Adaptive Thermal Management of Implantable Device

Ruizhi Chai, Ying Zhang

***Abstract*—The temperature increase in the surrounding tissue caused by the operation of neural prosthesis has raised growing concern as the device becomes more powerful. In this paper, a real-time adaptive thermal management method is developed for implantable devices to optimize their operation while maintaining a safe operating temperature. The method adopts a simplified thermal model for supporting real-time control and updates the model parameters online using a proposed recursive multi-step prediction error minimization method (RMSPEM). The performance of the developed thermal management method is evaluated using simulation studies, and the results show that it can achieve longer operation time and better overall performance while maintaining the safe operating temperature. In addition, the developed thermal model is validated using both the COMSOL simulation and an *in vitro* experiment.**

***Keywords*—Implantable device, thermal management, neuroprosthetics.**

I. Introduction

With the increase of functionality and complexities of implantable medical devices, the temperature rise in the tissue around the implants has become a new challenge due to the detrimental effects of exposing tissue to a temperature a few degree Celsius above the normal body temperature. For example, it is shown that a temperature increase greater than 1 °C could cause damage to the brain tissue [1].

Many studies have been conducted to investigate the thermal effect of implantable devices. In [2], various origins of temperature increase and the possible methods to model them are studied for a dual-unit retinal prosthesis. The results show that, in addition to the power dissipation of the implanted microchip, the telemetry coil, the stimulating electrodes, and the electromagnetic fields induced in the body could also contribute to the temperature increase, which can be modeled based on the Pennes bioheat equation [3]. The thermal effect of an integrated 3-D Utah electrode array (UEA) device implanted in the brain is investigated in [4] by experiments and numerical simulations that uses finite element analysis (FEA) to solve the Pennes bioheat equation. The results indicate that the numerical simulation and experimental measurements are in good agreement. Henry *et al.* [5] study the thermal effect of brain implants with a focus on light-emitting diode.

The commonly used method to guarantee thermal safety of the implantable device is combining numerical method with in vivo experiments during the design phase [6], [7], [8]. The maximum power dissipation is determined based on the maximum tolerable temperature, which is then specified as a design constraint. However, the performance of the device is often limited due to the fact that an over-conservative design is chosen in many cases to account for the uncertainties of the actual operation.

Several researchers have conducted research related to the online thermal management. Communication scheduling methods are developed for biosensor network applications to prevent the temperature from increasing above the safe limit [9], [10], [11]. For example, a method for selecting the network leader that communicates with the base station is developed by Tang *et al.* [12] based on the leadership rotation history and the location of the next leader. These work only consider the overheating caused by communication and focuses on reducing overheating by communication scheduling. Wentz *et al.* [13] developed a wireless supercapacitor-based headborne device, which stimulates brain cells with an LED array. To solve the overheating issue, the LEDs are shutdown when the temperatures reach a predefined threshold. As a primitive method of thermal management, this method can not optimize the operation of the device. A similar approach is also employed in Luo *et al.* [14]. But instead of a temperature sensor, a thermoresponsive micro circuit breaker is used to protect the device from overheating. Moreover, Krishna *et al.* [15] provides a comprehensive survey of the thermal management of cyber-physical systems. However, to the authors' knowledge, there have been few studies addressing the dynamic thermal management problem of the implantable device.

In the previous work [16], a joint power and thermal management method was developed to achieve three goals: sustainable operation, performance maximization, and thermal safety. However, the method was developed for a device exposed to air and the thermal effect of the body was not considered.

Herein, we investigate the dynamic thermal management problem of implantable devices, with a focus on neural prosthesis. Neural prosthesis are implantable devices that interact with the central nervous system or the peripheral nervous system and help to restore motor, sensory, or cognitive functionality that may have been damaged as a result of an injury or disease. This particular example is representative since it includes most of the potential causes of thermal dissipation: a microchip that could dissipate relatively large power, a telemetry system, and a potentially large number of stimulating electrodes. For this device, consuming power as low as possible is of importance from two standpoints. First, power budget for the implantable device operation is limited, whether they are battery operated or telemetry operated. Second, from the biological point of view, the heat generated by the circuits

R. Chai and Y. Zhang are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, 30332 USA. E-mail: rchai3@gatech.edu, yzhang@gatech.edu.

could potentially cause damage to the surrounding tissue. However, the requirement for high performance generally leads to large power consumption. As a result, there exists a trade-off between the performance of the device and its power consumption.

An adaptive thermal management method for neural prosthesis is proposed in this paper to maintain safe operating temperature and maximize system performance. To support the investigation, a COMSOL model of the UEA is first built to study the heating effect of the device on the surrounding tissue based on the Pennes bioheat equation. A simplified thermal model is then developed to facilitate real-time control. The model parameters are updated online with the proposed multi-step temperature prediction method (RMSPEM) given the real-time temperature measurement. Based on the analysis of the potential heat sources of neural prosthesis, the system model in terms of power consumption of each heat source is developed for thermal management using model predictive control (MPC). By adjusting the working status of all the components, the MPC controller maximizes the device performance while maintaining safe operation. The safe operation is guaranteed by enforcing the maximum allowable temperature with respect to the body temperature. The system performance is formulated as the objective function of the MPC problem. The proposed method is then extended to achieve joint thermal and power management by incorporating the energy storage model.

The remainder of the paper is organized as follows. Following the introduction, Section II presents the system model of implantable device with neural prosthesis as an example. The multi-step thermal effect prediction method is proposed in Section III. The adaptive thermal management method based on MPC is formulated in Section IV. Section V presents the simulation studies of the proposed method. At last, the paper is concluded in Section VI.

II. System Description and Model

In the work of Wentz *et al.* [13], the system components of a neural prosthesis are presented, which is relatively representative and therefore adopted for investigation in this paper. The neural prosthesis consists of optics module, radio module, power module, and motherboard module. The optics module holds up to 16 LEDs and is surgically affixed to the skull of a rat, while the remainder of the device is attached to the optics module via a low insertion force connector. Microelectrode array is another widely used tool for deep brain stimulation. In this paper, we focus on the implanted 3-D micro-electrodes, specifically, the UEA, since its thermal effect has been studied in previous literature [4] through experiment. The radio module mediates control command from a computer or laptop. The motherboard module contains the microcontroller and the power circuitry. The power module includes a supercapacitor and an antenna for power reception. Some neural prosthesis don't have energy storage devices, but for applications with large current requirement, energy storage device is often adopted. The system diagram employed in this paper is shown in Figure 1, with the energy storage device taken into consideration.

Since the temperature increase is caused by the power dissipated in the circuitry, the thermal effect can be investigated via modeling the power flow of the system shown in Figure 1. The whole hardware/software system is powered by the power received via the wireless interface. Here we use wireless interface to model the power antenna and radio chipset, both of which could potentially contribute to the heating of surrounding tissue. Energy storage device, such as battery [17] or supercapacitor [18], is used to store the extra energy received by the system with some efficiency η. The direct usage of the received power is free of energy losses. The software system and the controller are running on the microprocessor, which resides on the motherboard. We also assume that there is a temperature sensor connected to the motherboard. The temperature sensor can be placed at the surface of the electrode array and is used to provide the feedback information for the thermal management method. The electrode array is implanted at the surface of the brain tissue. Depending on the placement of each component, microprocessor, wireless interface, and electrode array may contribute to the temperature increase in the surrounding tissue.

We assume the power consumption of the microprocessor and electrode array can be adapted online. For the microprocessor, this can often be achieved by using various dynamic power management (DPM) techniques like DVS, or by adjusting the ratio of active and inactive time for some microprocessors with only rudimentary DPM capabilities [19]. The power consumption of the electrode array can be adjusted by controlling the number of stimulation channels and the stimulation pulse train [4][13]. The stimulation is, in many cases, achieved through current-mode stimulation pulse. The current intensity, the phase and interphase durations, as well as the stimulation frequency are the parameters related to the power consumption. Furthermore, the performance of each component is proportional to its power consumption. For the micro-electrode array, the number of channels and the pulse train impact the device performance, such as the visual characteristics (diameter, depth, brightness, and duration) in the case of retinal prosthesis [20]. The power consumption of the components must also be greater than certain threshold to guarantee minimum operation. The transmitted power is also assumed to be adjustable. The designed controller adapts the working status of each component to optimize the device operation based on the information like task requirement, temperature readings and the energy level of the storage device.

III. Thermal Effect Modeling

A. Thermal Dynamics Analysis

Heat transfer from the implantable device to its surrounding tissue can be modeled using the famous Pennes bioheat equation [3], by

$$\rho C\frac{\partial T(\mathbf{x},t)}{\partial t} = \nabla\cdot(k\nabla T(\mathbf{x},t)) + A_0 - B_0(T(\mathbf{x},t)-T_b) + \rho SAR + P_{electronics}, \tag{1}$$

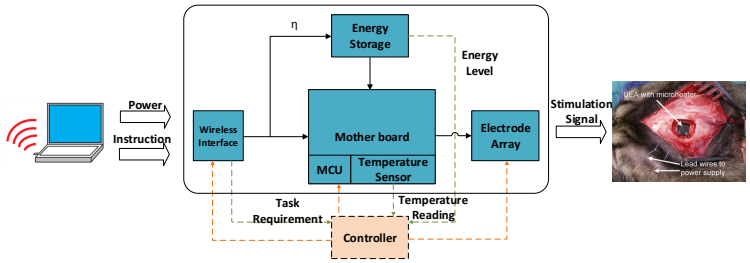
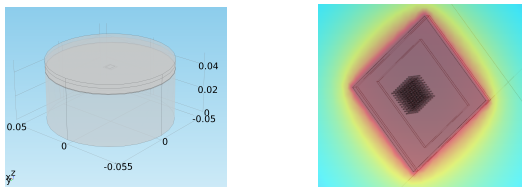


Fig. 1. Illustration of the system diagram (the picture of UEA is taken from [4])



(a) (b)

Fig. 2. Illustration of the developed COMSOL model (a) cylinderical human brain model with height and radius being 50 mm. The thickness of the scalp and skull are 3 and 5 mm respectively. (b) the UEA model with 10×10 micro-electrodes and a $7.88\ mm \times 7.53\ mm$ chip.

where T denotes the temperature in the tissue, $\mathbf{x}$ is the spatial coordinates, t represents the time, ρ and C are the tissue density and specific heat, respectively. $\nabla \cdot (k\nabla T(\mathbf{x}, t))$ models the thermal diffusion with k representing the thermal conductivity of tissue. $B_0(T(\mathbf{x}, t) - T_b)$ models the effect of blood perfusion and T_b is the temperature of blood. A_0 is the rate of the heat generation per unit volume of tissue by metabolism. ρSAR represents the heating effect due to the electromagnetic field and SAR is the abbreviation of specific absorption rate. $P_{electronics}$ is the power density of the implanted electronics.

The computational software COMSOL (Comsol Inc., Burlington, MA) that implements the FEA method can be used to solve the Pennes bioheat equation. The model of a UEA presented in [4] is implemented using the COMSOL Multiphysics software as shown in Figure 2. All the model parameters are adopted from [4]. According to [4], the simulation results based on the numerical solution of the Pennes bioheat equation and the experimental measurements are in good agreement. In this paper, we use this experimentally validated numerical model to demonstrate the effectiveness of the proposed thermal management method.

Numerous methods have been proposed to solve (1) more efficiently, including Finite Difference Time Domain (FDTD) methods [21], fourier transform based methods [22], [23], hybrid alternating-direction implicit (ADI) approach [24], and a method of superposition from separate sources combined with model simplification [25]. However, all these methods rely on sampling the temperature value in the simulation domain as they evolve in time. Their time complexity and space complexity make them unsuitable for real-time thermal management.

B. Simplified Thermal Model

A simplified thermal model is developed in this paper to facilitate the real-time management. To simplify the thermal model, the effect of SAR, blood perfusion and metabolism can be neglected from (1) and modeled as disturbance. For an implantable device like the one in [13] with power transfer frequency being 120 kHz and a low power UEA [4] with peak power consumption being 65 mW, the modeling error caused by this simplification is not very significant [26]. Moreover, the term $\nabla \cdot (k\nabla T(\mathbf{x}, t))$ in (1) can be replaced by $k\nabla^2 T(\mathbf{x}, t))$ for homogeneous materials. This resulting in a second order parabolic partial differential equation:

$$\rho C \frac{\partial T(\mathbf{x}, t)}{\partial t} = k\left(\frac{\partial^2 T(\mathbf{x}, t)}{\partial x^2} + \frac{\partial^2 T(\mathbf{x}, t)}{\partial y^2} + \frac{\partial^2 T(\mathbf{x}, t)}{\partial z^2}\right) + P_{electronics}. \quad (2)$$

The heat flow described by this differential equation has a similar form as that for electrical current, and there is a well known duality between them. More specifically, the heat flow (W) passing through a thermal resistor (°C/W) is equivalent to the electrical current (A) through an electrical resistance (Ohm), and the temperature difference (°C) corresponds to voltage difference (V). The heat absorbtion phenomenon can be denoted as the thermal equivalent capacitance (J/°C). The thermal conduction phenomenon can be represented as thermal equivalent resistance (K·m^2/W). The simplified thermal model is shown in [27].

Let P represent the power dissipation of the implantable device and T_0 represents the tissue temperature. R_t is the equivalent thermal resistance and C_t is the equivalent thermal capacitance. Then the temperature T of the tissue that has direct contact with the implantable device corresponds to the voltage of C_t, which can be calculated as

$$C_t \frac{dT}{dt} = \frac{T_0 - T}{R_t} + P. \quad (3)$$

The temperature can be solved numerically as

$$T(t) = (1 - \frac{\Delta t}{R_t C_t})T(t-1) + \frac{\Delta t}{R_t C_t} T_0 + \frac{\Delta t}{C_t} P(t-1). \quad (4)$$

In which, Δt is the step size.

Let $\Delta T(t) = T(t) - T_0$, then

$$\Delta T(t) = (1 - \frac{\Delta t}{R_t C_t})\Delta T(t-1) + \frac{\Delta t}{C_t} P(t-1). \quad (5)$$

Therefore, the complex thermal dynamics of the implantable device can be approximated by a first order linear function. The effect of the metabolism, blood perfusion and electromagnetic field are considered as disturbance to the system. With the following online multistep prediction method, this linear function can capture the fundamental dynamics of thermal effect. If, for some cases, the metabolism, blood perfusion and electromagnetic field have more significant impact on the temperature increase in the body, a linear model with higher order could be used to model the thermal effect.

C. Online Multistep Prediction Method

A suitable model for generating accurate output predictions within a horizon is crucial to achieve high performance closed-loop control. The methods to obtain such models are often referred to as MPC relevant identification methods [28]. Compared to standard prediction error methods (PEM), MPC relevant identification methods often have better long term prediction performance.

Based on the iterative single step prediction error method (ISSPEM) proposed by Farina *et al.* [29], a recursive multi-step prediction error method (RMSPEM) is developed to generate the optimal prediction model for temperature increase within a horizon of k time steps. The method takes the power consumption and temperature measurements as input, then calculates the coefficients of the following linear temperature model (6) that minimizes the prediction error.

$$\Delta T(t) = a\Delta T(t-1) + bP(t-1). \quad (6)$$

The discrete time transfer function corresponding to (6) is

$$G(z) = \frac{bz^{-1}}{1 - az^{-1}}. \quad (7)$$

The numerator is denoted as $N_G(z) = bz^{-1}$. The denominator is denoted as $D_G(z) = 1 - az^{-1}$. $\theta = (a, b)^T$ represent the vector of model parameters.

1) batch MSPEM: Let's first derive the multi-step prediction error (MSPEM) cost function and its optimization procedure for a batch of data. The MSPEM cost function represents the average prediction error within a horizon of k time steps.

Let $\Delta T(t+k|t)$ be the value of temperature increase predicted by iterating k times the recursive equation of (6), as a function of $\Delta T(\cdot)$ up to time t and $P(\cdot)$ up to time $t+k$. It can be represented as [30]:

$$\Delta T(t+k|t) = R_k(z)\Delta T(t) + E_k(z)N_G(z)P(t+k) \quad (8)$$

where $R_k(z)$ and $E_k(z)$ can be calculated as:

$$R_k(z) = a^k \quad (9)$$

and

$$E_k(z) = \sum_{i=0}^{k-1} a^i z^{-i}. \quad (10)$$

Let

$$\phi_k(t) = \begin{pmatrix} \Delta T(t) \\ P(t+k-1) \\ \vdots \\ P(t) \end{pmatrix}, \quad (11)$$

and

$$\Theta_k(\theta) = \begin{pmatrix} a^k \\ b \\ ab \\ \vdots \\ a^{k-1}b \end{pmatrix}. \quad (12)$$

The k-step ahead predictor can be reformulated as:

$$\Delta T(t+k|t) = \phi_k(t)^T \Theta_k(\theta). \quad (13)$$

For a batch of N measurements of temperature increase and power consumption values, define

$$Y_N^k = \begin{pmatrix} \Delta T(1+k) \\ \vdots \\ \Delta T(N) \end{pmatrix}, \quad (14)$$

and

$$\Phi_k = \begin{pmatrix} \phi_k(1)^T \\ \vdots \\ \phi_k(N-k)^T \end{pmatrix}. \quad (15)$$

The cost function of MSPEM is defined as

$$J_{MP}^N(k) = \frac{1}{k}\sum_{j=1}^{k} J_P^N(j), \quad (16)$$

in which

$$\begin{aligned} J_P^N(j) &= \frac{1}{N-j}\sum_{t=j+1}^{N} [\Delta T(t) - \Delta T(t|t-j)]^2 \\ &= \frac{1}{N-j}(Y_N^j - \Phi_j\Theta_j(\theta))^T(Y_N^j - \Phi_j\Theta_j(\theta)). \end{aligned} \quad (17)$$

This cost function can be minimized using the standard Newton method [30]. The optimal model parameters can be estimated iteratively as:

$$\theta_{j+1} = \theta_j - (\frac{\partial^2}{\partial\theta^2} J_{MP}^N(j)\Big|_{\theta=\theta_j})^{-1} \nabla_\theta J_{MP}^N(j)\Big|_{\theta=\theta_j}, \quad j = 1, \ldots, k-1. \quad (18)$$

The Hessian of $J_{MP}^N(j)$ can be calculated as:

$$\frac{\partial^2}{\partial\theta^2} J_{MP}^N(j)\Big|_{\theta=\theta_j} = \frac{1}{j}\sum_{s=1}^{j} \frac{\partial^2}{\partial\theta^2} J_P^N(s), \quad (19)$$

and

$$\frac{\partial^2}{\partial\theta^2} J_P^N(s) \approx \frac{2}{N-s}\nabla_\theta\Theta_s(\theta)\Phi_s^T\Phi_s\nabla_\theta\Theta_s(\theta)^T. \quad (20)$$

For the sake of convenience, we define

$$Q_N^s = \nabla_\theta\Theta_s(\theta)\Phi_s^T\Phi_s\nabla_\theta\Theta_s(\theta)^T. \quad (21)$$

The gradient of $J_{MP}^N(j)$ can be expressed as follows:

$$\nabla_\theta J_{MP}^N(j)\Big|_{\theta=\theta_j} = \frac{1}{j}\sum_{s=1}^{j}\nabla_\theta J_P^N(s) \tag{22}$$

and

$$\nabla_\theta J_P^N(s) = \frac{2}{N-s}\nabla_\theta\Theta_s(\theta)(\Phi_s^T\Phi_s\Theta_s(\theta) - \Phi_s^T Y_N^s). \tag{23}$$

In (20) and (23), Θ_j and its gradient $\nabla_\theta\Theta_j$ can be updated iteratively over j:

$$\Theta_{j+1} = W_j\Theta_j, \tag{24}$$

$$\nabla_\theta\Theta_{j+1} = \nabla_\theta\Theta_j {W_j}^T + ([1\ \mathbf{0}_{1,1+j}]\Theta_j){H_j}^T. \tag{25}$$

in which,

$$W_j = \left(\begin{pmatrix} a \\ \mathbf{0}_{j,1} \\ b \end{pmatrix}, \begin{pmatrix} \mathbf{0}_{1,j+1} \\ I_{j+1} \\ \mathbf{0}_{1,j+1} \end{pmatrix}\right), \tag{26}$$

and

$$H_j = \left(\begin{pmatrix} I_1 \\ \mathbf{0}_{j+2,1} \end{pmatrix}, \begin{pmatrix} \mathbf{0}_{j+1,2} \\ I_2 \end{pmatrix}\right). \tag{27}$$

And the initial value can be chosen as $\Theta_1 = \theta$, $\Delta_\theta\Theta_1 = I_2$.

2) RMSPEM: Let's now reformulate the optimization procedure of MSPEM into a recursive form so that each time when a new temperature measurement and power consumption value are recorded, the prediction model could be updated accordingly.

Suppose the temperature measurements and the power consumption data are recorded sequentially. The Nth tuple contains the measurement of temperature increase and the power consumption at Nth discrete time instant.

Let

$$R_j^N = \Phi_j^T\Phi_j = \sum_{s=1}^{N-j}\phi_j(s)\phi_j(s)^T, \tag{28}$$

and

$$K_j^N = {Y_N^j}^T\Phi_j = \sum_{s=1}^{N-j}\Delta T(s+j)\phi_j(s)^T. \tag{29}$$

R_j^N can be calculated recursively as

$$R_j^N = R_j^{N-1} + \phi_j(N-j)\phi_j(N-j)^T. \tag{30}$$

On the right hand side of (30), R_j^{N-1} can be determined with all the data up to $N-1$th discrete time instant, and $\phi_j(N-j)$ contains all the data up to Nth discrete time instant.

Similarly, K_j^N can be calculated recursively as

$$K_j^N = K_j^{N-1} + \Delta T(N)\phi_j(N-j)^T. \tag{31}$$

In which, K_j^{N-1} can be determined with all the data up to $N-1$th discrete time instant. $\Delta T(N)$ is the measurement available at Nth discrete time instant. $\phi_j(N-j)$ requires the data up to Nth discrete time instant.

Furthermore, given the saved R_j^{N-1} and K_j^{N-1}, to calculate R_j^N and K_j^N, only the temperature measurements from time $N-k$ to time N and the power consumption values from time $N-k$ to $N-1$ are needed. In practice, the temperature measurements and power consumption values could be saved in a FIFO queue of length $k+1$ and k respectively.

What is to be noticed is that, for $j = 1, \ldots, k$, the corresponding R matrix and K matrix have to be saved seperately. Each time, when a new tuple is recorded, R matrix and K matrix are updated using (30) and (31) for each j.

By using the recursive update of (30) and (31), the estimation of $\frac{\partial^2}{\partial\theta^2}J_P^N(s)$ and $\nabla_\theta J_P^N(s)$ can be reformulated as:

$$\frac{\partial^2}{\partial\theta^2}J_P^N(s) \approx \frac{2}{N-s}\nabla_\theta\Theta_s(\theta)R_s^N\nabla_\theta\Theta_s(\theta)^T, \tag{32}$$

$$\nabla_\theta J_P^N(s) = \frac{2}{N-s}\nabla_\theta\Theta_s(\theta)(R_s^N\Theta_s(\theta) - K_s^N). \tag{33}$$

Again, Θ_j and its gradient $\nabla_\theta\Theta_j$ can be updated using (24) and (25).

The algorithm is summarized in Algorithm 1.

Algorithm 1 RMSPEM method

Require: Previously obtained parameter vector θ_{pre}, Input Queue U, Output Queue Y, R, K
$\theta = \theta_{pre}$;
$\Theta = \theta_{pre}$;
$\nabla_\theta\Theta = I_2$;
Set μ; {Set learning rate}
$Q_N = 0$;
$P_N = 0$;
for j=1:k **do**
  Initialize ϕ_j; {Using input-output data from U and Y}
  $R[j] = R[j] + \phi_j\phi_j^T$;
  $K[j] = K[j] + \phi_j Y[end]$;
  $Q_N = \frac{j-1}{j}Q_N + \frac{1}{j}\nabla_\theta\Theta R[j]\nabla_\theta\Theta^T$;
  $P_N = \frac{j-1}{j}P_N + \frac{1}{j}\nabla_\theta\Theta(K[j] - R[j]\Theta))$;
  $\theta = \theta + \mu {Q_N}^{-1}P_N$; {Parameter updating}
  Calculate W_j, H_j;
  $\nabla_\theta\Theta = \nabla_\theta\Theta W_j^T + ([1, \mathbf{0}_{1\times j}]\Theta)H_j^T$;
  $\Theta = W_j\Theta$;
end for

3) Preprocessing: Note that the Newton method is only valid if matrix Q_N^j is nonsingular, $j = 1, \ldots, k$. This can be interpreted as a generalized identifiability condition.

For matrix Q_N^j to be full rank, it requires that

(a) $\nabla_\theta\Theta_j$ is full column rank (FCR),
(b) R_j^N is non-singular.

The first condition is investigated by Farian *et al.* [29]. As for the second condition, this requires the system input meets the excitation condition. This means, in practice, a preprocessing using the batch MSPEM is necessary before implementing the recursive MSPEM. The batch MSPEM is used to obtain the nonsingular matrix R_j^N for each j and the matrix K_j^N. These matrices along with the initial estimate of model parameter are then passed to the recursive MSPEM for online identification. The mechanism of the proposed method is shown in Fig 3.

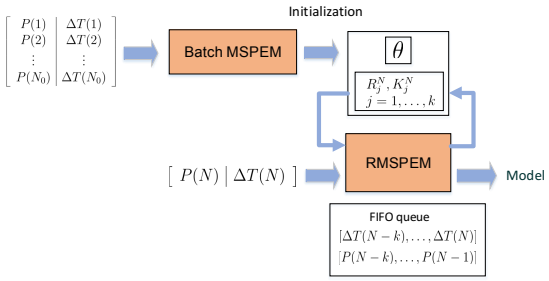


Fig. 3. Illustration of the proposed method

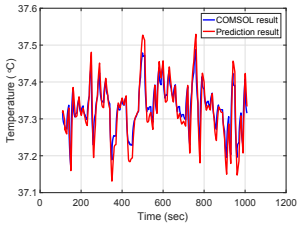
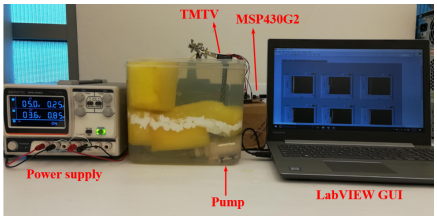
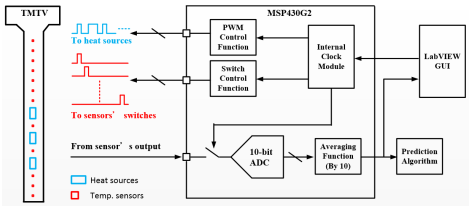


Fig. 4. Comparison between results of the developed prediction model and that of the COMSOL simulation model



(a)



(b)

(c)

Fig. 5. (a)The developed hardware testing system. (b) Hardware diagram. (c) The developed TMTV system. [31]

4) Algorithm validation: To demonstrate the effectiveness of the proposed thermal modeling techniques, the prediction model (6) with the RMSPEM updating procedure is compared with the COMSOL model described in Section III-A. A probe for measuring the temperature is placed at $(x,\ y,\ z) = (0,\ 0,\ 0.042)$, which is the position below the heat source. The COMSOL simulation is conducted for 1000 seconds. The power dissipation value of the UEA at each second is generated randomly following a Gaussian distribution in order to satisfy the identifiability condition. The generated power dissipation values are then constrained within $[0, 0.02]$ mW. The temperature measurements are recorded and converted into the temperature increase values with respect to body temperature, then stored along with the generated power dissipation value at the same time instant.

For the RMSPEM, the prediction window is set to be 10 steps with step size being 1 second. The first 100 seconds is used for the preprocessing. After the input and output queue are filled up at 110 seconds, the RMSPEM updates the parameters of the simplified thermal model according to the temperature increase values obtained by COMSOL. Then the updated model is used to predict the temperatures at 10 seconds later via the k-step ahead predictor (8). This prediction is compared with the results obtained from COMSOL. The comparison results are shown in Figure 4.

This comparison indicates that the thermal dynamics can be predicted by the simplified linear model and RMSPEM with relatively high accuracy. The Mean Square Error is about 8.04×10^{-4} °C. Thus we use this thermal modeling method to generate an accurate thermal model for the MPC.

Moreover, P in (6) represents the power consumption of the implantable devices, instead of the power dissipation. Generally, only part of the consumed power is dissipated as heat. This ratio can be determined by many factors. In practical applications, it may be difficult to track how much power is dissipated. One advantage of the proposed method is that this ratio can be learned adaptively according to the temperature feedback. Therefore, we can use the power consumption in (6) and regard the ratio as being already represented by coefficient b. Furthermore, the power consumption of embedded system is often controlled by adjusting the working mode of the device. Therefore, we can even use the controllable working mode to replace the power consumption P in (6) and the relation between the working mode and the power consumption can be learned online during the operation of the device. In general, by employing the simplified linear model and RMSPEM, more robust temperature control can be achieved, therefore enabling safe operation of implantable devices.

A hardware testing system [31] is also built to emulate the thermal effect of neural prosthesis. The system uses a custom designed temperature monitoring and management test vehicle (TMTV) with heat sources and temperature sensors to emulate the implanted electronics and a water circulation system to emulate the blood perfusion effect. A TI MSP430G2 board acts as the middleware between the TMTV and PC. It controls the operation of TMTV and sends the temperature measurements to PC, which is then processed by the LabView front end. Figure 5 demonstrates the developed hardware testing system.

We use this testing system to evaluate the prediction accuracy of the simplified thermal model. More specifically, we generate 2000 staircase PWM signals and apply the PWM

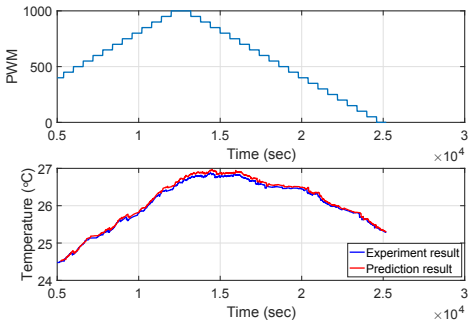


Fig. 6. Experiment result of the simplified thermal model.

signals to the heat sources on TMTV with a step size of 10 seconds. The first 500 temperature measurements are used for preprocessing and the rest are compared with the temperature predicted 10 steps ahead by the simplified thermal model. The comparison results are shown in Figure 6. It is shown that when the input PWM signal has a similar form of what produced by the MPC, the simplified thermal model is able to capture the temperature variation.

IV. Thermal Management with MPC

Based on the thermal model developed in the previous section, system state space model can be built according to the application requirements. The system state variables often include all the factors that are related to the system performance. For example, as in [16], the system state variables can include the temperature increase with respect to body temperature and the energy level of the storage device. In this paper, we assume that the battery is adopted as the storage device and the model in [33] is modified to obtain the battery state space model.

Based on [33], let the battery state variables be $x = x_1$ and $y = x_0 + x_1$, the discrete-time state space equation can be modified as

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} 1-\lambda_1\Delta t & 0 \\ -\lambda_1\Delta t & 1 \end{pmatrix} \begin{pmatrix} x(t-1) \\ y(t-1) \end{pmatrix} + \begin{pmatrix} \frac{2}{9} \\ \frac{3}{\alpha} \end{pmatrix} i(t-1), \tag{34}$$

in which $i(t-1) = \frac{P(t-1)}{V_t(t-1)}$ represents the current input, V_t is the terminal voltage of the battery and P represents the output power. $y(t)$ is defined to be 0 when the battery is at full capacity and 1 when the battery is dead. In practice, the terminal voltage of the battery can be measured using an ADC pin of the microprocessor. The reason for this modification is that the the extended quadratic programming problem with the state variables in the whole horizon can be solved very efficiently according to [34].

The control input of the system are the power consumption of the UEA and the microcontroller, and the received power of wireless interface. They are represented as P_{uea}, P_{mcu}, and P_{coil} respectively. By adjusting the control input, the state of charge of the battery can be controlled. From the thermal perspective, the power consumption of UEA is the only factor directly related to the heating of the implantable device as shown in the COMSOL model in Figure 2. In practical applications, the control input can be chosen as the adaptable system parameters that have direct relation to the power consumption of the components.

According to the system behavior predicted by the prediction model, the online controller computes the future control input that optimizes the system behavior in the future. In order to keep the control problem computationally tractable, the optimization is conducted for a finite horizon, leading to the concept of receding horizon optimization.

The objective function of neural prosthesis operation can be defined based on the application requirements. For example, the objective can be maintaining an optimal energy level r. We argue that, if the energy level is maintained at the optimal value, it achieves both continuous operation and maximum task performance, since the device has consumed all the energy received from the outside source. In general, it is not practical for a neural prosthesis to adjust its power budget to exactly maintain the optimal level during the operation. But the difference between the actual battery level and the optimal level can be minimized by adding $\sum_{t=1}^{N}(y(t)-r)^2$ into the objective function. By minimizing the objective function, the neural prosthesis operates around the optimal level.

Furthermore, in a system composed of several functioning units, the performance of a specific component can be maximized by adding its corresponding power consumption into the objective function. Assume the optimal working mode of a component has a power consumption of P_r and its power consumption at time t is $P(t)$, then $\sum_{t=1}^{N}(P(t)-P_r)^2$ can be added to the objective function.

One of the most important advantage of MPC is its ability to handle system constraints in a straight forward manner. The system state space model is specified as the constraint of the optimization problem. So is the thermal safety constraint. A maximum allowable temperature ΔT_{max} can be added as a constraint to prevent the hot spot from heating too much. The power consumption of the microprocessor and the UEA must not exceed the maximum allowable value. The received power must be less than a certain threshold, which ensures there is no safety issues.

The optimization problem is solved over a horizon of future k time steps, which starts at current time t. But only the control input corresponding to t in the solution sequence is actually applied to the system. The remaining control inputs are discarded. At the next time step, a new optimization problem based on the new measured data is solved over a shifted prediction horizon. At every time step, the control input applied to the system depends on the most recent measurements.

Since the thermal model and the battery model are all linear models, and all the constraints are linear. The resulting extended optimization problem with state variables in the whole horizon is a quadratic programming (QP) problem, which can be solved efficiently using methods like [32]. Moreover, the developed prediction model shows little correlation between different states and the objective function is composed of the terms that are only related to one state variable or control input, therefore methods proposed in [34] can be used to greatly reduce the computational cost, and the optimization problem

can be solved in real time.

To prevent the effect of external disturbance and model deviation caused by the simplified thermal model, the linear model (6) is updated at every time instance. The RMSPEM algorithm calculates the optimal coefficient for (6) given the temperature measurements and the control inputs in previous time steps.

However, the model update doesn't have to be executed at each time step. A possible way to reduce the computational cost is to call the RMSPEM only when the error of the predicted temperature and the actual measurement becomes larger than a certain threshold.

V. Simulation Studies

The developed thermal management method is investigated with two simulation studies of the neural prosthesis. The COMSOL model developed in Section III-A was used to emulate the thermal effect of the neural prosthesis. It was then used in conjunction with Matlab to perform real-time simulation. We implemented MPC for exemplary case studies using the CVX, a package for specifying and solving convex programs [35].

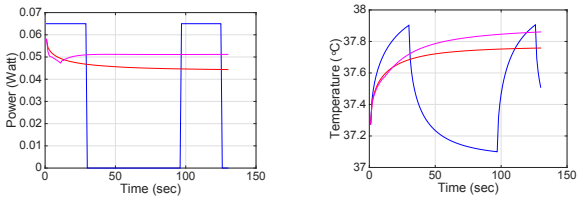
A. Adaptive Thermal Management

The first simulation study was designed to emulate the thermal effect of a wireless neural prosthesis that doesn't have an energy storage device. The inductive coil and microprocessor are planted on the surface of the scalp. Only the UEA is implanted in the brain tissue, which is connected to other parts of the system through wire. The prosthesis stimulates different layers of the brain using the UEA, whose performance is proportional to its power consumption during normal operation. The UEA has a power consumption within the range between 0.04 W and 0.065 W during active operation, and its power consumption is assumed to be 0 W while being put into the sleep mode. A ratio of 65% of the consumed power is assumed to dissipate into the surrounding tissue and turned into heat. The maximum allowable temperature of the surrounding tissue is set to be 0.9 °C above the body temperature.

The following MPC is designed to maximize the device performance and maintain safe operation via regulating the power consumption of the UEA. In this example, the thermal safety is achieved by adding a constraint that limits the maximum allowable temperature. The performance optimization is achieved by maximizing the power consumption of UEA. The quadratic program of the MPC can be formulated as

$$\begin{aligned} \min \quad & \sum_{t=1}^{k} (P_{max\_uea} - P_{uea})^2 \\ \text{s. t.} \quad & \Delta T(t) = a\Delta T(t-1) + bP_{uea}(t) \\ & \Delta T(t) \leq \Delta T_{max} \\ & 0 \leq P_{uea} \leq P_{max\_uea} \end{aligned} \tag{35}$$

In which, P_{uea} denotes the power consumption of the UEA and P_{max_uea} is its maximum value. $\Delta T(t)$ represents the temperature increase with respect to body temperature. ΔT_{max} is the maximum allowable temperature increase. The receding horizon for the MPC is chosen to be 10 time steps. At each time step, after the quadratic program is solved, P_{uea} is set to be 0 if the calculated value is less than 0.04 W, since the minimum operational power consumption of UEA is 0.04 W.



(a) (b)

Fig. 7. (a) UEA power consumption of all three cases. (b) Temperature measurement of all three cases. The blue curve represents the Case 1. The red curve represents the Case 2. The magenta curve represents the Case 3.

To investigate the performance of the developed method, three different cases were designed. The first case uses a method adopted in a previously published system [13] to control the operation of the UEA. With the method, the UEA is set to the maximum working mode at the beginning and turned to the sleep mode once the temperature reaches 37.9 °C. After the temperature drops below 37.1 °C, the UEA is then set back to the maximum working mode. It repeats this process during the operation, so the UEA is either on or off according to the temperature measurements. The second case uses the batch MSPEM to identify the thermal model in advance and then uses the MPC to calculate the desired power consumption at run time. The third case combines the batch MSPEM preprocessing with the RMSPEM and adjusts the thermal model parameter based on the real-time temperature measurements. The RMSPEM uses a prediction horizon of 10 time steps, which is the same as that of the receding optimization horizon of MPC. Results of these cases are shown in Figure 7, which indicates all three cases are able to maintain safe operation.

The three cases are also evaluated in terms of operation time and the square of difference between the calculated power consumption and the maximum power consumption. The operation time measures the total time when the power consumption of the UEA is greater than 0.04 W. The operation time of the three cases are 58 seconds, 130 seconds, and 130 seconds respectively. The square of difference with respect to the maximum power consumption are 0.3042, 0.40475, and 0.0259 respectively. Moreover, to quantify the complexity of the RMSPEM, the total execution time of RMSPEM for 130 seconds is recorded, which is 0.2593 seconds on a desktop computer with an i7-3770 CPU and 16 GB of RAM.

It is shown that by dynamically adjusting the working status of the UEA, the operation time can be significantly prolonged. The cases of MPC with real-time updating and the one with only batch pre-processing all maintain operation during the entire simulation time. Case 1 that adopted the On/Off approach has an operation time of less than half of the total time. The overall device performance is improved with

the adaptive management methods by exploring the acceptable operational working status. Moreover, MPC with RMSPEM updating (Case 3) shows better performance compared to MPC with batch MSPEM pre-processing (Case 2). This is because the accuracy of the prediction model is significantly improved by adjusting model parameter based on the real-time measurements. The RMSPEM provides feedback information for the MPC and therefore achieves the closed-loop control, which helps to prevent model deviation through time.

B. Joint Thermal and Power Management

The second simulation study adds the energy storage device into the prediction model and studies the joint power and thermal management of a battery powered neural prosthesis. The neural prosthesis is assumed to have the same layout as in the previous study. The battery has a charging efficiency of 0.8. The optimal energy level of the battery is assumed to be 0.5 for the task. The controller tracks the optimal energy level of the battery by adjusting the received power of the inductive coil P_{coil}, the power consumption of the main board P_{main} and the power consumption of the UEA P_{uea}. Moreover, noise is added to the measured output. The noise of battery level and temperature are generated with autoregressive autoregressive with exogenous (ARARX) [36] noise model.

The goal of the joint thermal and power management is to achieve sustainable operation, performance maximization, and thermal safety. The sustainable operation and performance maximization are guaranteed by designing the objective function of the quadratic programming. The thermal safety during operation is ensured by adding a constraint that limits the maximum temperature. The designed quadratic programming can be formulated as

$$\begin{aligned} \min \quad & \sum_{t=1}^{k}[(y(t)-r)^2+\gamma(P_{max\_uea}-P_{uea})^2] \\ \text{s. t.} \quad & Battery\ equation\ (34) \\ & Temperature\ equation\ (6) \\ & 0 \le y(t) \le 1 \\ & \Delta T(t) \le \Delta T_{max} \\ & P_{min\_main} \le P_{main} \le P_{max\_main} \\ & P_{min\_coil} \le P_{coil} \le P_{max\_coil} \\ & 0 \le P_{uea} \le P_{max\_uea} \end{aligned} \tag{36}$$

In which, r represents the optimal energy level and γ is the relative weight.

The receding horizon is chosen to be 10 time steps. As previously stated, P_{uea} is set to be 0 when the solution is less than 0.04 W. The battery terminal voltage is assumed to be a constant value 3.3 V given the fact that it varies little when there is no significant change in the state of charge. The battery parameter α is set to be 2.8, and λ is set to be 0.04.

The results of the joint thermal and power management are shown in Figure 8. The power consumption of the mainboard, the UEA and the received power from inductive coil are all within the specified range. The power consumption of the UEA is maximized to improve the performance of neural prosthesis. The energy level of the battery tracks the reference value of 0.5 despite the noise. The temperature of the surrounding tissue is below the safe threshold 37.9 °C during the entire simulation time, so the thermal safety constraint is satisfied.

VI. Conclusion

With implantable devices becoming more and more powerful, the heating caused by its operation has drawn growing concern. It is reported that for a neural implant a temperature increase of more than 1 °C could damage the surrounding brain tissue. In this paper, an adaptive thermal management method is developed to control the heating effect of the implantable device, with a neural prosthesis as an example. The developed method aims to fill in the gap of real time thermal management as opposed to the methods adopted in previous literatures that limit the working status of stimulating electrodes in the design phase. More specifically, a simplified thermal model is proposed to support the real time control and an online model parameter estimation method (RMSPEM) is developed to update the model parameters based on the real time temperature measurements. This modeling method is validated with a COMSOL Multiphysics thermal model and an *in vitro* experimental system. Based on the thermal model, the model predictive control is introduced to solve the thermal management problem. Simulation results based on the COMSOL model indicate that the developed method achieves longer operation time while maintaining safe operating temperature. Furthermore, based on the application requirements, the developed framework can be extended to incorporate the model of energy storage device and an example of joint thermal and power management is demonstrated.

Acknowledgment

This research was supported in part by the National Science Foundation under Grants ECCS-1711447 and CNS-1253390. The authors would also like to thank Dr. Ghovanloo, Yen-Pang Lai, and Wen Sun for providing help to build the hardware testing system.

References

[1] J. C. LaManna, K. A. McCracken, M. Patil, and O. J. Prohaska, "Stimulus-activated changes in brain tissue temperature in the anesthetized rat," *Metabolic brain disease*, vol.4, no.4, pp.225-237, 1989.

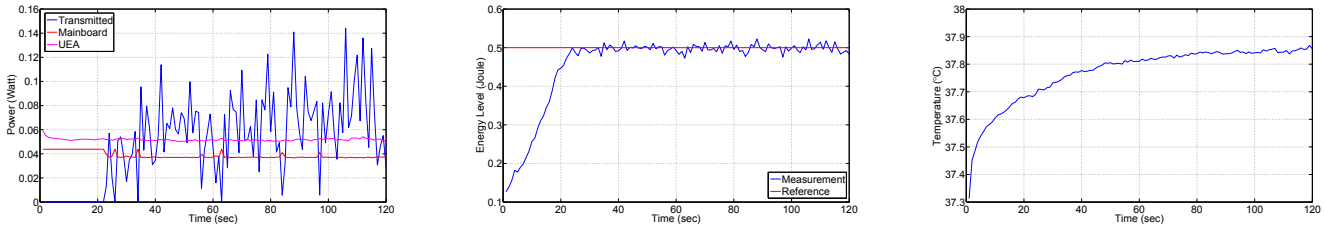
[2] G. Lazzi, "Thermal effects of bioimplants," *IEEE Engineering in Medicine and Biology Magazine*, vol.24, no.5, pp.75-81, 2005.

[3] H. Pennes, "Analysis of tissue and arterial blood temperatures in the resting human forearm," *Journal of applied physiology* vol.1, no.2, pp.93-122, 1948.

[4] S. Kim, P. Tathireddy, R. A. Normann, and F. Solzbacher, "Thermal impact of an active 3-D microelectrode array implanted in the brain," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol.15, no.4, pp.493-501, 2007.

[5] R. Henry, and V. Guruviah, "Review of thermal management of an LED for brain implants," *Optical And Microwave Technologies*. Springer, Singapore, 125-137, 2018.

[6] E. Katz, "Implantable Bioelectronics - Editorial Introduction," *Implantable Bioelectronics*, 1-5, 2014.



(a) (b) (c)

Fig. 8. (a) The power consumption of each component. (b) The battery energy level. (c) Temperature of the surrounding tissue.

[7] N. L. Opie, A. N. Burkitt, H. Meffin, and D. B. Grayden, "Heating of the eye by a retinal prosthesis: modeling, cadaver and in vivo study," *IEEE Transactions on Biomedical Engineering*, vol.59, no.2, pp.339-345, 2012.

[8] K. Silay, C. Dehollain, and M. Declercq, "Numerical thermal analysis of a wireless cortical implant with two-body packaging," *BioNanoScience*, vol.1, no.3, pp.78-88, 2011.

[9] C. H. W. Oey, and S. Moh, "A survey on temperature-aware routing protocols in wireless body sensor networks," *Sensors*, vol.13, no.8, pp.9860-9877, 2013.

[10] R. Kamal, and C. S. Hong, "Fault tolerant virtual backbone for minimum temperature in in vivo sensor networks," *2012 IEEE International Conference on Communications (ICC)*. 2012.

[11] M. Monowar, and F. Bajaber, "On designing thermal-aware localized QoS routing protocol for in-vivo sensor nodes in wireless body area networks," *Sensors*, vol.15, no.6, pp.14016-14044, 2015.

[12] Q. Tang, N. Tummala, S. Kumar S. Gupta, and L. Schwiebert, "Communication scheduling to minimize thermal effects of implanted biosensor networks in homogeneous tissue," *IEEE Transactions on Biomedical Engineering*, vol.52, no.7, pp.1285-1294, 2005.

[13] C. T. Wentz, J. G. Bernstein, P. Monahan, A. Guerra, A. Rodriguez, and E. S. Boyden, "A wirelessly powered and controlled device for optical neural control of freely-behaving animals," *Journal of neural engineering*, vol.8, no.4, 046021, 2011.

[14] Y. Luo, M. Dahmardeh, and K. Takahata, "Biocompatible circuit-breaker chip for thermal management of biomedical microsystems," *Journal of Micromechanics and Microengineering*, vol.25, no.5, p.055002, 2015.

[15] C. M. Krishna, and I. Koren, "Thermal-aware management techniques for cyber-physical systems," *Sustainable Computing: Informatics and Systems*, no.15, pp.39-51, 2017.

[16] R. Chai, Y. Zhang, and M. Ghovanloo, "Joint power and thermal management for implantable devices," *2015 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, pp.1-4, 2015.

[17] F. Zhang, and Z. Shi, "Optimal and adaptive battery discharge strategies for Cyber-Physical Systems," *IEEE Conference on Decision and Control*, pp.6232-6237, 2009.

[18] R. Chai, and Y. Zhang, "A Practical Supercapacitor Model for Power Management in Wireless Sensor Nodes," *IEEE Transactions on Power Electronics*, vol.30, no.12, pp. 6720-6730, 2015.

[19] P. M. Hettiarachchi, N. Fisher, M. Ahmed, L. Y. Wang, S. Wang, and W. Shi, "A design and analysis framework for thermal-resilient hard real-time systems," *ACM Transactions on Embedded Computing Systems (TECS)*, vol.13, no.5s, pp.146, 2014.

[20] S. Ethier, and M. Sawan, "Exponential current pulse generation for efficient very high-impedance multisite stimulation," *IEEE transactions on biomedical circuits and systems*, vol.5, no.1, pp.30-38, 2011.

[21] S. C. DeMarco, G. Lazzi, W. Liu, J. D. Weiland, and M. S. Humayun, "Computed SAR and thermal elevation in a 0.25-mm 2-D model of the human eye and head in response to an implanted retinal stimulator-Part I: Models and methods," *IEEE Transactions on Antennas and Propagation*, vol.51, no.9, pp.2274-2285, 2003.

[22] J. L. Dillenseger, and S. Esneault, "Fast FFT-based bioheat transfer equation computation," *Computers in Biology and Medicine*, vol.40, no.2, pp.119-123, 2010.

[23] G. Carluccio, D. Erricolo, S. Oh, and C. M. Collins, "An approach to rapid calculation of temperature change in tissue using spatial filters to approximate effects of thermal conduction," *IEEE Transactions on Biomedical Engineering* vol.60, no.6, pp.1735-1741, 2013.

[24] V. Singh, A. Roy, R. Castro, et al, "On the thermal elevation of a 60-electrode epiretinal prosthesis for the blind," *IEEE transactions on biomedical circuits and systems*, vol.2, no.4, pp.289-300, 2008.

[25] S. K. Das, S. T. Clegg, and T. V. Samulski, "Computational techniques for fast hyperthermia temperature optimization," *Medical physics*, vol.26, no.2, pp.319-328, 1999.

[26] W. M. Reichert, "Indwelling neural implants: strategies for contending with the in vivo environment," *CRC Press*, 2007.

[27] M. Pedram, and S. Nazarian, "Thermal modeling, analysis, and management in VLSI circuits: Principles and methods," *Proceedings of the IEEE*, vol.94, no.8, pp.1487-1501. 2006.

[28] R. A. Romano, A. S. Potts, and C. Garcia, "Model Predictive Control Relevant Identification," *INTECH Open Access Publisher*, 2012.

[29] M. Farina, and L. Piroddi, "Simulation error minimization identification based on multi-stage prediction," *International Journal of Adaptive Control and Signal Processing*, vol.25, no.5, pp.389-406, 2011.

[30] L. Ljung, "System Identification: Theory For the User," *PrenticeHall*, Upper Saddle River, N.J. USA, 2 edition, 1999.

[31] R. Chai, Y. Lai, W. Sun, M. Ghovanloo, and Y. Zhang, "Online Predictive Modeling for the Thermal Effect of Implantable Devices," *2018 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, 2018.

[32] Y. Lu, D. Li, Z. Xu, and Y. Xi, "Convergence analysis and digital implementation of a discrete-time neural network for model predictive control," *IEEE Transactions on Industrial Electronics*, vol.61, no.12, pp.7035-7045, 2014.

[33] F. Zhang, Z. Shi, and W. Wolf, "A dynamic battery model for co-design in cyber-physical systems," *29th IEEE International Conference on Distributed Computing Systems Workshops*, pp.51-56, 2009.

[34] Y. Wang, and S. Boyd, "Fast model predictive control using online optimization," *IEEE Transactions on Control Systems Technology*, vol.18, no.2, pp.267-278, 2010.

[35] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.0 beta," http://cvxr.com/cvx, September 2013.

[36] O. Nelles, "Nonlinear system identification: from classical approaches to neural networks and fuzzy models," *Springer Science & Business Media*, 2013.