2008 American Control Conference
Westin Seattle Hotel, Seattle, Washington, USA
June 11-13, 2008

D-Optimal Trajectory Design of Heterogeneous Mobile Sensors for Parameter Estimation of Distributed Systems

Christophe Tricaud Maciej Patan Dariusz Uciński YangQuan Chen

Abstract— **An approach is proposed to joint optimization of trajectories and measurement accuracies of mobile nodes in a sensor network collecting measurements for parameter estimation of a distributed parameter system. The problem is cast as maximization of the log-determinant of the information matrix associated with the estimated parameters over the set of all feasible information matrices, which yields a formulation in terms of convex optimization. This makes it possible to employ powerful tools of optimum experimental design theory to characterize the optimal solution and adapt the Wynn-Fedorov algorithm to construct its numerical approximation. As a crucial subtask in each iteration, a nontrivial optimal control problem must be solved, which is accomplished using the MATLAB PDE toolbox and the RIOTS_95 optimal control toolbox which handles various constraints imposed on the sensor motions. The effectiveness of the method is illustrated with a numerical example regarding a two-dimensional diffusion equation.**

I. INTRODUCTION

States in distributed parameter systems (DPS's), i.e., systems described by partial differential equations (PDEs), vary both spatially and temporally, but it is generally impossible to measure them over the whole spatial domain. Consequently, we are faced with the design problem of how to locate a limited number of measurement sensors so as to obtain as much information as possible about the process at hand. The location of sensors is not necessarily dictated by physical considerations or by intuition and, therefore, some systematic approaches should still be developed in order to reduce the cost of instrumentation and to increase the efficiency of identifiers.

Although it is well-known that the estimation accuracy of DPS parameters depends significantly on the choice of sensor locations, there have been relatively few contributions to the experimental design for those systems. The importance of sensor planning has been recognized in many application domains, e.g., regarding air quality monitoring systems, groundwater-resources management, recovery of valuable minerals and hydrocarbon, model calibration in meteorology and oceanography, chemical engineering, hazardous environments and smart materials [1]–[10]. Over the past years, increasingly thorough research on the development of strategies for efficient sensor placement has been observed (for reviews, see papers [11], [12] and comprehensive monographs [10], [13]). Nevertheless, much still has to be done in this respect, particularly in the light of recent advances in wireless sensor networks [14]–[19].

Nowadays, mobile platforms for sensors are available (mobile robots or unmanned air vehicles) which offer an appealing alternative to common stationary sensors with fixed positions in space [15]–[18], [20]. The complexity of the resulting design problem is expected to be compensated by a number of benefits. Specifically, sensors are not assigned to fixed positions which are optimal only on the average, but are capable of tracking points which provide at a given time instant best information about the parameters to be identified. Consequently, by actively reconfiguring a sensor system we can expect the minimal value of an adopted design criterion to be lower than the one for the stationary case. Areas of direct application of such mobile sensing techniques include air pollutant measurements in the environment obtained from monitoring cars moving in an urban area, or atmospheric variables acquired using instruments carried in a satellite or aircraft [21]. Low-cost mobile platforms with wireless communications capabilities for sensor networks are now available. They get cheaper and cheaper, and more advanced ones are under development. With a group of such autonomous vehicles equipped with sensors, we can enhance the performance of the measurements.

The number of publications on optimized mobile observations for parameter estimation is limited. In the seminal article [22], Rafajłowicz considers the D-optimality criterion and seeks an optimal time-dependent measure, rather than the trajectories themselves. On the other hand, Uciński [10], [23], [24], apart from generalizations of Rafajłowicz's results, develops some computational algorithms based on the Fisher information matrix. He reduces the problem to a state-constrained optimal-control one for which solutions are obtained via the methods of successive linearizations which is capable of handling various constraints imposed on sensor motions. In turn, in [25] Uciński and Chen attempted to properly formulate and solve the time-optimal problem for moving sensors which observe the state of a DPS so as to estimate some of its parameters.

In the literature on mobile sensors, it is most often assumed that the optimal measurement problem consists in the design of trajectories of a given number of identical sensors. In this paper, we formulate it in quite a different manner. First of all, apart from sensor controls and initial positions, the number of sensors constitutes an additional design variable. Additionally, we allow for different measurement accuracies of individual sensors, which are quantified by weights steering the corresponding measurement variances. This leads to a much more general formulation which most often produces an uneven allocation of experimental effort between different sensors. The corresponding solutions could then be implemented on a sensor network with heterogeneous mobile nodes. It turns out that these solutions can be determined using convex optimization tools commonly used in optimum experimental design [26]–[28]. As a result, much better accuracies of the parameter estimates can be achieved.

II. OPTIMAL SENSOR LOCATION PROBLEM

Let $\Omega \subset \mathbb{R}^n$ be a bounded spatial domain with sufficiently smooth boundary Γ, and let $T = (0, t_f]$ be a bounded time interval. Consider a distributed parameter system (DPS) whose scalar state at a spatial point $\boldsymbol{x} \in \bar{\Omega} \subset \mathbb{R}^n$ and time instant $t \in \bar{T}$ is denoted by $y(\boldsymbol{x}, t)$. Mathematically, the system state is governed by the partial differential equation (PDE)

$$\frac{\partial y}{\partial t} = \mathcal{F}\big(\boldsymbol{x}, t, y, \boldsymbol{\theta}\big) \quad \text{in } \Omega \times T, \tag{1}$$

where $\mathcal{F}$ is a well-posed, possibly nonlinear, differential operator which involves first- and second-order spatial derivatives and may include terms accounting for forcing inputs specified *a priori*. The PDE (1) is accompanied by the appropriate boundary and initial conditions

$$\mathcal{B}(\boldsymbol{x}, t, y, \boldsymbol{\theta}) = 0 \qquad \text{on } \Gamma \times T, \tag{2}$$

$$y = y_0 \qquad \text{in } \Omega \times \{t = 0\}, \tag{3}$$

This work is supported in part by the NSF International Research and Education in Engineering (IREE) grant #0540179. Christophe Tricaud is supported by Utah State University President Fellowship (2006-2007)

C. Tricaud and Y. Q. Chen are with the Department of Electrical and Computer Engineering, Utah State University, 4160 Old Main Hill, Utah State University, Logan, UT 84322-4160, USA ctricaud@cc.usu.edu, yqchen@ieee.org

M. Patan and D. Uciński are with the Institute of Control and Computation Engineering, University of Zielona Góra, ul. Podgórna 50, 65–246 Zielona Góra, POLAND {M.Patan, D.Ucinski}@issi.uz.zgora.pl

respectively, $\mathcal{B}$ being an operator acting on the boundary Γ and $y_0 = y_0(\boldsymbol{x})$ a given function. Conditions (2) and (3) complement (1) such that the existence of a sufficiently smooth and unique solution is guaranteed. We assume that the forms of $\mathcal{F}$ and $\mathcal{B}$ are given explicitly up to an m-dimensional vector of unknown constant parameters $\boldsymbol{\theta}$ which must be estimated using observations of the system. The implicit dependence of the state y on the parameter vector $\boldsymbol{\theta}$ will be be reflected by the notation $y(\boldsymbol{x}, t; \boldsymbol{\theta})$.

We assume that the vector $\theta \in \mathbb{R}^m$ is to be estimated from measurements made by N moving sensors over the observation horizon T. We call $\boldsymbol{x}^j : T \to \Omega_{\rm ad}$ the trajectory of the j-th sensor, where $\Omega_{\rm ad} \subset \Omega \cup \Gamma$ is a compact set representing the area where measurements can be made. The observations are of the form

$$z^j(t) = y(\boldsymbol{x}^j(t), t) + \varepsilon(\boldsymbol{x}^j(t), t), \quad t \in T, \quad j = 1, \dots, N, \quad (4)$$

where ε constitutes the measurement noise which is assumed to be is zero-mean, Gaussian, spatial uncorrelated and white [29]–[31], i.e.,

$$\mathrm{E}\big\{\varepsilon(\boldsymbol{x}^j(t), t)\varepsilon(\boldsymbol{x}^i(\tau), \tau)\big\} = \delta_{ji}\delta(t-\tau)\frac{\sigma^2}{p_j}, \quad (5)$$

where σ^2/p_j defines the intensity of the noise, σ^2 is a constant term, p_j stands for a positive scaling factor, δ_{ij} and $\delta(\,\cdot\,)$ standing for the Kronecker and Dirac delta functions, respectively. Although white noise is a physically impossible process, it constitutes a reasonable approximation to a disturbance whose adjacent samples are uncorrelated at all time instants for which the time increment exceeds some value which is small compared with the time constants of the DPS. The white-noise assumption is consistent with most of the literature on the subject.

Note that instead of several mobile sensors whose accuracies are characterized by the same variance σ^2, we use sensors for which the variance of measurement errors is σ^2/p_j. This means that a large weight p_j indicates that the j-th sensor guarantees more precise measurements than sensors with lower weight values. With no loss of generality, we assume that the weights p_j satisfy the following normalization condition:

$$\sum_{j=1}^{N} p_j = 1, \qquad p_j \geq 0, \quad j = 1, \dots, N, \quad (6)$$

i.e., they belong to the probability simplex.

In the presented framework, the parameter identification problem is usually formulated as follows: Given the model (1)–(3) and the outcomes of the measurements z^j along the trajectories $\boldsymbol{x}^j$, $j = 1, \dots, N$, determine an estimate $\widehat{\boldsymbol{\theta}} \in \Theta_{\rm ad}$ ($\Theta_{\rm ad}$ being the set of admissible parameters) which minimizes the generalized output least-squares fit-to-data functional given by [30], [32]

$$\widehat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\vartheta} \in \Theta_{\rm ad}} \sum_{j=1}^{N} p_j \int_T \big[z^j(t) - y(\boldsymbol{x}^j(t), t; \boldsymbol{\vartheta})\big]^2 \,\mathrm{d}t \quad (7)$$

where y now solves (1)–(3) for $\boldsymbol{\theta}$ replaced by $\boldsymbol{\vartheta}$.

We feel, intuitively, that the parameter estimate $\widehat{\boldsymbol{\theta}}$ depends on the number of sensors N, the trajectories $\boldsymbol{x}^j$ and the associated weights p_j since the right-hand side of eqn. (7) does it. This fact suggests that we may attempt to select these design variables so as to produce best estimates of the system parameters after performing the actual experiment. Note that the weights p_j can be interpreted here as sensor costs, which are inversely proportional to the variances of the corresponding measurement errors introduced by them. The weights must sum up to unity, which means that our budget on the experiment is fixed. Then the problem is how to spend it, i.e., how many and how accurate sensors to buy so as to get the most accurate parameter estimates while assuming that their trajectories are also going to be optimized in a way.

To form a basis for the comparison of different design settings, a quantitative measure of the 'goodness' of particular settings is required. A logical approach is to choose a measure related to the expected accuracy of the parameter estimates to be obtained from the data collected (note that the design is to be performed off-line, before taking any measurements). Such a measure is usually based on the concept of the *Fisher Information Matrix* (FIM) [9], [22] which is widely used in optimum experimental design theory for lumped systems [26]–[28]. When the time horizon is large, the nonlinearity of the model with respect to its parameters is mild and the measurement errors are independently distributed and have small magnitudes, the inverse of the FIM constitutes a good approximation of the covariance matrix for the estimate of $\boldsymbol{\theta}$ [26]–[28].

The FIM has the following representation [10], [29]:

$$\boldsymbol{M} = \sum_{j=1}^{N} p_j \int_T \boldsymbol{g}(\boldsymbol{x}^j(t), t)\boldsymbol{g}^{\mathsf{T}}(\boldsymbol{x}^j(t), t)\,\mathrm{d}t, \quad (8)$$

where

$$\boldsymbol{g}(\boldsymbol{x}, t) = \nabla_{\boldsymbol{\vartheta}} y(\boldsymbol{x}, t; \boldsymbol{\vartheta})\big|_{\boldsymbol{\vartheta}=\boldsymbol{\theta}^0} \quad (9)$$

denotes the vector of the so-called *sensitivity coefficients*, $\boldsymbol{\theta}^0$ being a prior estimate to the unknown parameter vector θ [10], [23].

The sought optimal design settings can be found by maximizing some scalar function Ψ of the information matrix. The introduction of the design criterion permits to cast the sensor location problem as an optimization problem, and the criterion itself can be treated as a measure of the information content of the observations. Several choices exist for such a function [26]–[28] and the most popular one is the D-optimality criterion

$$\Psi[\boldsymbol{M}] = \log\det(\boldsymbol{M}). \quad (10)$$

Its use yields the minimal volume of the confidence ellipsoid for the estimates. In what follows, we shall restrict our attention to this optimality criterion.

III. MOBILE SENSOR MODEL

A. Node Dynamics

Although measurement accuracies may vary from sensor to sensor, we assume that all sensors are conveyed by identical vehicles whose motions are described by

$$\dot{\boldsymbol{x}}^j(t) = \boldsymbol{f}(\boldsymbol{x}^j(t), \boldsymbol{u}^j(t)) \quad \text{a.e. on } T, \quad \boldsymbol{x}^j(0) = \boldsymbol{x}_0^j \quad (11)$$

where a given function $\boldsymbol{f} : \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}^n$ is required to be continuously differentiable, $\boldsymbol{x}_0^j \in \mathbb{R}^n$ defines an initial sensor configuration, and $\boldsymbol{u}^j : T \to \mathbb{R}^r$ is a measurable control function which satisfies

$$\boldsymbol{u}_l \leq \boldsymbol{u}^j(t) \leq \boldsymbol{u}_u \quad \text{a.e. on } T \quad (12)$$

for some constant vectors $\boldsymbol{u}_l$ and $\boldsymbol{u}_u$, $j = 1, \dots, N$.

For each $j = 1, \dots, N$, given any initial position $\boldsymbol{x}_0^j$ and any control function, there is a unique absolutely continuous function $\boldsymbol{x}^j : T \to \mathbb{R}^n$ which satisfies (11) a.e. on T. In what follows, we will call it the state trajectory corresponding to $\boldsymbol{x}_0^j$ and $\boldsymbol{u}^j$.

B. Pathwise State Constraints

In reality, some restrictions on the motions are inevitably induced. First of all, all sensors should stay within the admissible region $\Omega_{\rm ad}$ where measurements are allowed. We assume that it is a compact set defined as follows:

$$\Omega_{\rm ad} = \{\boldsymbol{x} \in \Omega \cup \Gamma \mid b_i(\boldsymbol{x}) \leq 0,\, i = 1, \dots, I\} \quad (13)$$

where b_i's are given continuously differentiable functions. Accordingly, the conditions

$$b_i(\boldsymbol{x}^j(t)) \leq 0, \quad \forall\, t \in T \quad (14)$$

must be fulfilled, where $1 \leq i \leq I$ and $1 \leq j \leq N$.

C. Parameterization of Vehicle Controls

From now on we make the assumption that the controls of the available vehicles can be represented in parametric form

$$\boldsymbol{u}^j(t) = \boldsymbol{\eta}(t, \boldsymbol{a}^j), \quad t \in T, \tag{15}$$

where $\boldsymbol{\eta}$ denotes a given function such that $\boldsymbol{\eta}(\,\cdot\,, \boldsymbol{a}^j)$ is continuous for each fixed $\boldsymbol{a}^j$ and $\boldsymbol{\eta}(t, \,\cdot\,)$ is continuous for each fixed t, the constant parameter vector $\boldsymbol{a}^j$ ranging over a compact set $A \subset \mathbb{R}^q$. An exemplary parameterization can rely on using B-splines as employed in numerous optimal control solvers, e.g., RIOTS_95 described later.

Based on this parameterization, we can define the mapping $\boldsymbol{\chi}$ which assigns every $\boldsymbol{c}^j = (\boldsymbol{x}_0^j, \boldsymbol{a}^j) \in \Omega_{\rm ad} \times A$ a trajectory $\boldsymbol{x}^j = \boldsymbol{\chi}(\boldsymbol{c}^j)$ through solving (11) for the initial position $\boldsymbol{x}_0^j$ and control defined by (15).

Since only the controls and trajectories satisfying the imposed constraints are interesting, we introduce the set

$$C = \big\{\boldsymbol{c} = (\boldsymbol{x}_0, \boldsymbol{a}) \in A \times \Omega_{\rm ad} : \boldsymbol{\eta}(\,\cdot\,, \boldsymbol{a}) \text{ satisifes (12)}, \ \boldsymbol{\chi}(\boldsymbol{c}) \text{ satisfies (14)}\big\} \tag{16}$$

and assume that it is nonempty. A trivial verification shows that C is also compact.

Given N sensors, we thus obtain trajectories $\boldsymbol{x}^j$ corresponding to vectors $\boldsymbol{c}^j \in \mathbb{R}^{n+q}$, $j = 1, \ldots, N$. The FIM can then be rewritten as

$$\boldsymbol{M}(\xi_N) = \sum_{j=1}^{N} p_j \int_T \boldsymbol{g}(\boldsymbol{x}(t), t)\boldsymbol{g}^{\mathsf{T}}(\boldsymbol{x}(t), t)\Big|_{\boldsymbol{x}=\boldsymbol{\chi}(\boldsymbol{c}^j)} \mathrm{d}t. \tag{17}$$

where, for simplicity of notation, we represent the decision variables as the table

$$\xi_N = \begin{Bmatrix} \boldsymbol{c}^1, & \boldsymbol{c}^2, & \ldots, & \boldsymbol{c}^N \\ p_1, & p_2, & \ldots, & p_N \end{Bmatrix}. \tag{18}$$

Applying the terminology of optimum experimental design, we call this table a *discrete design*, while $\boldsymbol{c}^1, \ldots, \boldsymbol{c}^N$ are termed the *support points* and $p_1, \ldots, p_N$ are referred to as the corresponding *weights*.

Observe that a design ξ_N can be interpreted as a discrete probability distribution on a finite subset of C, cf. (6). As is standard in optimum experimental design theory, we can extend this idea and think of a design as a probability measure ξ for all Borel sets of C including single points. With respect to such a modification, we can define the FIM analogous to (17) for a design ξ:

$$\boldsymbol{M}(\xi) = \int_C \boldsymbol{\Upsilon}(\boldsymbol{c})\, \xi(\mathrm{d}\boldsymbol{c}), \tag{19}$$

where

$$\boldsymbol{\Upsilon}(\boldsymbol{c}) = \int_T \boldsymbol{g}(\boldsymbol{x}(t), t)\boldsymbol{g}^{\mathsf{T}}(\boldsymbol{x}(t), t)\Big|_{\boldsymbol{x}=\boldsymbol{\chi}(\boldsymbol{c})} \mathrm{d}t. \tag{20}$$

The integration in (19) is to be understood in the Lebesgue-Stieltjes sense. This leads to the so-called *continuous* designs which constitute the basis of the modern theory of optimal experiments and originate in seminal works by Kiefer and Wolfowitz [33]. It turns out that such an approach drastically simplifies the design and the remainder of the paper is devoted to this issue.

IV. CHARACTERIZATION OF OPTIMAL SOLUTIONS

For clarity, we adopt the following notational conventions: Here and subsequently, we will use the symbol $\Xi(C)$ to denote the set of all probability measures on C. Let us also introduce the notation $\mathfrak{M}(C)$ for the set of all admissible information matrices, i.e.,

$$\mathfrak{M}(C) = \big\{\boldsymbol{M}(\xi) : \xi \in \Xi(C)\big\} \tag{21}$$

Then we may redefine an optimal design as a solution to the optimization problem

$$\xi^\star = \arg \max_{\xi \in \Xi(C)} \Psi[\boldsymbol{M}(\xi)]. \tag{22}$$

The theoretical results presented in this section constitute straightforward adaptations of their counterparts of Chapter 3 in [10]. We begin with certain convexity and representation properties of $M(\xi)$.

Lemma 1: For any $\xi \in \Xi(C)$ the information matrix $\boldsymbol{M}(\xi)$ is symmetric and non-negative definite.

Lemma 2: $\mathfrak{M}(C)$ is compact and convex.

Lemma 3: For any $\boldsymbol{M}_0 \in \mathfrak{M}(C)$ there always exists a purely discrete design ξ of the form (18) with no more than $m(m+1)/2+1$ support points such that $\boldsymbol{M}(\xi) = \boldsymbol{M}_0$. If $\boldsymbol{M}_0$ lies on the boundary of $\mathfrak{M}(C)$, then the number of support points is less than or equal to $m(m+1)/2$.

The above lemma makes it justified to restrict our attention only to discrete designs with a limited number of supporting points, so the introduction of continuous designs being probability measures for all Borel sets of C, which may seem at first sight a superfluous complication, is solely a technicality. On the other hand, it greatly simplifies the reasoning and does lead to very tangible results.

The next result provides a characterization of the optimal designs.

Theorem 1: We have the following properties:

(i) An optimal design exists which is discrete and comprises no more than $m(m+1)/2$ support points (i.e., one less than predicted by Lemma 3).
(ii) The set of optimal designs is convex.
(iii) A design $\xi^\star$ is optimal iff

$$\max_{\boldsymbol{c} \in C} \varphi(\boldsymbol{c}, \xi^\star) = m, \tag{23}$$

where

$$\varphi(\boldsymbol{c}, \xi) = \mathrm{trace}[\boldsymbol{M}^{-1}(\xi)\boldsymbol{\Upsilon}(\boldsymbol{c})] \tag{24}$$

(iv) For any purely discrete optimal design $\xi^\star$, the function $\varphi(\,\cdot\,, \xi^\star)$ has value zero at all support points.

It is now clear that the function φ is of paramount importance in our considerations, as it determines the location of the support points in the optimal design $\xi^\star$ (they are situated among its points of global maximum). Moreover, given any design ξ, it indicates points at which a new observation contributes to the greatest extent. Indeed, adding a new observation atomized at a single point $\boldsymbol{c}^+$ amounts to constructing a new design

$$\xi^+ = (1 - \lambda)\xi + \lambda\xi_{\boldsymbol{c}^+} \tag{25}$$

for some $\lambda \in (0, 1)$. If λ is sufficiently small, then it may be concluded that

$$\Psi[\boldsymbol{M}(\xi^+)] - \Psi[\boldsymbol{M}(\xi)] \approx \lambda\varphi(\boldsymbol{c}^+, \xi) \tag{26}$$

i.e., the resulting increase in the criterion value is approximately equal to $\lambda\varphi(\boldsymbol{c}^+, \xi)$.

Analytical determination of optimal designs is possible only in simple situations and for general systems it is usually the case that some iterative design procedure will be required. The next theorem, called the *equivalence theorem*, is useful in the checking for optimality of designs.

Theorem 2: The following characterizations of an optimal design $\xi^\star$ are equivalent in the sense that each implies the other two:

(i) the design $\xi^\star$ maximizes $\Psi[\boldsymbol{M}(\xi)]$,
(ii) the design $\xi^\star$ minimizes $\max_{\boldsymbol{c} \in C} \varphi(\boldsymbol{c}, \xi)$, and
(iii) $\max_{\boldsymbol{c} \in C} \varphi(\boldsymbol{c}, \xi^\star) = m$.

All the designs satisfying (i)–(iii) and their convex combinations have the same information matrix $\boldsymbol{M}(\xi^\star)$.

The above results provide us with tests for the optimality of designs. In particular,

1. If the sensitivity function $\varphi(\boldsymbol{c}, \xi)$ is less than or equal to m for all $\boldsymbol{c} \in C$, then ξ is optimal.

2. If the sensitivity function $\varphi(\boldsymbol{c},\xi)$ exceeds m, then ξ is not optimal.

The interesting thing about these results is that in addition to revealing striking minimax properties of optimal designs, they also provide sequential numerical design algorithms. The underlying idea is quite simple. Suppose that we have an arbitrary (non-optimal) design ξ_k obtained after k iteration steps. Further, let $\varphi(\,\cdot\,,\xi_k)$ attain its maximum (necessarily $> m$) at $\boldsymbol{c}=\boldsymbol{c}_k^0$. Then the design

$$\xi_{k+1}=(1-\lambda_k)\xi_k+\lambda_k\xi_{\boldsymbol{c}_k^0} \tag{27}$$

(here $\xi_{\boldsymbol{c}_k^0}$ stands for the unit-weight design concentrated at $\boldsymbol{c}_k^0$) leads to an increase in the value of $\Psi[\boldsymbol{M}(\xi_{k+1})]$ for a suitably small λ_k. This follows since the derivative with respect to λ_k is positive, i.e.,

$$\frac{\partial}{\partial\lambda_k}\Psi[\boldsymbol{M}(\xi_{k+1})]\Big|_{\lambda_k=0^+}=m-\varphi(\boldsymbol{c}_k^0,\xi_k)>0 \tag{28}$$

The steps in using the outlined gradient method can be briefly summarized as follows [26], [27], [34], [35]:

Step 1. Guess a discrete non-degenerate starting design measure ξ_0 (we must have $\det(\boldsymbol{M}(\xi_0))\neq 0$). Choose some positive tolerance $\epsilon\ll 1$. Set $k=0$.

Step 2. Determine $\boldsymbol{c}_k^0=\arg\max_{\boldsymbol{c}\in C}\varphi(\boldsymbol{c},\xi_k)$. If $\varphi(\boldsymbol{c}_k^0,\xi_k)<m+\epsilon$, then *STOP*.

Step 3. For an appropriate value of $0<\lambda_k<1$, set

$$\xi_{k+1}=(1-\lambda_k)\xi_k+\lambda_k\xi_{\boldsymbol{c}_k^0}$$

increment k by one and go to Step 2.

In the same way as for the classical first-order algorithms in common use in optimum experimental designs for many years, it can be shown that the above algorithm converges to an optimal design provided that the sequence $\{\lambda_k\}$ is suitably chosen. For example, the choices which satisfy one of the conditions below will yield the convergence:

(i) $\lim_{k\to\infty}\lambda_k=0$, $\sum_{k=0}^{\infty}\lambda_k=\infty$ (Wynn's algorithm),

(ii) $\lambda_k=\arg\min_{\lambda}\Psi[(1-\lambda)\boldsymbol{M}(\xi_k)+\lambda\boldsymbol{M}(\xi_{\boldsymbol{c}_k^0})]$ (Fedorov's algorithm),

Computationally, Step 2 is of crucial significance but at the same time it is the most time-consuming step in the algorithm. Complications arise, among other things, due to the necessity of calculating a global maximum of $\varphi(\,\cdot\,,\xi_k)$ which is usually multimodal (getting stuck in one of local maxima leads to precocious termination of the algorithm). Therefore, while implementing this part of the computational procedure an effective global optimizer is essential.

V. OPTIMAL CONTROL FORMULATION OF THE SEARCH FOR THE CANDIDATE SUPPORT POINT

Step 2 of the Wynn-Fedorov algorithm of the previous section necessitates determination of $\arg\max_{\boldsymbol{c}\in C}\varphi(\boldsymbol{c},\xi_k)$. This formulation can be interpreted as a finite-dimensional approximation to the following optimization problem: Find the pair $(\boldsymbol{x}_0,\boldsymbol{u})$ which maximizes

$$\begin{aligned}J(\boldsymbol{x}_0,\boldsymbol{u})&=\operatorname{trace}\Big[\boldsymbol{M}^{-1}(\xi^k)\int_T\boldsymbol{g}(\boldsymbol{x}(t),t)\boldsymbol{g}^{\mathsf{T}}(\boldsymbol{x}(t),t)\,\mathrm{d}t.\Big]\\&=\int_T\boldsymbol{g}^{\mathsf{T}}(\boldsymbol{x}(t),t)\boldsymbol{M}^{-1}(\xi^k)\boldsymbol{g}(\boldsymbol{x}(t),t)\,\mathrm{d}t\end{aligned} \tag{29}$$

over the set of feasible pairs

$$\begin{aligned}\mathcal{P}=\big\{(\boldsymbol{x}_0,\boldsymbol{u})\mid u:T\to\mathbb{R}^r\text{ is measurable,}\\ \boldsymbol{u}_l\le\boldsymbol{u}(t)\le\boldsymbol{u}_u\ \text{ a.e. on } T,\ \boldsymbol{x}_0\in\Omega_{\mathrm{ad}}\big\},\end{aligned} \tag{30}$$

subject to the pathwise state inequality constraints (14).

Evidently, its high nonlinearity excludes any possibility of finding closed-form formulas for its solution. Accordingly, we must resort to numerical techniques. A number of possibilities exist in this respect [36], [37], but since this problem is already in canonical form, we can solve it using one of the existing packages for numerically solving dynamic optimization problems, such as RIOTS_95 [38], DIRCOL [39] or MISER [40]. In our implementation, we employed the first of them, i.e., RIOTS_95, which is designed as a MATLAB toolbox written mostly in C and running under Windows 98/2000/XP and Linux. It provides an interactive environment for solving a very broad class of optimal control problems. The users's problems can be prepared purely as M-files and no compiler is required to solve them. To speed up the solution process, the functions defining the problem can be coded in C and then compiled and linked with some pre-built linking libraries. The implemented numerical methods are supported by the theory outlined in [36], which uses the approach of consistent approximations. Systems dynamics can be integrated with fixed step-size Runge-Kutta integration, a discrete-time solver or a variable step-size method. The software automatically computes gradients for all functions with respect to the controls and any free initial conditions. The controls are represented as splines, which allows for a high-degree of function approximation accuracy without requiring a large number of control parameters. There are three main optimization routines, each suited for different levels of generality, and the most general is based on sequential quadratic programming methods (it was also used in our computations reported in the next section).

Note that in RIOTS_95 the controls are internally approximated by linear, quadratic or cubic splines, and this immediately defines the parameterization (15).

VI. ILLUSTRATIVE EXAMPLE

In this section, we use demonstrative example to illustrate our method. We consider the two-dimensional diffusion equation

$$\frac{\partial y}{\partial t}=\nabla\cdot(\kappa\nabla y)+F \tag{31}$$

for $x\in\Omega=(0,1)^2$ and $t\in[0,1]$, subject to homogeneous zero initial and Dirichlet boundary conditions, where $F(x,t)=20\exp(-50(x_1-t)^2)$. The spatial distribution of the diffusion coefficient is assumed to have the form

$$\kappa(x_1,x_2)=\theta_1+\theta_2x_1+\theta_3x_2. \tag{32}$$

In our example, we select the initial estimates of the parameter values as $\theta_1^0=0.1$, $\theta_2^0=-0.05$ and $\theta_3^0=0.2$, which are assumed to be nominal and known prior to the experiment. The excitation function F in (31) simulates a source with a vertical line support along the x_2-axis, which moves like a plane wave with constant speed from the left to the right boundary of Ω within the observation interval $[0,1]$.

The determination of the Fisher information matrix for a given experiment requires the knowledge of the vector of the sensitivity coefficients $\boldsymbol{g}=\operatorname{col}[g_1,g_2,g_3]$ along sensor trajectories. The FIM can be obtained using the direct differentiation method [10] by solving the following system of PDEs:

$$\begin{aligned}\frac{\partial y}{\partial t}&=\nabla\cdot(\kappa\nabla y)+F,\\ \frac{\partial g_1}{\partial t}&=\nabla\cdot\nabla y+\nabla\cdot(\kappa\nabla g_1),\\ \frac{\partial g_2}{\partial t}&=\nabla\cdot(x_1\nabla y)+\nabla\cdot(\kappa\nabla g_2),\\ \frac{\partial g_3}{\partial t}&=\nabla\cdot(x_2\nabla y)+\nabla\cdot(\kappa\nabla g_3),\end{aligned} \tag{33}$$

in which the first equation represents the original state equation and the next three equations are obtained from the differentiation of the first equation with respect to the parameters θ_1, θ_2 and θ_3, respectively. The initial and Dirichlet boundary conditions for all the four equations are homogeneous.

The system (33) has been solved numerically using the routines from MATLAB PDE toolbox and stored g_1, g_2 and g_3 interpolated

at the nodes of a rectangular grid in a four-dimensional array (we applied uniform partitions using 21 grid points per each spatial dimension and 31 points in time), cf. Appendix I in [10] for details. Since values of g_1, g_2 and g_3 may have been required at points which were not necessarily nodes of that grid, the relevant interpolation was thus performed using cubic splines in space (for this purpose MATLABs procedure `interp2` has been applied) and linear splines in time. Since, additionally, the derivatives of g with respect to spatial variables and time were required during the trajectory optimization process, these derivatives were approximated numerically using the central difference formula.

Next, we used RIOTS_95 to determine D-optimal sensor trajectories in accordance with the Wynn-Fedorov algorithm. The dynamics follow the simple model

$$\dot{\boldsymbol{x}}^j(t) = \boldsymbol{u}^j(t), \quad \boldsymbol{x}^j(0) = \boldsymbol{x}_0^j, \tag{34}$$

and additional constraints

$$|u_i^j(t)| \leq 0.7, \quad \forall t \in T, \quad i = 1, \ldots, 6 \tag{35}$$

restricting the maximum sensor velocity components were imposed on the controls. Our goal is to design their trajectories so as to obtain possibly the best estimates of θ_1, θ_2 and θ_3.

A program was implemented using a low-cost PC (AMD Athlon 3800+, 2GB RAM) running on Windows XP and MATLAB 701 (R2006a). We run the program twice with 4 iterations and 200 randomly chosen initial positions for each iteration. Each run took between 10 and 45 seconds for each initial position. This is necessary if we wish to get an approximation to a global maximum in Step 2 of the Wynn-Fedorov algorithm. This is a trade-off between the computation time and the number of possible initial positions.

Figs. 1 and 3 present the results obtained for these two simulations. The initial sensor positions are marked with open circles, and the sensors positions at the consecutive points of the time grid are marked with dots. When available, weights are inserted inside the figures, each weight being positioned by its respective trajectory.

The first run gives two different trajectories with weights of 0.54807 and 0.45193. Based on the generalized weighted LS criterion each weight can be interpreted in terms of an experimental cost, which is inversely proportional to the variance of the observation error along a given trajectory. Thus we may think of the weights as the cost related e.g. to the sensitivity of the measurement devices. Following this interpretation, we should spent approx. 55% of total experimental costs to assure more accurate sensor for the first trajectory, and approx. 45% to the second trajectory, which requires less sensitive sensor. In contrary, the second run results in three distinct trajectories with weights of 0.44464, 0.34726 and 0.2081 (cf. Fig. 3). However, combining second and third trajectories together with the total weight 0.55536, we can observe that this solution is quite similar to the previous one with only two distinct sensor paths. The differences can be explained in terms of the suboptimality of the solutions for the internal problem in Step. 2 of the Wynn-Fedorov algorithm (in order to assure the compromise between the computational burden and the quality of solution, in practice we are satisfied with fairly good approximation of global optimum). Thus, in both simulations we come up with only different suboptimal solutions to our problem, but with acceptable quality in the practical sense. The obtained Fisher information matrices are

$$M_{(1)} = \begin{pmatrix} 124.3815 & 68.0614 & 25.7666 \\ 68.0614 & 41.5653 & 13.4240 \\ 25.7666 & 13.4240 & 8.7691 \end{pmatrix} \tag{36}$$

$$M_{(2)} = \begin{pmatrix} 130.0149 & 72.3503 & 26.6154 \\ 72.3503 & 44.2181 & 14.1798 \\ 26.6154 & 14.1798 & 8.6267 \end{pmatrix} \tag{37}$$

with the criterion values Ψ equal to 7.4888 and 7.3672, respectively.

For comparison, we also present the results obtained using the technique described in [10] for D-optimum trajectories of moving sensors. This strategy is similar to ours but does not use weights in the computation of the FIM (or more precisely, the weights are fixed and assumed to be equal for each trajectory). Results are shown in Fig. 2 (2 sensors) and Fig. 4 (3 sensors).

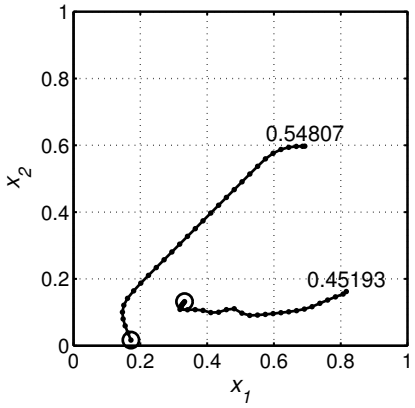

Fig. 1. Optimal trajectory of 2 sensors using weighted D-optimality criterion ($\Psi = 7.4888$)

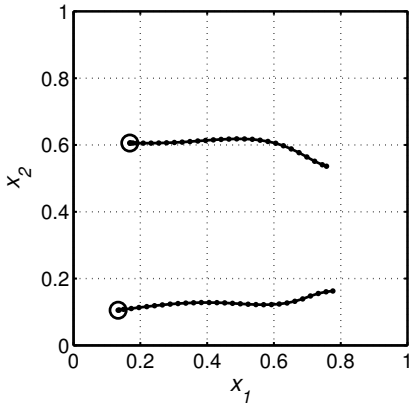

Fig. 2. Optimal trajectory of 2 sensors using standard D-optimality criterion ($\Psi = 7.4017$)

VII. CONCLUSION

The results contained in this paper show that some well-known methods of optimum experimental design for linear regression models can be applied to the setting of the mobile sensor trajectory design problem for parameter estimation of DPS's in case we wish to simultaneously optimize the number of sensors and their trajectories, as well as to optimally share the experimental effort. The latter is understood here as allowing for different measurement accuracies of individual sensors, which are quantified by weights steering the corresponding measurement variances. This leads to a much more general setting which most frequently produces an uneven allocation of experimental effort between different sensors. This remains in contrast with the existing approaches. The corresponding solutions could then be implemented on a sensor network with heterogeneous mobile nodes. In the paper we demonstrate that these solutions can be determined using convex optimization tools commonly employed in optimum experimental design and show how to apply numerical tools of algorithmic optimal control to support the determination of the optimal solutions.

REFERENCES

[1] A. Nehorai, B. Porat, and E. Paldi, "Detection and localization of vapor-emitting sources," *IEEE Transactions on Signal Processing*, vol. 43, no. 1, pp. 243–253, 1995.

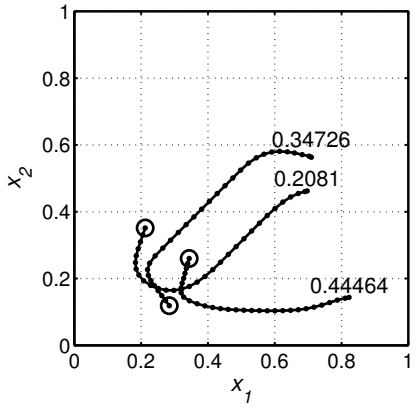

Fig. 3. Optimal trajectory of 3 sensors using weighted D-optimality criterion ($\Psi = 7.3672$)

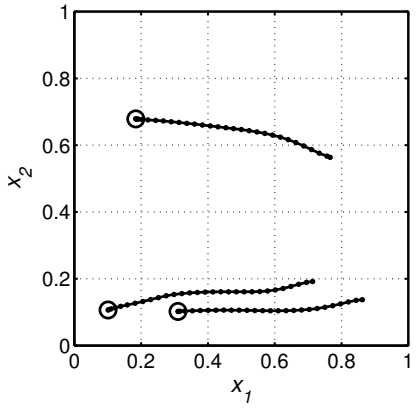

Fig. 4. Optimal trajectory of 3 sensors using standard D-optimality criterion ($\Psi = 7.4959$)

[2] B. Porat and A. Nehorai, "Localizing vapor-emitting sources by moving sensors," *IEEE Transactions on Signal Processing*, vol. 44, no. 4, pp. 1018–1021, 1996.

[3] A. Jeremić and A. Nehorai, "Design of chemical sensor arrays for monitoring disposal sites on the ocean floor," *IEEE Transactions on Oceanic Engineering*, vol. 23, no. 4, pp. 334–343, 1998.

[4] ——, "Landmine detection and localization using chemical sensor array processing," *IEEE Transactions on Signal Processing*, vol. 48, no. 5, pp. 1295–1305, 2000.

[5] I. M. Navon, "Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography," *Dynamics of Atmospheres and Oceans*, vol. 27, pp. 55–79, 1997.

[6] D. N. Daescu and I. M. Navon, "Adaptive observations in the context of 4D-Var data assimilation," *Meteorology and Atmospheric Physics*, vol. 85, pp. 205–226, 2004.

[7] P. D. Christofides, *Nonlinear and Robust Control of PDE Systems: Methods and Applications to Transport-Reaction Processes*, ser. Systems & Control: Foundations & Applications. Boston: Birkhäuser, 2001.

[8] H. T. Banks, R. C. Smith, and Y. Wang, *Smart Material Structures: Modeling, Estimation and Control*, ser. Research in Applied Mathematics. Paris: Masson, 1996.

[9] N.-Z. Sun, *Inverse Problems in Groundwater Modeling*, ser. Theory and Applications of Transport in Porous Media. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1994.

[10] D. Uciński, *Optimal Measurement Methods for Distributed-Parameter System Identification*. Boca Raton, FL: CRC Press, 2005.

[11] C. S. Kubrusly and H. Malebranche, "Sensors and controllers location in distributed systems — A survey," *Automatica*, vol. 21, no. 2, pp. 117–128, 1985.

[12] M. van de Wal and B. de Jager, "A review of methods for input/output selection," *Automatica*, vol. 37, pp. 487–510, 2001.

[13] D. Uciński, *Measurement Optimization for Parameter Estimation in Distributed Systems*. Zielona Góra: Technical University Press, 1999.

[14] F. Zhao and L. J. Guibas, *Wireless Sensor Networks: An Information Processing Approach*. Amsterdam: Morgan Kaufmann Publishers, 2004.

[15] C. G. Cassandras and W. Li, "Sensor networks and cooperative control," *European Journal of Control*, vol. 11, no. 4–5, pp. 436–463, 2005.

[16] B. Sinopoli, C. Sharp, L. Schenato, S. Schaffert, and S. S. Sastry, "Distributed control applications within sensor networks," *Proceedings of the IEEE*, vol. 91, no. 8, pp. 1235–1246, 2003.

[17] C.-Y. Chong and S. P. Kumar, "Sensor networks: Evolution, opportunities, and challenges," *Proceedings of the IEEE*, vol. 91, no. 8, pp. 1247–1256, 2003.

[18] D. Culler, D. Estrin, and M. Srivastava, "Overview of sensor networks," *IEEE Computer*, vol. 37, no. 8, pp. 41–49, 2004.

[19] N. Jain and D. P. Agrawal, "Current trends in wireless sensor network design," *International Journal of Distributed Sensor Networks*, vol. 1, pp. 101–122, 2005.

[20] P. Ögren, E. Fiorelli, and N. E. Leonard, "Cooperative control of mobile sensor networks: Adaptive gradient climbing in a distributed environment," *IEEE Transactions on Automatic Control*, vol. 49, no. 8, pp. 1292–1302, 2004.

[21] K. Nakano and S. Sagara, "Optimal measurement problem for a stochastic distributed parameter system with movable sensors," *International Journal of Systems Science*, vol. 12, no. 12, pp. 1429–1445, 1981.

[22] E. Rafajłowicz, "Optimum choice of moving sensor trajectories for distributed parameter system identification," *International Journal of Control*, vol. 43, no. 5, pp. 1441–1451, 1986.

[23] D. Uciński, "Optimal sensor location for parameter estimation of distributed processes," *International Journal of Control*, vol. 73, no. 13, pp. 1235–1248, 2000.

[24] D. Uciński and J. Korbicz, "Optimal sensor allocation for parameter estimation in distributed systems," *Journal of Inverse and Ill-Posed Problems*, vol. 9, no. 3, pp. 301–317, 2001.

[25] D. Uciński and Y. Chen, "Time-optimal path planning of moving sensors for parameter estimation of distributed systems," in *Proc. 44th IEEE Conference on Decision and Control, and the European Control Conference 2005,* Seville, Spain, 2005, published on CD-ROM.

[26] V. V. Fedorov and P. Hackl, *Model-Oriented Design of Experiments*, ser. Lecture Notes in Statistics. New York: Springer-Verlag, 1997.

[27] É. Walter and L. Pronzato, *Identification of Parametric Models from Experimental Data*, ser. Communications and Control Engineering. Berlin: Springer-Verlag, 1997.

[28] A. C. Atkinson and A. N. Donev, *Optimum Experimental Designs*. Oxford: Clarendon Press, 1992.

[29] Z. H. Quereshi, T. S. Ng, and G. C. Goodwin, "Optimum experimental design for identification of distributed parameter systems," *International Journal of Control*, vol. 31, no. 1, pp. 21–29, 1980.

[30] S. Omatu and J. H. Seinfeld, *Distributed Parameter Systems: Theory and Applications*, ser. Oxford Mathematical Monographs. New York: Oxford University Press, 1989.

[31] M. Amouroux and J. P. Babary, "Sensor and control location problems," in *Systems & Control Encyclopedia. Theory, Technology, Applications*, M. G. Singh, Ed. Oxford: Pergamon Press, 1988, vol. 6, pp. 4238–4245.

[32] H. T. Banks and K. Kunisch, *Estimation Techniques for Distributed Parameter Systems*, ser. Systems & Control: Foundations & Applications. Boston: Birkhäuser, 1989.

[33] J. Kiefer and J. Wolfowitz, "Optimum designs in regression problems," *The Annals of Mathematical Statistics*, vol. 30, pp. 271–294, 1959.

[34] S. M. Ermakov, Ed., *Mathematical Theory of Experimental Design*. Moscow: Nauka, 1983, (In Russian).

[35] E. Rafajłowicz, *Algorithms of Experimental Design with Implementations in MATHEMATICA*. Warsaw: Akademicka Oficyna Wydawnicza PLJ, 1996, (In Polish).

[36] E. Polak, *Optimization. Algorithms and Consistent Approximations*, ser. Applied Mathematical Sciences. New York: Springer-Verlag, 1997.

[37] W. A. Gruver and E. Sachs, *Algorithmic Methods in Optimal Control*. London: Pitman Publishing Limited, 1980.

[38] A. L. Schwartz, E. Polak, and Y. Chen, *A Matlab Toolbox for Solving Optimal Control Problems. Version 1.0 for Windows*, May 1997. [Online]. Available: http://www.schwartz-home.com/˜adam/RIOTS/

[39] O. von Stryk, *User's Guide for DIRCOL, a Direct Collocation Method for the Numerical Solution of Optimal Control Problems. Version 2.1*, Fachgebiet Simulation und Systemoptimierung, Technische Universität Darmstadt, Nov. 1999. [Online]. Available: http://www.sim.informatik.tu-darmstadt.de/index/leftnav.html.en

[40] L. S. Jennings, M. E. Fisher, K. L. Teo, and C. J. Goh, *MISER 3: Optimal Control Software, Version 2.0. Theory and User Manual*, Department of Mathematics, University of Western Australia, Nedlands, 2002. [Online]. Available: http://www.cado.uwa.edu.au/miser/