



# Observer design for a nonlinear heat equation: Application to semiconductor wafer processing

Alexander Schaum<sup>a</sup>, Stefan Koch<sup>b,\*</sup>, Martin Kleindienst<sup>c</sup>, Markus Reichhartinger<sup>d</sup>, Thomas Meurer<sup>a</sup>, Jaime A. Moreno<sup>e,2</sup>, Martin Horn<sup>b,1</sup>

<sup>a</sup> Chair of Automation and Control, Kiel University, Kiel, Germany

<sup>b</sup> Christian Doppler Laboratory for Model-Based Control of Complex Test Bed Systems, Institute of Automation and Control, Graz University of Technology, Graz, Austria

<sup>c</sup> Lam Research AG, Villach, Austria

<sup>d</sup> Institute of Automation and Control, Graz University of Technology, Graz, Austria

<sup>e</sup> Instituto de Ingeniería, Universidad Nacional Autónoma de México, D.F., México

## ARTICLE INFO

### Article history:

Received 6 April 2022

Received in revised form 5 September 2022

Accepted 6 September 2022

Available online 5 October 2022

### Keywords:

Nonlinear heat equation

Pointwise measurement injection observer

Input-to-state stability

Wafer processing

## ABSTRACT

In this paper, the problem of observer design for a class of 1D nonlinear heat equations with pointwise in-domain temperature measurements is addressed. A pointwise measurement injection observer is designed and the robust convergence of its estimation error in presence of bounded distributed perturbations is established by verifying input-to-state stability. The obtained convergence conditions express the underlying interplay between heat conduction and radiation and include specific dependencies on the sensor locations which are the main degrees of freedom in the design approach. The theoretical results are experimentally validated on a semiconductor wafer processing unit.

© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The increasing demand for semiconductor components continues worldwide, as more and more computing power and data storage space are required in all areas of our life. The fast technological progress and the imposed requirements on device performance create new challenges for the semiconductor industry. The production of modern microchips involves hundreds of process steps such as photolithography, ion implantation, and etching. These steps are repeated and carried out cyclically. For the production of complex integrated circuits, single-wafer processing is preferred over batch processes, as these processes enable much finer features.

Heat and mass transfer mechanisms frequently occur in a number of silicon wafer production steps and, in particular, the

temperature of the process fluids or materials involved often has a significant influence on the quality of the final product. In single wafer spin clean or wet chemical etching it is often necessary to heat up the wafer to a predefined temperature, e.g., to remove condensation from the wafer surface prior processing or to chemically treat the surface of silicon wafers with highly reactive gases. The latter belong to the so-called rapid thermal processing (RTP), which typically requires high temperatures, often in the range of 300–400 °C or even higher [1,2]. The former process, i.e. the removal of condensation, takes place at lower temperatures, usually around 150 °C, and is known as baking. The heating is achieved by halogen lamps or by a large number high power LEDs placed beneath the wafer [3,4]. To avoid thermal stress in the wafer during heat up, large temperature gradients in the wafer need to be avoided. This issue requires precise temperature control. However, the establishment of a feedback temperature controller is challenging as in most applications the contactless in-situ measurement of the entire surface temperature is not available. Thermal imaging cameras, for example, fail to precisely measure the temperature of low-doped wafers, due to their low emissivity. For most RTP or low temperature thermal processes, the temperature cannot be measured at all or only pointwise on the wafer surface.

However, the design and implementation of a feedback controller often requires full state information, i.e. the availability of the entire radial wafer surface temperature or the temperature

\* Corresponding author.

E-mail addresses: [alsc@tf.uni-kiel.de](mailto:alsc@tf.uni-kiel.de) (A. Schaum), [stefan.koch@tugraz.at](mailto:stefan.koch@tugraz.at) (S. Koch), [martin.kleindienst@lamresearch.com](mailto:martin.kleindienst@lamresearch.com) (M. Kleindienst), [markus.reichhartinger@tugraz.at](mailto:markus.reichhartinger@tugraz.at) (M. Reichhartinger), [tm@tf.uni-kiel.de](mailto:tm@tf.uni-kiel.de) (T. Meurer), [JMorenoP@ii.unam.mx](mailto:JMorenoP@ii.unam.mx) (J.A. Moreno), [martin.horn@tugraz.at](mailto:martin.horn@tugraz.at) (M. Horn).

<sup>1</sup> The financial support by the Austrian Federal Ministry for Digital and Economic Affairs, the National Foundation for Research, Technology and Development and the Christian Doppler Research Association is gratefully acknowledged.

<sup>2</sup> Financial support by DGAPA-UNAM, PAPIIT IN 102121.

measured at several points along the radius of the wafer, see, e.g., [5,6]. Thus, the realization of such controllers in a production tool often requires the implementation of a state observer, i.e., an estimator that provides the spatial and temporal evolution of the wafer surface temperature from available measurements.

The temperature of the wafer surface depends on time and space and thus is a distributed parameter system (DPS). The dynamic behavior of such a system are governed by partial differential equations (PDEs). When following a model based design paradigm, an observer for a DPS can, in principle, either be designed based on an approximation via a lumped parameter system or by directly using the PDE. The first approach, which also is referred to as early lumping, is generally accompanied by a loss of relevant information about the system dynamics. This issue is delicate in that stability results obtained for finite-dimensional approximations do not necessarily hold for the PDE model. Furthermore, the finite-dimensional approximations typically are of high system order. The order of the model determines in many design approaches also the order of the observer. For finite-dimensional approximations of PDEs, this can lead to high-order algorithms demanding large computational effort.

In recent years, late lumping has been extensively developed as an intriguing alternative to early lumping design as the above mentioned drawbacks do not exist. The approximation, which is finally required for the implementation in a real-time setup is carried out at the stage of implementation. As long as the approximation yields stable and robustness preserving results, robust stability obtained for the original PDE model also hold true for the lumped model. The observer and controller design can thus be carried out exploiting the particular PDE structure, independent of the final approximation technique. For linear DPSs the associated theory for 1D spatial domains has been well-developed, e.g., in [7–9]. Particular approaches to be mentioned here include modal (or spectral) decompositions [10,11], backstepping [9,12–14] and high-gain observers [15]. The backstepping approach has also been extended to several spatial dimensions [16,17]. For semilinear and nonlinear PDE models literature is more sparse. In [16,18] an extended Luenberger observer design has been proposed. Variable structure estimation schemes have been developed, e.g., in [19]. Observers based on nonlinear evolution equations and absolute stability have been investigated in [20, 21]. For transport–reaction systems with unknown reaction rate asymptotic observers have been addressed in [22]. High-gain observers have been used in [23] and matrix inequality-based designs have been studied in [11,24,25]. The backstepping approach has been further extended for some classes of semilinear and quasilinear systems [26–30]. Dissipativity-based observer design approaches have been discussed in [31,32]. Eventhough these results show the great potential of applying late-lumping design approaches, they all involve a substantial amount of preliminary analysis and design steps requiring a deep knowledge of PDE theory. A rather simple design approach, requiring only essential knowledge about PDEs is the pointwise measurement injection observer design that has been proposed in [33] for 1D semilinear heat equations and has been extended in [34,35] to classes of semilinear parabolic systems, and in [36] to a class of 1D parabolic transport–reaction systems with unknown inputs. The design resembles a reduced-order observation scheme from finite-dimensional systems [37,38] in which the measurement information is imposed in form of an algebraic constraint. This approach has already been used for nonlinear heat equations in [39] considering a single in-domain measurement and a perfect, unperturbed model. In [39] the observer design has been carried out after applying a Kirchhoff state transformation and the temperature estimate is obtained after retransformation.

In this paper, an observer for systems governed by a nonlinear perturbed 1D heat equation in cylindrical coordinates with in-domain measurements is designed. The design of the observer is required for the estimation of the temperature of silicon wafers in semiconductor production, its application is obviously not limited to this specific example. The proposed observer is an extension of the pointwise measurement injection observer [33,34,36,39] and takes into account different nonlinearities. Imposing practically reasonable assumptions on the system dynamics, the observer estimation error converges to zero exponentially which is formally proven by Lyapunov techniques. The observer is robust with respect to pointwise disturbances acting directly at the sensor location. In the presence of non-vanishing bounded distributed disturbances, the estimation error dynamics are shown to be input-to-state stable. Finally, the proposed observer scheme is validated experimentally on a semiconductor processing tool. To that end a mathematical model of the process is derived. It is based on the model proposed in [4,40] where focus is placed on modeling the input shape functions relating the electrical power supplied to the heating device, i.e., the actuator with the heat flux density introduced to the wafer. The tool used for validation is equipped with a thermographic camera capable of measuring the entire wafer surface temperature of wafers with a high dopant level. Having a distributed measurement of the radial temperature enables computing the spatio temporal evolution of the estimation error. Furthermore, in principle, any number of pointwise sensors can be emulated with this setup. The experimental results confirm the theoretical findings, show that the assumptions placed in the stability analysis are reasonable, and confirm that a good trade-off between convergence speed and implementation effort is achieved.

In contrast to [39] in the present paper (i) multiple in-domain temperature measurements are considered, (ii) an unperfect model with distributed perturbation is considered and robust convergence is established in the sense of input-to-state stability, (iii) the design is carried out in the original coordinates and the Kirchhoff transformation is employed only for the convergence assessment and (iv) the theoretical results are experimentally validated on a semiconductor wafer processing unit. Thus the results of the present paper generalize, extend and experimentally validate the preliminary ones in [39].

The paper is structured as follows. In Section 2 the problem statement is presented. In Section 3 the pointwise measurement injection observer is designed. In Section 4 sufficient conditions for the observer estimation error convergence are established. In Section 5, the observer is used to estimate the temperature of a silicon wafer, for which a mathematical model of the process is first derived. Experimental results are presented at the end of this section. Section 6 concludes the paper.

## Notation

The absolute value of a scalar  $c$  is denoted by  $|c|$ . For functions  $u, v \in L^2(a, b)$  a weighted inner product is denoted by  $\langle u, v \rangle_z = \int_a^b zu(z)v(z)dz$  and the induced norm by  $\|v\|_z = \sqrt{\langle v, v \rangle_z}$ . The Sobolev space  $H^2(a, b)$  is accordingly defined as  $H^2(a, b) = \{v \in L^2(a, b) \mid v^{(k)} \in L^2(a, b) \text{ for } k = 0, 1, 2\}$  where  $v^{(k)} = \frac{\partial^k}{\partial z^k} v(z)$ . Matrices and vectors are denoted by boldface letters.

## 2. Nonlinear heat equation and problem formulation

Consider the following nonlinear heat equation (see, e.g., [41])

$$\rho c_p(T) \partial_t T = \frac{1}{r} \partial_r [rk(T) \partial_r T] + \gamma(T) + \sum_{v=1}^q b_v u_v + w \quad (1a)$$

for  $t > 0$ ,  $r \in (0, R)$ , temperature  $T(r, t)$ , with distributed perturbation  $w$ , boundary conditions

$$\partial_r T(0, t) = \partial_r T(R, t) = 0 \quad (1b)$$

for  $t > 0$  and initial temperature profile

$$T(r, 0) = T_0(r), \quad r \in [0, R]. \quad (1c)$$

Furthermore  $m \in \mathbb{N}$  pointwise temperature measurements

$$y_i(t) = T(\varrho_i, t), \quad i = 1, \dots, m, \quad t \geq 0 \quad (1d)$$

with the measurement locations  $\varrho_i \in (0, R)$  are considered. In the above equations the parameter  $\rho$  is the density,  $c_p \in C^1([T_{\min}, \infty))$  is the heat capacity,  $k \in C^1([T_{\min}, \infty))$  represents the heat conductivity. The function  $\gamma$  encompasses (nonlinear) source terms due to convective heat transfer and thermal radiation and it is assumed that  $\gamma \in C^1([T_{\min}, \infty))$  with bounded derivative  $\gamma'$ , so that  $\gamma$  is Lipschitz continuous. The spatially distributed disturbance  $w(\cdot, t) \in L^2(0, R)$  is bounded so that

$$\sup_{t \geq 0} \max_{r \in (0, R)} |w(r, t)| \leq w^+ \quad (1e)$$

holds true. The characteristic shape functions of the  $q \in \mathbb{N}$  actuators with input signal  $u_\nu$ , are denoted by  $b_\nu \in L^2([0, R])$ ,  $\nu = 1, \dots, q$ . A specific application example will be discussed in Section 5.

In the following the existence of a unique solution  $T : [0, \infty) \times [0, R] \rightarrow \mathbb{R}$  for which (1) holds true for all initial profiles  $T_0 \in L^2(0, R)$  is assumed.

The goal of the subsequent observer design is to provide an estimate  $\hat{T}(\cdot, t)$  of the temperature profile  $T(\cdot, t)$  at time  $t \geq 0$  by combining the pointwise measurements  $y_i(t) = T(\varrho_i, t)$ ,  $i = 1, \dots, m$  and the model (1), so that the associated observation error

$$\tilde{T}(\cdot, t) = \hat{T}(\cdot, t) - T(\cdot, t) \quad (2)$$

is input-to-state stable (ISS) [42–44], i.e., there exist comparison functions  $\alpha \in \mathcal{KL}$  and  $\beta \in \mathcal{K}_\infty$  so that

$$\|\tilde{T}(\cdot, t)\|_r \leq \alpha(\|\tilde{T}_0\|_r, t) + \beta(w^+). \quad (3)$$

### 3. Pointwise measurement injection observer

In this section, the pointwise measurement injection observer design proposed in [33,34,36,39] is followed to design an exponentially convergent observer for the system (1).

The observer is set up as a copy of the plant (1) driven by the measurement signals at the measurement points, i.e.,

$$\rho c_p(\hat{T}) \partial_t \hat{T} = \frac{1}{r} \partial_r \left[ rk(\hat{T}) \partial_r \hat{T} \right] + \gamma(\hat{T}) + \sum_{\nu=1}^q b_\nu u_\nu \quad (4a)$$

for  $t > 0$ ,  $r \in (0, R) \setminus \{\varrho_1, \dots, \varrho_m\}$ ,

$$\partial_r \hat{T}(0, t) = \partial_r \hat{T}(R, t) = 0 \quad (4b)$$

for  $t > 0$  and

$$\hat{T}(r, 0) = \hat{T}_0(r), \quad r \in [0, R] \quad (4c)$$

driven by the measurements at  $r = \varrho_i$  with

$$\hat{T}(\varrho_i, t) = y_i(t), \quad t \geq 0, \quad i = 1, \dots, m. \quad (4d)$$

It can be clearly seen that the only design degrees of freedom in this observer structure are the measurement locations entering in the algebraic conditions (4d).

With the observation error  $\tilde{T} = \hat{T} - T$  defined in (2), the direct pointwise injection of the measurement signal at the  $m$  measurement points (4d) implies that

$$\tilde{T}(\varrho_i, t) = 0, \quad t \geq 0, \quad i = 1, \dots, m. \quad (5)$$

Denoting  $\varrho_0 = 0$ ,  $\varrho_{m+1} = R$  the constraints (5) actually imply that the observation error can be decomposed over the  $m+1$  intervals

$$J_1 = [0, \varrho_1], \quad J_j = (\varrho_{j-1}, \varrho_j], \quad j = 2, \dots, m+1, \quad (6a)$$

according to

$$\tilde{T}(r, t) = \begin{cases} \tilde{T}_j(r, t), & r \in J_j, \quad j \in \{1, \dots, m+1\} \\ 0, & r = \varrho_i, \quad i \in \{1, \dots, m\}. \end{cases} \quad (6b)$$

**Remark 1.** Note that for simplicity no measurements at the boundary are considered. As will become clear in the sequel the results can be easily adapted to include boundary measurements.

In each interval  $J_j$  the error dynamics is governed by a PDE

$$\begin{aligned} \partial_t \tilde{T}_j = & \frac{1}{\rho c_p(T + \tilde{T}_j)} \left( \frac{1}{r} \partial_r \left( rk(T + \tilde{T}_j) \partial_r (T + \tilde{T}_j) \right) + \gamma(T + \tilde{T}_j) \right. \\ & \left. + \sum_{\nu=1}^q b_\nu u_\nu \right) \\ & - \frac{1}{\rho c_p(T)} \left( \frac{1}{r} \partial_r \left( rk(T) \partial_r (T) \right) + \gamma(T) + \sum_{\nu=1}^q b_\nu u_\nu + w \right) \end{aligned} \quad (7a)$$

for  $r \in J_j$ ,  $t > 0$ ,  $j = 1, \dots, m+1$  with boundary conditions

$$\partial_r \tilde{T}_1(0, t) = 0, \quad \tilde{T}_1(\varrho_1, t) = 0 \quad (7b)$$

$$\tilde{T}_j(\varrho_i, t) = 0, \quad j = 2, \dots, m, \quad i = j-1, j \quad (7c)$$

$$\tilde{T}_{m+1}(\varrho_m, t) = 0, \quad \partial_r \tilde{T}_{m+1}(R, t) = 0. \quad (7d)$$

Note that, due to the nonlinearity, the error dynamics explicitly depends on the (unknown) solution  $T : [0, R] \times [0, \infty) \rightarrow [T_{\min}, \infty)$ .

In the subsequent section it is shown that by adequately choosing the measurement locations  $\varrho_i$ ,  $i = 1, \dots, m$  the observation error (2) robustly converges in the norm  $\|\cdot\|_r$ .

**Remark 2.** In the case that a boundary measurement is considered at  $r = 0$ , in (7b) the condition changes to  $\tilde{T}_1(0, t) = 0$ . For a boundary measurement at  $r = R$  correspondingly the condition (7d) changes to  $\tilde{T}_m(R, t) = 0$ .

### 4. Convergence assessment

In this section, the convergence of the observation errors of the proposed pointwise measurement injection observer (4) is established in dependency of the measurement locations using Lyapunov's direct method. For the convergence analysis the following assumption is made.

**Assumption 1.** The thermal diffusivity  $\alpha = \frac{k(T)}{\rho c_p(T)}$  is constant for  $T \in [T_{\min}, \infty)$ . This assumption goes along with the standard assumptions in the literature on nonlinear heat conduction processes and is fulfilled with sufficient accuracy in many application scenarios [39,41,45,46].

**Remark 3.** Note that for the implementation of the observer the thermal diffusivity  $\alpha$  does not have to be constant and can depend on the temperature. The same holds for the density and the specific heat capacity.

**Theorem 1 (Proof in Appendix).** Consider the nonlinear heat Eq. (1) with a bounded disturbance  $\sup_{t \geq 0} \max_{z \in (0, R)} |w(r, t)| \leq w^+$ ,  $m$  measurements at  $\varrho_i \in (0, R)$ ,  $i = 1, \dots, m$  and the pointwise measurement injection observer (4). Let  $\alpha > 0$  be the thermal

diffusivity,  $L_\gamma$  denote the Lipschitz constant of  $\gamma$  in (1) and let the heat conductivity  $k$  be positive and bounded. In consequence the function

$$\psi(\tilde{T}; T) = \int_T^{T+\tilde{T}} k(\tau) d\tau, \quad (8)$$

is invertible and its inverse  $\psi^{-1}$  is Lipschitz continuous with Lipschitz constant  $L_{\psi^{-1}}$ . Consider further the operators  $\mathcal{A}_j$  defined by

$$\mathcal{A}_j \epsilon_j = \frac{\alpha}{r} \partial_r (r \partial_r \epsilon_j), \quad j = 1, \dots, m+1 \quad (9)$$

$$\mathcal{D}(\mathcal{A}_j) = \begin{cases} \{v \in H^2([0, \varrho_1]) \mid \partial_r v(0) = 0, \quad j = 1 \\ v(\varrho_1) = 0\}, \\ \{v \in H^2([\varrho_j, \varrho_{j+1}]) \mid v(\varrho_j) = 0 \quad j = 2, \dots, m \\ = v(\varrho_{j+1})\}, \\ \{v \in H^2([\varrho_m, R]) \mid v(\varrho_m) = 0, \quad j = m+1 \\ \partial_r v(R) = 0\}, \end{cases}$$

with eigenvalues  $\lambda_n(\mathcal{A}_j)$ ,  $n \in \mathbb{N}$ . Then the associated observation error  $\tilde{T}$  defined in (2) is ISS in the sense of (3) if the measurement locations are chosen such that

$$\kappa > \bar{L} + 1/2 \quad (10)$$

with  $\bar{L} = \alpha L_\gamma L_{\psi^{-1}}$  and  $\kappa$  given by

$$-\kappa = \max_{j=1, \dots, m+1} \sup_{n \in \mathbb{N}} \lambda_n(\mathcal{A}_j). \quad (11)$$

**Remark 4.** As noted above the main degree of freedom in the design is the choice of the number and location of the sensors. This choice influences the value of  $\kappa$  in the preceding considerations. If for a given number of sensors it is not possible to achieve the inequality (10) then one can consider additional sensors, which directly yields a larger value of  $\kappa$ .

**Remark 5.** If measurements on the boundary are considered the main difference in the preceding considerations consists in the different domains of the operators  $\mathcal{A}_j$ ,  $j = 1, m$  for which the modified boundary conditions (Dirichlet instead of Neumann) have to be considered in the definition of their respective domains  $\mathcal{D}(\mathcal{A}_j)$ . Note that this will lead to different dominant eigenvalues  $\lambda_n(\mathcal{A}_j)$ ,  $j = 1, m$  and thus will influence the convergence condition. In general, in comparison between Neumann and Dirichlet boundary conditions the convergence condition improves for the latter ones.

**Remark 6.** The consideration of Assumption 1 could be leveraged by setting  $\alpha(T) = \bar{\alpha} + \tilde{\alpha}(T)$  with a constant (e.g., mean) value  $\bar{\alpha}$  and the term  $\tilde{\alpha}$  accounting for deviations from  $\bar{\alpha}$ . Setting up the observer with  $\bar{\alpha}$  would lead to an additional error term proportional to  $\tilde{\alpha}(T)$ , which could be included into the nonlinear and perturbation part in the above analysis. From a practical point of view this is anyway not necessary, as commented in Assumption 1.

## 5. Application to silicon wafer processing

In this section, the proposed observer scheme is applied for the purpose of wafer temperature estimation. After a brief introduction to the process under consideration, a mathematical model, in accordance with the generic model presented in Section 2, is derived. An experimental evaluation of the proposed observer rounds up this section.

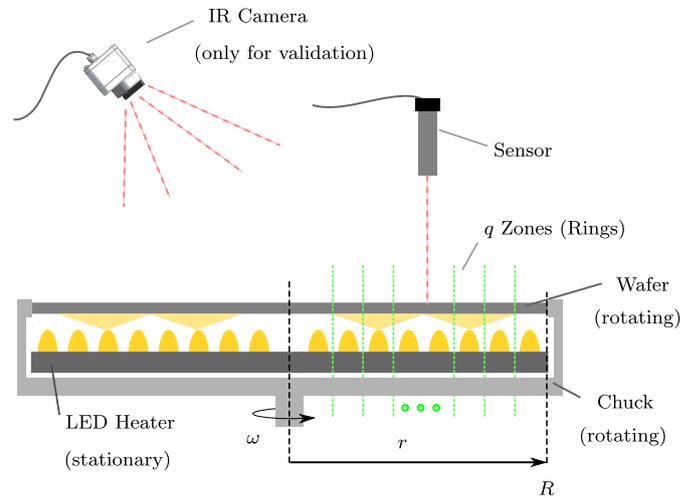


Fig. 1. Schematic of the considered silicon wafer production process.

### 5.1. Process description

A schematic drawing of the considered setup is shown in Fig. 1. The silicon wafer is held by a so-called wafer-chuck which rotates with a process specific speed. While the chuck, and thus the wafer rotates, the bottom side of the wafer is heated contactlessly via a stationary heating plate. In the system at hand the heating plate consists of more than 1000 high power LEDs which are grouped into  $q = 20$  concentrically arranged rings equipped with  $p_\mu$ ,  $\mu = 1, \dots, q$ , LEDs. All  $p_\mu$  LEDs within one ring are synchronized, i.e., they are actuated simultaneously and with the same electrical power, whereas each ring can be actuated individually. This design enables targeted heating with respect to the radial coordinate. Furthermore, the temperature is measured contact-free at selected points on the wafer's surface.

For validation of this observer a thermographic camera is installed. The camera is capable of measuring the entire surface temperature of wafers with a high dopant level. To design the observer a model describing the dynamics of the outlined process is derived in a first step.

### 5.2. Mathematical model

The temperature of the wafer, denoted by  $T(\mathbf{x}, t)$  for  $\mathbf{x} = [x \ y]^T \in \bar{\Omega} = \{\mathbf{x} \in \mathbb{R}^2 \mid \|\mathbf{x}\| \leq R\}$ ,  $t > 0$ , and  $T : \bar{\Omega} \times [0, \infty) \rightarrow [0, \infty)$  satisfies (see, e.g., [41,46])

$$\rho c_p(T) \partial_t T = \text{div}(k(T) \text{grad } T) + \Gamma(T) + \dot{W} \quad (12)$$

for  $\mathbf{x} \in \Omega = \{\mathbf{x} \in \mathbb{R}^2 \mid \|\mathbf{x}\| < R\}$ ,  $t > 0$

$$\partial_{\mathbf{x}} T \cdot \mathbf{n} = 0 \quad (13)$$

for  $\mathbf{x} \in \delta\bar{\Omega}$ ,  $t > 0$  with  $\mathbf{n}$  being the outward normal unit vector along the boundary  $\delta\bar{\Omega} = \bar{\Omega} \setminus \Omega$ , and

$$T(\mathbf{x}, 0) = T_0(\mathbf{x}) \quad (14)$$

for  $\mathbf{x} \in \bar{\Omega}$ . The function  $\Gamma \in C^1([0, \infty))$  encompasses convective heat transfer and thermal radiation and  $\dot{W} \in L^2([0, R])$  denotes the heat source, i.e., volumetric heat introduced to the system by the LED heater.

To model the heat source, consider the heat flux density  $\dot{q}_{v,\mu}^0$  introduced by the  $\mu$ th LED of the  $v$ th ring. It is described mathematically by the two-dimensional Gaussian distribution

$$\dot{q}_{v,\mu}^0(\mathbf{x}) = \frac{\eta P_{\text{el}} \mu_{v,\mu}^0}{2\pi \sqrt{\det(\delta)}} e^{-\frac{1}{2}(\mathbf{x} - \bar{\mathbf{x}}_{v,\mu})^T \delta^{-1} (\mathbf{x} - \bar{\mathbf{x}}_{v,\mu})} \quad (15)$$

where  $P_{el}$  is the electrical power supplied to the LED,  $u_{v,\mu}^0 \in [0, 1]$  is the actuating signal,  $\eta$  is an efficiency factor, and  $\bar{\mathbf{x}}_{v,\mu} = [\bar{x}_{v,\mu} \bar{y}_{v,\mu}]^T$  with  $\bar{x}_{v,\mu}$  and  $\bar{y}_{v,\mu}$  denoting the position of the LED. The position of the LED is expressed in polar coordinates

$$\bar{x}_{v,\mu} = \bar{r}_v \cos(\bar{\phi}_\mu), \quad (16a)$$

$$\bar{y}_{v,\mu} = \bar{r}_v \sin(\bar{\phi}_\mu), \quad v = 1 \dots q, \quad \mu = 1 \dots p_v \quad (16b)$$

where  $\bar{r}_v$  and  $\bar{\phi}_\mu$  are distance and angle w.r.t. the center of the LED, respectively. The matrix  $\delta$  describes the radiation pattern of the LED. Assuming symmetric radiation, i.e., matrix  $\delta$  satisfies  $\delta = \delta^T \mathbf{I}$ ,  $\delta \in \mathbb{R} \setminus \{0\}$  where  $\mathbf{I} \in \mathbb{R}^2$  is the identity matrix, provides for the simplification

$$\dot{q}_{v,\mu}^0(\mathbf{x}) = \frac{\eta P_{el} u_{v,\mu}^0}{2\pi \delta^2} e^{-\frac{1}{2\delta^2} [(x-\bar{x}_{v,\mu})^2 + (y-\bar{y}_{v,\mu})^2]}. \quad (17)$$

Furthermore, assuming angularly homogeneous initial temperature profiles and taking into account that all actuators are synchronized along circles, i.e.,  $u_{v,\mu}^0 = u_v$  angularly homogeneous profiles are obtained, i.e.,

$$T(\mathbf{x}_i, t) = T(\mathbf{x}_j, t) \quad \text{for all } \mathbf{x}_i, \mathbf{x}_j : \|\mathbf{x}_i\| = \|\mathbf{x}_j\|$$

and  $t \geq 0$ . The overall heat flux density introduced by one entire LED ring therefore is written as

$$\dot{q}_v(\mathbf{x}) = \sum_{\mu=1}^{p_v} \dot{q}_{v,\mu}^0(\mathbf{x}) \quad (18)$$

which, using (17), yields

$$\dot{q}_v(\mathbf{x}) = \frac{\eta P_{el} u_v}{2\pi \delta^2} \sum_{\mu=1}^{p_v} e^{-\frac{1}{2\delta^2} [(x-\bar{x}_{v,\mu})^2 + (y-\bar{y}_{v,\mu})^2]}. \quad (19)$$

Fig. 2 shows the simulated heat flux density when one single LED is actuated,<sup>3</sup> whereas Fig. 3 shows the heat flux density when all LEDs in one ring are actuated. In cylindrical coordinates and considering (16a), (16b) gives

$$\begin{aligned} \dot{q}_v(\mathbf{x}(r, \phi)) &= \frac{\eta P_{el} u_v}{2\pi \delta^2} \sum_{\mu=1}^{p_v} e^{-\frac{1}{2\delta^2} [r^2 + \bar{x}_{v,\mu}^2 + \bar{y}_{v,\mu}^2 - 2r(\bar{x}_{v,\mu} \cos(\phi) + \bar{y}_{v,\mu} \sin(\phi))]} \\ &= \frac{\eta P_{el} u_v}{2\pi \delta^2} \sum_{\mu=1}^{p_v} e^{-\frac{1}{2\delta^2} [r^2 + \bar{r}_v^2 - 2r\bar{r}_v \cos(\phi - \bar{\phi}_{v,\mu})]}. \end{aligned} \quad (20)$$

As the LEDs are packed quite dense within one ring and, assuming a sufficiently large value of  $\delta$ , the variation of the heat flux density along the angular coordinate is small, see Fig. 3. This justifies to approximate the heat flux density (20) by its average (along the angular coordinate) which yields

$$\begin{aligned} \dot{q}_v(r) &= \frac{1}{2\pi r} \int_0^{2\pi} \dot{q}_v(\mathbf{x}(r, \phi)) r d\phi \\ &= \frac{1}{2\pi r} \frac{\eta P_{el} u_v}{2\pi \delta^2} \int_0^{2\pi} \sum_{\mu=1}^{p_v} e^{-\frac{1}{2\delta^2} [r^2 + \bar{r}_v^2 - 2r\bar{r}_v \cos(\phi - \bar{\phi}_{v,\mu})]} r d\phi \\ &= \frac{\eta P_{el} u_v}{(2\pi \delta)^2} \sum_{\mu=1}^{p_v} \int_0^{2\pi} e^{-\frac{1}{2\delta^2} [r^2 + \bar{r}_v^2 - 2r\bar{r}_v \cos(\phi - \bar{\phi}_{v,\mu})]} d\phi \\ &= \frac{\eta P_{el} u_v}{(2\pi \delta)^2} \sum_{\mu=1}^{p_v} e^{-\frac{1}{2\delta^2} (r^2 + \bar{r}_v^2)} \int_0^{2\pi} e^{\frac{r\bar{r}_v}{\delta^2} \cos(\phi - \bar{\phi}_{v,\mu})} d\phi \\ &= \frac{\eta P_{el} u_v}{2(\pi \delta)^2} \sum_{\mu=1}^{p_v} e^{-\frac{1}{2\delta^2} (r^2 + \bar{r}_v^2)} \int_0^\pi e^{\frac{r\bar{r}_v}{\delta^2} \cos(\phi)} d\phi \end{aligned} \quad (21)$$

<sup>3</sup> Note that this is for illustration purposes only as the actuation of one single LED is not possible with the actual actuator.

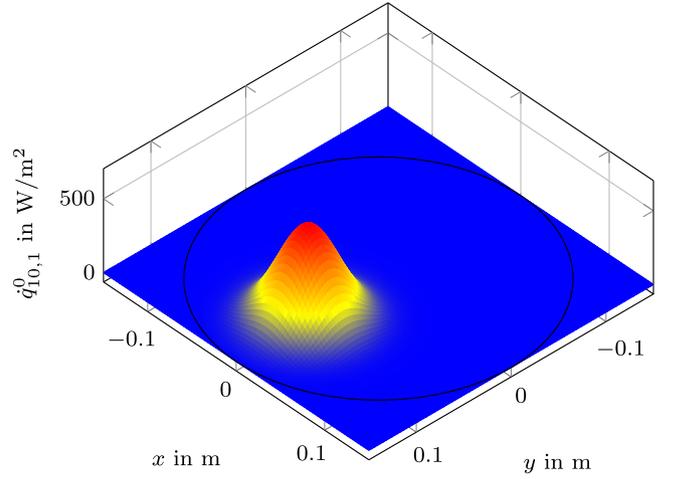


Fig. 2. One single LED actuated with  $u_{10,1}^0 = 1$ .

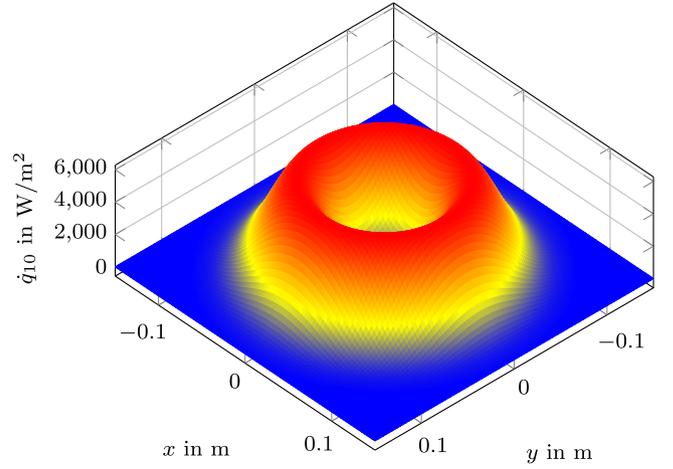


Fig. 3. Entire LED ring actuated with  $u_{10} = 1$ .

where the last line is due to periodicity of the integrand. Note that the integral in the above function, scaled by  $1/\pi$ , is an integral representation of the modified Bessel function of the first kind (in the following denoted by  $I_0$ ), i.e.,

$$\int_0^\pi e^{\frac{r\bar{r}_v}{\delta^2} \cos(\phi)} d\phi = \pi I_0 \left( \frac{r\bar{r}_v}{\delta^2} \right), \quad (22)$$

see, [47]. Substituting (22) into (21) yields

$$\begin{aligned} \dot{q}_v(r) &= \frac{\eta P_{el} u_v}{2\pi \delta^2} \sum_{\mu=1}^{p_v} I_0 \left( \frac{r\bar{r}_v}{\delta^2} \right) e^{-\frac{1}{2\delta^2} (r^2 + \bar{r}_v^2)} \\ &= \frac{\eta P_{el} u_v}{2\pi \delta^2} p_v I_0 \left( \frac{r\bar{r}_v}{\delta^2} \right) e^{-\frac{1}{2\delta^2} (r^2 + \bar{r}_v^2)}. \end{aligned} \quad (23)$$

Introducing the scaled Bessel function

$$I_0^*(\xi) = I_0(\xi) e^{-|\operatorname{Re}\{\xi\}|} \quad (24)$$

leads to

$$\dot{q}_v(r) = \frac{\eta P_{el} u_v}{2\pi \delta^2} p_v I_0^* \left( \frac{\bar{r}_v r}{\delta^2} \right) e^{-\frac{(r-\bar{r}_v)^2}{2\delta^2}}. \quad (25)$$

Note that the averaged heat flux density only depends on the radial coordinate. The averaged heat flux densities given in (25) are plotted in Fig. 5 for the parameters provided in Table 1. The first ring is located at  $\bar{r}_1 = 16$  mm and the following rings are

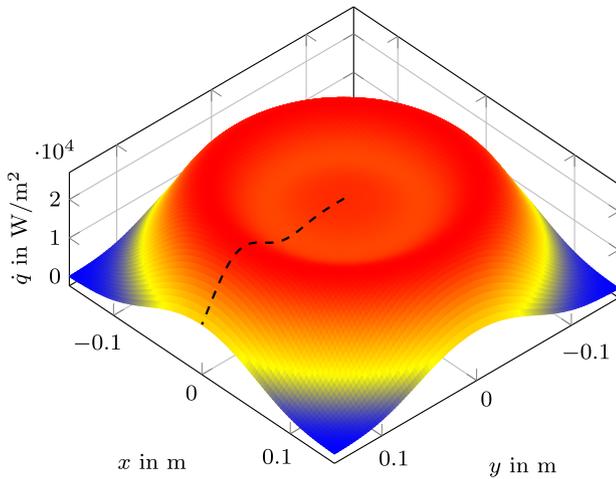


Fig. 4. Randomly actuated LED heater.

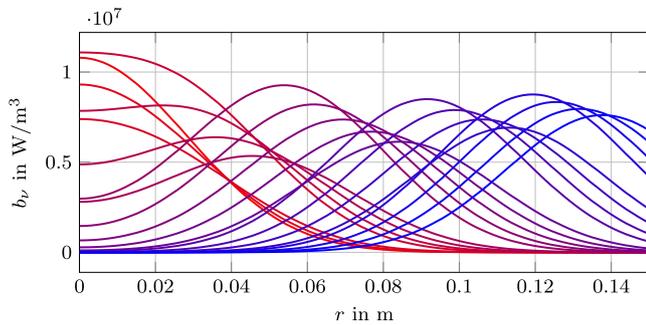


Fig. 5. Input shape functions.

distributed (almost) equidistantly over the radius of the wafer. The number of LEDs per ring is the same for certain rings, i.e.  $p_v = m_1$  for  $v \in \{1, 2, 3\}$ ,  $p_v = m_2$  for  $v \in \{4, 5, 6, 7\}$ ,  $p_v = m_3$  for  $v \in \{8, \dots, 12\}$ ,  $p_v = m_4$  for  $v \in \{13, \dots, 16\}$  and  $p_v = m_1$  for  $v \in \{17, \dots, 20\}$ . Due to confidentiality reasons, specific parameters of the LED heating device, such as  $\bar{r}_v$ ,  $p_v$ ,  $\eta$  and  $P_{el}$  cannot be given here. The overall heat flux density is depicted in Fig. 4 for a random actuation  $u_v \in [0, 1]$ . The computed averaged heat flux density, i.e.,  $\sum_{v=1}^q \dot{q}_v(r)$  with  $\dot{q}_v(r)$  given in (25) is plotted in the same figure (black dashed line) and shows good agreement.

In view of these considerations, assuming an initial temperature profile  $T_0$  which is homogeneous in the angular direction, i.e.  $T_0(\mathbf{x}) = T_0(r)$  for all  $r = \|\mathbf{x}\| \in [0, R]$ , the dynamics of the system can be modeled in cylindrical coordinates. Thus the system model can be represented by (1) with the volumetric heat flux taking the form

$$\dot{W} = \frac{1}{h} \sum_{v=1}^q \dot{q}_v(r) = \sum_{v=1}^q b_v(r) u_v(t) \quad (26)$$

where  $h$  is the thickness of the wafer. The input shape functions are obtained by substituting the averaged heat flux density (25) into (26) and solving for  $b_v(r)$  yielding

$$b_v(r) = \frac{\eta P_{el}}{2\pi \delta^2 h} p_v I_0^* \left( \frac{\bar{r}_v r}{\delta^2} \right) e^{-\frac{(r-\bar{r}_v)^2}{2\delta^2}}. \quad (27)$$

It can be seen in Fig. 5, that for rings equipped with the same number of LEDs, the maximum of the heat flux density decreases for rings having larger diameters.

Table 1  
System parameters.

Parameter	Description	Value	Unit
$T_r$	Ambient temperature	293.15	K
$h$	Wafer thickness	$1550 \cdot 10^{-6}$	m
$R$	Wafer radius	0.15	m
$\sigma_{sb}$	Stefan–Boltzmann constant	$5.6704 \cdot 10^{-8}$	W/(m <sup>2</sup> K <sup>4</sup> )
$a$	Heat transfer coefficient	4.7	W/(m <sup>2</sup> K)
$\delta$	Variance	0.015	m
$\varepsilon$	Total emissivity	0.95	–

The heat losses are considered in  $\gamma(T)$  in the PDE (1) and modeled by

$$\gamma(T) = \frac{a}{h} (T_r - T) + 2 \frac{\varepsilon \sigma_{sb}}{h} (T_r^4 - T^4) \quad (28)$$

which takes into account convective and radiative heat transfer where  $T_r$  denotes the ambient temperature,  $a$  is a heat transfer coefficient and  $\sigma_{sb}$  represents the Stefan–Boltzmann constant. The parameter  $\varepsilon$  is the total emissivity. Note that according to (28) the function  $\gamma$  is locally Lipschitz continuous.

The dependency of the specific heat capacity  $c_p$  as well as the thermal conductivity  $k$  on the wafer's temperature  $T$  are modeled by

$$c_p(T) = 703 + \frac{255 \left( \frac{T}{300} \right)^{1.85} - 1}{\left( \frac{T}{300} \right)^{1.85} + \frac{255}{703}}, \quad (29)$$

and

$$k(T) = 150 \left( \frac{T}{300} \right)^{-1.3}, \quad (30)$$

respectively, see, e.g., [48,49]. The constant system parameters are summarized in Table 1.

### 5.3. Experimental validation of the observer

The observer for the considered process is given by (4) with (27), (28), (29) and (30). For the implementation of the observer there are in principal two approaches: to directly discretize the observer, or, to apply in advance the so-called Kirchhoff transformation to simplify the governing PDE of the observer, see, e.g., [39,50]. The latter approach yields an equivalent observer which is simpler to discretize in space. The estimated temperature is obtained by applying the inverse Kirchhoff transformation to the estimate of the transformed observer. The results presented in the following have been derived by exploiting the second approach. In the implementation of the observer, the temperature dependencies of the thermal conductivity and the heat capacity are taken into account according to (29) and (30), respectively. The observer has been discretized by replacing the spatial derivatives by finite differences where the number of segments is set to 50. The resulting system of ODEs, i.e.,  $\dot{\vartheta}(t) = \mathbf{f}(\vartheta(t), \mathbf{u}(t))$  where the state vector  $\vartheta$  includes the estimated temperature (in transformed coordinates) at the grid points and  $\mathbf{u}$  denotes the (transformed) external inputs to the observer, has been discretized in time by applying the implicit Euler scheme. Thus, one obtains the difference equation  $\vartheta_k = \vartheta_{k-1} + T_s \mathbf{f}(\vartheta_k, \mathbf{u}_k)$  for the estimated temperature at time step  $k = 1, 2, 3, \dots$ . The time step is set to  $T_s = 0.2$  s which is equal to the sampling time used to record the measured temperature. The finite difference equations were solved in MATLAB using the build in nonlinear equations solver *fsolve*. As mentioned above, for the validation of the proposed observer a thermographic camera was installed. In this experiment, the pointwise measurements are emulated by taking only selected points from the camera image.

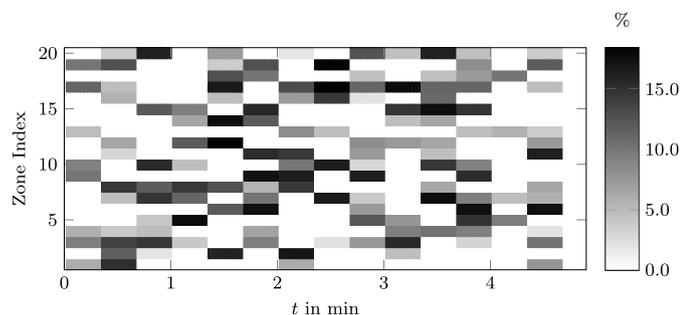


Fig. 6. System excitation.

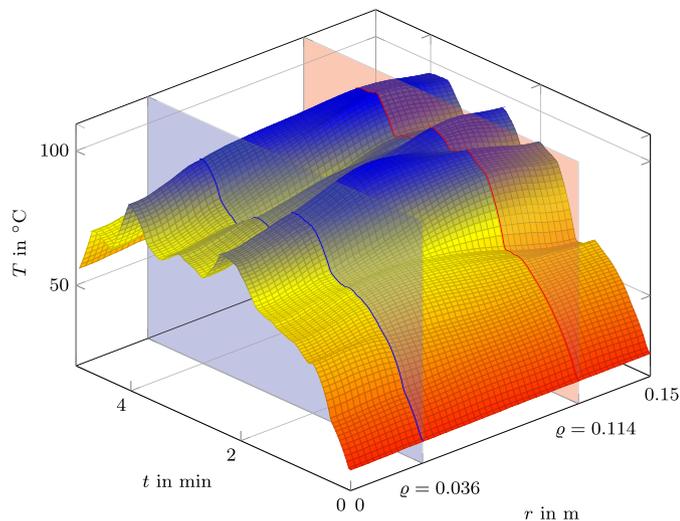


Fig. 7. Measured temperature.

During the experiment the heater is excited randomly. The actuating signal is illustrated in Fig. 6. The plot shows the heater power in percent for all 20 rings over time where, as can be seen in the bar beside the plot, dark gray fields correspond to larger heating power.

The plot in Fig. 7 shows the corresponding spatial and temporal evolution of the radial temperature measured by the thermographic camera. Two experiments are carried out. In the first experiment only one temperature measurement is considered in the observer whereas in the second experiment two pointwise measurements are taken into account. For the first experiment the measurement is taken at  $\varrho_1 = 0.036$  m. The temporal evolution of the temperature at this point is depicted as blue line in Fig. 7. In the second experiment an additional measurement at  $\varrho_2 = 0.114$  m is assumed. The temperature at this point is drawn as red colored line in Fig. 7. It is noteworthy that the observer is executed offline which allows to take the same measurement set in both experiments.

The resulting spatial and temporal evolution of the estimation error for the first experiment is depicted in Fig. 8. For visualization purposes the data have been downsampled (in the time coordinate). The observer was initialized with  $\hat{T}_0 = 0$ . It can be seen, that the estimation error converges to a vicinity of zero as ensured by the discussed ISS property. Thus, the observer provides accurate estimates of the wafer temperature. The result with the additional measurement is given in Fig. 9. Compared to the first experiment, the convergence is significantly faster. This can be seen in Fig. 10 which shows the evolution the estimation error norm for both experiments. In either case, the norm of the estimation error converges towards zero and remains

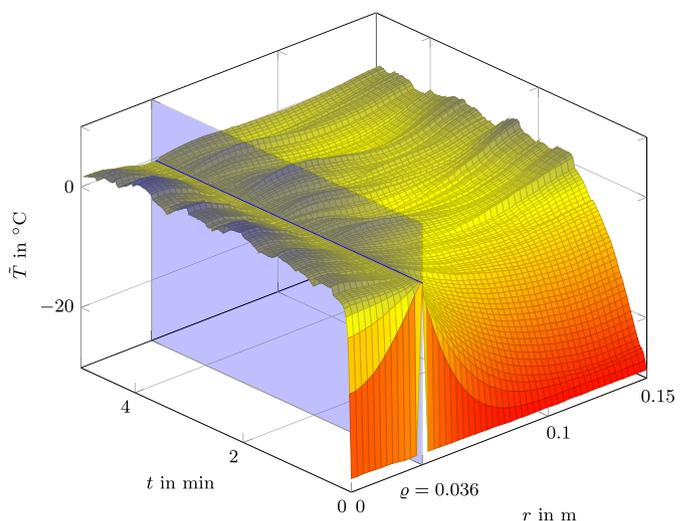


Fig. 8. Estimation error with one sensor.

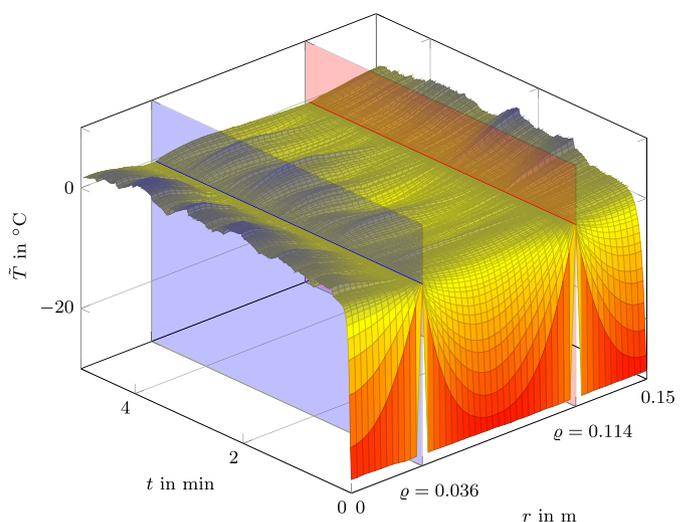


Fig. 9. Estimation error with two sensors.

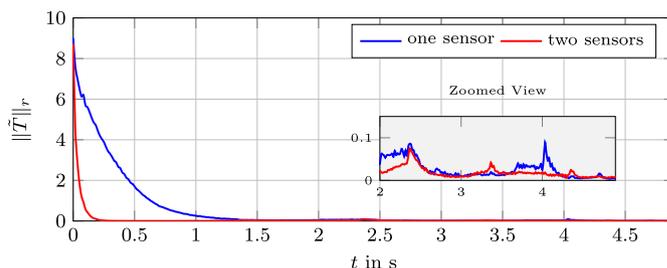


Fig. 10. Norm of the estimation error.

bounded as demonstrated in Section 4. However, after transients have vanished, the estimation error is in the same range for both experiments and the additional second measurement does not significantly improve the estimation accuracy. This is mainly attributed to the fact that model uncertainties and disturbances act in a distributed manner over the entire spatial domain.

**Table 2**  
Comparison of numerical schemes (one sensor).

Numerical scheme	Spatial segments	Time step (s)	RMS error (°C)
Explicit, variable step (ODE45)	25	–	3.21
Explicit Euler	25	0.0001	3.21
Implicit Euler	25	0.2	3.22
Explicit, variable step (ODE45)	50	–	3.28
Explicit Euler	50	0.0001	3.28
Implicit Euler	50	0.2	3.29
Explicit, variable step (ODE45)	100	–	3.44
Explicit Euler	100	0.0001	3.44
Implicit Euler	100	0.2	3.45

Table 2 compares the root mean square (RMS) estimation error<sup>4</sup> for different numerical schemes and with finer and wider spatial grid, respectively. For the simulation with the variable step size solver and the forward Euler scheme, where the step size is smaller than the sampling time  $T_s = 0.2$  s, the inputs are kept constant over the sampling interval. It can be seen, that these variations have a small impact on the result (as long as the time step is kept small).

## 6. Conclusion

A pointwise measurement injection observer was designed for a 1D nonlinear heat conduction problem considering in-domain measurements. In particular the heat equation in polar coordinates with nonlinear dependency of the thermal conductivity on temperature as well as nonlinear heat loss terms is considered. The problem is motivated by silicon wafer production processes that require estimation of the wafers surface temperature from point-wise measurements. The input-to-state stability of dynamics of the observation error in presence of distributed bounded disturbances is established using an ISS Lyapunov functional in combination with the nonlinear Kirchhoff transformation. Explicit conditions for ISS are established in terms of the sensor locations and the Lipschitz conditions for the involved nonlinearities under the mild and practically feasible assumption of constant thermal diffusivity. The observer has been tested using measurements from a real silicon wafer production tool. The experimental results agree with theoretical findings and show the effectiveness of the observer.

Future studies will combine the proposed observer scheme with state-feedback control algorithms.

## CRediT authorship contribution statement

**Alexander Schaum:** Methodology, Conceptualization, Writing – original draft, Writing – review & editing, Formal analysis. **Stefan Koch:** Methodology, Conceptualization, Visualization, Writing – original draft, Writing – review & editing, Data curation, Resources, Investigation, Formal analysis, Software. **Martin Kleindienst:** Methodology, Conceptualization, Writing – review & editing, Data curation, Resources, Investigation, Formal analysis, Software. **Markus Reichhartinger:** Conceptualization, Writing – review & editing, Formal analysis. **Thomas Meurer:** Conceptualization, Supervision, Writing – review & editing, Formal analysis. **Jaime A. Moreno:** Conceptualization, Writing – review & editing, Formal analysis. **Martin Horn:** Conceptualization, Supervision, Project administration, Funding acquisition, Writing – review & editing, Formal analysis.

<sup>4</sup> For this investigation the observer has been initialized with the measured temperature over the entire spatial domain.

## Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Martin Kleindienst reports a relationship with Lam Research AG that includes: employment.

## Data availability

The data that has been used is confidential.

## Appendix. Proof of Theorem 1

Consider the positive definite candidate Lyapunov functional

$$V(\tilde{T}) = \frac{1}{2} \int_0^R r \left( \int_T^{T+\tilde{T}} k(\tau) d\tau \right)^2 dr. \quad (A.1)$$

The rate of change of  $V$  along solutions of the error dynamics (7) are given by

$$\frac{d}{dt} V(\tilde{T}) = \int_0^R r \left( \int_T^{T+\tilde{T}} k(\tau) d\tau \right) \left( k(T + \tilde{T}) \partial_t (T + \tilde{T}) - k(T) \partial_t T \right) dr$$

where, according to (6b) and (7) it holds that

$$\begin{aligned} & k(T + \tilde{T}) \partial_t (T + \tilde{T}) - k(T) \partial_t T \\ &= \frac{k(T + \tilde{T})}{\rho c_p (T + \tilde{T})} \left( \frac{1}{r} \partial_r (rk(T + \tilde{T}) \partial_r (T + \tilde{T})) + \gamma(T + \tilde{T}) \right. \\ &\quad \left. + \sum_{v=1}^q b_v u_v \right) \\ &\quad - \frac{k(T)}{\rho c_p (T)} \left( \frac{1}{r} \partial_r (rk(T) \partial_r (T)) + \gamma(T) \right. \\ &\quad \left. + \sum_{v=1}^q b_v u_v + w \right). \end{aligned}$$

Notice that according to Assumption 1

$$\alpha = \frac{k(T + \tilde{T})}{\rho c_p (T + \tilde{T})} = \frac{k(T)}{\rho c_p (T)} = \text{const.}$$

and that for any constant  $T_{\min}$  it holds, by virtue of the Leibnitz integral rule, that

$$\begin{aligned} k(T) \partial_r T &= \partial_r \int_{T_{\min}}^T k(\tau) d\tau, \\ k(T + \tilde{T}) \partial_r (T + \tilde{T}) &= \partial_r \int_{T_{\min}}^{T+\tilde{T}} k(\tau) d\tau. \end{aligned} \quad (A.2)$$

Hence it follows that

$$k(T + \tilde{T}) \partial_t (T + \tilde{T}) - k(T) \partial_t T$$

$$\begin{aligned}
 &= \alpha \left( \frac{1}{r} \partial_r \left( r \partial_r \int_{T_{\min}}^{T+\tilde{T}} k(\tau) d\tau \right) + \gamma(T + \tilde{T}) + \sum_{v=1}^q b_v u_v \right) \\
 &\quad - \alpha \left( \frac{1}{r} \partial_r \left( r \partial_r \int_{T_{\min}}^T k(\tau) d\tau \right) + \gamma(T) \right. \\
 &\quad \left. + \sum_{v=1}^q b_v u_v + w \right) \\
 &= \frac{\alpha}{r} \partial_r \left( r \partial_r \int_{T_{\min}}^{T+\tilde{T}} k(\tau) d\tau - r \partial_r \int_{T_{\min}}^T k(\tau) d\tau \right) \\
 &\quad + \alpha \left( \gamma(T + \tilde{T}) + \sum_{v=1}^q b_v u_v - \gamma(T) \right. \\
 &\quad \left. - \sum_{v=1}^q b_v u_v - w \right) \\
 &= \frac{\alpha}{r} \partial_r \left( r \partial_r \int_T^{T+\tilde{T}} k(\tau) d\tau \right) + \varphi(\tilde{T}; T) - \alpha w
 \end{aligned}$$

with

$$\begin{aligned}
 \varphi(\tilde{T}; T) &:= \alpha \left( \gamma(T + \tilde{T}) - \gamma(T) \right), \\
 \varphi(0; T) &= 0, \quad \forall T \in [T_{\min}, \infty).
 \end{aligned} \tag{A.3}$$

Accordingly it holds that

$$\begin{aligned}
 \frac{d}{dt} V(\tilde{T}) &= \\
 &\int_0^R r \left( \int_T^{T+\tilde{T}} k(\tau) d\tau \right) \left( \frac{\alpha}{r} \partial_r \left( r \partial_r \int_T^{T+\tilde{T}} k(\tau) d\tau \right) \right. \\
 &\left. + \varphi(\tilde{T}; T) - \alpha w \right) dr
 \end{aligned}$$

Note that (A.2) actually gives rise to the so-called Kirchhoff transformation [41, Chapt. 7]

$$\phi(T) := \int_{T_{\min}}^T k(\tau) d\tau. \tag{A.4}$$

Define

$$\epsilon(\tilde{T}; T) := \int_T^{T+\tilde{T}} k(\tau) d\tau, \quad \epsilon_j(\tilde{T}_j; T) := \int_T^{T+\tilde{T}_j} k(\tau) d\tau, \tag{A.5}$$

for  $j = 1, \dots, m + 1$ . Recall the function  $\psi$  defined in (8). Note that (see, e.g., [39])  $\phi$  and thus  $\psi$  are both invertible and further  $\psi^{-1}$  is Lipschitz continuous with Lipschitz constant  $L_{\psi^{-1}}$  as long as  $k$  is strictly positive and bounded. In consequence it holds true that

$$\tilde{T} = \psi^{-1}(\epsilon), \quad \tilde{T}_j = \psi^{-1}(\epsilon_j). \tag{A.6}$$

Considering no measurement at  $r = 0$  one can thus write

$$\begin{aligned}
 \frac{d}{dt} V(\tilde{T}) &= \int_0^R r \left( \frac{\alpha}{r} \partial_r (r \partial_r \epsilon) + \varphi(\tilde{T}; T) - \alpha w \right) dr \\
 &= \sum_{j=0}^m \int_{\epsilon_j}^{\epsilon_{j+1}} r \epsilon_j \left( \frac{\alpha}{r} \partial_r (r \partial_r \epsilon_j) + \varphi(\tilde{T}_j; T) - \alpha w \right) dr.
 \end{aligned}$$

Recall the definition of the operators  $\mathcal{A}_j$ ,  $j = 1, \dots, m + 1$  from (9). Note that according to [51] all the  $-\mathcal{A}_j$  are Sturm-Liouville operators with eigenvalues  $-\lambda_n(\mathcal{A}_j)$   $n \in \mathbb{N}$  satisfying  $0 > \lambda_1 > \lambda_2 > \dots$  and  $\lim_{n \rightarrow \infty} \lambda_n(\mathcal{A}_j) = -\infty$  for all  $j = 1, \dots, m + 1$ .

Furthermore, let  $u, v \in L^2([\epsilon_j, \epsilon_{j+1}])$  and equip  $L^2([\epsilon_j, \epsilon_{j+1}])$  with the inner product

$$\langle u, v \rangle_{r,j} = \int_{\epsilon_j}^{\epsilon_{j+1}} ru(r)v(r)dr$$

and induced norm

$$\|v\|_{r,j} = \sqrt{\langle v, v \rangle_{r,j}}.$$

For all  $v \in \mathcal{D}(\mathcal{A}_j)$  it holds that

$$\langle v, \mathcal{A}_j v \rangle_{r,j} \leq \sup_{n \in \mathbb{N}} \lambda_n(\mathcal{A}_j) \|v\|_{r,j}^2. \tag{A.7}$$

Let  $\kappa$  be the maximum of the suprema of the eigenvalues  $\lambda_n(\mathcal{A}_j)$ , i.e.,

$$-\kappa = \max_{j=1, \dots, m+1} \sup_{n \in \mathbb{N}} \lambda_n(\mathcal{A}_j). \tag{A.8}$$

It holds that  $\kappa > 0$  and accordingly that for  $w = 0$

$$\begin{aligned}
 \frac{d}{dt} V &= \sum_{j=1}^{m+1} \int_{\epsilon_{j-1}}^{\epsilon_j} \left( r \epsilon_j \mathcal{A}_j \epsilon_j + r \epsilon_j \varphi(\tilde{T}_j; T) \right) dr \\
 &= \sum_{j=1}^{m+1} \langle \epsilon_j, \mathcal{A}_j \epsilon_j \rangle_{r,j} + \sum_{j=1}^{m+1} \int_{\epsilon_{j-1}}^{\epsilon_j} r \epsilon_j \varphi(\tilde{T}_j; T) dr \\
 &\leq \sum_{j=1}^{m+1} \sup_{n \in \mathbb{N}} \lambda_n(\mathcal{A}_j) \|\epsilon_j\|_{r,j}^2 + \int_0^R r \epsilon |\varphi(\tilde{T}; T)| dr \\
 &\leq -\kappa \sum_{j=1}^{m+1} \int_{\epsilon_{j-1}}^{\epsilon_j} r \epsilon_j^2 dr + \int_0^R r |\epsilon| |\varphi(\tilde{T}; T)| dr \\
 &\leq -\kappa \int_0^R r \epsilon^2 dr + \int_0^R r |\epsilon| |\varphi(\tilde{T}; T)| dr.
 \end{aligned}$$

Furthermore, notice that the functional  $V$  can be directly expressed in terms of the transformed observation error  $\epsilon(\tilde{T}; T)$  defined in (A.5) and it holds true that

$$\begin{aligned}
 V(\tilde{T}) &= \frac{1}{2} \int_0^R r \epsilon^2(\tilde{T}; T) dr = \frac{1}{2} \|\epsilon(\tilde{T}; T)\|_r^2, \\
 a \|\tilde{T}\|_r^2 &\leq V(\tilde{T}) \leq b \|\tilde{T}\|_r^2
 \end{aligned} \tag{A.9}$$

for all  $a \leq (1/2L_{\psi^{-1}})^2$  and  $b \geq (L_{\psi})^2/2$ , since this implies that

$$a \|\tilde{T}\|_r^2 = a \|\psi^{-1}(\epsilon)\|_r^2 \leq a (L_{\psi^{-1}})^2 \|\epsilon\|_r^2 \leq \frac{1}{2} \|\epsilon\|_r^2 = V(\tilde{T})$$

and

$$V(\tilde{T}) = \frac{1}{2} \|\epsilon\|_r^2 = \frac{1}{2} \|\psi(\tilde{T})\|_r^2 \leq \frac{1}{2} (L_{\psi})^2 \|\tilde{T}\|_r^2 \leq b \|\tilde{T}\|_r^2.$$

Recall that  $\gamma$  is Lipschitz continuous and let  $L_{\gamma}$  denote its Lipschitz constant. In consequence, recalling the definition (A.3) of  $\varphi$  it follows that for all  $\tilde{T} = \psi^{-1}(\epsilon)$  and  $T \in [T_{\min}, \infty)$  one has

$$|\varphi(\tilde{T}; T)| \leq \alpha L_{\gamma} |\tilde{T}| = \alpha L_{\gamma} |\psi^{-1}(\epsilon)| \leq \alpha L_{\gamma} L_{\psi^{-1}} |\epsilon| = \bar{L} |\epsilon|$$

with  $\bar{L} = \alpha L_{\gamma} L_{\psi^{-1}}$ , so that

$$\begin{aligned}
 \frac{d}{dt} V &\leq -\kappa \int_0^R r \epsilon^2 dr + \int_0^R r |\epsilon| |\varphi(\tilde{T}; T)| dr \\
 &\leq -2\kappa V + \bar{L} \int_0^R r |\epsilon|^2 dr \\
 &= -2\kappa V + \bar{L} \|\epsilon\|_r^2 = -2(\kappa - \bar{L}) V.
 \end{aligned}$$

Note that this property together with (A.9) implies the exponential stability of  $\tilde{T} = 0$  in the norm  $\|\cdot\|_r$  as long as  $\kappa > \bar{L}$  and  $w = 0$  (cp. the complementary analysis in [39]). For a bounded

disturbance  $0 \neq w$  with bound  $w^+$  defined in (1e) it further holds true that

$$\begin{aligned} \frac{d}{dt} V(\tilde{T}) &\leq -(\kappa - \bar{L}) \|\epsilon\|_r^2 - \int_0^R r \epsilon \alpha w dr \\ &\leq -(\kappa - \bar{L}) \|\epsilon\|_r^2 + \frac{1}{2} (\|\epsilon\|_r^2 + \alpha^2 (w^+)^2) \\ &= -\left(\kappa - \bar{L} - \frac{1}{2}\right) \|\epsilon\|_r^2 + \frac{\alpha^2}{2} (w^+)^2. \end{aligned}$$

This property together with (A.9) implies that  $V$  is an ISS Lyapunov functional [42–44] and thus the ISS (see (3)) of the observation error dynamics (7) follows for all  $\kappa > \bar{L} + 1/2$ , as stated in the theorem.  $\square$

## References

- [1] R.B. Fair, *Rapid Thermal Processing: Science and Technology*, Academic Press, 2012.
- [2] V.E. Borisenko, P.J. Hesketh, *Rapid Thermal Processing of Semiconductors*, Springer Science & Business Media, 2013.
- [3] A. Emami-Naeini, J.L. Ebert, D. de Roover, R.L. Kosut, M. Dettori, L.M.L. Porter, S. Ghosal, Modeling and control of distributed thermal systems, *IEEE Trans. Control Syst. Technol.* 11 (5) (2003) 668–683.
- [4] M. Kleindienst, M. Reichhartinger, M. Horn, F. Staudegger, Observer-based temperature control of an LED heated silicon wafer, *J. Process Control* 70 (2018) 96–108.
- [5] C.D. Schaper, M.M. Mosehi, K.C. Saraswat, T. Kailath, Modeling, identification, and control of rapid thermal processing systems, *J. Electrochem. Soc.* 141 (11) (1994) 3200.
- [6] E. Dassau, B. Grosman, D.R. Lewin, Modeling and temperature control of rapid thermal processing, *Comput. Chem. Eng.* 30 (4) (2006) 686–697.
- [7] R.F. Curtain, H. Zwart, *An Introduction to Infinite-Dimensional Linear Systems Theory*, Springer, 1995.
- [8] M. Tucsnak, G. Weiss, *Observation and Control of Operator Semigroups*, Birkhäuser-Verlag, Basel, 2009.
- [9] M. Krstic, A. Smyshlyaev, *Boundary Control of PDEs: A Course on Backstepping Designs*, SIAM, 2008.
- [10] R.F. Curtain, Finite-dimensional compensator design for parabolic distributed parameter systems with point sensors and boundary input, *IEEE Trans. Autom. Control* 27 (1982) 98–104.
- [11] Y. Yang, S. Djuljevic, Linear matrix inequalities (LMIs) observer and controller design synthesis for parabolic PDE, *Eur. J. Control* 20 (2014) 227–236.
- [12] A. Smyshlyaev, M. Krstic, Backstepping observers for a class of parabolic pdes, *Syst. Control Lett.* 54 (2005) 613–625.
- [13] T. Meurer, A. Kugi, Tracking control for boundary controlled parabolic PDEs with varying parameters: Combining backstepping and differential flatness, *Automatica* 45 (2009) 1182–1194.
- [14] J. Deutscher, Backstepping design of robust output feedback regulators for boundary controlled parabolic PDEs, *IEEE Trans. Automat. Control* 61 (2016) 2288–2294.
- [15] C. Kitsos, G. Besancon, C. Prieur, High-gain observer for  $3 \times 3$  linear heterodirectional hyperbolic systems, *Automatica* 129 (2021) 109607.
- [16] T. Meurer, Control of Higher Dimensional PDEs, in: *Communication and Control Engineering*, Springer, 2013.
- [17] L. Jadachowski, T. Meurer, A. Kugi, Backstepping observers for linear PDEs on higher-dimensional spatial domains, *Automatica* 51 (2015) 85–97.
- [18] T. Meurer, On the extended Luenberger observer for semilinear distributed-parameter systems, *IEEE Trans. Autom. Control* 58 (7) (2013) 1732–1743.
- [19] M. Ksouri, O. Boubaker, J.-P. Babary, Variable structure estimation and control of nonlinear distributed parameter bioreactors, in: *Proc. IEEE International Conference on Systems, Man, and Cybernetics*, 1998, pp. 3770–3774.
- [20] P. Ligarius, J.F. Couchouron, Asymptotic observers for a class of evolution operators, a nonlinear approach, *C. R. Acad. Sci.* 1 324 (3) (1997) 355–360.
- [21] R.F. Curtain, M.A. Demetriou, K. Ito, Adaptive compensators for perturbed positive real infinite dimensional systems, *Int. J. Appl. Math. Comput. Sci.* 4 (13) (2003) 441–452.
- [22] D. Dochain, State observers for tubular reactors with unknown kinetics, *J. Process Control* 10 (2000) 259–268.
- [23] C. Kitsos, G. Besançon, C. Prieur, High-gain observer design for a class of quasi-linear integro-differential hyperbolic systems—application to an epidemic model, *IEEE Trans. Automat. Control* 67 (1) (2022) 292–303.
- [24] G. Hagen, I. Mezic, Spillover stabilization in finite-dimensional control and observer design for dissipative evolution equations, *SIAM J. Control Optim.* 42 (2) (2003) 746–768.
- [25] A. Schaum, J.A. Moreno, E. Fridman, J. Alvarez, Matrix Inequality-Based observer design for a class of distributed transport-reaction systems, *Int. J. Robot. Nonlinear Control* 24 (16) (2013).
- [26] R. Vazquez, M. Krstic, Control of 1-D parabolic PDEs with Volterra nonlinearities. Part I: Design, *Automatica* 44 (2008) 2778–2790.
- [27] R. Vazquez, M. Krstic, Control of 1-D parabolic PDEs with Volterra nonlinearities. Part II: Analysis, *Automatica* 44 (2008) 2791–2803.
- [28] A. Hasan, Backstepping boundary control for semilinear parabolic PDEs, in: *Proceedings of the IEEE 54th Annual Conference on Decision and Control, CDC*, 2015, pp. 2513–2518.
- [29] L. Hu, R. Vazquez, F.D. Meglio, M. Krstic, Boundary exponential stabilization of 1-dimensional inhomogeneous quasi-linear hyperbolic systems, *SIAM J. Control Optim.* 57 (2) (2019) 963–998.
- [30] A. Hasan, Output feedback stabilization of semilinear parabolic PDEs using backstepping, 2016, arXiv:1612.03867.
- [31] A. Schaum, T. Meurer, J.A. Moreno, Dissipative observers for coupled diffusion-convection-reaction systems, *Automatica* 94 (2018) 307–314.
- [32] A. Schaum, *Dissipativity in Control Engineering. Applications to Finite and Infinite Dimensional Systems*, DeGruyter, 2021.
- [33] A. Schaum, J.A. Moreno, J. Alvarez, T. Meurer, A simple observer scheme for a class of 1-D semi-linear parabolic distributed parameter systems, in: *Proceedings European Control Conference*, 2015, pp. 49–54.
- [34] A. Schaum, J. Alvarez, T. Meurer, J.A. Moreno, State-estimation for a class of tubular reactors using a pointwise innovation scheme, *J. Process Control* 60 (2017) 104–114.
- [35] P. Feketa, A. Schaum, T. Meurer, Distributed parameter state estimation for the gray-scott reaction-diffusion model, *Systems* 9 (4) (2021).
- [36] A. Schaum, An unknown input observer for a class of diffusion-convection-reaction systems, *At-Automatisierungstechnik* 66 (7) (2018) 548–557.
- [37] D.G. Luenberger, Observing the state of a linear system, *IEEE Trans. Mil. Electron.* 8 (2) (1964) 74–80.
- [38] D.G. Luenberger, An introduction to observers, *IEEE Trans. Autom. Control* 16 (6) (1971) 596–602.
- [39] A. Schaum, S. Koch, M. Reichhartinger, T. Meurer, J.A. Moreno, M. Horn, Nonlinear observer design for a 1D heat conduction process, in: *Proceedings of the 60th IEEE Conference on Decision and Control (CDC)*, Austin, Texas, 2021, pp. 1149–1154.
- [40] M. Kleindienst, M. Reichhartinger, S. Koch, M. Horn, A distributed-parameter approach for the surface temperature estimation of an LED heated silicon wafer, *IEEE Trans. Semicond. Manuf.* 32 (4) (2019) 559–565.
- [41] M.J. Latif, *Heat Conduction*, third ed., Springer-Verlag, Berlin Heidelberg, 2009.
- [42] E.D. Sontag, On the input-to-state stability property, *Eur. J. Control* 1 (1) (1995) 24–36.
- [43] E.D. Sontag, Y. Wang, On characterizations of the input-to-state stability property, *Systems Control Lett.* 24 (5) (1995) 351–359.
- [44] A. Mironchenko, C. Prieur, Input-to-state stability of infinite-dimensional systems: recent results and open questions, *SIAM Rev. Soc. Ind. Appl. Math.* 62 (3) (2020) 529–614.
- [45] S. Lin, H.-S. Chu, Application of inverse problem algorithm for temperature uniformity in rapid thermal processing, *Thin Solid Films* 42 (2002) 280–289.
- [46] R. Kakoschke, E. Bußmann, H. Föll, Modelling of wafer heating during rapid thermal processing, *Appl. Phys. A* 50 (1990) 141–150.
- [47] M. Abramowitz, I.A. Stegun, *Handbook of Mathematical Functions - with Formulas, Graphs, and Mathematical Tables*, Dover Publications, Inc., New York, 1965.
- [48] S. Tiwari, *Compound Semiconductor Device Physics*, Academic Press, Amsterdam, Boston, 2013.
- [49] V. Palankovski, R. Schultheis, S. Selberherr, Simulation of power heterojunction bipolar transistors on gallium arsenide, *IEEE Trans. Electron Devices* 48 (6) (2001) 1264–1269.
- [50] D. Benameur, A. Maidi, S. Djennoune, J.-P. Corriou, Observer design for a nonlinear diffusion system based on the Kirchhoff transformation, *Int. J. Dyn. Control* 6 (1) (2018) 154–166.
- [51] C. Delattre, D. Dochain, J. Winkin, Sturm-Liouville systems are Riesz-spectral systems, *Int. J. Appl. Math. Comput. Sci.* 13 (2003) 481–484.