Optimal Design of Multivariate Sensors

Antonio Bicchi*
Gaetano Canepa

Centro "E.Piaggio",
Università di Pisa

Address: Via Diotisalvi 2,
56100 PISA, ITALY

* Corresponding Author. Email: bicchi@vm.cnuce.cnr.it

Short title: Optimal Design of Multivariate Sensors.
Classification number: 062D, 0670

Abstract

In this paper the design of sensing systems for the measurement of multiple physical quantities related to a dynamical system is considered. A multivariate sensor comprises several simple transducers, each measuring a scalar quantity that comes from the combination of the components of the quantity to be measured. From the collection of measurements of single transducers at different times, the desired information is extracted by analog or digital processing. Besides the choice of technological characteristics of the transducers to be employed, the designer of multivariate sensors is usually allowed some freedom in choosing the number of transducers, their arrangement in the system, and the time scheduling of their measurements. These choices are the subject of optimal policies in the design phase, whose goal is to maximize some performance (or minimize some cost) criterion. We survey some of the existing approaches to optimal design of multivariate sensors, according to the different types of systems they are applied to. Two examples of optimal sensor design are discussed as an illustration of the methods.

1 Introduction

Complex sensing systems are increasingly being employed in sophisticated systems to handle a large variety of data from the system and extracting the desired pieces of information. Applications range from large-scale process control systems to biomedical monitoring devices, and to robotics, to name a few. In general, the quantities of interest are embedded in a dynamic relationship, and basic transducers are available whose scalar measurements come from the combination of some or all of the components to be measured. The sensing system comprises the collection of such basic transducers, and of the processing means for extracting from their data the information that is ultimately desired. Accordingly, the task of the designer of a multivariate sensor is twofold: on one side, there is the choice of the technological realization of basic transducers; on the other, the organization of transducers in a system involving decisions about their number, location, scheduling, etc. Both aspects aim at the same goal of maximizing sensitivity and minimizing errors. To a large extent, however, the technological and methodological aspects of the design process can be separately analyzed. This paper is focussed on the latter problem.

This paper consists of a tutorially-oriented survey of problems and methods that may occur to scientists involved in the design of multivariate sensory systems, and hinges upon the point that optimal design of sensors for a large variety of applications (state estimation for continuous or discrete-time systems, lumped or distributed parameter systems, calibration, parameter identification, etc.) can be cast in the common format of a nonlinear programming problem, "*Find the set of r design variables ξ, that maximizes a generalized performance index $I(\xi)$ subject to the constraint $\xi \in \Omega \subset \mathbb{R}^r$* ". Methods for solving the problem above are available in the literature and implemented in many mathematical software libraries and packages. We therefore focus our attention on the correct formulation of convenient performance indices, whose maximization corresponds to optimization of sensor accuracy. It will be shown that a general performance index can be defined on the basis of the *Fisher information matrix* associated with the system.

The paper is organized as follows: in section 2 the background of several types of systems for which the sensor design problem may be posed to the designer are reviewed, and the role of the information matrix is illustrated. In section 3 techniques for extracting the desired information out of raw data from transducers are reviewed, with the purpose of providing insight in the sensor design process. Section 4 is dedicated to illustrating design

performance indices that can be used for optimizing the sensor arrangement. In section 5, the optimal design of the sensing system for detecting the deformations of a simple elastic beam is reported as an illustration; a more realistic application to the design of force/torque sensors for robotic manipulators is finally illustrated.

2 General framework

The problem of sensor design may arise in connection with a wide variety of systems, with different structures and characteristics. However, a general framework exists in which most sensor design problems can be studied. For many of the systems considered in this paper, we will refer to a *measurement equation* of the form

$$\mathbf{Y} = \mathbf{M}\mathbf{X} \tag{1}$$

where $\mathbf{X}$ is the vector (state) to be measured, $\mathbf{Y}$ is a data record from measurements, and $\mathbf{M}$ is the measurement operator, that in general depends upon the sensor design in a known relationship.

A first, requisite in the design of sensors for a given system is to guarantee that different states of the system are distinguishable from data: mathematically, that the nullspace of $\mathbf{M}$ is void.

Equation (1) is usually overconstrained: i.e., it consists of more equations than unknowns. Measurement noise and inaccurate modelling of the system lead to perturbations of the measurement equation, and consequently to the inconsistence of (1). Such perturbations will be modelled according to their effects on either measurement data, δ_Y, or on the measurement operator, δ_M. Assuming additive modelling errors, the perturbed measurement equation can be written as

$$\mathbf{Y} = (\mathbf{M} + \delta_M)\mathbf{X} + \delta_Y \tag{2}$$

Under reasonable assumptions, the measurement error δ_Y and the modelling error δ_M are "small" compared to other terms in (2). In solving (2), however, those perturbations are propagated to the solution $\mathbf{X}$ in a way that is basically dictated by $\mathbf{M}^{-1}$, which in turn intrinsically depends on the sensor design. In general, different algorithms will lead to different approximations of the solution. The second goal for the sensor designer, therefore, consists of obtaining the least propagation of measurement and modelling errors in the output, irrespective of the algorithm adopted for the solution. This concept of "absolute" accuracy is well captured in the so-called *Fisher information*

matrix. In the rest of this section the embodiment of the concept of measurement equation and information matrix in some frequently encountered types of system will be introduced.

2.1 State estimation in linear systems

Let the system whose states are to be measured be described by the linear, continuous-time, autonomous differential equation

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{G}\mathbf{u}(t) \quad (3)$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{v}(t) \quad (4)$$

where $\mathbf{x}$, $\mathbf{y}$, $\mathbf{u}$, and $\mathbf{v}$ are n, m, q, and $m-$dimensional vectors, respectively, describing the states, outputs, process (driving) noise, and measurement noise; let the matrices $\mathbf{A}$, $\mathbf{G}$, $\mathbf{C}$ have suitable dimensions. Equation (4) models the multivariate sensory subsystem: the sensor output is related to the system states through the output matrix $\mathbf{C}$, and is affected by measurement noise $\mathbf{v}$. The purpose of sensors is to obtain information on the current state of the system, $\mathbf{x}(t)$, from the measured output function $\mathbf{y}(t)$.

If the number m of (independent) measurements available at any instant of time is equal to, or larger than, the dimension n of the state vector, the measurement equation (1) is easily recovered by putting $\mathbf{Y} = \mathbf{y}(t)$, $\mathbf{M} = \mathbf{C}$, $\mathbf{X} = \mathbf{x}(t)$. Measurement errors are $\delta_Y = \mathbf{v}(t)$ in this case, and perturbations δ_M are due to inaccuracies in modelling the output matrix $\mathbf{C}$. Since in this case the dynamic equation (3) plays no role, this will be referred to as a "static" sensor.

Most often, however, fewer measurements are instantaneously available than state variables, and the dynamics (3) have to be taken into account. The most straightforward approach for reconstructing the state from the outputs in a time-invariant system (where $\mathbf{A}$, $\mathbf{G}$, and $\mathbf{C}$ are constant) consists of repeatedly deriving the measured outputs. From (3) and (4), in the absence of noise ($\mathbf{u}(t) \equiv 0$, $\mathbf{v}(t) \equiv 0$), we have

$$\begin{aligned} \mathbf{y} &= \mathbf{C}\mathbf{x}(t), \\ \dot{\mathbf{y}} &= \mathbf{C}\mathbf{A}\mathbf{x}(t), \\ &\cdots \\ \mathbf{y}^{(p-1)} &= \mathbf{C}\mathbf{A}^{p-1}\mathbf{x}(t). \end{aligned}$$

These equations can be put in the measurement equation form by putting

$$\mathbf{Y} = \left[\mathbf{y}^T \dot{\mathbf{y}}^T \ldots \mathbf{y}^{(p-1)T}\right]^T;$$

$$\mathbf{M} \;=\; \left[\mathbf{C}^T, \mathbf{A}^T\mathbf{C}^T,\; \cdots,\; \mathbf{A}^{(p-1)T}\mathbf{C}^T\right]^T.$$

The measurement matrix $\mathbf{M}$ in this case is what is called, in system-theoretic language, the *observability matrix* (of the second kind) of the system. In the absence of noise, then, the problem of state reconstruction is equivalent to the solution of a linear system of $p \cdot m$ equations in n unknowns. A unique solution can be found if and only if the rank of $\mathbf{M}$ is equal to n: in that case, the system is said to be *completely observable*.

The first requirement in the design of sensors for a given system, i.e. in the choice of $\mathbf{C}$ given $\mathbf{A}$, in this case translates to guaranteeing that the observability condition is met. Classical results of linear system theory find direct application here. From the Cayley-Hamilton theorem, observability can be decided using no more output derivatives than are the unknown states, $p \leq n$. For non-pathological systems, the observability requirement is not very hard to meet: even with a single basic transducer and any t, a sensor design $\mathbf{C}$ can always be found such that the state is completely observable if $\mathbf{A}$ is *non-derogatory* (i.e. if the degree p of the minimal polynomial of $\mathbf{A}$ such that $\mathbf{A}^p + a_1\mathbf{A}^{p-1} + \ldots + a_p = 0$ is equal to the state dimension, $p = n$). In general, the use of $m \geq m_g$ transducers is required completely to observe a dynamic system whose maximum geometric eigenvalue multiplicity is m_g. In the following it is assumed that the above necessary conditions on the number of basic transducers for complete observability of the system are met.

State reconstruction using the derivatives of measurements cannot be applied to time-varying systems, and even for time-invariant systems it is highly unpractical due to the inaccuracy of derivative operations in the presence of noise. A more general method for state reconstruction is based on the estimation of the state at the initial time t_0. In fact, assuming perfect knowledge of the system model, the unperturbed state $\mathbf{x}(t)$ at any time t can be formally reconstructed from knowledge of $\mathbf{x}(t_0) = \mathbf{x}_0$ at time t_0 by using the solution of the differential equation (3)

$$\mathbf{x}(t) = \Phi(t, t_0)\mathbf{x}_0, \tag{5}$$

that is expressed in terms of the *state-transition matrix* $\Phi(t, t_0)$. By measuring the corresponding output function at time instants τ between t_0 and t we have

$$\mathbf{y}(\tau) = \mathbf{C}(\tau)\Phi(\tau, t_0)\mathbf{x}_0$$

and, multiplying both sides by $\Phi^T(\tau, t_0)\mathbf{C}^T(\tau)\mathbf{W}(\tau)$ (here $\mathbf{W}(\tau)$ is a positive definite $m \times m$ weighting matrix) and integrating over $[t_0, t]$, we obtain

$$\mathcal{Y}(t) = \mathcal{M}(t)\, \mathbf{x}_0, \tag{6}$$

where

$$\mathcal{Y}(t) = \int_{t_0}^{t} \Phi^T(\tau, t_0)\mathbf{C}^T(\tau)\mathbf{W}(\tau)\mathbf{y}(\tau)d\tau; \tag{7}$$

$$\mathcal{M}(t) = \int_{t_0}^{t} \Phi^T(\tau, t_0)\mathbf{C}^T(\tau)\mathbf{W}(\tau)\mathbf{C}(\tau)\Phi(\tau, t_0)d\tau. \tag{8}$$

Whenever we are able to evaluate the solution (5), therefore, we may also solve the linear system (6) of n equations for the n unknowns $\mathbf{x}_0$. This can be done if the $n \times n$ matrix $\mathcal{M}$ is invertible. Note that invertibility of $\mathcal{M}$ does not depend on the weight matrix, as long as $\mathbf{W}$ is positive definite for all t. If $\mathbf{W}$ is the identity matrix for all t, $\mathcal{M}$ is called the *observability matrix* (of the first kind) for the system (3)-(4). More generally, (8) presents an instance of the so-called *Fisher information matrix* associated with the given system and sensor, provided that $\mathbf{W}(t)$ is chosen corresponding to the statistic properties of the measurement process (this will be discussed in section 3).

2.2 Calibration Problems

The calibration process of an instrument, consisting in the estimation of a constant n-dimensional parameter vector $\mathbf{x}$ by using an m-dimensional measurement vector $\mathbf{y}$, can be modelling in the framework described above by simply letting $\mathbf{A} = \mathbf{G} = 0$. If $m < n$, calibration is only possible with a time varying observation matrix $\mathbf{C}(t)$. This is the case, for instance, in the calibration of an inertial navigation system [Friedland, 1977]. By integrating over time, or considering discrete measurements as discussed in the subsections above, the calibration problem is again reduced to the solution of an overconstrained measurement equation.

2.3 Discrete time linear systems

In practical applications, the continous time estimation described above may not be feasible, and interest is usually focussed on measurements of the system states taken at discrete instants t_k, $k = 0, 1, 2, \ldots$. Further justification for considering discrete-time sensor models is provided by an

important result of Mehra [1976], stating that for any continuous sensor system design, there exists a discrete design with a finite-length record of measurements that provides the same information on the system (see section 4). The *difference equation* describing the system (3) is in this case

$$\begin{aligned} \mathbf{x}_{k+1} &= \Phi_k \mathbf{x}_k + \Gamma \mathbf{w}_k & (9) \\ \mathbf{y}_k &= \mathbf{C}_k \mathbf{x}_k + \mathbf{v}_k & (10) \end{aligned}$$

where

$$\begin{aligned} \Phi_k &= \Phi(t_{k+1}, t_k) \\ \Gamma \mathbf{w}_k &= \int_{t_k}^{t^{k+1}} \Phi(t_{k+1}, \tau) \mathbf{G}(\tau) \mathbf{u}(\tau) d\tau. \end{aligned}$$

Representing a record of q measurements taken at times t_k, $k = 0, 1, \ldots, q-1$ by a column vector $\mathbf{Y}$, a measurement equation is obtained in the form (1) by putting $\mathbf{X} = \mathbf{x}_0$, and

$$\mathbf{M} = \left[\mathbf{C}_0^T, \Phi_0^T \mathbf{C}_1^T, \Phi_0^T \Phi_1^T \mathbf{C}_2^T, \ \cdots, \ \prod_{j=0}^{q-1} \Phi_j^T \mathbf{C}_q^T \right]^T$$

In the presence of driving and measurement noise, perturbations are introduced as in (2),

$$\mathbf{Y} = (\mathbf{M} + \delta_M)\mathbf{X} + \delta_Y,$$

where

$$\begin{aligned} \delta_Y &= \delta_w + \delta_v \\ \delta_w &= \left[0, (\Gamma_0 \mathbf{w}_0)^T \mathbf{C}_1^T, \cdots, \left(\sum_{j=0}^{q-1} \prod_{i=j+1}^{q-1} \Phi_i \Gamma_j \mathbf{w}_j \right)^T \mathbf{C}_q^T \right]^T \\ \delta_v &= \left[\mathbf{v}_0^T \mathbf{v}_1^T \ldots \mathbf{v}_{q-1}^T \right]^T, \end{aligned}$$

and the modelling error δ_M reflects the fact that the system model (9) and (10) are only known within a limited accuracy.

2.4 Distributed Parameter Systems

If the nature of the dynamical system to which sensors are to be attached is continuous (a distributed parameter system, DPS) rather than representable

by a finite state vector (as in a lumped parameter system, LPS), the ordinary differential equation description (3), (4) must be replaced by a partial differential equation description. For a linear time-invariant DPS, the same formal appearance of differential and output equations (3), (4) can be retained, with the understanding that now $\mathbf{x}$ is infinite-dimensional (i.e. an element of a function space, usually a separable Hilbert space), and that $\mathbf{A}$ represents a differential operator which is supposed to be regular and to generate a semi-group $e^{\mathbf{A}t}$ $(t \geq 0)$. In these hypotheses, solution (5) is applicable to DPS as well. As for the observability of the system state, the relationship (6) holds with the understanding

$$\begin{aligned} \mathcal{Y} &= \int_0^t e^{\mathbf{A}^*\tau}\mathbf{C}^*\mathbf{y}(\tau)d\tau; \\ \mathcal{M} &= \int_0^t e^{\mathbf{A}^*\tau}\mathbf{C}^*\mathbf{W}\mathbf{C}e^{\mathbf{A}\tau}d\tau \end{aligned}$$

(here "*" indicates the adjoint operator). Note that, notwithstanding the resemblance of the formulae in the LPS and DPS case, subtle system-theoretic differences do occur (e.g., in the definition of *complete observability*; also, the information matrix $\mathcal{M}$ may be unbounded for DPS). For a detailed reference on DPS theory, see e.g. Curtain and Pritchard [1978].

Distributed parameter systems usually allow the sensor designer a much richer variety of choices than is affordable in LPS. In relation with the spatial domain where the system is modelled, the shape and extension of the areas where sensing is performed and the distribution of sensing "action"(sensitivity) over such areas are subject to design, along with the location of sensing zones and their number. An arrangement of sensors (possibly including "pointwise" transducers inside or on the boundaries of the system spatial domain) is said "strategic" for the system if the observability condition is fulfilled. El Jay and Pritchard [1987] showed that, for a given sensor location and area, there always exists a distribution of action that makes the sensors strategic; the converse also holds. In practice, it turns out that it is always possible to observe a system even by one sensor (or by $m \leq m_a$ sensors, where m_a is the maximum multiplicity of the eigenfunctions of $\mathbf{A}$) [El Jay, 1991].

For almost every practical purpose, DPS models are dealt with by approximation methods, that ultimately lead to a finite dimensional dynamic system of the type (3)-(4). This can be done either by replacing partial differential equations with their finite-differences approximations, or by using the so-called N-modal approximation. The latter method, that is the

most widely used in optimal sensor design literature (see e.g. Carotenuto and Raiconi [1980]), involves the truncation of the infinite Fourier series expansion of the states and of the observations in its first N terms, according to the increasing order of the partial differential operator eigenvalues. The truncation can be applied at different stages of the optimal design procedure. However, in most cases the final stage of design optimization is applied to a finite-dimensional system whose main characteristics are described by the associated *Fisher information matrix*. Other approaches, such as that of Curtain and Ichikawa [1978] who generalize the information matrix approach to an infinite dimensional "information operator", fall outside the scope of this paper. For a review of optimal sensor placement techniques in DPS, see e.g. Kubrusly and Malebranche [1985].

2.5 Nonlinear Systems

When dealing with nonlinear systems of the type

$$\dot{\mathbf{x}} = f(\mathbf{x}, \mathbf{w}, t), \quad 0 \leq t \leq T \tag{11}$$

with measurements taken at discrete time instants t_k,

$$\mathbf{y}_k = h(\mathbf{x}, t_k) + \mathbf{v}_k, \quad k = 1, 2, \ldots, N$$

the problem of state reconstruction is much more complex than for linear systems, and specific assumptions usually are needed on the class of nonlinearities involved. However, it can be shown that, regardless of the actual reconstruction method used, an upper bound on the observation accuracy can be provided in terms of the *Fisher information matrix* associated with the system. The matrix can be derived, according to Kosut et al. [1982], as

$$\mathcal{M} = \sum_{k=1}^{N} \mathbf{H}_k^T \mathbf{W}_k \mathbf{H}_k$$

where $\mathbf{W}_k$ is the inverse of the covariance matrix associated with $\mathbf{v}_k$,

$$\mathbf{H}_k = \left(\frac{\partial h}{\partial \mathbf{x_0}} \right)_{t=t_k} = \left(\frac{\partial h}{\partial \mathbf{x}} \, \frac{\partial x}{\partial \mathbf{x_0}} \right)_{t=t_k},$$

and the state sensitivity matrix $\frac{\partial \mathbf{x}}{\partial \mathbf{x_0}}$ is the solution of

$$\frac{d}{dt} \left(\frac{\partial \mathbf{x}}{\partial \mathbf{x_0}} \right) = \frac{\partial f}{\partial \mathbf{x}} \, \frac{\partial \mathbf{x}}{\partial \mathbf{x_0}}, \qquad \left(\frac{\partial \mathbf{x}}{\partial \mathbf{x_0}} \right)_{t=0} = \mathbf{I}_n$$

2.6 Other Applications

Techniques for optimal design of sensors in both dynamical and static systems have been developed and used in many more fields than can be accounted for in this paper. One field where an interesting body of theory has been developed is optimal design of inputs (or "experiments") for model identification. The problem of optimally planning experiments for inferring unknown parameters has been extensively treated in statistics, see e.g. Fedorow [1972], and in the automatic control and system theoretic literature [Mehra, 1974; Goodwin and Payne, 1977; see also Ljung, 1987]. Optimal design of sensors for identification has been particularly well studied in the context of distributed parameter systems (see e.g. the survey paper of Kubrusly and Malebranche [1985]). The optimal sensor location problem has been studied for detecting sensor or actuator failures [Watanabe *et al.*, 1985], and more generally for detecting changes in the dynamical behavior of systems [Basseville *et al.*, 1987].

Furthermore, the optimal location of sensing or detecting devices has been studied with reference to radar system design [Banach and Cunningham, 1988], oceanic engineering [Wilson, 1988], to robotics [Cameron and Durrant-White, 1990; Menq *et al.*, 1989; Schroer *et al.*, 1992], and more generally to sensor data fusion problems [Clark and Yuille, 1990].

3 Solving the Measurement Equation

In the section 2, it has been shown that a wide variety of problems in multivariate sensing can be formulated in terms of the usually overconstrained measurement equation (1),

$$\mathbf{Y} = \mathbf{MX}.$$

Considering only the fundamental linear, finite-dimensional case, the kernel of the measurement operator is a $d \times m$ matrix $\mathbf{M}$. We proceed to summarize the basic tools used to solve linear measurement equations.

3.1 Least Squares Solution

The method consists of fitting data $\mathbf{Y}$ with a hyperplane in the model space, trying to minimize the sum of squared misfit errors. The idea has been applied to virtually every scientific field since circa 1800, and the point in presenting it here is to emphasize some aspects related with the measurement of physically dimensioned quantities. The *weighted least squares* solution to

(1) is defined as the vector $\mathbf{X}_{ls}$ such that the vector of deviations, $\mathbf{Y}-\mathbf{M}\mathbf{X}_{ls}$, is minimum in the $\mathbf{W}$-*weighted 2-norm* sense, i.e.

$$\mathbf{X}_{ls} = \arg\min_{\mathbf{X}} (\mathbf{Y}-\mathbf{M}\mathbf{X})^T \mathbf{W} (\mathbf{Y}-\mathbf{M}\mathbf{X}).$$

The least-squares estimate can be obtained through elementary calculus as

$$\mathbf{X}_{ls} = (\mathbf{M}^T\mathbf{W}\mathbf{M})^{-1}\mathbf{M}^T\mathbf{W}\mathbf{Y}.$$

The numerical values of the weight matrix are subject to the constraint that $\mathbf{W}$ is symmetric and positive definite (no meaning can be associated to the least-squares problem otherwise). The physical dimensions of the entries in $\mathbf{W}$ are also constrained: they must be such that the sum of weighted squares to be minimized is carried over physically homogeneous quantities. If $[[\alpha]]$ denotes the physical dimension of the quantity α, then the dimension of the weights must satisfy $[[\mathbf{W}_{i,j}]] = [[k]]\,[[\mathbf{Y}_i]]^{-1}[[\mathbf{Y}_j]]^{-1}$, where $[[k]]$ are arbitrary dimensions. Without loosing in generality, it can be assumed that norms are always adimensional, i.e. that $[[k]] = 1$. Note also that $\mathbf{M}$ has physical dimensions, since $[[\mathbf{M}_{i,j}]] = [[\mathbf{Y}_i]][[\mathbf{X}_j]]^{-1}$. As a consequence of the physical dimensionality of the problem, bare (not weighted) least squares hardly ever make sense in practical problems. Even though one may have no idea of the right numerical weights to use, and would like to use say the identity weight matrix $\mathbf{I}$, one must be careful to attach to its elements the correct physical dimensions: this will prevent different results to derive from the same problem when described with different metric units.

3.2 Bayesian Inference

The least-squares method is a purely deterministic approach to the solution of the measurement problem. A fundamental drawback of the method is in the choice of weights: the method does not provide hints so as how to choose $\mathbf{W}$. Yet, the weight matrix deeply affects the final solution. The intuition is that weights should be chosen to take into greater account accurate measurements and discard noisy data. However, the idea can only be formalized in a probabilistic framework, such as that provided by Bayes' theory of inference. The basis of the Bayesian approach to the solution of measurement equations is in defining the statistical properties of the *data space* and of the *model space*. For a finite-dimensional linear measurement system, these are vectorial spaces (Y and X, respectively) of suitable dimensions (both algebrical and physical), where vectors $\mathbf{Y}$ and $\mathbf{X}$ take their actual values,

and where probability density functions (p.d.f.) are defined for the variables of interest. The *a priori* state of information consists in a p.d.f. defined over the model space X, $f_{prior}(X)$, which models any knowledge one may have on the system model parameters independently from the present act of measurement, due e.g. to physical insight or to independent measurements carried out previously. In the formation of measurement data, two information sources are to be considered, i.e. the forward solution of the physical model, and the act of measuring itself. The state of information on the experimental uncertainties in measurement outputs can be modelled by means of a p.d.f. $f_{exp}(\mathbf{Y})$ over Y (this is usually provided by the instrument supplier), while modelling errors (due e.g. to imperfect knowledge of $\mathbf{M}$ in (1)) can be represented by a conditional p.d.f. $f_{mod}(\mathbf{Y}|\mathbf{X})$ in the data space Y (or, more generally according to Tarantola [1987], by a joint p.d.f $f_{mod}(\mathbf{Y}, \mathbf{X})$ over $X \times Y$).

Fusing the different information in an estimate of $\mathbf{X}$ leads to a *posterior* p.d.f over X, that is described by Bayes' formula

$$f_{post}(\mathbf{X}) = f(\mathbf{X}|\mathbf{Y}) = \alpha_b \; f_{prior}(\mathbf{X}) \int_Y f_{exp}(\mathbf{Y}) f_{mod}(\mathbf{Y}|\mathbf{X}) d\mathbf{Y}, \qquad (12)$$

where α_b is a normalization factor such that $\int_X f_{post}(\mathbf{X}) d\mathbf{X} = 1$. The process of information fusion is described in figure 1, adapted from Tarantola [1987], with reference to the most general case where the measurement equation $\mathbf{Y} = \mathbf{M}(\mathbf{X})$ is nonlinear.

Although the posterior p.d.f. on the model space represents the most complete description of the state of information on the quantity to be measured one may wish, a final decision on what is the "best" estimate of $\mathbf{X}$ needs usually be taken. Several possibilities arise in general, depending upon the adopted criterion of optimality:

- the *maximum a posteriori* estimate (MAP) corresponds to choosing the model value $\mathbf{X}_{map}$ with largest posterior p.d.f, such that $f_{post}(\mathbf{X}_{map}) \geq f_{post}(\mathbf{X})$ for all $\mathbf{X} \in X$;

- the *maximum likelyhood* estimate (MLE) maximizes the probability of the observed value of the output, given the candidate solution $\mathbf{X}$, so that $f(\hat{\mathbf{Y}}|\mathbf{X}_{mle}) \geq f(\hat{\mathbf{Y}}|\mathbf{X})$, $\forall \mathbf{X} \in X$. If the a priori p.d.f. $f_{prior}(\mathbf{X})$ is non-informative (in most cases that means uniform over X), MAP and MLE estimates coincide;

- the *minimum variance* estimate (MVE) (or *minimum mean square*, MMSE) minimizes the weighted ℓ_2-norm of deviations, i.e. the quan-

tity $\int_X (\hat{\mathbf{X}} - \mathbf{X})^T \mathbf{W} (\hat{\mathbf{X}} - \mathbf{X})\ f(\mathbf{X}|\mathbf{Y}) d\mathbf{X}$. The MVE can be shown to be independent from the (arbitrary, positive definite) matrix $\mathbf{W}$, and equals the *conditional mean* of $\mathbf{X}$ given $\mathbf{Y}$.

Figure 1-e illustrates these estimates. While very little can be said in general about the performance of such estimators, well known particularizations apply under certain assumptions on the prior distributions:

- If both $f_{prior}(\mathbf{X})$ and $f_{exp}(\hat{\mathbf{Y}}|\tilde{\mathbf{Y}})$ are generalized Gaussian distributions of order p, and *if modelling errors are negligible*, the MAP estimator can be easily evaluated from the measurement output and the prior expectation (see for instance Tarantola [1987]). For $p = 1$, the MAP estimator equals the *least-absolute-value* estimate (in the weighted ℓ_1-norm sense), minimizing the quantity $\sum_i \frac{|(\mathbf{MX})_i - \hat{\mathbf{Y}}_i|}{w_{d,i}} + \sum_j \frac{|\mathbf{X}_j - \mathbf{X}_{prior,j}|}{w_{m,j}}$. For $p = \infty$, the MAP estimate corresponds to the *minimax-absolute-value* estimate (in the weighted ℓ_∞-norm sense), that minimizes

$$\max \left\{ \max_i \frac{|(\mathbf{MX})_i - \hat{\mathbf{Y}}_i|}{w_{d,i}} ; \max_j \frac{|\mathbf{X}_j - \mathbf{X}_{prior,j}|}{w_{m,j}} \right\}.$$

- If a normal distribution (an order-2 Gaussian) can be assumed for all prior information, the MAP estimate enjoys many useful properties. First (and perhaps most importantly for the problem of optimal sensor design), since the convolution (in (12)) of two Gaussian distributions is Gaussian, the modelling and experimental errors in measurements simply combine by addition of the covariance matrices of experimental and modelling errors, $\mathbf{C}_Y = \mathbf{C}_{exp} + \mathbf{C}_{mod}$ [Tarantola, 1987]. Roughly speaking, errors in the knowledge of $\mathbf{M}$ (calibration errors) can be ignored provided that experimental errors in $\mathbf{Y}$ are suitably increased. This result holds for nonlinear sensor models as well. For linear models, the a posteriori p.d.f. is also Gaussian, the MVE and MAP estimates coincide and evaluate to

$$\hat{\mathbf{X}} = \mathbf{C}_{post}(\mathbf{M}^T \mathbf{C}_Y^{-1} \mathbf{Y} + \mathbf{C}_{prior}^{-1} \mathbf{X}_{prior}), \tag{13}$$

$$\mathbf{C}_{post} = (\mathcal{M} + \mathbf{C}_{prior}^{-1})^{-1}, \tag{14}$$

where $\mathcal{M}$, the *Fisher information matrix* for linear measurement equations, is defined as

$$\mathcal{M} = \mathbf{M}^T \mathbf{C}_Y^{-1} \mathbf{M}. \tag{15}$$

Solution (13) is the minimizer of a quadratic form over the product space $Y \times X$ (*misfit* function), and is apparently a generalization of least squares, where covariance matrices are chosen as weights (covariance matrices inherently satisfy the physical dimensions constraints above discussed). As a final remark, the Gauss-Markov theorem [Rao, 1973] ensures that the estimate (13) is the best linear unbiased estimate (BLUE) in the minimum-variance sense even for non-Gaussian a priori distributions. This result may seem to indicate some "absolute optimality" of the least-squares estimate. However, the MVE of a non-Gaussian distribution may not be a significant estimate, as apparent in figure 1-e. This is the case for instance when a few measurements are grossly in error (*outliers*): the MVE in this case can provide meaningless results. This fact is sometimes used to point out the *lack of robustness* of the MVE.

3.3 Recursive Methods

Most often sensors that are part of dynamical systems must provide new updates of model estimates without referring to the whole story of sensed data. A generalized Gaussian posterior distribution over X corresponding to a given record of data, for instance, is completely described by its mean and dispersion estimators (covariance). When a new datum is available, all prior information can be extracted from those statistics. A method that do not use prior information explicitly, but through its statistics only, is called *recursive*. The *Kalman filter* for state estimation in discrete-time systems (9) is the recursive implementation of the MVE solution above discussed. Its continuous-time version is the optimal (in the MV sense) observer for a linear system subject to uncorrelated, zero-mean, Gaussian white noise disturbances. The state estimate of a Kalman filter for the system (3)-(4), where the driving and measurement noise processes have the (possibly time-varying) spectral density matrices $\mathbf{Q}$ and $\mathbf{R}$, respectively, evolves from the initial guess $\hat{\mathbf{x}}(0) = \hat{\mathbf{x}}_0$ according to

$$\dot{\hat{\mathbf{x}}} = \mathbf{A}\hat{\mathbf{x}} + \mathbf{P}\mathbf{C}^T\mathbf{R}^{-1}(\mathbf{y} - \mathbf{C}\hat{\mathbf{x}}),$$

where the estimate covariance matrix $\mathbf{P}$ obeys the Riccati differential equation:

$$\begin{aligned} \dot{\mathbf{P}} &= \mathbf{A}\mathbf{P} + \mathbf{P}\mathbf{A}^T + \mathbf{G}\mathbf{Q}\mathbf{G}^T - \mathbf{P}\mathbf{C}^T\mathbf{R}^{-1}\mathbf{C}\mathbf{P} \qquad (16) \\ \mathbf{P}(0) &= \mathbf{P}_0. \end{aligned}$$

The role played by the sensor design (i.e., the choice of $\mathbf{C}$), and also by its time scheduling (which affects $\mathbf{R}(t)$, as discussed by Mehra [1976]), in the evolution of the estimate covariance $\mathbf{P}$, is not easily analyzed directly from (16). Design (and scheduling) performance indices have therefore to be based on either a steady-state solution of (16), which exists and is unique for completely controllable and observable time-invariant systems, as for instance in Arbel [1982]; or, more generally, on bounds on $\mathbf{P}$. The Cramèr-Rao [Rao, 1973] bounds are written in the linear case as

$$[\mathcal{M}_R(t_0,t)+\mathcal{N}_Q^{-1}(t_0,t)]^{-1} \leq \mathbf{P}(t_0,t) \leq \mathcal{M}_R^{-1}(t_0,t)+\mathcal{N}_Q(t_0,t) \qquad (17)$$

where $\mathcal{M}_R$, the $\mathbf{R}(t)$-weighted observability matrix according to (8), is the *Fisher information matrix* for this problem, and

$$\mathcal{N}_Q = \int_{t_0}^{t} \Phi(t,\tau)\mathbf{G}(\tau)\mathbf{Q}(\tau)\mathbf{G}^T(\tau)\Phi^T(t,\tau)d\tau$$

is independent of the sensor design and scheduling. According to this observation, Mehra [1976] proposes to consider the minimization of $\mathcal{M}_R^{-1}$ as an instrument to minimize $\mathbf{P}$. This is further justified by the fact that, in the absence of process noise ($\mathbf{Q}=0$) and of prior information ($\mathbf{P}_0^{-1}=0$), the Riccati equation solution is exactly $\mathbf{P}(t_0,t)=\mathcal{M}_R^{-1}(t_0,t)$.

Cramèr-Rao bounds on estimate covariance for non-linear systems have also been applied to optimal sensor design problems (see e.g. Kosut *et al.* [1982]). In the context of non-linear systems, however, it must be emphasized that minimum-variance estimates do not enjoy the properties that make them desirable for linear systems, and MVE-based optimal sensor design is often questionable.

4 Design Performance Indices

The core of sensor optimizing design is in the choice of suitable performance indices whose minimization (or maximization) over the set of allowed design parameters provides the sought optimum. Such indices must conform to two fundamental requisites: they must be capable of embodying the desirable properties of the sensor, and they must be computationally efficient. The second requirement obviously follows from their use in practical optimization algorithms, whose numerical convergence properties and time requirements strongly depend on the index.

4.1 Linear-Gaussian hypothesis

From the assumption that white, zero-mean, Gaussian noise processes act on a continuous or discrete time linear system, it is possible to obtain a rather complete description of how design parameters affect the overall reconstruction accuracy. In such a case, in fact, the posterior p.d.f. in the model space X is Gaussian itself, with mean value and covariance operator as given by (13). A particular estimate $\hat{\mathbf{X}}$ from a given data record is a random variable itself, which is asymptotically normally distributed about the "true" value $\mathbf{X}_0$ with covariance $\mathbf{C}_{post}$. A useful interpretation of that result can be given in terms of the so-called *confidence ellipsoids*. To do so, let us first assume that the measurement equation only contains physically adimensional quantities, and that no prior information is available. The confidence ellipsoids are subsets of X described in terms of the *Fisher information matrix* $\mathcal{M} = \mathbf{M}^T \mathbf{C}_Y^{-1} \mathbf{M}$ as

$$E_\alpha = \left\{ \hat{\mathbf{X}} \in X | (\mathbf{X}_0 - \hat{\mathbf{X}})^T \mathcal{M} (\mathbf{X}_0 - \hat{\mathbf{X}}) \leq \alpha^2 \right\}$$

The probability that the MVE solution falls inside E_α is a tabulated function of α. A confidence interval for the i-th component of $\hat{\mathbf{X}}$ can be easily derived from its standard deviation $\mathcal{M}_{ii}^{-1/2}$ and statistical tables (see figure 2). The shape of the ellipsoids of the family depends on the eigenstructure of the posterior covariance operator. If $\mathcal{M} = \mathbf{U}\Lambda\mathbf{U}^T$ is the diagonal decomposition of $\mathcal{M}$, the length of the major axis of the ellipsoid is the inverse of the square root of the minimum eigenvalue $\Lambda_{1,1}$ and lies in the direction of the corresponding eigenvector $\mathbf{U}_1$ (the first column of $\mathbf{U}$); the minor axis length is the inverse of the square root of the maximum eigenvalue of $\mathcal{M}$, and so forth (see figure 2). Accordingly, the direction of the major axis corresponds to the least estimate accuracy in X. Other parameters such as componentwise Cramèr-Rao bounds CR_i and "geometric dilution of precision" (GDOP) can be obtained from confidence ellipsoids (see e.g. Maine and Iliff [1981]).

In the general case that priori information is available, and that physical dimensions are present, rather than with the eigenstructure of the information matrix, one should deal with the generalized eigenvalue-eigenvector problem $\mathcal{M}\mathbf{v} = \lambda \mathbf{C}_{prior}^{-1} \mathbf{v}$. The direction of the major axis of the corresponding ellipsoid is the combination of unknowns which is resolved best from the given measurements. Equivalently, one may refer to the *normalized measurement equations*,

$$\bar{\mathbf{Y}} = \bar{\mathbf{M}} \bar{\mathbf{X}} \tag{18}$$

where the coordinates are changed according to

$$\bar{\mathbf{Y}} = \mathbf{T}_Y \mathbf{Y}; \tag{19}$$
$$\bar{\mathbf{X}} = \mathbf{T}_X \mathbf{X}; \tag{20}$$
$$\bar{\mathbf{M}} = \mathbf{T}_Y \mathbf{M} \mathbf{T}_X^{-1}; \tag{21}$$

The linear transformations $\mathbf{T}_Y$, $\mathbf{T}_X$ are chosen so that, in the new coordinates, the prior information covariance matrix and the measurement covariance matrix are identity matrices of suitable order; moreover, the information matrix is diagonal. Such transformations can be calculated as follows:

$$\begin{aligned} \mathbf{T}_Y &= \mathbf{C}_Y^{-1/2} \\ \mathbf{T}_X &= \mathbf{S}^T \mathbf{C}_{prior}^{-1/2}, \end{aligned}$$

where $\mathbf{S}\bar{\mathbf{C}}_{post}\mathbf{S}^T$ is a diagonal decomposition of $\mathbf{C}_{prior}^{-1/2}\mathbf{C}_{post}\mathbf{C}_{prior}^{-T/2}$. Normalized measurement equations only involve adimensional quantities, and have the desirable property of being unique and invariant with changes of variables. Moreover, the confidence ellipsoid geometry is completely described by the diagonal elements of $\mathcal{M}$. Finally, note that also the posterior covariance matrix $\bar{\mathbf{C}}_{post}$ is diagonal in the new coordinates.

From the properties of confidence ellipsoids, it appears that virtually any reasonable design performance index for a linear-Gaussian system sensor can be related to the ellipsoid shape and volume. Natural candidates for maximization are the determinant (proportional to the product of the eigenvalues), the trace (sum of the eigenvalues), and the minimum eigenvalue of $\mathcal{M}$. Müller and Weber [1972] embedded the above indices in a general form

$$I_s = \left(\frac{1}{n} \mathrm{tr} \mathcal{M}^s \right)^{1/s}, \qquad s \geq 0,\ n = \text{dimension of } X$$

which specializes as

$$\begin{aligned} \lim_{s \to 0} I_s &= \sqrt[n]{\mathsf{det}\ \mathcal{M}} \quad \text{(determinant index)} \\ I_1 &= \frac{n}{\mathsf{trace}(\mathcal{M}^{-1})} \quad \text{(trace index)} \\ \lim_{s \to \infty} I_s &= \lambda_{min}(\mathcal{M}) \quad \text{(min eigenvalue index).} \end{aligned}$$

Obviously, constraints on feasible design parameters are in order in any genuine engineering problem, that make the above maximization problems practically meaningful.

Let us first consider the optimal allocation in time of sensing accuracy for continuous time linear systems (measurement scheduling problem). Assume that the overall amount of accuracy we can use for estimating the initial state by means of measurements between times t_0 and t_1 is bounded by $\int_{t_0}^{t_1} \mathbf{R}^{-1}(t)dt \leq A$. As already mentioned, Mehra [1976] has shown that, among all feasible schedules $\mathbf{R}(t)$ that optimize a given index, there is a discrete schedule $\mathbf{R}^*(t) = \mathbf{R}_i\delta(t - t_i)$ consisting of measurements at a finite number of instants t_i only, with $t_0 \leq t_i \leq t_1$ and $\sum \mathbf{R}_i^{-1} = A$. Roughly speaking, this result allows us to replace a continuous-time measurement process such as (4) with a discrete process (10) for analysis purposes. The same author presents an algorithm for choosing the optimal discrete schedule.

A general formulation of the optimization of sensor design with respect to a set of parameters ξ (e.g., the physical parameters of location and orientation of transducers in the system) that enter in the measurement matrix $\mathbf{C}$ through functions $\mathbf{C}_{i,j} = f_{i,j}(\xi)$ and are subject to constraints $g_k(\xi) \leq 0$, is not amenable to analysis, and is only tractable by numerical optimization methods (see example 2 below). Note that the problem will in general have a multiplicity of locally optimal solutions, and that randomized search algorithms (such as simulated annealing or Monte Carlo methods) may recommend for such cases. In the assumption that the elements of the measurement matrix can be chosen directly by the designer, and that the design constraints can be formulated as a norm condition on the measurement matrix itself, as e.g. $\|\mathbf{C}\| \leq 1$, the solution of the optimization problem has been shown by Müller and Weber [1972] to correspond to the solution of a nonlinear eigenvalue problem. Simple results are only available for single-input, single-output, time-invariant systems; even then, however, numerical analysis techniques are ultimately required to obtain the optimal design [Mehra, 1976]. It is important to note that a given optimal design under an index I_{s_1} is not in general optimal also under I_{s_2} [Müller and Weber, 1972].

4.2 Worst-Case Design

The design indices based on the linear-Gaussian hypotheses described in the above section are well established design tools for multivariate sensors that obey the hypotheses. Two fundamental theorems of statistics, namely the central limit teorem and the Gauss-Markov theorem, guarantee to some extent the extension of those methods to cases that depart from the strict

hypotheses, due e.g. to some degree of nonlinearity. However, there do exist cases where other indices may be considered by the designer. One motivation is found in high-responsibility systems subject to spurious disturbances (outliers), where the designer's primary concern is to minimize worst-case sensing errors rather than error statistics. As already noted, MV estimators may lack robustness to outliers, and related optimal design criteria may not be the preferable choice.

Explicit worst-case/deterministic bounds on the errors, by which a system's states or parameters are estimated, are also required in conjunction with some specific controller designs for the system, such as the currently popular robust H_∞ controllers. Although the theory of optimal worst-case/deterministic estimation and/or identification is much more recent than that of MVE, several important results have appeared in literature (see e.g. Belforte *et al.* [1987], Tempo [1988], and Helmicki *et al.* [1991]). However, application of H_∞ techniques to optimal design of sensors have not received much attention to date. Finally, note again that in the presence of non-Gaussian modelling errors the MVE approach and related optimization results are not really founded on a justifiable rationale; however, the simple results for worst-case optimal design of sensors that are described in the following apply to that case directly. Considering the normalized measurement equation (18) and omitting overbars for simplicity, let $\mathbf{X}_0$ be the corresponding "true" solution, and let $\mathbf{X}$ be the least-squares solution of the same equation with modelling and measurement errors added:

$$\mathbf{Y} = (\mathbf{M} + \delta_M)\ \mathbf{X} + \delta_Y.$$

We are interested in providing a bound on the maximum difference between $\mathbf{X}_0$ and $\mathbf{X}$, given bounds on the largest values attainable by perturbations $\delta_\mathbf{Y}$ and δ_M. The problem is clearly related to perturbation theory in numerical computation, in particular to Wilkinson's theory of condition [1965]. Among the various slightly different results available from the literature, we briefly describe one from Golub and VanLoan [1989] because of its simplicity and the insight it provides.

Assume that $\mathbf{M}$ is full rank, that $\|\mathbf{M}\mathbf{X}_0 - \mathbf{Y}\| = 0$, and that

$$\epsilon = max \left\{ \frac{\|\delta_M\|}{\|\mathbf{M}\|}, \frac{\|\delta_Y\|}{\|\mathbf{Y}\|} \right\} < \frac{\sigma_{min}(\mathbf{M})}{\sigma_{max}(\mathbf{M})},$$

where $\sigma_{min}(\mathbf{M})$ and $\sigma_{max}(\mathbf{M})$ are the minimum and maximum singular values of the measurement matrix $\mathbf{M}$, respectively. Note that in normalized

equations these correspond to the square root of the eigenvalues of the *Fisher information matrix* $\mathcal{M}$. Moreover, put $\mathbf{E} = \delta_M/\epsilon$ and $\mathbf{e} = \delta_Y/\epsilon$, and let $0 < s < \epsilon$ parametrize a class of measurement equations with increasing perturbations, whose least-squares solution $\mathbf{X}(s)$ is given by

$$(\mathbf{M} + s\mathbf{E})^T \ (\mathbf{M} + s\mathbf{E}) \ \mathbf{X}(s) = (\mathbf{M} + s\mathbf{E})^T \ (\mathbf{Y} + s\mathbf{e}),$$

Differentiating this solution with respect to s at $s = 0$ we obtain

$$\mathbf{E}^T \ \mathbf{M} \ \mathbf{X}(0) + \mathbf{M}^T \ \mathbf{E} \ \mathbf{X}(0) + \mathbf{M}^T\mathbf{M} \ \frac{d\mathbf{X}}{ds}\bigg|_{s=0} = \mathbf{E}^T\mathbf{Y} + \mathbf{M}^T\mathbf{e},$$

and, from the series expansion about $\mathbf{X}(0) = \mathbf{X}_0$,

$$\mathbf{X} = \mathbf{X}(\epsilon) = \mathbf{X}_0 + \epsilon \ \frac{d\mathbf{X}(s)}{ds}\bigg|_{s=0} + O(\epsilon^2).$$

Using the above assumptions and taking norms, we obtain

$$\begin{array}{rcl} \|\mathbf{X}_0 - \mathbf{X}\| & \leq & \epsilon\|(\mathbf{M}^T \ \mathbf{M})^{-1}\mathbf{M}^T\| \ (\|\mathbf{e}\| + \|\mathbf{E}\| \ \|\mathbf{X}_0\|) + O(\epsilon^2) \\ & = & \sigma_{min}^{-1}(\mathbf{M}) \ (\|\delta_Y\| + \|\delta_M\| \ \|\mathbf{X}_0\|) + O(\epsilon^2). \end{array} \tag{22}$$

In terms of relative errors we have the well known result

$$\frac{\|\mathbf{X}_0 - \mathbf{X}\|}{\|\mathbf{X}_0\|} \leq 2\kappa\epsilon + O(\epsilon^2), \tag{23}$$

where $\kappa = \sigma_{max}(\mathbf{M})/\sigma_{min}(\mathbf{M})$ is the *condition number* of $\mathbf{M}$, i.e. the square root of the condition number of the *Fisher information matrix* associated with the system. Note that a sharper result that avoids the dependency on ϵ being "small enough" can be obtained at the price of a more involved proof (see e.g. Stewart [1977]). Thus, in the worst case, relative errors on measurements and modelling are propagated to the solution by a factor $\kappa(\mathbf{M})$ (by definition, $\kappa \geq 1$). Note that this "worst case" bound is a pessimistic estimate of errors in the sense that it has zero mathematical likelihood of being attained. However, there do exist cases when the bound is essentially reached, as shown in the following example:

$$\mathbf{M} = \begin{bmatrix} 0.9826 & 0.6515 \\ 0.7227 & 0.07269 \\ 0.7534 & 0.6316 \end{bmatrix}, \mathbf{Y} = \begin{bmatrix} 0.708805 \\ 0.394791 \\ 0.584581 \end{bmatrix},$$

with perturbations

$$\delta_M = 10^{-3} \begin{bmatrix} -0.200244 & -0.121044 \\ 1.26751 & 0.766189 \\ -0.613201 & -0.370671 \end{bmatrix}, \delta\mathbf{Y} = 10^{-3} \begin{bmatrix} 0.140798 \\ -0.891222 \\ 0.43116 \end{bmatrix}.$$

In this case, $\kappa(\mathbf{M}) = 4.75231$, $\epsilon = 10^{-3}$ and $\frac{\|\mathbf{X}_0 - \mathbf{X}\|}{\|\mathbf{X}_0\|} = 0.00950462 = 2\kappa\epsilon$.

Note that the absolute worst-case bound (22) suggests to maximizing the minimum eigenvalue of the *Fisher information matrix* of the sensor, that is also one of the design performance indices obtained from statistic analysis in the linear-Gaussian hypothesis. Minimization of the condition number, on the other hand, minimizes the worst-case relative error (analogous in a sense to a noise-to-signal ratio). In terms of relative errors, it may happen that adding measurements is not rewarding from an overall accuracy point of view, and it might even be counterproductive. This is shown in the following most simple example. Let the measurement matrix be the 2 by 2 identity matrix, that is obviously best conditioned ($\kappa = 1$), and assume that a measurement is added that is equal to the second one. The new measurement matrix and a possible vector of measurements are

$$\mathbf{M} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}, \ \mathbf{Y} = \begin{bmatrix} 0 \\ \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}.$$

The worst-case perturbations are in this case

$$\delta\mathbf{M} = 10^{-3} \begin{bmatrix} 0 & -\sqrt{2}/2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \ 10^{-3} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

Hence, in this case $\kappa(\mathbf{M}) = \sqrt{2}$, $\epsilon = 10^{-3}$ and $\frac{\|\mathbf{X}_0 - \mathbf{X}\|}{\|\mathbf{X}_0\|} = 2\kappa\epsilon = 2\ 10^{-3}\ \sqrt{2}$.

The seemingly contradictory fact that a criterion based on the condition number may be lowered by adding measurements, while more measurements can only increase the information on the system, reflects the geometric, rather than statistical, nature of condition number criteria. Sensors are optimal with respect to the condition number criterion if every state is equally well resolved from the measurement, irrespective of the absolute level of accuracy. This characteristic of the criterion allows in some cases decomposition of the optimization problem in two phases, dedicated to maximization of relative and absolute errors, respectively (see example 2).

5 Design Examples

Example 1

Consider the design of a sensing system for estimating the state of a thin bar pinned at its extremities (see figure 3).

The beam dynamics represent a simple example of distributed parameter system, and are modelled by the Euler-Bernoulli partial differential equation

$$m\frac{\partial^2 u(x,t)}{\partial t^2} - 2\xi\frac{\partial^2}{\partial x^2}\frac{\partial u(x,t)}{\partial t} + EI\frac{\partial^4 u(x,t)}{\partial x^4} = 0$$

for $0 \leq x \leq L$ and $t \geq 0$, where $u(x,t)$ is the transverse displacement of the beam and the applied force distribution is zero. For convenience, put the beam parameters m (mass per unit length), I (moment of inertia), E (modulus of elasticity), and L (length between supports) to unity, while the damping coefficient is $\xi = 0.05$. Position transducers are to be optimally placed at locations x_i, $0 \leq x_i \leq 1$ along the beam to reconstruct the state of the system. For the boundary conditions imposed by pinned supports, modal frequencies are $\omega_k = (k\pi)^2$ and modal shapes are $\phi_k(x) = sin(k\pi x)$. If a 3-modal approximation is adopted, a finite-dimensional dynamic model can be written as

$$\begin{aligned} \dot{\mathbf{v}}(t) &= \mathbf{A}\mathbf{v}; \\ \mathbf{y} &= \mathbf{C}\mathbf{v} \end{aligned}$$

where

$$\mathbf{v} = \begin{bmatrix} u_1(t) \\ u_2(t) \\ u_3(t) \\ \dot{u}_1(t) \\ \dot{u}_1(t) \\ \dot{u}_1(t) \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 0 & \mathbf{I}_3 \\ -\Lambda^2 & -2\xi\Lambda \end{bmatrix}$$

$$\Lambda = \begin{bmatrix} \pi^2 & 0 & 0 \\ 0 & 4\pi^2 & 0 \\ 0 & 0 & 9\pi^2 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} \phi_1(x_1) & \phi_2(x_1) & \phi_3(x_1) & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \phi_1(x_m) & \phi_2(x_m) & \phi_3(x_m) & 0 & 0 & 0 \end{bmatrix}.$$

From the above continuous-time approximation a discrete-time model is then obtained according to (9) and (10), with constant sampling time $T_s = 1/40$ sec, as $\Phi = e^{\mathbf{A}T_s}$. Sensor measurements and model are affected by normally distributed zero-mean errors with covariance matrix $\mathbf{C}_Y$. Assuming that

the measurement equation is in normalized form, i.e. that a coordinate transform has been applied such that $\mathbf{C}_Y = \mathbf{I}_m$, optimization criteria will be applied to the measurement matrix $\mathbf{M} = \left[\mathbf{C}^T, \Phi^T\mathbf{C}^T, \ \ldots \ , \Phi^{5T}\mathbf{C}^T\right]^T$.

Let first consider the optimal location of a single sensor ($m = 1$). In figure 4 the determinant, trace, minimum singular value, and inverse of the condition number of the information matrix are plotted versus sensor position along the beam. Owing to symmetry, only one half of the beam is considered; the ordinate scales have been normalized for comparison purposes. It clearly results that the complete observability condition is met by any sensor location different from the zeros (nodes) of the mode shapes ϕ_i (located at $\mathbf{x} = 1/3$, $\mathbf{x} = 1/2$, and $\mathbf{x} = 2/3$). This is in accordance to the results of El Jay and Pritchard [1987], by which a single sensor can be "strategic" if the maximum multiplicity of the eigenvalues is one as in this case (one sensor may be insufficient if rigid body modes are present). It can be observed from figure 4 that the optimal sensor location varies with the index considered. Corresponding data are reported in table 6.

In figure 5 plots of the four indices are reported on a mesh of locations for two position sensors placed on one half of the beam length. Optimal design data are reported in table 6. For identical locations of the two sensors on the beam (i.e., along the diagonals of the mesh plots), the indices are exactly twice those obtained with one sensor (except the inverse of condition number one that does not change), and follow the same curves shown in figure 4. The increase in accuracy is much larger if different locations are chosen.

Example 2

As a second example, the results of the optimized design of a more complex device, a six-axis force-torque sensor, is reported. The design of multi-axis force sensors, i.e. instruments for measuring several (up to 6) components of force and torque simultaneously, is relevant to different fields such as wind-tunnel testing, adaptive control of machines and thrust-stand testing of rocket engines. Starting from the mid seventies, applications to robotics and telemanipulation focussed much attention. A particular application to the design of miniaturized robotic contact sensors to fit the phalanges of a robot hand is described in Bicchi [1992], and briefly reported here.

Force-torque sensing is considered here as a static problem, where the relationship between the unknowns (six normalized components of the resultant fiorce and torque applied on the sensor) and the measurements (m

transducers, usually strain-gauges) is linear. Small deviation from linearity can actually be accomodated for by nonlinear compensation techniques [Hirose and Yoneda, 1990]; for simplicity however, nonlinear compensation is neglected in the optimal design described in the following. Design parameters directly affect the elements of the measurement matrix $\mathbf{M} = \mathbf{C}$. The mechanical structure of the sensor is sketched in figure 6: it consists of a hollow, thin-walled cylinder. Strain-gauges are applied on the external surface of the cylinder. The cylinder dimensions, the position of the gauges on its structure and their orientation are the variables to be optimized. Figure 6 shows the $\{O, \zeta_1, \zeta_2, \zeta_3\}$ reference frame in which the components of the load applied to the sensor extremities are expressed. The axis ζ_1 is placed along the cylinder axis. The position of the i^{th} gauge is uniquely determined by the cylindrical coordinates of its center point, ζ_{1_i} and θ_i, and by the angle ϕ_i formed by the gauge axis with the cylinder axis. The design variables are three for each gauge, plus the cylinder radius and wall thickness, for a total of 20.

The simplicity of structure allows the evaluation of the entries of the measurement matrix $\mathbf{C}$ by means of simple relations of elastic beam theory [Timoshenko, 1965]. Note that the computation of $\mathbf{C}$ has to be done at each iteration, so that the use of more demanding techniques as finite element methods (FEM) is unexpedient. The strain measured by the i^{th} gauge is

$$v_i = \sum_{j=1,6} C_{i,j} p_j,$$

where p_j is the value of the j^{th} load component normalized with respect to its nominal maximum value, $p_{j,max}$. Elements $C_{i,j}$ are as follows:

$$\begin{aligned}
C_{i,1} &= p_{1,max} W_n (\cos^2 \phi_i - \nu \sin^2 \phi_i); \\
C_{i,2} &= p_{2,max} \left[W_f \zeta_i \cos\theta_i (\cos^2 \phi_i - \nu \sin^2 \phi_i) + W_s \sin\theta_i \sin 2\phi_i \right]; \\
C_{i,3} &= p_{3,max} \left[W_f \zeta_i \sin\theta_i (\cos^2 \phi_i - \nu \sin^2 \phi_i) + W_s \cos\theta_i \sin 2\phi_i \right]; \\
C_{i,4} &= p_{4,max} W_t \sin 2\phi_i; \\
C_{i,5} &= p_{5,max} W_f \cos\theta_i (\cos^2 \phi_i - \nu \sin^2 \phi_i); \\
C_{i,6} &= p_{6,max} W_f \sin\theta_i (\cos^2 \phi_i - \nu \sin^2 \phi_i);
\end{aligned}$$

where ν is the Poisson's ratio for the structure material, and the moduli W

are defined as

$$\begin{aligned} W_n &= \frac{1}{2\pi RsE}; \\ W_f &= \frac{2W_n}{R}; \\ W_s &= 2(1+\nu)W_n; \\ W_t &= \frac{(1+\nu)W_n}{R}, \end{aligned}$$

E is Young's modulus, R is the cylinder radius and s its wall thickness. The relations above are valid in the assumption $s \ll R$: in this case, $1/s$ is a scale factor of $C_{i,j}$, that does not affect its structure. In other words, the relative accuracy of the sensor does not depend on the wall thickness; s can be chosen independently of other design variables, and made such that minimum sensitivity and maximum strain level requirements are met. A worst-case optimal design of the sensor can then be obtained through minimazion of the condition number index, and sensitivity adjusted with s.

In the reported example, other design constraints for the sensor to fit a robotic hand fingertip were considered:

$$\begin{aligned} 0 &> |\zeta_i - \zeta_j| - l_{max}, \\ R_{max} &> R > R_{min} > 2s, \end{aligned}$$

where $l_{max} = 15mm$, $R_{max} = 6mm$. It is assumed that every strain gauge exhibit equal noise statistics: the measurement noise covariance matrix is therefore diagonal with equal elements, and does not affect the design. To comply with severe size requirements, the minimum necessary number of transducers (i.e., 6) have been used on the sensor.

Two criteria have been considered for optimization: maximization of the minimum singular value of the information matrix, and minimization of its condition number. The former criterion is valid both in the linear-Gaussian framework, and in the absolute worst-case framework. The latter can address the relative worst-case case. As a consequence of its sensitivity to the absolute scaling of the information matrix, the singular value criterion tends to push the design at the lower bounds of the allowed ranges of wall thickness and cylinder radius. It was decided then to use the two methods in a cascaded fashion, using first the condition number minimization to find a reasonably "balanced" design. The condition number criterion is indifferent to the wall thickness parameter, but sets an optimal cylinder radius, which is used in the subsequent minimum singular value optimization.

The singular values and the condition number of the compliance matrix are complex functions of the $6 \times 3 + 1 = 19$ design variables whose minima cannot be found analytically. A numerical algorithm (a version of Powell's and Broyden's methods [Dahlquist, 1974]) has been employed. The values of design variables corresponding to the lowest of the relative minima found in repeated trials are given in table 3 and 4, corresponding to minimizing the condition number and to maximizing σ_{min}, respectively. The singular value decompositions of corresponding compliance matrices are reported in tables 5 and 6, respectively.

These results show that a fairly well conditioned design can be obtained with the proposed method. Note that, according to table 6, the ratio between the maximum and minimum sensitivity in the final design is 2.16. Minimum sensitivity occurs approximately corresponding to the p_1 component of the load, acting along the cylinder axis, that is intuitively the stiffest direction for the sensor.

6 Conclusions

In this paper an introduction and overview of methods for the optimal design of multivariate sensors has been presented. An attempt has been made at unifying the treatment of different sensing problems in a general framework, where common features of solution methods and optimization mathematical criteria could be emphasized. A central role in the problem is played by the eigenstructure of the *Fisher information matrix* associated with the systems at hand. Indications have been given that care must be taken in dealing with the physical dimensions of vector variables. Also, a glimpse into the worst-case optimization methods has been provided, along with an application to a practical device for robotic sensing.

Acknowledgments

This work is supported by Progetto Finalizzato Robotica CNR, Italy, Contracts No. 92.01079.PF67 and 92.01047.PF67 (University of Pisa) .

References

Arbel, A.: "Sensor placement in optimal filtering and smoothing problems", IEEE Trans. Automat. Contr., vol. AC-27, pages 94-8, Feb. 1982.

Banach, M., and Cunningham, J.: "Synthesis of arbitrary and conformal arrays using non-linear optimization techniques", Proc. IEEE National Radar Conf. (USA), pages 38-43, Apr.1988.

Curtain, R.F., and Pritchard, A.J.: "Infinite Dimensional Linear Systems Theory", Lecture Notes in Control and Information Science, vol.8, Springer-Verlag, Berlin 1978.

M. Basseville, A. Benveniste and G. V. Moustakides and A. Rougée: "Optimal sensor location for detecting changes in dinamical behavior", IEEE Trans. Automat. Contr., vol. AC-32, pages 1067-75, Dec. 1987.

Belforte, G., Bona, B., and Frediani, S.: "Optimal Sampling Schedule for parameter estimation of linear models with unknown but bounded measurement errors", IEEE Trans. Automat. Contr., vol. AC-32, ppl79-182, 1987.

A. Bicchi: "A Criterion for Optimal Design of Multiaxis force Sensors", Journal of Robotics and Autonomous Systems, vol.10, no. 4, 1992.

Cameron, A., and Durrant-White, H.: "A Bayesian Approach to Optimal Sensor Placement", Int. J. Robotics Res., vol. 9, pages 70-88, Oct. 1990.

Clark, J.J., and Yuille, A. L.: "Data Fusion for Sensory Information Processing Systems", Kluwer Academic Publishers, 1990.

Carotenuto, L., and Raiconi, G.: "Identifiability and identification of a Galerkin approximation for a class of distributed parameter systems", Int. J. Syst. Sci., 11, 1035, 1980.

Curtain, R.F., and Ichikawa, A.: "Optimal Location of Sensors for Filtering for Distributed Systems", In A.Ruberti (Ed.), Distributed Parameter Systems: Modelling and Identification, Lecture Notes in Control and Inf. Sci., vol.1, Springer-Verlag, Berlin, 1978.

G. Dahlquist and **r**A. Björk: "Numerical Methods", Prentice Hall, 1974.

El Jay, A.: "Distribuited System Analysis Via Sensors and Actuators", Sensors and Actuators A, vol. 29, pages 1-11, Sep. 1991.

El Jay, A., and Pritchard, A.J.: "Actuators and Sensors in Distributed Systems", Int. J. of Control, 46, pp. 1139-1153, 1987.

Fedorow, V.V.:"Theory of Optimal Experiments", New York, Academic Press, 1972.

Friedland, B.: "On the Calibration Problem", IEEE Trans. Automatic Control, vol.AC22, no.6, Dec. 1977.

Goodwin, G.C., and Payne, R.L.: "Dynamic System Identification: Experiment Design and Data Analysis", Academic Press, New York, 1977.

Golub, G.H., and VanLoan, C.F.: "Matrix Computations", John Hopkins University Press, Baltimore, MD, 1989.

Helmicki, A.J., Jacobson, C.A., and Nett, C.N.: "Control-oriented system identification: A worst-case/deterministic approach in H_∞", IEEE Trans. Automat. Contr., pp.1163-1176, Oct 1991.

Hirose, S. and Yoneda, K.: "Development of optical 6-axial force sensor ans its signal calibration considering non-linear interference", Proc. IEEE Intl. Conf. Robotics & Automation, pp. 46-53, Cincinnati, May 13-18, 1990.

Kosut, R.L., Arbel, A. and Messler, K.M.: "Optimal Sensor Design for State Reconstruction", IEEE Trans. on Automatic Control, vol. AC-27, n.1, pp. 242-4, Feb. 1982.

Kubrusly, C.S., and Malebranche, H.: "Sensors and Controllers Location in Distributed Systems - a Survey", Automatica, vol.21, pp.117-128, 1985.

Ljung, L.: "System Identification: Theory for the User", Prentice-Hall, 1987.

Maine R.E. and Iliff, K.W.: "Formulation and Implementation of a Practical Algorithm for Parameter Estimation with Process and Measurement Noise". SIAM J. Appl. Math., vol. 41, no.3, pp.558-579, 1981.

Mehra, R.K. : "Optimal Input Signals for parameter estimation in dynamic systems - Survey and new results", IEEE Transactions on Automatic Control, vol. AC-19, No.9, Dec. 1974.

Mehra, R.K.: "Optimization of Measurement Schedules and Sensor Designs for Linear Dynamic Systems", IEEE Transactions on Automatic Control, vol. AC-21, No.1, Feb. 1976.

Menq, C.H., Borm, J.H. and Lai, J.Z.: "Identification and observability measure of a basis set of error parameters in robot calibration", J. Mechanisms, Transmission, and Automation in Design, vol. 111, pp. 513-518, 1989.

Müller, P.C., and Weber, H.I.: " Analysis and Optimization of Certain Qualities of Controllability and Observability for Linear Dynamical Systems", Automatica, vol.8, pp.237-246, 1972.

Rao, C.R.: "Linear Statistical Inference and Its Applications", Wiley, New York, 1973.

Schroer, K., Uhl, L., Albright, S., and Huttenhofer, M.: "Ensuring solvability and analyzing results of the nonlinear robot calibration problem", Proc. Second Intl. Symposium on Measurement and Control in Robotics, pp. 851-858, Tsukuba Science City, Japan, Nov. 16-19, 1992.

Stewart, G.W.: "On the perturbation of pseudo-inverses, projections, and linear least squares problems", SIAM Review, vol.14, no.4, 1977.

Tarantola, A.: "Inverse Problem Theory", Elsevier, Amsterdam, 1987.

Tempo, R.: "Robust estimation and filtering in the presence of bounded noise", IEEE Trans. on Automatic Control, vol. AC-33, n.9, pp. 864-7, Sept. 1988.

Timoshenko, S., and Young, D.H.: "Theory of Structures," McGraw Hill, NY, 1965.

Watanabe, K., Sasaki, M., and Himmelblau, D.M.: "Determination of Optimal Measuring Sites for Fault Detection of Nonlinear Systems", Int. J. Syst. Sci., vol.16, Nov. 1985.

Wilkinson, J.H.: "The Algebraic Eigenvalue Problem" Clarendon Press, Oxford, UK, chap.4 pp.189ff, 1965.

Wilson,, B.B.: "Method for Optimizing Environmental Observing Networks", Proc. Oceans '88, pages 629-31, Oct.1988.

Captions

Figure 1: The process of Bayesian inference (adapted from Tarantola [1987]). (a) A priori information on the model space, $f_{prior}(\mathbf{X})$, and information on experimental data $f_{exp}(\mathbf{Y})$ are independent, and combine in the joint p.d.f. $f_{joint}(\mathbf{Y}, \mathbf{X})$, (b). Information on modelling is represented by $f_{mod}(\mathbf{Y}, \mathbf{X})$ (c). The conjunction of $f_{joint}(\mathbf{Y}, \mathbf{X})$ and $f_{mod}(\mathbf{Y}, \mathbf{X})$ is $f_{post}(\mathbf{Y}, \mathbf{X})$ (d). The marginal p.d.f.'s $f_{post}(\mathbf{X})$ and $f_{post}(\mathbf{Y})$ (e) can be obtained directly from $f_{post}(\mathbf{Y}, \mathbf{X})$. Different estimators can be applied to these results, as illustrated in (e).
Figure 2: A two-dimensional confidence ellipsoid
Figure 3: Flexible bar pinned at its extremities and its first three bending modes.
Figure 4: Sensor design performance indices plotted vs. the position of one sensor on a pinned flexible bar.
Figure 5: Sensor design performance indices plotted vs. the position of two sensors on a pinned flexible bar.
Figure 6 A miniaturized force/torque sensor with cylindrical beam structure.
Table 1: Optimal sensor location according to different performance indices.
Table 2: Optimal sensor location according to different performance indices in a two d.o.f. design.
Table 3: Optimal design variables for the six-axis miniaturized force sensor, corresponding to the nominal load: $p_1 = p_2 = p_3 = 10N$; $p_4 = 60Nmm$; $p_5 = p_6 = 85Nmm$. Criterion used: minimization of condition number. Optimal condition number = 2.16; minimum singular value = 2.65.
Table 4: Optimal design for the sensor of table 3, using the criterion of maximization of minimum singular value. Minimum singular value = 2.8; corresponding condition number = 2.75.
Table 5: Singular values and corresponding vectors of the sensor compliance matrix for the optimal set of design variables listed in table 3.
Table 6: Singular values and corresponding vectors of the sensor compliance matrix for the optimal set of design variables listed in table 4.

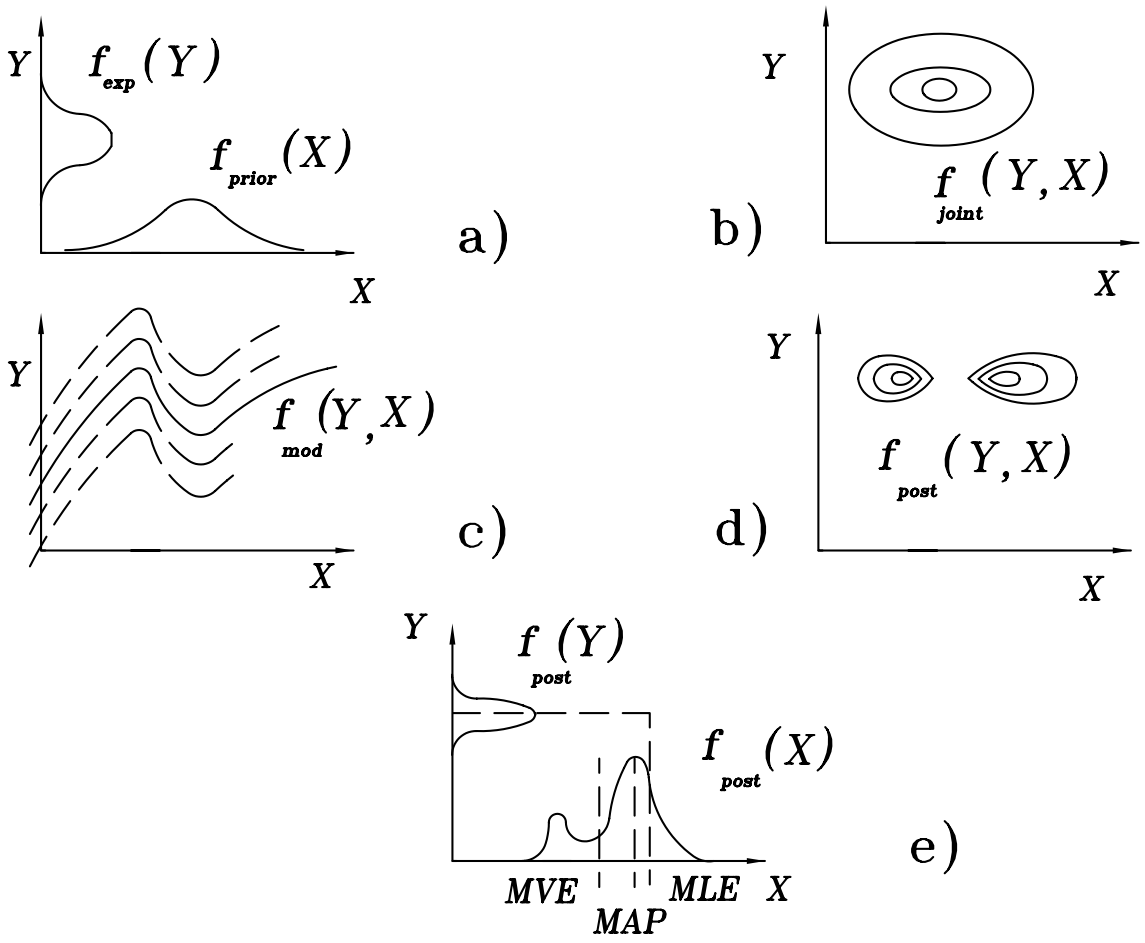

Y
$f_{exp}(Y)$
$f_{prior}(X)$
X
a)
Y
$f_{joint}(Y,X)$
X
b)
Y
$f_{mod}(Y,X)$
X
c)
Y
$f_{post}(Y,X)$
X
d)
Y
$f_{post}(Y)$
$f_{post}(X)$
MVE
MAP
MLE
X
e)

Figure 1

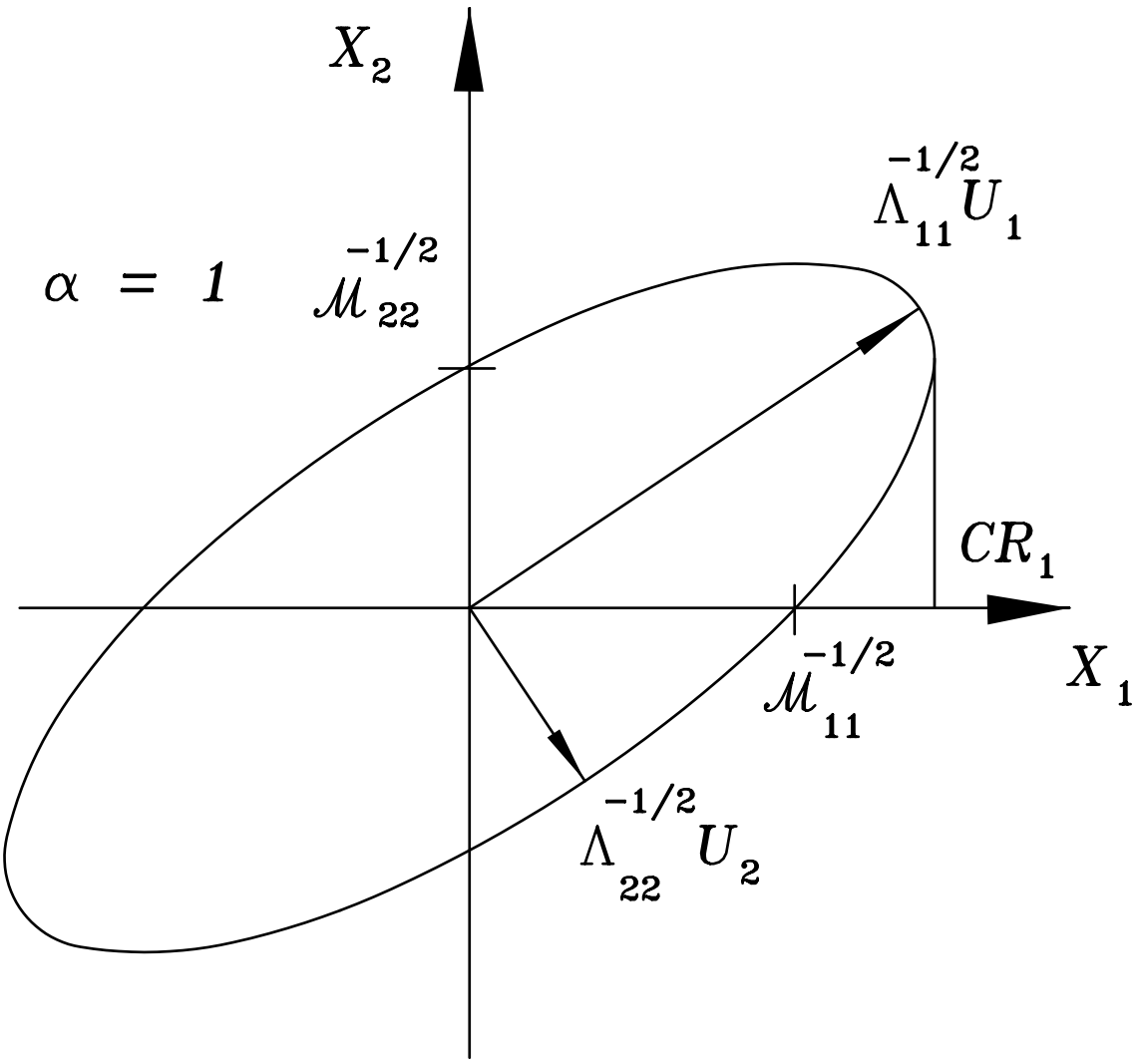

X_2
$\alpha = 1$
$\mathcal{M}_{22}^{-1/2}$
$\Lambda_{11}^{-1/2} U_1$
CR_1
X_1
$\mathcal{M}_{11}^{-1/2}$
$\Lambda_{22}^{-1/2} U_2$

Figure 2

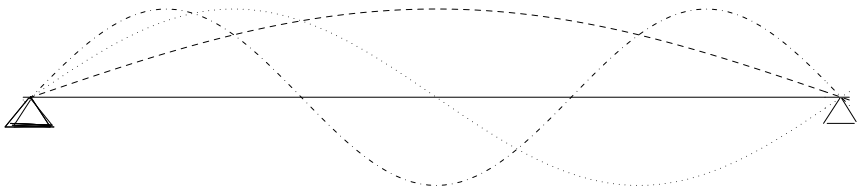


Figure 3

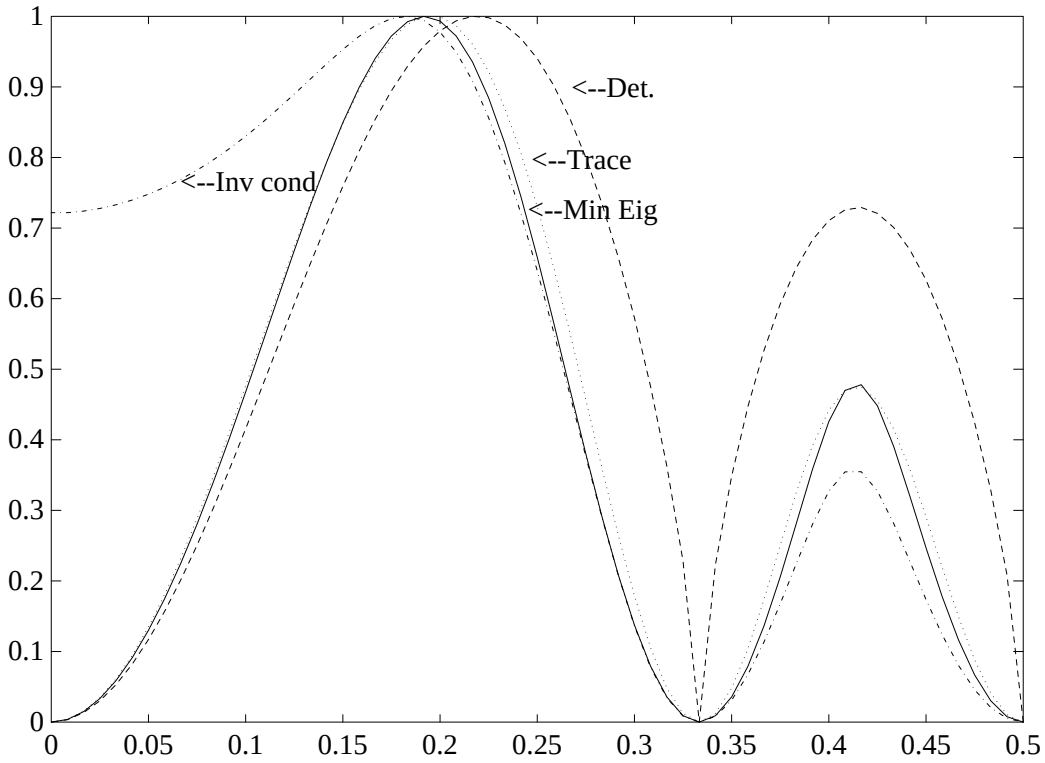
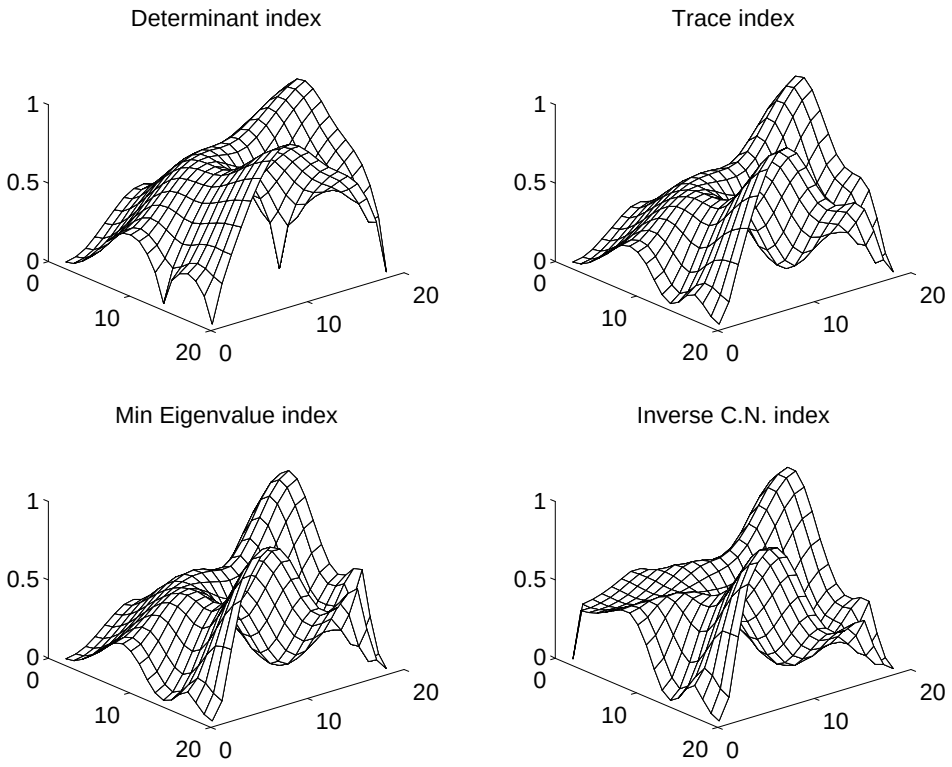

<--Det.
<--Trace
<--Min Eig
<--Inv cond
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5

Figure 4

Figure 5


Determinant index
Trace index
Min Eigenvalue index
Inverse C.N. index
1
0.5
0
0
10
20
0
10
20

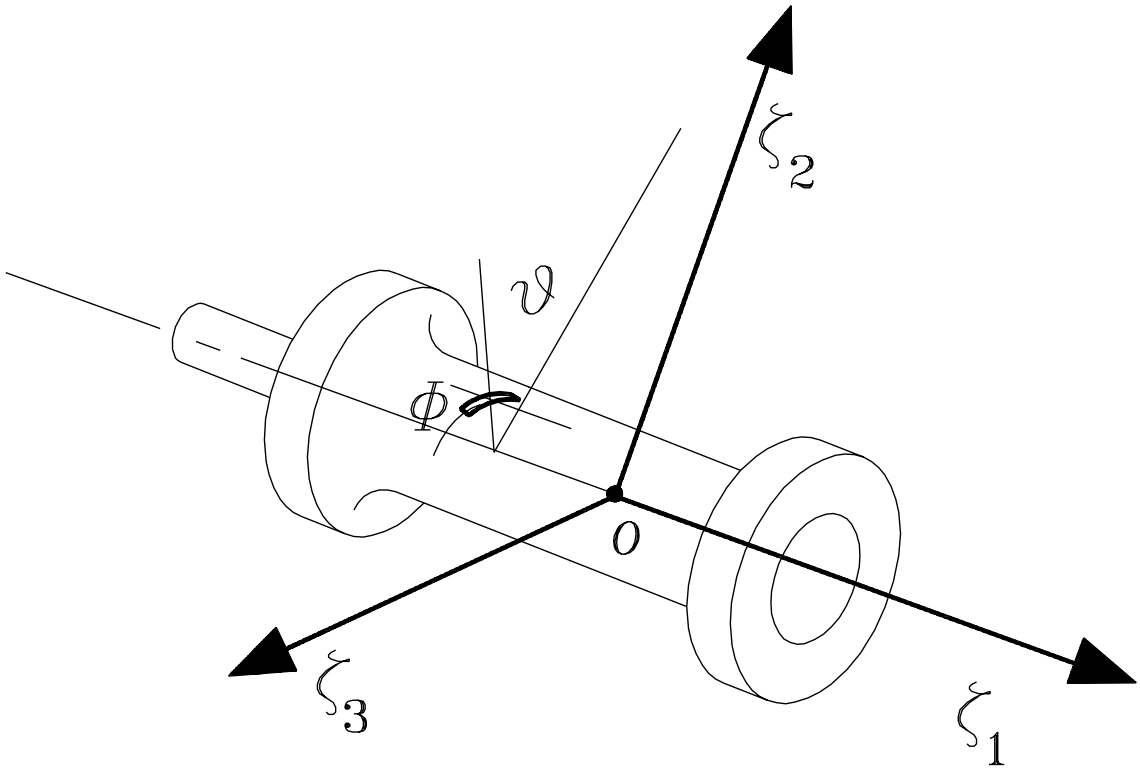


Figure 6

	Determinant	Trace	Min. Eigenvalue	Inverse of C. N.
Optimal Sensor location	13/60	23/120	23/120	11/60
Index value	$2.8\ 10^{-2}$	$4.3\ 10^{-4}$	$9.7\ 10^{-5}$	$3.7\ 10^{-5}$

Table 1

	Determinant	Trace	Min. Eigenvalue	Inverse of C. N.
Optimal Sensor location	2/9 1/2	7/36 1/2	1/6 1/2	5/36 1/2
Index value	0.08	$1.7\ 10^{-3}$	$4\ 10^{-4}$	$0.8\ 10^{-4}$

Table 2

Variable Gauge no.	ϕ(rad)	θ(rad)	ζ(mm)	R(mm)	s(mm)
1	0.24	-0.18	-1.3	5.6	0.1
2	-0.52	2.1	-5.8		
3	0.57	3.7	-2.6		
4	-0.14	4.2	-1.1		
5	-0.55	0.1	3.1		
6	0.29	1.9	4.3		

Table 3

Variable Gauge no.	ϕ(rad)	θ(rad)	ζ(mm)	R(mm)	s(mm)
1	0.21	-1.3	-0.14	5.6	0.1
2	-0.49	1.2	-6.76		
3	0.35	3.0	-4.0		
4	-0.57	4.53	-1.1		
5	0.24	0.0	3.4		
6	-0.12	2.1	3.4		

Table 4

	Singular Structure					
Singular values	γ_1 5.73	γ_2 5.73	γ_3 4.56	γ_4 3.03	γ_5 2.65	γ_6 2.65
Singular vectors	$\mathbf{u}_1$	$\mathbf{u}_2$	$\mathbf{u}_3$	$\mathbf{u}_4$	$\mathbf{u}_5$	$\mathbf{u}_6$
	0.00	0.00	0.00	0.06	-0.06	0.99
	-0.01	-0.07	0.10	-0.41	-0.90	-0.03
	0.03	0.06	0.94	-0.24	0.20	0.03
	0.09	-0.12	0.31	0.87	-0.35	-0.07
	-0.91	0.40	0.04	0.11	-0.07	0.00
	0.41	0.90	-0.03	0.05	-0.10	0.00

Table 5

	Singular Structure					
Singular values	γ_1 7.71	γ_2 6.06	γ_3 4.66	γ_4 3.61	γ_5 2.80	γ_6 2.80
Singular vectors	$\mathbf{u}_1$	$\mathbf{u}_2$	$\mathbf{u}_3$	$\mathbf{u}_4$	$\mathbf{u}_5$	$\mathbf{u}_6$
	0.02	0.00	0.00	0.00	0.10	0.99
	-0.56	-0.48	0.31	0.60	0.07	0.00
	-0.33	0.19	0.75	-0.55	0.01	0.00
	-0.10	-0.05	-0.03	-0.01	-0.99	0.10
	-0.21	0.85	0.04	0.47	-0.03	0.01
	0.72	-0.04	0.59	0.35	-0.09	0.00

Table 6