamics of the model. The covariance matrices of the noise vectors in Eqs. (6) and (11) are defined by

$$\left.\begin{aligned} \boldsymbol{S}_y^2 &= \mathscr{E}\{\boldsymbol{v}(i)\boldsymbol{v}^{\mathrm{T}}(i)\} = \sigma_y^2\boldsymbol{I}\\ \boldsymbol{S}_u^2 &= \mathscr{E}\{\boldsymbol{v}_u(i)\boldsymbol{v}_u^{\mathrm{T}}(i)\} = \sigma_u^2\boldsymbol{I}\\ &\mathscr{E}\{v_a^2(i)\} = \sigma_a^2\end{aligned}\right\} \tag{12}$$

Replacing Eq. (11) in Eq. (6), the following calibration equation is obtained

$$\begin{aligned}&(\boldsymbol{I} - \boldsymbol{H})(\boldsymbol{y}(i+1) - \boldsymbol{v}(i+1) - \boldsymbol{y}(i) + \boldsymbol{v}(i)) =\\ &\quad \boldsymbol{F}(\boldsymbol{y}(i) - \boldsymbol{s} - \boldsymbol{v}(i)) + \boldsymbol{F}_a(y_a(i) - s_a - v_a(i)) +\\ &\quad \boldsymbol{F}_u(\boldsymbol{y}_u(i) - \boldsymbol{v}_u)\end{aligned} \tag{13}$$

The parameters to be estimated are the differential biases

$$\Delta s = s_2 - s_1,\ \Delta s_a = s_a - s_1 \tag{14}$$

to be collected into $\boldsymbol{q}_s^{\mathrm{T}} = [\Delta s \quad \Delta s_a]$, and the vector $\boldsymbol{q}_p$ in Eq. (10). Eq. (13) can be rewritten in a compact form by separating parameters, known terms and errors as follows

$$(\boldsymbol{I} - \boldsymbol{H})\Delta\boldsymbol{y}(i) = \boldsymbol{W}(i)\boldsymbol{q} + \boldsymbol{e}(i) \tag{15}$$

To this end, let us firstly define the measurement differences

$$y_{21} = y_2 - y_1,\ y_{a1} = y_a - y_1,\ y_{a2} = y_a - y_2 \tag{16}$$

Then, the matrix $\boldsymbol{W}$, the vector $\Delta\boldsymbol{y}$, the parameter vector $\boldsymbol{q}$ and the error e in Eq. (15) are found to be

$$\left.\begin{aligned} &\Delta\boldsymbol{y}(i) = \boldsymbol{y}(i+1) - \boldsymbol{y}(i)\\ &\boldsymbol{W}(i) =\\ &\begin{bmatrix}y_{21} & 0 & y_{a1} & 0 & y_{u1} & 0 & -q_1 & -q_{a1}\\ 0 & -y_{21} & 0 & y_{a2} & 0 & y_{u2} & q_2 + q_{a2} & -q_{a2}\end{bmatrix}\\ &\boldsymbol{q}^{\mathrm{T}} = [\boldsymbol{q}_p^{\mathrm{T}} \quad \boldsymbol{q}_s^{\mathrm{T}}]\\ &\boldsymbol{e}(i) = (\boldsymbol{I} - \boldsymbol{H})(\boldsymbol{v}(i+1) - \boldsymbol{v}(i)) -\\ &\qquad \boldsymbol{F}\boldsymbol{v}(i) - \boldsymbol{F}_a v_a(i) + \boldsymbol{F}_u\boldsymbol{v}_u(i)\end{aligned}\right\} \tag{17}$$

Only differential biases Δs and Δs_a, and not the absolute s and s_a, enter the unknown vector $\boldsymbol{q}$, in accordance with the concept of relative calibration. The vector size is $m=8$. Clearly Eq. (17) is a bilinear form in the unknown parameters. Bilinear identification usually calls for bilinearity in the output and input signals[13-14].

1.2 Differential steady state equation

The main goal of a relative calibration procedure is to estimate Δs under uncertain dynamics. It is of interest to compare Eq. (13) to steady-state conditions. Assuming $\dot{\boldsymbol{\theta}}\cong 0$ in Eq. (3) and equal commands $u_1=u_2$, one obtains

$$\left.\begin{aligned}\theta_2-\theta_1 &=-\frac{\Delta g_a(\theta_a-\underline{\theta})}{g+g_a/2}+\frac{P_{2,\max}v_{u2}-P_{1,\max}v_{u1}}{2g+g_a}+\eta\\ \underline{\theta} &=(\theta_1+\theta_2)/2\end{aligned}\right\} \tag{18}$$

where η accounts for deviations from the steady state, and Δg_a is the unknown deviation of g_{a1} and g_{a2} from the nominal value g_a. Taking the difference $y_{21}=y_2-y_1$ in Eq. (11) and replacing $\theta_2-\theta_1$ with Eq. (18), one finds the calibration equation

$$y_{21}=\Delta s-\Delta g_a\frac{y_a-\underline{y}+\Delta s_a-\Delta s/2+v_a-\underline{v}}{g+g_a/2}+\frac{P_{2,\max}v_{u2}-P_{1,\max}v_{u1}}{2g+g_a}+v_2-v_1+\eta \tag{19}$$

where $\underline{y}$ and $\underline{v}$ denote mean values like $\underline{\theta}$ in Eq. (18). The second term in the right-hand side of Eq. (19) with the unknown Δg_a polarizes the estimate of Δs. For such a reason, $y_a-\underline{y}$ is usually regulated to be sufficiently small. To show consistency and efficiency of dynamic calibration, the worst-case of unregulated $\mathbf{y}_a$ is considered.

2 Calibration equations and identifiability

2.1 Static calibration

Static calibration is obtained from Eq. (19) which is rewritten in discrete time as

$$y_{21}(i_k)=\boldsymbol{W}(i)\,\boldsymbol{p}+e_{21}(i_k)+\eta_{21}(i_k) \tag{20}$$

where $i_k(i_k=0,1,\cdots,N_k-1)$ refers to appropriate steady state intervals $k(k=0,\cdots,K-1)$ where Eq. (20) holds. The total measurement size is

$$M=\sum_{k=0}^{K-1}N_k\ll N \tag{21}$$

which is usually much less than the available data length N.

Assuming $|\Delta g_a|\ll 2g$ the entries of Eq. (20) hold

$$\left.\begin{aligned}\boldsymbol{W} &=\begin{bmatrix}1 & \Delta y=y_a-\underline{y}\end{bmatrix}\\ \boldsymbol{p}^{\mathrm{T}} &=\begin{bmatrix}\Delta s & -\Delta g_a/(g+g_a/2)\end{bmatrix}\end{aligned}\right\} \tag{22}$$

and

$$\left.\begin{aligned}\eta_{21} &=\eta-\Delta g_a(\Delta s_a-\Delta s/2)/(g+g_a/2)\\ e_{21} &\cong(P_{2,\max}v_{u2}-P_{1,\max}v_{u1})/(2g+g_a)+v_2-v_1\end{aligned}\right\} \tag{23}$$

Assuming covariance matrices as in Eq. (12), and $P_{k,\max}\cong P_{\max}$, $k=1,2$, the noise covariance becomes

$$\sigma_{21}^2=\mathcal{E}\{e_{21}^2\}=2\left(\frac{P_{\max}}{2g+g_a}\right)^2\sigma_u^2+2\sigma_y^2 \tag{24}$$

Then, splitting Δy in Eq. (22) into the mean $\Delta\underline{y}$ and the alternate component $\Delta\tilde{y}$, i. e. into $\Delta y=\Delta\underline{y}+\Delta\tilde{y}$, denoting the RMS of the alternate component with $\sigma_{\Delta a}^2$, and the form factor as $\varphi_{\Delta y}=\sigma_{\Delta y}/\Delta\underline{y}$, the Cramer-Rao matrix results

$$\boldsymbol{C}^2\cong\sigma_{21}^2(\boldsymbol{W}^{\mathrm{T}}\boldsymbol{W})^{-1}\cong\frac{\sigma_{21}^2}{N}\begin{bmatrix}1+\varphi_{\Delta y}^{-2} & -\Delta\underline{y}\,\sigma_{\Delta y}^{-2}\\ -\Delta\underline{y}\,\sigma_{\Delta y}^{-2} & \sigma_{\Delta y}^{-2}\end{bmatrix} \tag{25}$$

Assuming stationary zero-mean noise e_{21} in Eq. (20), calibration is obtained through ordinary least squares as

$$\hat{\boldsymbol{p}}=(\boldsymbol{W}^{\mathrm{T}}\boldsymbol{W})^{-1}\boldsymbol{W}^{\mathrm{T}}y_{21} \tag{26}$$

Neglecting the noise component of $y_a-\underline{y}$ in $\boldsymbol{W}$, the polarization of the differential bias can be shown to be approximated by the mean value of η_{21} in Eq. (23), namely

$$\mathcal{E}\{\Delta\hat{s}\}-\Delta s\cong\mathcal{E}\{\eta_{21}\} \tag{27}$$

which is a combination of η and of the unknown conductance uncertainty Δg_a. The variance of $\mathcal{E}\{\Delta\hat{s}\}-\Delta s$ tends to be larger than the entry (1,1) in Eq. (25) because of η, as confirmed by Monte Carlo runs in Section 4.

A goal of dynamic calibration is to provide unbiased and efficient estimates of the differential bias without wasting measurement data.

2.2 Dynamic calibration

Upon collection of N measurement pairs

$\Delta \boldsymbol{y}(i)$, $i=0,\cdots,N-1$, equation Eq. (15) is rewritten in the vector form

$$\Delta \boldsymbol{y}(\boldsymbol{q}) = \boldsymbol{W}(\boldsymbol{q})\boldsymbol{q} + \boldsymbol{e}(\boldsymbol{q}) \tag{28}$$

where $\dim \Delta \boldsymbol{y} = 2N$. Dependence of $\boldsymbol{W}$ and of its entries on $\boldsymbol{q}$ will be usually dropped to simplify notations. Splitting the two components of $\Delta \boldsymbol{y}(i)$ and $\boldsymbol{e}(i)$ in Eq. (28) into separate vectors $\Delta \boldsymbol{y}_j$ and $\boldsymbol{e}_j$, $j=1,2$, Eq. (28) can be block partitioned into

$$\Delta \boldsymbol{y} = \begin{bmatrix} \Delta \boldsymbol{y}_1 \\ \Delta \boldsymbol{y}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{W}_1 \\ \boldsymbol{W}_2 \end{bmatrix} \boldsymbol{q} + \begin{bmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \end{bmatrix} \tag{29}$$

and the blocks of $\boldsymbol{W}$ can be arranged as

$$\boldsymbol{W} = \begin{bmatrix} \boldsymbol{W}_1 \\ \boldsymbol{W}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{y}_{21} & 0 & \boldsymbol{y}_{a1} & 0 & \boldsymbol{y}_{u1} & 0 & -q_1\boldsymbol{\eta} & -q_{a1}\boldsymbol{\eta} \\ 0 & -\boldsymbol{y}_{21} & 0 & \boldsymbol{y}_{a2} & 0 & \boldsymbol{y}_{u2} & (q_2+q_{a2})\boldsymbol{u} & -q_{a2}\boldsymbol{\eta} \end{bmatrix} \tag{30}$$

The vectors $\boldsymbol{y}_x(x=21,a1,a2,u1,u2)$ in Eq. (30) have N components $\boldsymbol{y}_x(i)$, and all components of $\boldsymbol{\eta}$ are unitary.

Eq. (28) is bilinear because $\boldsymbol{W}$ is a function of $\boldsymbol{q}$, and the same holds for $\Delta \boldsymbol{y}$ and $\boldsymbol{e}$. The latter is also correlated between successive samples. As a result, the error covariance matrix $\boldsymbol{S}_e^2$ is non diagonal and band-type as

$$\boldsymbol{S}_e^2(\boldsymbol{q}) = \begin{bmatrix} \boldsymbol{S}^2 & \boldsymbol{R} & \cdots & 0 \\ \boldsymbol{R}^{\mathrm{T}} & \boldsymbol{S}^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{S}^2 \end{bmatrix} \tag{31}$$

Using Eqs. (12) and (8) the sub-matrices in Eq. (31) hold

$$\left.\begin{aligned} \boldsymbol{S}^2(\boldsymbol{q}) &= \sigma_y^2(\boldsymbol{I}-\boldsymbol{H})(2\boldsymbol{I}+\boldsymbol{F}\boldsymbol{F}^{\mathrm{T}})(\boldsymbol{I}-\boldsymbol{H})^{\mathrm{T}} + \\ &\quad \sigma_a^2\boldsymbol{F}_a\boldsymbol{F}_a^{\mathrm{T}} + \sigma_u^2\boldsymbol{F}_u\boldsymbol{F}_u^{\mathrm{T}} \\ \boldsymbol{R}(\boldsymbol{q}) &= -\sigma_y^2(\boldsymbol{I}-\boldsymbol{H}+\boldsymbol{F}) \end{aligned}\right\} \tag{32}$$

The treatment of Eq. (28) is based on the following assumption.

Assumption 2.1 The noise variances σ_y^2 and σ_a^2 are of the same magnitude and denoted with σ^2. The norm of the command noise covariance $\sigma_u^2\boldsymbol{F}_u\boldsymbol{F}_u^{\mathrm{T}}$ in Eq. (32) is much lower than σ^2. The sampling time T in Eq. (8) is selected such that the entries of $\boldsymbol{F}$, $\boldsymbol{F}_a$ and $\boldsymbol{F}_u$ are sufficiently smaller than unit.

Asumption 2.1 implies that matxices in Eq. (32) can be simplified as follows

$$\left.\begin{aligned} &\boldsymbol{S}^2(\boldsymbol{q}) = \sigma^2(2\boldsymbol{I}+o(|\boldsymbol{q}|^2)) \cong 2\sigma^2\boldsymbol{I},\ o(|\boldsymbol{q}|^2) \ll 1 \\ &\boldsymbol{R}(\boldsymbol{q}) = -\sigma^2(\boldsymbol{I}+o(|\boldsymbol{q}|)) \cong -\sigma^2\boldsymbol{I},\ o(|\boldsymbol{q}|) \ll 1 \\ &\boldsymbol{H} \cong 0 \end{aligned}\right\} \tag{33}$$

The first consequence of Eq. (33) is that $\boldsymbol{S}_e^2$ can be rearranged into a block-diagonal matrix

$$\boldsymbol{S}_e^2 \cong \begin{bmatrix} \boldsymbol{S}_0^2 & 0 \\ 0 & \boldsymbol{S}_0^2 \end{bmatrix},\ \boldsymbol{S}_0^2 = \sigma^2 \begin{bmatrix} 2 & -1 & \cdots & 0 \\ -1 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 2 \end{bmatrix} \tag{34}$$

The second consequence goes to the entries of $\boldsymbol{W}$. The neglected covariance terms in Eq. (33) derive from noisy variables entering $\boldsymbol{W}$. Therefore, Assumption 2.1 allows treating matrix $\boldsymbol{W}$ as noise-free. The third issue calls for a design of T capable of satisfying Eq. (33). Computation of the Cramer-Rao bound would show that calibration efficiency could be improved to some extent by increasing the discrete-time poles q_x, $x=1,2,a$ in Eq. (9), which can be achieved by a larger T. Thus a design trade-off can be looked for, as mentioned in Section 4.

2.3 Identifiability conditions

Analysis of the matrix $\boldsymbol{W}$ in Eqs. (28) and (30) reveals identifiability conditions. Since the theorem to be proven calls for time-varying measurements (equivalent to persistent excitation in identification[2]), identifiability conditions are proved assuming noise-free measurements in $\boldsymbol{W}$. Indeed, noisy measurements in $\boldsymbol{W}$ improve time variability (noise is employed to guarantee persistent excitation). Thus identifiability conditions that are proved assuming noise-free $\boldsymbol{W}$ may be referred to as "robust".

Theorem 2.1 A necessary and sufficient condition for the parameter vector $\boldsymbol{q}$ to be "robustly" identifiable from Eq. (28) is that at least one of the pairs (q_1, q_{a2}), (q_2, q_{a1}) and (q_{a2}, q_{a1}) is nonzero, and each triple of noise-free measurements $\{\boldsymbol{y}_{21}, \boldsymbol{y}_{a1}, \boldsymbol{y}_{u1}\}$ and $\{\boldsymbol{y}_{21}, \boldsymbol{y}_{a2}, \boldsymbol{y}_{u2}\}$ contains time varying and linear independent signals.

Proof The first statement derives by taking four blocks of Eq. (17) containing rows at

successive times, say i, $i+1$, $i+2$, $i+3$, and building the relevant $m\times m$ matrix $\boldsymbol{W}(i, i+3)$, which has the same form as of Eq. (30), but the vectors of Eq. (30) are replaced by four dimensional sub-vectors. Since the det$\boldsymbol{W}(i, i+3)$ is a minor of $\boldsymbol{W}$, nonzero minor guarantees rank $\boldsymbol{W}=m$. The determinant holds

$$\det\boldsymbol{W}(i, i+3) = (q_1 q_{a2} + q_2 q_{a1} + q_{a1} q_{a2})d \quad (35)$$

where d only depends on the four-dimensional triples $\{\boldsymbol{y}_{21}, \boldsymbol{y}_{a1}, \boldsymbol{y}_{u1}\}$ and $\{\boldsymbol{y}_{21}, \boldsymbol{y}_{a2}, \boldsymbol{y}_{u2}\}$. Assuming without loss of generality the pair (q_1, q_{a2}) is zero, and reordering columns and rows, $\boldsymbol{W}(i, i+3)$ can be made block-diagonal as follows

$$\left.\begin{aligned}\boldsymbol{W}(i, i+3) &= \begin{bmatrix}\boldsymbol{W}_1 & 0\\ 0 & \boldsymbol{W}_2\end{bmatrix}\\ \boldsymbol{W}_1 &= \begin{bmatrix}\boldsymbol{y}_{21}+\Delta s\boldsymbol{\eta} & \boldsymbol{y}_{a1}+\Delta s_a\boldsymbol{\eta} & \boldsymbol{y}_{u1} & -q_{a1}\boldsymbol{\eta}\end{bmatrix}\end{aligned}\right\} \quad (36)$$

where $\boldsymbol{W}_2$ has a similar form to $\boldsymbol{W}_1$. The necessity for each triple $\{\boldsymbol{y}_{21}, \boldsymbol{y}_{a1}, \boldsymbol{y}_{u1}\}$ and $\{\boldsymbol{y}_{21}, \boldsymbol{y}_{a2}, \boldsymbol{y}_{u2}\}$ to contain linearly independent and time-varying signals emerges from Eq. (36). Sufficiency is proved by reducing variability to the form

$$\boldsymbol{W}_1 = \begin{bmatrix} y_{21}(i)+\Delta s & y_{a1}(i)+\Delta s_a & y_{u1}(i) & -q_{a1}\\ y_{21}(i+1)+\Delta s & y_{a1}(i+1)+\Delta s_a & y_{u1}(i) & -q_{a1}\\ y_{21}(i+1)+\Delta s & y_{a1}(i)+\Delta s_a & y_{u1}(i+1) & -q_{a1}\\ y_{21}(i)+\Delta s & y_{a1}(i+1)+\Delta s_a & y_{u1}(i+1) & -q_{a1}\end{bmatrix} \quad (37)$$

Indeed, the determinant of Eq. (37) holds

$$\Delta\boldsymbol{W}_1 = -q_{a1}(y_{u1}(i) - y_{u1}(i+1)) \times (y_{21}(i) - y_{21}(i+1))(y_{a1}(i) - y_{a1}(i+1)) \quad (38)$$

and is non zero if and only if the command y_{u1} and the differential temperatures y_{21} and y_{a1} are time-varying and linearly independent.

3 Weighted least squares

3.1 Maximum likelihood

Assuming the error $\boldsymbol{e}$ in Eq. (28) is zero-mean Gaussian distributed with the covariance matrix $\boldsymbol{S}_e^2(\boldsymbol{q})$ in Eq. (31), the log-likelihood function of $\boldsymbol{e}$, given $\boldsymbol{q}$, is found to be

$$-\log L_e(\boldsymbol{e}, \boldsymbol{q}) = -\log\det\boldsymbol{S}_e(q) + \frac{1}{2}(\Delta\boldsymbol{y}(\boldsymbol{q}) - \boldsymbol{W}(\boldsymbol{q})\boldsymbol{q})^{\mathrm{T}}\boldsymbol{S}_e^{-2}(\boldsymbol{q})(\Delta\boldsymbol{y}(\boldsymbol{q}) - \boldsymbol{W}(\boldsymbol{q})\boldsymbol{q}) \quad (39)$$

The functional (39) is not in the Gauss-Markov form because of the unknown covariance $\boldsymbol{S}_e^2(\boldsymbol{q})$ and of the parameter-dependent matrix $\boldsymbol{W}(\boldsymbol{q})$. To approximate Eq. (39) with a nonlinear weighted least squares (WLS) functional, a value $\underline{\boldsymbol{q}}$ of the parameter vector is assumed to be known. The uncertainty $\Delta\boldsymbol{q}$ in $\boldsymbol{q}=\underline{\boldsymbol{q}}+\Delta\boldsymbol{q}$ becomes the unknown parameter vector to be identified.

Assumption 2.1 in Section 2 allows to approximate the submatrices in Eq. (32) of the covariance matrix as constant matrices as in Eq. (33), or more accurately as

$$\left.\begin{aligned}\boldsymbol{S}^2(\boldsymbol{q}) &= \boldsymbol{S}^2(\underline{\boldsymbol{q}}) + \sigma^2 o(|\Delta\boldsymbol{q}|^2) \cong \boldsymbol{S}^2(\underline{\boldsymbol{q}}) = \underline{\boldsymbol{S}}^2\\ \boldsymbol{R}(\boldsymbol{q}) &= \boldsymbol{R}(\underline{\boldsymbol{q}}) + \sigma^2 o(|\Delta\boldsymbol{q}|) \cong \boldsymbol{R}(\underline{\boldsymbol{q}}) = \overline{\boldsymbol{R}}\end{aligned}\right\} \quad (40)$$

Thus det$\boldsymbol{S}_e$ becomes known, and det$\boldsymbol{S}_e$ can be pushed out of the log-likelihood function which latter simplifies as follows

$$-\log L_e(\boldsymbol{q}) \cong J(\boldsymbol{q}) = \frac{1}{2}(\Delta\boldsymbol{y}(\boldsymbol{q}) - \boldsymbol{W}(\boldsymbol{q})\boldsymbol{q})^{\mathrm{T}} \cdot \boldsymbol{S}_e^{-2}(\underline{\boldsymbol{q}})(\Delta\boldsymbol{y}(\boldsymbol{q}) - \boldsymbol{W}(\boldsymbol{q})\boldsymbol{q}) \quad (41)$$

The functional (41) is in the WLS form.

One can now compute and set to zero the gradient vector of (41), namely

$$\nabla J(\boldsymbol{q}) = \boldsymbol{g}(\boldsymbol{q}) = \frac{\partial(\boldsymbol{W}\boldsymbol{q})^{\mathrm{T}}}{\partial\boldsymbol{q}}\boldsymbol{S}_e^{-2}(-\Delta\boldsymbol{y}+\boldsymbol{W}\boldsymbol{q}) = 0 \quad (42)$$

where $\partial(\boldsymbol{W}\boldsymbol{q})/\partial\boldsymbol{q}$ denotes the $2N\times m$ Jacobian matrix

$$\frac{\partial(\boldsymbol{W}\boldsymbol{q})}{\partial\boldsymbol{q}} = \boldsymbol{W} + \left[\frac{\partial\boldsymbol{W}}{\partial q_1}\boldsymbol{q} \quad \cdots \quad \frac{\partial\boldsymbol{W}}{\partial q_m}\boldsymbol{q}\right] = \boldsymbol{W} + \Delta\boldsymbol{W} = \boldsymbol{W}_t \quad (43)$$

The second term $\Delta\boldsymbol{W}$ in Eq. (43), which is dependent on $\boldsymbol{q}$, can be shown to be a function of the unknown differential biases in Eq. (14), and the total matrix $\boldsymbol{W}_t$ contains unbiased measurements.

Employing the Jacobian matrix (43) and the following weighted matrices and vectors

$$\left.\begin{aligned}&\boldsymbol{U}_t = \boldsymbol{S}_e^{-1}\boldsymbol{W}_t,\ \Delta\boldsymbol{U} = \boldsymbol{S}_e^{-1}\Delta\boldsymbol{W}\\ &\Delta\boldsymbol{y}_w = \boldsymbol{S}_e^{-1}\Delta\boldsymbol{y} = \boldsymbol{U}\boldsymbol{q} + \boldsymbol{e}_w,\ \boldsymbol{e}_w = \boldsymbol{S}_e^{-1}\boldsymbol{e}\end{aligned}\right\} \quad (44)$$

the gradient Eq. (42) can be rewritten into the "modified" WLS equations

$$\left.\begin{aligned}&\boldsymbol{U}_t^{\mathrm{T}}(-\Delta\boldsymbol{y}_w + (\boldsymbol{U}_t - \Delta\boldsymbol{U})\boldsymbol{q}) = 0\\ &\hat{\boldsymbol{q}} = \boldsymbol{F}^{-1}\boldsymbol{U}^{\mathrm{T}}(\Delta\boldsymbol{y}_w + \Delta\boldsymbol{U}\boldsymbol{q})\end{aligned}\right\} \quad (45)$$

Theorem 2.1 provides the conditions for the Fisher information matrix (parameter dependent)

$$\mathbf{F}(\boldsymbol{q}) = \mathbf{U}_t^{\mathrm{T}}(\boldsymbol{q})\mathbf{U}_t(\boldsymbol{q}) \tag{46}$$

which has been employed in Eq. (45), to be invertible. The solution $\hat{\boldsymbol{q}}$ in Eq. (45) is the parameter estimate to be achieved iteratively as outlined in Section 3.2. The meaning of Eq. (45) is as follows: the measurement vector $\Delta\mathbf{y}_w$ is made unbiased by $\Delta\mathbf{U}\boldsymbol{q}$. Known or zero differential bias would turn Eq. (45) into a classical WLS equation, because of zero $\Delta\mathbf{W}$.

3.2 Least squares iterative solution

The modified WLS Eq. (45) is nonlinear in the parameter vector $\boldsymbol{q}$ and must be solved as an iterative least squares problem. The recursive equation at the kth step, $k=0,1,\cdots$ derives from Eq. (45) and takes the form

$$\left.\begin{aligned}&\hat{\boldsymbol{q}}(k+1) = \mathbf{F}_k^{-1}\mathbf{U}_{t,k}^{\mathrm{T}}(\Delta\mathbf{y}_w + \Delta\mathbf{U}_k\hat{\boldsymbol{q}}(k)),\ \hat{\boldsymbol{q}}(0) = \underline{\boldsymbol{q}}\\&\mathbf{F}_k = \mathbf{U}_{t,k}^{\mathrm{T}}\mathbf{U}_{t,k},\ \mathbf{U}_{t,k} = \mathbf{U}_t(\hat{\boldsymbol{q}}(k)),\ \Delta\mathbf{U}_k = \Delta\mathbf{U}(\hat{\boldsymbol{q}}(k))\end{aligned}\right\} \tag{47}$$

where $\mathbf{F}_k$ is the Fisher information matrix (46) computed for $\boldsymbol{q}=\hat{\boldsymbol{q}}(k)$. Consider the expression of $\Delta\mathbf{y}_w$ in Eq. (44) and substitute it in Eq. (47). The latter equation converts into the classical recursive equation

$$\left.\begin{aligned}&\hat{\boldsymbol{q}}(k+1) = \boldsymbol{q}(k) + \mathbf{B}(k)\boldsymbol{e}_w(k) = \hat{\boldsymbol{q}}(k) - \mathbf{F}_k^{-1}\mathbf{g}(k)\\&\hat{\boldsymbol{q}}(0) = \underline{\boldsymbol{q}},\ \mathbf{B}(k) = \mathbf{F}_k^{-1}\mathbf{U}_{t,k}^{\mathrm{T}}\end{aligned}\right\} \tag{48}$$

where $-\mathbf{g}(k)=\mathbf{U}_{t,k}^{\mathrm{T}}\boldsymbol{e}_w(k)$ is the negative gradient in Eq. (42) computed at $\boldsymbol{q}=\boldsymbol{q}(k)$ and $\mathbf{F}_k$ is positive definite for $N\geqslant N_{\min}\geqslant m$ such that $\mathbf{F}_k$ is full rank for any $k\geqslant 0$. Thus Eq. (48) is a Newton-Raphson algorithm in the average because $\mathbf{F}_k$ can be proven to be the expected value of the Hessian matrix. Convergence in the mean is expressed by the following Lemma.

Lemma 3.1 The iterative Eq. (48) decreases the expected value of the functional (41) at each step as soon as

$$\mathrm{Tr}\mathscr{E}\{(\mathbf{U}_{k+1}-\mathbf{U}_k)^{\mathrm{T}}(\mathbf{U}_{k+1}-\mathbf{U}_k)\mathbf{F}_k^{-1}\} < N \tag{49}$$

Condition (49) is always satisfied, since the left-hand side does not depend on N, being the ratio of two Fisher matrices. Convexity in the mean of Eq. (41), that can be proven, and Lemma 3.1 imply that Eq. (48) quickly converges to the argument of the minimum of the log-likelihood functional J in Eq. (40). Monte Carlo simulations have proven that a few iterations are sufficient (<5). Whether the argument of the minimum equals or not the unknown parameter value $\boldsymbol{q}$ is a problem of consistency.

3.3 Consistency and efficiency

To investigate polarization and covariance, let us rewrite the solution (47) in terms of the unknown "true" $\boldsymbol{q}$ as follows

$$\hat{\boldsymbol{q}} = \boldsymbol{q} + \mathbf{F}^{-1}(\boldsymbol{q})\mathbf{U}_t^{\mathrm{T}}(\boldsymbol{q})\boldsymbol{e}_w(\boldsymbol{q}) \tag{50}$$

In the case of a noise-free $\mathbf{U}_t(\boldsymbol{q})$, as implied by Assumption 2.1 in Section 2, the estimate is unbiased since $\mathscr{E}\{\boldsymbol{e}_w\}=\mathscr{E}\{\boldsymbol{e}\}=0$. To be complete, the noise-free assumption is abandoned by considering that $\mathbf{U}_t(\boldsymbol{q})$ is a "random" matrix because of the measurement noise. By separating the noisy component $\mathbf{V}_t$ as $\mathbf{U}_t=\mathbf{U}_{t0}+\mathbf{V}_t$, inversion in Eq. (50) can be approximated to the first order as

$$\left.\begin{aligned}&\mathbf{F}^{-1} = (\mathbf{U}_t^{\mathrm{T}}\mathbf{U}_t)^{-1} \cong \mathbf{F}_0^{-1}(I - 2\mathbf{V}_t^{\mathrm{T}}\mathbf{U}_{t0}\mathbf{F}_0^{-1})\\&\mathbf{F}_0 = \mathbf{U}_{t0}^{\mathrm{T}}\mathbf{U}_{t0}\end{aligned}\right\} \tag{51}$$

Eq. (50), rewritten with the help of Eq. (51), becomes

$$\hat{\boldsymbol{q}} \cong \boldsymbol{q} + \mathbf{F}_0^{-1}((\mathbf{U}_{t0}+\mathbf{V}_t)^{\mathrm{T}} - 2\mathbf{V}_t^{\mathrm{T}}\mathbf{U}_{t0}\mathbf{F}_0^{-1}\mathbf{U}_{t0}^{\mathrm{T}})\boldsymbol{e}_w(\boldsymbol{q}) \tag{52}$$

The expected value of Eq. (52) is approximated by

$$\mathscr{E}\{\hat{\boldsymbol{q}}\} \cong \boldsymbol{q} + \mathscr{E}\{\mathbf{F}_0^{-1}\mathbf{V}_t^{\mathrm{T}}(\mathbf{I} - 2\mathbf{U}_{t0}\mathbf{F}_0^{-1}\mathbf{U}_{t0}^{\mathrm{T}})\boldsymbol{e}_w\} \tag{53}$$

and shows that $\mathscr{E}\{\hat{\boldsymbol{q}}\}$ is biased, since the second term in the right-hand side is not zero owing to the correlation between $\mathbf{V}_t$ and $\boldsymbol{e}_w$. Since $\mathbf{U}_{t0}\mathbf{F}_0^{-1}\mathbf{U}_{t0}^{\mathrm{T}}$ is a projection operator, one can assume the projected error is negligible, which simplifies Eq. (53) to

$$\left.\begin{aligned}&\mathscr{E}\{\hat{\boldsymbol{q}}\} \cong \boldsymbol{q} + \mathbf{F}_0^{-1}\mathscr{E}\{\mathbf{V}^{\mathrm{T}}\boldsymbol{e}_w\} = \boldsymbol{q} + \mathbf{F}_0^{-1}\boldsymbol{\varepsilon}_v\\&\boldsymbol{\varepsilon}_v^{\mathrm{T}} \cong N\sigma^2[-1\ \ -1\ \ -1\ \ -1\ \ 0\ \ 0\ \ 0\ \ 0]\end{aligned}\right\} \tag{54}$$

Moreover, since $\mathbf{F}_0$ can be shown to contain a factor $\gamma^2(N)\propto N^{\nu}$, $\nu>1$, it follows

$$\lim_{N\to\infty}\mathscr{E}\{\hat{\boldsymbol{q}}\} \cong \boldsymbol{q} \tag{55}$$

The following Theorem achieves one of the targets stated in Section 2.1.

Theorem 3.1　The estimate $\hat{\boldsymbol{q}}_s$ of the differential bias vector $\boldsymbol{q}_s$ is unbiased for a length $N \geqslant N_{\min} \geqslant m$ of the measurement vector in Eq. (28) such that $\boldsymbol{F}_k$ in Eq. (47) is full rank for any $k \geqslant 0$. The estimate $\hat{\boldsymbol{q}}$ is consistent.

Proof　The first part descends from the last two entries of Eq. (54) being zero. The second part descends from Eq. (55).

The second target aims at lowering the bias covariance with respect to static calibration in Section 2.1. Take the expected value of the estimate error in Eq. (50), i.e.

$$\mathcal{E}\{(\hat{\boldsymbol{q}}-\boldsymbol{q})(\hat{\boldsymbol{q}}-\boldsymbol{q})^{\mathrm{T}}\} = \mathcal{E}\{\boldsymbol{F}^{-1}(\boldsymbol{q})\boldsymbol{U}_t^{\mathrm{T}}(\boldsymbol{q})\boldsymbol{e}_w(\boldsymbol{q})\boldsymbol{e}_w^{\mathrm{T}}(\boldsymbol{q})\boldsymbol{U}_t(\boldsymbol{q})\boldsymbol{F}^{-1}(\boldsymbol{q})\} \tag{56}$$

Assuming that $\boldsymbol{U}_t(\boldsymbol{q})$ is noise-free, Eq. (56) reduces to the inverse of the Fisher matrix in Eq. (46), and the estimate $\hat{\boldsymbol{q}}$ can be said to be efficient. In the generic case, a first-order development as in Eq. (51) can be exploited leading to the Theorem.

Theorem 3.2　The estimate $\hat{\boldsymbol{q}}_s$ is asymptotically efficient.

The lack of space forbids further analysis of the Fisher matrix and comparison of dynamic and static covariance. Comparison is left to Monte Carlo runs in Section 4.

4　Simulated results

Two kinds of simulated Monte Carlo experiments have been performed in agreement with Tab. 1. (Ⅰ) Case 1. The same staircase power supply was provided to both sensors, increasing their temperature from about 290 to 350 K. Steps are long enough to reach steady state. Profiles of this kind are suitable to calibrate the sensor response in the whole measurement range[15]. Here a priori sensor calibration error has been assumed to be constant (bias). (Ⅱ) Case 2. Power has been supplied in the form of a square wave so as to improve identifiability in agreement with Theorem 2.1. The simulated parameters and their uncertainty are reported at the end of this section. The temperature profiles of both cases are shown in Fig. 2 and Fig. 3.

Tab. 1　Simulated parameters

No.	Parameter	Symbol	Unit	Value
0	Duration		s	5 250
1	Time unit	T	s	5
2	Sensor noise	σ_y	K	0.005
3	Power noise	σ_u		0.001
4	Power step			0.1
5	Power peak	$P_{k,\max}$	W	10
6	Conductance	g	W/K	1±20%
7	idem	g_a	W/K	0.1±20%
8	Asymmetry	Δg_a	W/K	0.005
9	Capacitance	C_1	J/K	20±20%
10	Capacitance	C_2	J/K	24±20%
11	True bias	Δs	K	−0.1
12	True bias	Δs_a	K	0.1
13	Known biases	$\underline{\Delta s}, \underline{\Delta s_a}$	K	0
14	Monte Carlo trials			500
15	Samples: dynamic case	N		1 050
16	Static case	M		140

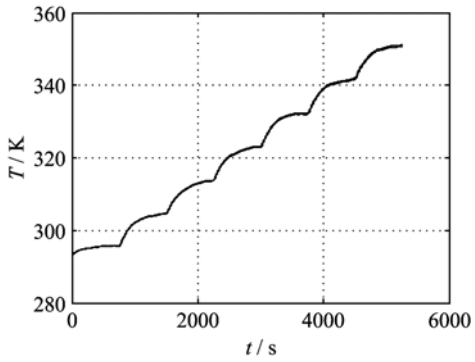

Fig. 2　Temperature profile of the Case 1 (staircase)

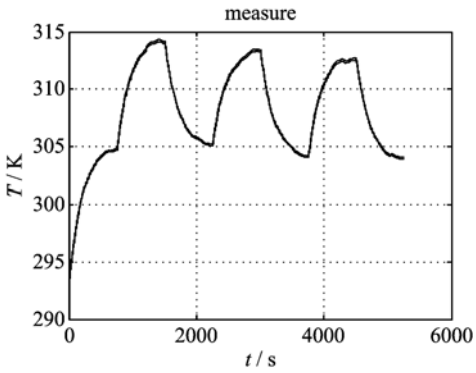

Fig. 3　Temperature profile of the Case 2 (square wave)

The differential measurements y_{21} and y_{a1} entering the calibration matrix $\boldsymbol{W}$ in Eq. (30) are shown in Figs. 4 and 5.

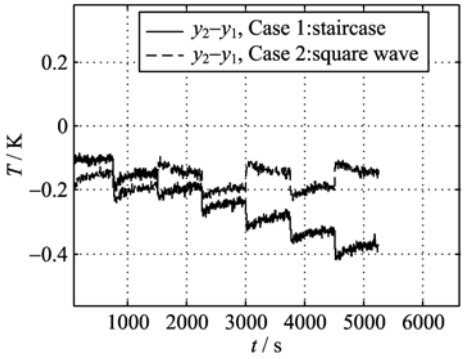

Fig. 4 Differential measurement of the sensors under calibration

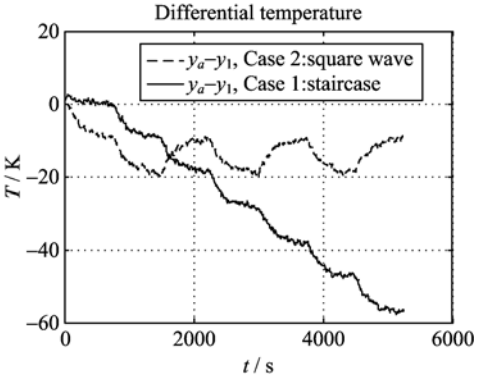

Fig. 5 Differential measurement: Ambient to sensor 1

In both cases, static and dynamic calibration have been performed. Dynamic calibration is performed employing all the set of measurements. In static calibration only a small interval of $N_k=20$ samples at the end of each step $k=0,\cdots,6$ is employed as shown in Fig. 6.

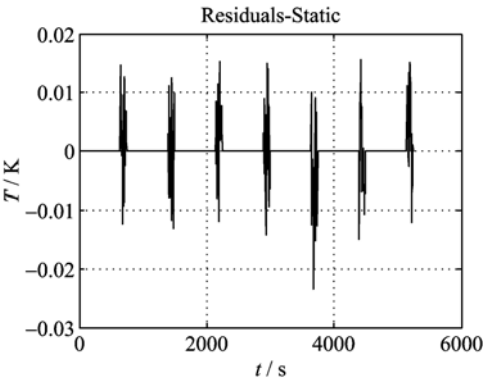

Fig. 6 Static calibration residuals in the Case 1: The measurement intervals are marked by non zero residuals

Mean value and RMS of the a posteriori calibration error

$$\hat{\boldsymbol{e}}_{\Delta s}=\Delta\hat{s}-\Delta s \tag{57}$$

are summarized in Tab. 2.

Tab. 2 Calibration of differential bias

No.	Case	Type	Mean/mK	RMS/mK
0	1, staircase	Static	−2.2	2.3
1		Dynamic	0.3	0.6
2	2, square wave	Static	5.1	4.2
3		Dynamic	<0.1	1.1

The histograms of the estimated bias $\Delta\hat{s}$ in Eq. (14), obtained from Monte Carlo trials, are shown in Figs. 7 and 8. Histograms from static calibration are plotted using a dashed line. Histogram ordinates have been converted to probability density forcing the underlying area to be unitary.

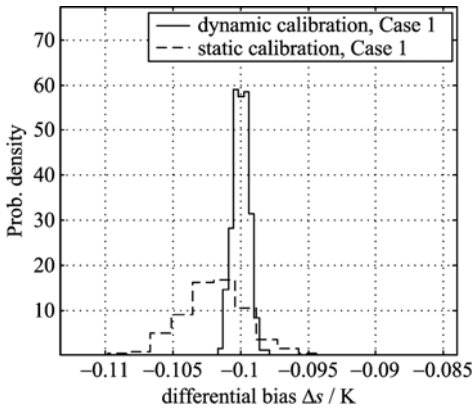

Fig. 7 Case 1: Histograms of static and dynamic calibrations

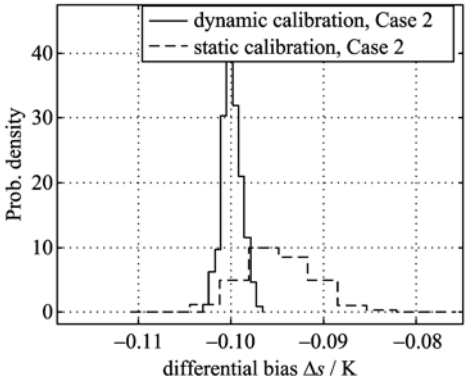

Fig. 8 Case 2: Histograms of static and dynamic calibrations

Results in Tab. 2 and comparison of Figs. 7 and 8 suggest that the staircase profile performs better under both the dynamic and static calibrations. This occurs because of the higher RMS of the input and measured signals as Fig. 2 to Fig. 5 show. As a confirmation of the Case 2 (square wave) and of the dynamic calibration, Fig. 9 shows a significant RMS reduction when the

input step fraction increases from 0.1 to 0.3 (dashed histogram). Improvement cannot be fully applied to static calibration since steady state conditions become remoter, and consequently polarization doubles the value of Tab. 2.

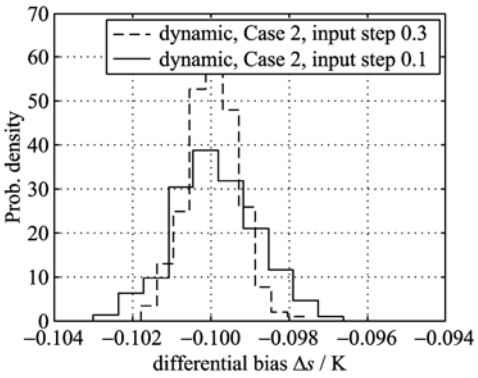

Fig. 9　Calibration improvement in the Case 2 due to input step increase

According to Tab. 2 static calibration suffers from a higher RMS with respect to dynamic calibration, which is due to a much smaller size of measurements (about 15%). Given the same input profile (either Case 1 or 2), RMS could be reduced by enlarging the measurement interval at the end of each input step, but to the detriment of polarization.

Optimization of the time unit T, as mentioned in Section 2, has been made through Monte Carlo trials. Given the measurement length N, the input profile in Tab. 1 and keeping fixed the correction matrix $\boldsymbol{H}$ in Eq. (7), increasing T reduces the calibration RMS to the detriment of polarization. This fact occurs because discrete-time matrices in Eq. (8) tend to become approximate as soon as $|\boldsymbol{H}|$ in Eq. (13) increases. Polarization can be appreciated from residuals as the latter ones are affected by the polarization of the parameter vector $\boldsymbol{q}_p$ in Eq. (10).

A pair of test results reported in Tab. 3 confirms the previous analysis and indicates that $T=2.5$ s is near optimal. The unbiased residuals of both sensors under the smaller time unit in Tab. 3 are shown in Fig. 10. The residual RMS of about 7 mK in Fig. 10 equals $\sqrt{2}\sigma_y$, where σ_y is the sensor noise standard deviation reported in Tab. 1. The sensor noise dominates e in Eq. (17) because of Assumption 2.1. Finally Cramer-Rao bound is compared to Monte Carlo RMS in Tab. 4.

Tab. 3　Time unit optimization (Case 1, dynamic)

No.	Parameter/mK	$T=5$ s		$T=2.5$ s	
		Mean	RMS	Mean	RMS
1	Bias error	−0.2	0.6	0.0	0.7
2	Residuals	−2.9	7.2	−0.5	7.1

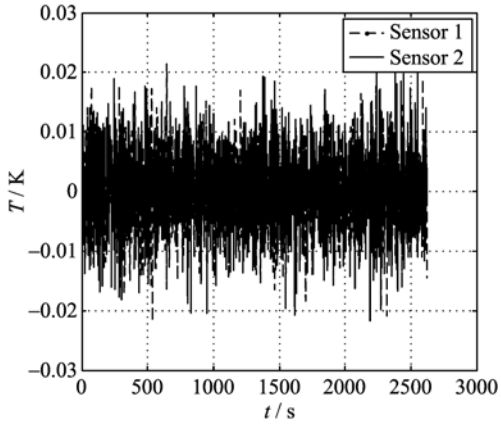

Fig. 10　Sensors 1 and 2 calibration residuals under the smaller time unit T in Tab. 3

Tab. 4　Cramer-Rao bound and Monte Carlo RMS

No.	Case	Type	Cramer-Rao/mK	RMS/mK
0	1, staircase	Static	1.1	2.3
1		Dynamic	0.6	0.6
2	2, square wave	Static	2.4	4.2
3		Dynamic	1.1	1.1

Cramer-Rao bounds in Tab. 4 have been obtained from Eq. (25) (static case) and by inverting the Fisher matrix in Eq. (46) (dynamic case). Both were adjusted through the RMS of the calibration residuals. Static calibration shows a large discrepancy because of the term η in Eq. (22) which expresses deviation from steady state conditions.

5　Conclusion

It has been shown how to use a dynamic model to improve differential calibration of the bias for temperature sensors against a steady-state calibration, given the same suite of measurements. The main concept is that the uncertainty on the

sensor bias estimation is due to asymmetries of the thermal link with the ambient and to a limited amount of data. Polarization can be eliminated and efficiency can be improved by including a thermodynamic model of the equipment, which allows using all the available data, including the equipment transient response and not only steady-state conditions. The dynamic model is constrained in the present formulation to have the same order of the sensor size. Such a model constraint, which simplifies equations, must be tested in the presence of modeling errors, a subject not treated here. A simple model (second order) has been used in the paper, but it can be extended to more complex apparatus. The proposed method has the advantage of being not limited to calibration wells, but it can be applied to any apparatus. To this end no specific condition has been demanded to the temperature of the surrounding chamber (ambient), and the latter can be left unregulated, as in the simulated trials.

Acknowledgments Part of this work has been funded by Regione Piemonte, Italy, under E2 project.

References

[1] Joint Committee for Guides in Metrology. International vocabulary of metrology — Basic and general concepts and associated terms [EB/OL]. http://www.highbeam.com/doc/1G1-190749475.html.

[2] Soderstrom T, Stoica P. System Identification[M]. New York: Prentice-Hall, 1989.

[3] Ljung L. System Identification, Theory for the User [M]. 2ed, New Jersey: Prentice-Hall, Upper Saddle Hall, 1999.

[4] Walter E, Pronzato L. Identification of Parametric Models from Experimental Data [M]. London: Springer, 1997.

[5] Ljung L. Approaches to identification of nonlinear systems[C]// Proceedings of the 29th Chinese Control Conference. Beijing, China, 2010.

[6] Joint Committee for Guides in Metrology, Evaluation of measurement data — Guide to the expression of uncertainty in measurement[EB/OL]. http://www.mendeley.com/research/evaluation-measurement-data-guide-expression-uncertainty-measurement/.

[7] Samaras N S, Simaan M S. Two-point boundary value temperature control of hot strip via water cooling[J]. ISA Transactions, 1997, 36(1): 11-20.

[8] Gumpher J, Bather W A, Wedel D. LPCVD silicon nitride uniformity improvement using adaptive real-time temperature control [J]. IEEE Transactions on Semiconductor Manufacturing, 2003, 16(1): 26-35.

[9] Roig J, Flores D, Cortes I, et al. Non-uniform temperature and heat generation in thin-film SOI LDMOS with uniform drift doping [J]. IEE Proceedings in Circuits, Devices and Systems, 2006, 153(1): 82-87.

[10] Ospina J, Canuto E, Molano-Jimenez A, et al. Multilayer control of an optical reference cavity for space applications[J]. IEEE Transactions on Industrial Electronics, 2010, 57(7): 2 507-2 518.

[11] Ospina J, Canuto E. Uncertainty on differential measurements and its reduction using the calibration by comparison method[J]. Metrologia, 2008, 45(4): 389-394.

[12] Kleiboer L, Havinga P. Calibration of sensors in sensor networks[C]// Proceedings of the 10th IEEE Conference on Emerging Technology and Factory Automation. Catania, Italy, 2005, 2: 519-526.

[13] Han S, Kim J, Sung K. Extended generalized total least squares method for the identification of bilinear systems[J]. IEEE Transactions on Signal Processing, 1996, 44(4): 1 015-1 018.

[14] Juang J N. Continuous-time bilinear system identification[J]. Nonlinear Dynamics, 2005, 39(1): 79-94.

[15] Aggarwal P, Syed Z, El-Sheimy N. Thermal calibration of low cost MEMS sensors for land vehicle navigation system[C]// Proceeding of the 67th IEEE Vehicular Technology Conference. Marina Bay, Singapore: IEEE Press, 2008: 2 859-2 863.

中国科学技术大学学报

Zhongguo Kexue Jishu Daxue Xuebao

（月刊，1965年创刊）

第42卷 第5期（总第229期）2012年5月出版

Journal of University of Science and Technology of China

(Monthly, Started in 1965)

Vol.42 No.5 (Serial No.229) May 2012

主管单位 中国科学院

主办单位 中国科学技术大学

主 编 何多慧

编辑出版者 中国科学技术大学学报编辑部
（安徽省合肥市金寨路96号，230026
电话：0551-3600717,3601961,3607694）

印 刷 者 中国科学技术大学印刷厂

国内发行 安徽省邮局

国外发行 中国国际图书贸易总公司
（北京399信箱，100044）

订 购 处 全国各地邮局

Administered by Chinese Academy of Sciences

Sponsored by University of Science and Technology of China

Editor-in-Chief HE Duohui

Edited and Published by Editorial Department of Journal of University of Science and Technology of China (Hefei 230026,P.R.China)

Tel 86-551-3600717,3601961,3607694; Fax 86-551-3606707

E-mail just@ustc.edu.cn; http://just.ustc.edu.cn

Printed by Printing House of University of Science and Technology of China

Distributed by China International Book Trading Corporation
(P.O. Box 399 Beijing 100044, China)

ISSN 0253-2778

中国标准刊号 ISSN 0253-2778 CN 34-1054/N 邮发代号 国内 26-31 国外 M527 定价 30.00元/期 360.00元/年