HAL open science

More Safe Optimal Input Signals for Parameter Estimation of Linear Systems Described by ODE

Ewaryst Rafajlowicz, Wojciech Rafajlowicz

To cite this version:

Ewaryst Rafajlowicz, Wojciech Rafajlowicz. More Safe Optimal Input Signals for Parameter Estimation of Linear Systems Described by ODE. 26th Conference on System Modeling and Optimization (CSMO), Sep 2013, Klagenfurt, Austria. pp.267-277, 10.1007/978-3-662-45504-3_26 . hal-01286434

HAL Id: hal-01286434

https://inria.hal.science/hal-01286434v1

Submitted on 10 Mar 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

More safe optimal input signals for parameter estimation of linear systems described by ODE*

Ewaryst Rafajłowicz and Wojciech Rafajłowicz

Institute of Computer Engineering Control and Robotics,
ewaryst.rafajlowicz@pwr.wroc.pl

Abstract. Our starting point is the ascertainment that D-optimal input signals recently considered by the same authors [12] can be too dangerous for applying them to real life system identification. The reason is that they grow too fast in time. In order to obtain more safe input signals, but still leading to a good estimation accuracy of parameter estimates, we propose a quality criterion that is a mixture of D-optimality and a penalty for too fast growth of input signals in time.
Our derivations are parallel to those in [12] up to a certain point only, since we obtain different optimality conditions in the form of an integral equation. We also briefly discuss a numerical algorithm for its solution.

Keywords: parameter estimation, optimal input signal, D-optimality, optimal experiment design, integral equations

1 Introduction

Our aim is to provide optimality conditions for optimal input signals that are D-optimal for parameter estimation in linear systems described by ordinary differential equations. In opposite to our paper [12], we put emphasis not only on parameter estimation accuracy, but also on safety of input signals. This leads to different optimality conditions than those obtained in [12]. They are derived from the variational approach and they are expressed in a convenient form of integral equations.

By the lack of space, we do not discuss the selection of input signal when a feedback is present. As it was demonstrated in [6] for systems described by ODE's – the presence of feedback can be beneficial. The result presented here can be generalized to the case with a feedback, using the results on output sensitivity to parameters for systems with feedback (see [13], but this is outside the scope of this paper.

As is known, problems of selecting input signal for parameter estimation in linear systems have been considered in seventies of 20-th century. The results were summarized in [9] and two monographs [5] and [19]. They have been concentrated mainly on the so-called frequency domain approach. This approach leads

* The paper was supported by the National Council for Research of Polish Government under grant 2012/07/B/ST7/01216, internal code 350914 of Wrocław University of Technology.

to a beautiful theory, which runs in paralel to the optimal experiment design theory (see [1]). The main message of this approach is that optimal input signal are linear combinations of a finite number of sinusoidal waves with precisely selected frequencies. We stress that sums of sinusoids can have unexpectedly large amplitudes, which can be dangerous for an identified systems. Notice that the frequency domain approach requires an infinite observation horizon. Here we assume a finite observation horizon, which leads to quite different results.

Research in this area was stalled for nearly 20 years. Recently, since the beginning of 21-st century, the interest of researchers has rapidly grown up (see [3, 4, 7, 8, 16, 17] for selected recent contributions).

Related problems of algorithms for estimating parameters in PDE's and selecting allocation of sensors are also not discussed (we refer the reader to [2, 15] and [10]).

2 Problem statement

System description. Consider a system described by ODE

$$\frac{d^r\, y(t)}{d\, t^r} + a_{r-1}\, \frac{d^{r-1}\, y(t)}{d\, t^{r-1}} + \ldots + a_0\, y(t) = a_r\, u(t), \quad t \in (0,\, T] \tag{1}$$

with zero initial conditions, where $y(t)$ is the output, $u(t)$ is the input signal.

The solution $y(t;\, \bar{a})$ of (1) depends on the vector $\bar{a} = [a_0,\, a_1, \ldots,\, a_r]^{tr}$ of unknown parameters. The impulse response (response to the Dirac Delta or Green's function) of ODE (1) will be denoted later by $g(t; \bar{a})$. Notice that $g(t; \bar{a}) = 0$, $t < 0$ and $y(t;\, \bar{a}) = \int_0^T g(t-\tau; \bar{a})\, u(\tau)\, d\tau$. Remark that it is not necessary to assume differentiability of y w.r.t. $\bar{a}$, because solutions of linear ODE's are known to be analytical functions of ODE parameters.

As is well known, differential sensitivity $y(t;\, \bar{a})$ to parameter changes, can be expressed and calculated in a number of ways (see [13] for a brief summary). For our purposes it is convenient to express it through $r \times 1$ vector of sensitivities $\bar{k}(t; \bar{a}) \stackrel{def}{=} \nabla_a g(t; \bar{a})$. Then,

$$\nabla_a y(t;\, \bar{a}) = \int_0^T \bar{k}(t-\tau; \bar{a})\, u(\tau)\, d\tau. \tag{2}$$

Observations and the estimation accuracy. Available observations have the form:

$$\Upsilon(t) = y(t;\, \bar{a}) + \varepsilon(t), \quad t \in [0,\, T],$$

where $\varepsilon(t)$ is zero mean, finite variance, uncorrelated, Gaussian white noise, more precisely, $\varepsilon(t)$ is implicitly defined by $d\, W(t) = \varepsilon(t)\, dt$, where $W(t)$ is the Wiener process.

It can be shown (see [5]), that the Fisher information matrix $\mathbf{M}_T(u)$ for estimating $\bar{a}$ has the form:

$$\mathbf{M}_T(u) = \int_0^T \nabla_a y(t; \bar{a})\, (\nabla_a y(t; \bar{a}))^{tr}\, dt,$$

where $\nabla_a y(t;\bar{a})$ depends on $u(.)$ through (2). From the linearity of (1) it follows that $\mathbf{M}_T(u)$ can be expressed as follows

$$\mathbf{M}_T(u) = \int_0^T \int_0^T H(\tau,\, \nu;\, \bar{a})\, u(\tau)\, u(\nu)\, d\tau\, d\nu,$$

where $H(\tau,\, \nu;\, \bar{a}) \stackrel{def}{=} \int_0^T \bar{k}(t-\tau;\bar{a})\, \bar{k}^{tr}(t-\nu;\bar{a})\, dt$.

From the Cramer-Rao inequality we know that for any estimator $\tilde{a}$ of $\bar{a}$ we have:

$$cov(\tilde{a}) \geq [\mathbf{M}_T(u)]^{-1}. \tag{3}$$

When observation errors are Gaussian and the minimum least squares estimator is used, then the equality in (3) is asymptotically attained. Thus, it is meaningful to minimize interpretable functions of $\mathbf{M}_T(u)$, e.g., $\min Det[\mathbf{M}_T(u)]^{-1}$ w.r.t. $u(.)$, under certain constraints on $u(.)$.

Problem formulation. Define

$$\mathcal{U}_0 = \left\{ u : \int_0^T u^2(t)\, dt \leq e_1,\ Det[\mathbf{M}_T(u)] > 0 \right\},$$

where $e_1 > 0$ is a level of available (or admissible) energy of input signals. Let $C_0(0,\, T)$ denote the space of all functions that are continuous in $[0,\, T]$.

In [12] the following problem has been considered. Find $u^* \in \mathcal{U}_0 \cap C_0(0,\, T)$ for which $\min_{u\in\mathcal{U}_0} Det\left[\mathbf{M}_T^{-1}(u)\right]$, or equivalently,

$$\max_{u\in\mathcal{U}_0} \log\left[Det\left(\mathbf{M}_T(u)\right)\right] \tag{4}$$

is attained. As it was demonstrated in [12] (see also comparisons at the end of this paper), signals that are optimal in the above sense can be even more dangerous than sums of sinusoids.

For these reasons, we consider here putting additional constraints on the system output.

Typical output constraints include:

$$\int_0^T y^2(t)\, dt \leq e_2, \quad y(t) = \int_0^T g(t-\tau, \bar{a})\, u(\tau)\, d\tau, \tag{5}$$

$$\int_0^T \dot{y}^2(t)\, dt \leq e_3, \quad \dot{y}(t) = \int_0^T g'(t-\tau, \bar{a})\, u(\tau)\, d\tau, \tag{6}$$

$$\int_0^T \ddot{y}^2(t)\, dt \leq e_4, \quad \ddot{y}(t) = \int_0^T g''(t-\tau, \bar{a})\, u(\tau)\, d\tau. \tag{7}$$

Later, we consider only (5), because results for (6) and (7) can be obtained from those for (5) by formal replacement of g by g' or g'', respectively.

Define $\mathcal{U}_1 = \left\{ u : \int_0^T u^2(t)\, dt \leq e_1,\ \int_0^T y^2(t)\, dt \leq e_2,\ Det[\mathbf{M}_T(u)] > 0 \right\}$, where $e_2 > 0$ is the admissible level of energy of output $y(t)$ signal that depends on $u(.)$ as in (5).

Now, our problem reads as follows: find $u^* \in \mathcal{U}_1 \cap C_0(0, T)$ for which

$$\max_{u \in \mathcal{U}_1} \log \left[Det \left(\mathbf{M}_T(u) \right) \right] \tag{8}$$

is attained.

In fact, in most cases only one of the constraints $\int_0^T u^2(t)\, dt \leq e_1$ and $\int_0^T y^2(t)\, dt \leq e_2$ is active, depending on the system properties and on values of e_1 and e_2. In practice, the following way of obtaining an input signal that is informative and simultaneously safe for the identified system can be proposed.

Upper level algorithm

1. For a given value of available input energy $e_1 > 0$ solve problem (4). Denote its solution by u_1^*, say. Set $\ell = 1$.
2. Calculate the output signal y_ℓ^*, corresponding to u_ℓ^*. If y_ℓ^* is safe for the system, then stop – u_ℓ^* is our solution. Otherwise, go to step 3.
3. Calculate $e_{trial} = \int_0^T [y_\ell^*(t)]^2\, dt$ and select e_2 less then e_{trial}, e.g., $e_2 = \theta\, e_{trial}$, where $0 < \theta < 1$.
4. Set $\ell = \ell + 1$ and solve problem (8). Denote its solution by u_ℓ^* and go to step 2. Notice that now the constraint for the output energy is active and the one on input energy is almost surely not active (see the discussion in the next section).

The curse of the lack of a priori knowledge As in optimum experiment design for nonlinear (in parameters) regression estimation (see [1, 11]), also here the optimal u^* depends on unknown $\bar{a}$. Furthermore, condition $\int_0^T [y(t)]^2\, dt \leq e_2$ also contains unknown $\bar{a}$. The ways of circumventing these difficulties have been discussed for a long time. The following ways are usually recommended (see [11]):

1. use "the nominal" parameter values for $\bar{a}$, e.g., values from previous experiments,
2. the "worst case" analysis, i.e. solve the following problem

$$\max_{u \in \mathcal{U}_1} \min_{\bar{a}} \log \left[Det \left(\mathbf{M}_T(u; \bar{a}) \right) \right],$$

   where $\mathbf{M}_T(u; \bar{a})$ is the Fisher information matrix, in which dependence on $\bar{a}$ is displayed,
3. the Bayesian approach: use prior distribution imposed on $\bar{a}$ and average $\mathbf{M}_T(u; \bar{a})$ with respect to it,
4. apply the adaptive approach of subsequent estimation and planning stages.

Later we use the "nominal" parameter values $\bar{a}$. Thus, we obtain locally optimal input signals, which are optimal in a vicinity of nominal parameters. We underline that the results are relevant also to the Bayesian and adaptive approaches in the sense that it is easy to obtain optimality conditions in a way that is similar to the one presented bellow.

3 Optimality conditions

Define the Lagrange function

$$L(u,\gamma) = \log\left[Det\left(\mathbf{M}_T(u)\right)\right] - \gamma_1\left(\int_0^T u^2(t)\,dt - e_1\right) - \gamma_2\left(\int_0^T y^2(t)\,dt - e_2\right),$$

where γ_1, γ_2 are the Lagrange multipliers and $\gamma = [\gamma_1,\, \gamma_2]^{tr}$.

Let $u^* \in \mathcal{U}_1 \cap C_0(0,\, T)$ be a solution of (8) problem and let $u_\epsilon(t) = u^*(t) + \epsilon\, f(t)$, where $f \in C_0(0,\, T)$ is arbitrary. Then, for the Gateaux differential of L we obtain

$$\frac{\partial L(u_\epsilon,\, \gamma)}{\partial\, \epsilon}\Big|_{\epsilon=0} = 2\int_0^T f(\nu)\left[\int_0^T \widehat{ker}(\tau,\, \nu,\, u^*)\, u^*(\tau)\, d\tau - \gamma_1\, u^*(\nu)\right] d\nu\,,$$

where, for $u \in \mathcal{U}_0$, we define:
a) $\widehat{ker}(\tau,\, \nu,\, u) \stackrel{def}{=} ker(\tau,\, \nu,\, u) - \gamma_2\, G(\tau,\, \nu)$,
b) $G(\tau,\, \nu) \stackrel{def}{=} \int_0^T g(t-\tau, \bar{a})\, g(t-\nu, \bar{a})\, dt$,
c) $ker(\tau,\, \nu,\, u) \stackrel{def}{=} trace\left[\mathbf{M}_T^{-1}(u)\, H(\tau,\, \nu,\, \bar{a})\right]$.
The symmetry of kernel $\widehat{ker}(\tau,\, \nu,\, u)$ was used in calculating $\frac{\partial L(u_\epsilon,\, \gamma)}{\partial\, \epsilon}\Big|_{\epsilon=0}$. Notice that $\widehat{ker}(\tau,\, \nu,\, u)$ depends also on γ_2 but this is not displayed in the notation.

If u^* is optimal, then for each $f \in C_0(0,\, T)$ we have $\frac{\partial L(u_\epsilon,\, \gamma)}{\partial\, \epsilon}\Big|_{\epsilon=0} = 0$. Then, by the fundamental lemma of the calculus of variation we obtain.

Proposition 1 *If u^* solves problem (8), then it fulfils the following integral equation*

$$\int_0^T \left[ker(\tau,\, \nu,\, u^*) - \gamma_2\, G(\tau,\, \nu)\right] u^*(\tau)\, d\tau = \gamma_1\, u^*(\nu), \quad \nu \in (0,\, T). \tag{9}$$

Notice that the constraints:
– input energy constraint $\int_0^T u^2(t)\, dt \le e_1$ and
– output energy constraint $\int_0^T y^2(t)\, dt \le e_2$
can be simultaneously active at the optimal solution u^* in a very special case. Namely, if

$$\int_0^T [u^*(t)]^2\, dt = e_1, \tag{10}$$

then we must also have

$$\int_0^T \int_0^T G(\tau,\, \nu)\, u^*(\tau)\, u^*(\nu)\, d\tau\, d\nu = e_2. \tag{11}$$

In other words, simultaneous equality of both constraints (10) and (11) is not a generic case and without loosing generality, we can consider only the cases when

only one of them[1] is active.

Case I – only input energy constraint active.
This case is exactly the one considered in our paper [12]. We provide a brief summary of the results and their extension. Notice that in this case $\gamma_2 = 0$ and we consider problem (4). For simplicity of formulas we set $e_1 = 1$.

Proposition 2 *If u^* solves problem (4), then it fulfils the following integral equation*

$$\int_0^T ker(\tau,\, \nu,\, u^*)\, u^*(\tau)\, d\tau = \gamma_1\, u^*(\nu), \quad \nu \in (0,\, T). \tag{12}$$

Furthermore, $\gamma_1 = r + 1 = dim(\bar{a})$ and this is the largest eigenvalue of the following <u>*linear*</u> *eigenvalue problem*[2]

$$\int_0^T ker(\tau,\, \nu,\, u^*)\, \phi(\tau)\, d\tau = \lambda\, \phi(\nu), \quad \nu \in (0,\, T). \tag{13}$$

Thus, the eigenfunction corresponding to the largest eigenvalue is a natural candidate for being the optimal input signal. In the case of multiple eigenvalues, one can consider all linear combinations of the eigenfunctions that correspond to the largest eigenvalue.

Proposition 3 *Condition (12) with $\gamma_1 = r + 1$ is sufficient for the optimality of u^* in problem (4).*

This result was announced in [12] under additional condition on T. It occurs that this condition can be removed. The proof is given in the Appendix.

Algorithm – Case I As one can notice, (13) is nonlinear, because of the dependence of $ker(\tau,\, \nu,\, u^*)$. The simplest algorithm that allows to circumvent this difficulty is the following:
Start with arbitrary $u_0 \in \mathcal{U}_0$, $||u_0|| = 1$ and iterate for $p = 0,\, 1, \ldots$

$$\hat{u}_{p+1}(\nu) = \frac{1}{r+1} \int_0^T ker(\tau,\, \nu,\, u_p)\, u_p(\tau)\, d\tau, \quad u_{p+1} = \hat{u}_{p+1}/||\hat{u}_{p+1}||$$

until $||u_{p+1} - u_p|| < \delta$, where $\delta > 0$ is a preselected accuracy.
The proof of convergence can be based on the fixed point theorem, but this is outside the scope of this paper.

This algorithm has been used for solving examples in the next section. It was convergent in several or at most several dozens of steps. The Nyström method with the grid step size 0.01 was used to approximate integrals.

[1] We exclude also the case that both of them are inactive, because if $u(t)$ is replaced by $\rho\, u(t)$ with $\rho > 1$, then $det[\mathbf{M}_T(\rho\, u)] > det[\mathbf{M}_T(\, u)]$.

[2] As is known [18], the eigenvalue problem for linear integral equation with nonnegative definite and symmetric kernel has the following solution: its eigenvalues are real and nonnegative, while the corresponding eigenfunctions are orthonormal in $L_2(0,\, T)$ (or can be orthonormalized, if there are multiple eigenvalues).

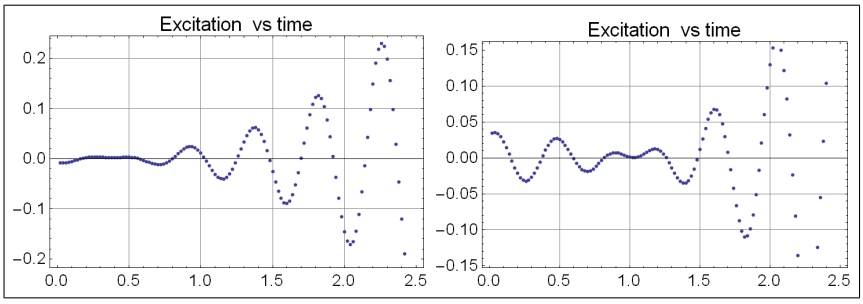

Fig. 1. Left panel – more aggressive input signal (no output constraints, right panel – more safe input signal (with output power constraint) obtained for a system described in Section 4.

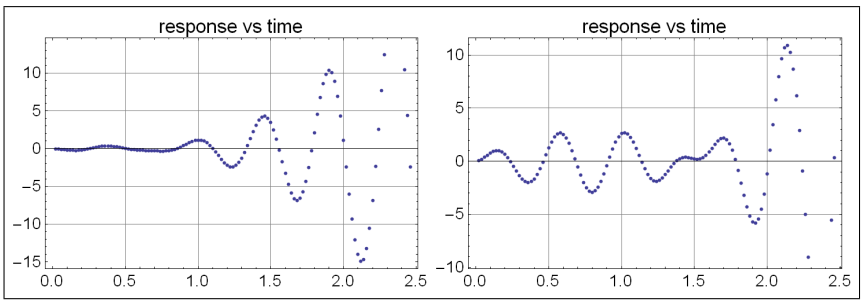

Fig. 2. Left panel – more aggressive output signal (no output constraints, right panel – more safe output signal (with output power constraint) obtained for a system described in Section 4.

Case II – only output energy constraint active.
If only the output energy constraint active then $\gamma_1 = 0$ and Proposition 1 immediately implies.

Proposition 4 *If u^* solves problem (4) with the input energy constraint inactive, then it fulfils the following integral equation for $\nu \in (0, T)$*

$$\int_0^T ker(\tau, \nu, u^*)\, u^*(\tau)\, d\tau = \gamma_2 \int_0^T G(\tau, \nu)\, u^*(\tau)\, d\tau. \tag{14}$$

Furthermore, $\gamma_2 = (r+1)/e_2$.

The last fact follows from the multiplication of the both sides of (14) by $u^*(\nu)$ and their subsequent integration. One should also observe that

$$\int_0^T \int_0^T G(\tau, \nu)\, u^*(\tau)\, u^*(\nu)\, d\tau\, d\nu = \int_0^T [y^*(t)]^2\, dt = e_2.$$

For solving (14) one can use iterations analogous to the Algorithm - Case I.

The discussion of sufficiency of (14) is outside the scope of this paper. It is however worth noticing that the following linear, generalized eigenvalue problem is associated with (14): find non-vanishing eigenfunctions and eigenvalues of the following equation:

$$\int_0^T ker(\tau,\, \nu,\, u^*)\, \psi(\tau)\, d\tau = \gamma_2 \int_0^T G(\tau,\, \nu)\, \psi(\tau)\, d\tau, \quad \nu \in (0,\, T). \tag{15}$$

4 Example

Consider the system $\ddot{y}(t) + 2\,\xi\,\dot{y}(t) + \omega_0^2\, y(t) = \omega_0\, u(t)$ with known resonance frequency ω_0 and unknown damping parameter ξ to be estimated, $\dot{y}(0) = 0$, $y(0) = 0$. Its impulse response has the form: $g(t;\xi) = \exp(-\xi\, t)\, \sin(\omega_0\, t)$, $t > 0$, while its sensitivity $k(t;\xi) = \frac{d\, g(t;\xi)}{d\, \xi}$ has the form: $k(t;\xi) = -t\, \exp(-\xi\, t)\, \sin(\omega_0\, t)$, $t > 0$.

Our starting point for searching informative but safe input signal for estimating ξ is $u_0(t) = 0.14\, sin(3\, t)$, $t \in [0,\, 2.5]$. It provides $\mathbf{M}_T = 0.013$ and $\int y^2 = 25.8$ at $\xi_0 = 0.1$, which is our nominal value. Firstly, problem (4) has been solved. The results are shown at the left panels of Fig. 1 and Fig. 2 for input signal and the corresponding output with $\int y^2 = 68.7$, respectively. This input signal is much more informative than u_0 – it provides $\mathbf{M}_T = 0.745$, but is too aggressive – its largest amplitude is about 0.25.

Then, according to the proposed methodology, the constraint $\int y^2 \leq 53.2$ has been added and problem (8) has been solved. The results are shown at the right panels of Fig. 1 and Fig. 2 for input signal and the corresponding output with $\int y^2 = 53.2$, respectively. This input signal is less aggressive – its maximum is 0.15 – and only slightly less informative $\mathbf{M}_T = 0.715$ than more aggressive input signal described above. The relative information efficiency of these two signals is 96% . Thus, we have reduced the largest amplitude by 60% , the output energy by 77% , while D-efficiency dropped by 4%. It is worth to mention that the more safe input signal provides 56 times better information content than our initial guess u_0.

Appendix

Proof of Proposition 3. Notice that $\gamma_2 = 0$ in this case. By direct calculations for $\frac{\partial L(u_\epsilon, \gamma)}{\partial\, \epsilon}$ we have the following expression:

$$\frac{\partial L(u_\epsilon, \gamma)}{\partial\, \epsilon} = 2 \int_0^T f(\nu) \left[\int_0^T ker(\tau, \nu, u_\epsilon)\, (u^*(\tau) + \epsilon\, f(\tau))\, d\tau - \gamma\, u_\epsilon(\nu) \right] d\nu$$

Before obtaining $\frac{\partial^2 L(u_\epsilon, \gamma)}{\partial\, \epsilon^2}$ it is worth considering $K_\epsilon(\tau,\, \nu,\, u_\epsilon) \stackrel{def}{=} \left.\frac{\partial ker(\tau, \nu, u_\epsilon)}{\partial\, \epsilon}\right|_{\epsilon=0}$

$$K_\epsilon(\tau,\, \nu,\, u_\epsilon) = \int_0^T \bar{k}^{tr}(t' - \tau; \bar{a})\, \left.\frac{\partial\, M_T^{-1}(u_\epsilon)}{\partial\, \epsilon}\right|_{\epsilon=0} \bar{k}(t' - \nu; \bar{a})\, dt' \tag{16}$$

Using the well known formula $\frac{\partial B^{-1}(\epsilon)}{\partial \epsilon} = -B^{-1}(\epsilon)\,\frac{\partial B(\epsilon)}{\partial \epsilon}\,B^{-1}(\epsilon)$, valid for differentiable matrix valued functions $B(\epsilon)$ that are nonsingular and symmetric, we obtain

$$\frac{\partial M_T^{-1}(u_\epsilon)}{\partial \epsilon}\Big|_{\epsilon=0} = -\,M_T^{-1}(u^*)\,\mathcal{G}(u^*,\,f)\,M_T^{-1}(u^*) \tag{17}$$

where $\mathcal{G}(u^*,\,f)$ is an $R \times R$ symmetric matrix defined as follows

$$\mathcal{G}(u^*,\,f) \stackrel{def}{=} \int_0^T \left[\int_0^T \bar{k}(t-\nu';\bar{a})\,u^*(\nu')\;d\nu' \int_0^T \bar{k}^{tr}(t-\tau';\bar{a})\,f(\tau')\,d\tau' + \right. \tag{18}$$

$$\left. \int_0^T \bar{k}(t-\nu;\bar{a})\,f(\nu)\;d\nu \int_0^T \bar{k}^{tr}(t-\tau;\bar{a})\,u^*(\tau)\,d\tau\right]\,dt$$

Define $\bar{f} = \int_0^T f(\nu)\,d\nu$, $\overline{f^2} = \int_0^T f^2(\nu)\,d\nu$. Differentiation of $L(u_\epsilon,\,\gamma_1)$ w.r.t. ϵ yields $\frac{\partial^2 L(u_\epsilon,\,\gamma_1)}{\partial \epsilon^2}\Big|_{\epsilon=0} =$

$$= 2\left(\bar{f}^2 - \gamma_1\,\overline{f^2}\right) + 2\int_0^T\int_0^T K_\epsilon(\tau,\,\nu,\,u^*)\,f(\nu)\,u^*(\tau)\,d\tau\,d\nu \tag{19}$$

Subsequent substitutions of (18) into (17) and then to (16) plus tedious calculations lead to the following expression for the second summand in (19)

$$-2\,trace\left[M_T^{-1}(u^*)\,Z\,\left(Z\;M_T^{-1}(u^*)\right)^{tr}\right], \tag{20}$$

where $Z \stackrel{def}{=} \int_0^T Y(t)\,F^{tr}(t)\,dt$, while $Y(t) \stackrel{def}{=} \int_0^T \bar{k}(t-\tau)\,u^*(\tau)\,d\tau$, $F(t) \stackrel{def}{=} \int_0^T \bar{k}(t-\nu)\,f(\nu)\,d\nu$. The matrix in the square brackets in (20) is nonnegative definite. Thus, the whole expression is negative or zero.

From the obvious inequality $\int_0^T \left(\sqrt{\gamma_1}\,f(t) - \bar{f}\right)^2\,dt \geq 0$ we immediately obtain $\left(\bar{f}^2 - \gamma_1\,\overline{f^2}\right) \leq (1 - T - 2\sqrt{\gamma_1})\,\bar{f}^2$ If $(1 - 2\sqrt{\gamma_1} - T) \leq 0$, then the expression $\left(\bar{f}^2 - \gamma_1\,\overline{f^2}\right)$ is nonpositive for all admissible non-constant functions $f \neq 0$, which yields that (19) is negative, proving sufficiency. It remains to be sure that $(1 - 2\sqrt{\gamma_1} - T) \leq 0$ or equivalently that

$$T \geq 1 - 2\sqrt{\gamma_1}. \tag{21}$$

Notice that $\gamma_1 = r + 1$. Thus, $1 - 2\sqrt{(r+1)}$ is always negative, even if only one parameter is estimated. Hence, condition (21) always holds for $T > 0$.

References

1. Atkinson A. C., Donev A. N., Tobias R. D.: *Optimum Experimental Designs, with SAS,* Oxford University Press Inc., New York 2007

2. Banks H. T., Kunisch K., Estimation Techniques for Distributed Parameter Systems. Birkhauser, Boston, 1989
3. Gevers M. et all, Optimal experiment design for open and closed-loop system identification. Communications In Information and Systems 11(3), 197-224, 2011
4. Gevers M., Identification for Control: From the Early Achievements to the Revival of Experiment Design. European Journal of Control 11, 1-18, 2005
5. Goodwin, G. C. and Payne, R. L. Dynamic System Identification. Experiment Design and Data Analysis, Mathematics in Science and Engineering. Academic Press, New York, 1977
6. Hjalmarsson H., Gevers M., De Bruyne F., For model-based control design, closed-loop identification gives better performance. Automatica 32, 1659–1673, 1996
7. Hjalmarsson H., System identification of complex and structured systems, European Journal of Control, 15, pp. 275310, 2009
8. Hjalmarsson H. and H. Jansson, Closed loop experiment design for linear time invariant dynamical systems via LMIs, Automatica, 44, pp. 623636., 2008
9. Mehra R. K. Optimal input signals for parameter estimation in dynamic systems. Survey and new results. IEEE Trans. Automat. Control AC-21, 5564, 1976
10. Patan M., Optimal sensor network scheduling in identification of distributed parameter systems, Springer-Verlag, Lecture Notes in Control and Information Sciences, 2012.
11. Pronzato L., Pazman A., Design of Experiments in Nonlinear Models, Lecture Notes in Statistics, vol. 212, Springer, 2013
12. Rafajlowicz E. and Rafajlowicz W. A variational approach to optimal input signals for parameter estimation in systems with spatio-temporal dynamics, Proceedings of the 10th Int. Workshop in Model-Oriented Design and Analysis / Eds. D. Uciski, A. C. Atkinson, M. Patan (eds.), Lagow, Poland, 2013, pp 219-227, Series" Contribution to Statistics, Springer, 2013
13. Rafajlowicz E. and Rafajlowicz W. Control of linear extended nD sytems with minimized sensitivity to parameter uncertainties, Multidimensional Systems and Signal Proc., vol. 24, pp. 637-656, 2013
14. Skubalska-Rafajłowicz E. and Rafajłowicz E., Sampling multidimensional signals by a new class of quasi-random sequences, Multidim. Syst. Sign. Process., published on-line, June 2010, DOI 10.1007/s11045-010-0120-5
15. Ucinski D., Optimal Measurement Methods for Distributed Parameter System Identification. CRC Press, London, New York, 2005
16. Valenzuela, P. and Rojas, C. and Hjalmarsson, H., Optimal input design for nonlinear dynamic systems: a graph theory approach, arXiv preprint arXiv:1310.4706, 2013
17. Wahlberg B., Hjalmarsson H., and Stoica P., On the Performance of Optimal Input Signals for Frequency Response Estimation, IEEE Trans. Automatic Contr., Vol. 57, NO. 3, pp 766-771, 2012
18. Yosida K., Functional Analysis. Springer-Verlag, 1981
19. Zarrop M.B., *Optimal experimental design for dynamic system identification*, Lecture Notes in Control and Information Science 21. Springer-Verlag, Berlin, Heildelberg, New York, 1979