

Technical University of Denmark



Optimal sensor configuration for complex systems

Sadegh, Payman; Spall, J. C.

Published in:

Intelligent Control (ISIC), 1998. Held jointly with IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA), Intelligent Systems and Semiotics (ISAS), Proceedings

Link to article, DOI:

[10.1109/ISIC.1998.713691](https://doi.org/10.1109/ISIC.1998.713691)

Publication date:

1998

Document Version

Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):

Sadegh, P., & Spall, J. C. (1998). Optimal sensor configuration for complex systems. In Intelligent Control (ISIC), 1998. Held jointly with IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA), Intelligent Systems and Semiotics (ISAS), Proceedings (pp. 376-380). IEEE. DOI: 10.1109/ISIC.1998.713691

DTU Library

Technical Information Center of Denmark

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Optimal Sensor Configuration for Complex Systems¹

PAYMAN SADEGH*, and JAMES C. SPALL†

* Dept. of Mathematical Modeling, Technical University of Denmark,
DK-2800 Lyngby, Denmark. e-mail: ps@imm.dtu.dk

† The Johns Hopkins University, Applied Physics Laboratory, Laurel,
MD 20723-6099, USA. e-mail: james.spall@jhuapl.edu.

Abstract

The paper considers the problem of sensor configuration for complex systems. Our approach involves definition of an appropriate optimality criterion or performance measure, and description of an efficient and practical algorithm for achieving the optimality objective. The criterion for optimal sensor configuration is based on maximizing the overall sensor response while minimizing the correlation among the sensor outputs. The procedure for sensor configuration is based on simultaneous perturbation stochastic approximation (SPSA). SPSA avoids the need for detailed modeling of the sensor response by simply relying on observed responses as obtained by limited experimentation with test sensor configurations. We will illustrate the approach with the optimal placement of acoustic sensors for signal detection in structures. This includes both a computer simulation study for an aluminum plate, and real experimentations on a steel I-beam.

KEY WORDS: Optimal sensor configuration, Detection, Complex systems, Stochastic approximation, SPSA, Acoustic sensors and signals.

1. Introduction

It is often of interest to maximize the amount of information provided by the experiments that are conducted to draw an inference upon a system. In this paper, we study optimization of sensor configurations. Configuration of sensors encompasses placement of sensors on an object, adjustment of sensor operating conditions (such as frequency response, potentiometer settings, pressure sensitivities), sensor orientations at fixed locations, etc. Similar to any other general optimization problem, our solution requires selection of a criterion or performance measure, and selection of an optimization algorithm.

We are interested in systems where the prior knowledge is sparse or too complex to readily offer useful models for solving the sensor configuration problem. We assume that it is possible to fix the sensors at a configuration of interest, and experiment with the real system, a physical prototype, or a computer simulation to generate a realization of the process

underlying the sensors responses for that sensor configuration. The application part of the paper, which concerns nondestructive evaluation of structures using acoustic emission (AE) (see e.g. [1, 2]), provides an example where such experimentations are possible. The application studies optimal placement of a number of acoustic sensors on the surface of a structure to provide measurements of AE signals caused by crack formation in the structure, under the realistic assumption that no reliable *a priori* model is available for the sensor placement problem. The relevant experimentations may be possible in a variety of other applications such as biomedical signal processing, routine monitoring of a system, vision-based quality monitoring of a production, and vibration monitoring of an object.

Assuming the availability of relevant data, a possible approach is experimental modeling of the response prior to the design of optimal sensor configuration, similar to the widely applied response surface methodology (see e.g. [3]). The methodology typically involves fitting models for the response as a function of design variables. The obvious advantage of the approach relative to those relying upon complex or oversimplified *a priori* models for response determination, is its reliance upon data. However, the modeling part of such approaches, except for trivial cases, requires detailed analytical and experimental work such as experiment design for modeling, determination of an appropriate model structure, parameter estimation, and model validation. For a discussion of similar modeling issues related to the response surface methodology see [3] and the references therein. Moreover, the approach is indirect, requiring an optimization phase in addition to the modeling and analysis phase. For instance, the response surface optimization method requires at least $p + 1$ measurements for model fitting at a point (or 2^p if the traditional factorial design is used) where p is the number of parameters being optimized. This is followed by a line search to obtain a new point, and the process may be repeated for numerous iterations. Finally, most modeling issues arising in connection with the indirect approach to an optimization problem pertain to the specific problem at hand, making the approach inappropriate for a unified treatment of the generic optimization.

Alternative experimental approaches concern application of *direct* algorithms, which rely only on (noisy) observations of the criterion and therefore circumvent the intermediate modeling phase. This makes the direct approach both convenient and easily extendable to a large variety of sen-

¹The first author's work was partly supported by the Danish Research Academy, grant S950029, during his stay at JHU/APL. James Spall's work is supported by U.S. Navy contract N00039-95-C-0002 and the JHU/APL IRAD program. The authors would like to thank Prof. James Spicer, JHU, Dept. of Materials Science, for his valuable comments and assistance regarding the I-beam experiments, and Dr. Jane Spicer, JHU/APL, for her valuable comments.

sensor configuration problems. Finite difference stochastic approximation (FDSA) is a traditionally used direct algorithm. FDSA is a stochastic approximation (SA) technique in the Kiefer-Wolfowitz setting [4] which uses finite differencing of the noisy observations of the criterion for gradient approximations [5]. In this way, the optimization parameters (in the context of sensor configuration design, the parameters are the spatial coordinates of the sensors and/or other sensor parameters) are *estimated* using an iterative procedure that consists of gradient approximation and parameter updating. Although FDSA is flexible in the respect that no gradient information is required, the practical implementability of the technique is questionable in high-dimensional problems. Recall that at each iteration, the basic two-sided FDSA requires $2p$ observations for gradient approximation where p is the number of parameters being optimized. For example, if the location of the sensors is of interest, the number of parameters is equal to the number of sensors times 2 or 3 depending on the relevant spatial coordinates. For high-dimensional cases, the large number of required observations, which involve *real measurements* on the system, limits the usefulness of the FDSA technique.

A relatively new SA algorithm in the Kiefer-Wolfowitz setting, simultaneous perturbation stochastic approximation or SPSA [6], has made solution to problems of this complexity possible. Similar to FDSA, SPSA approximates the gradient using only noisy observations of the criterion, i.e. SPSA is a direct algorithm. In contrast to FDSA, however, SPSA uses only *two* observations, obtained by *simultaneous* random perturbation of the parameters, to form a gradient approximation, which implies a p -fold savings in the number of observations per gradient approximation (relative to two sided FDSA). In [6] and [7], it is shown that under reasonably general conditions, the p -fold savings per gradient approximation translates directly into a p -fold savings in the *total* number of measurements for the complete optimization process. Since the SPSA approach requires no detailed modeling information and no experimental procedures beyond the standard input-output analysis of conventional experimental design methods (including response surface), it provides an efficient approach for solving the sensor configuration problems in a large variety of complex systems.

Although SPSA has several features that make it especially appropriate for the sensor configuration problem, let us briefly compare SPSA with other relevant optimization methods. Stochastic gradient techniques such as Robbins-Monro SA [5] require an input-output model of the process in order to obtain the noisy gradient observation (and deterministic gradient based techniques, such as Newton-Raphson or conjugate gradient, are even more inappropriate since they assume noise-free measurements of the gradient). Simulated annealing (see e.g. [8]), like SPSA, uses only measurements of the objective function (not the gradient). However, this algorithm is well-known to converge slowly and to be sensitive to noise in the function measurement (see e.g. [9, 10]). SPSA, on the other hand, converges relatively quickly and is explicitly designed to cope with noise in the function measurements. Finally, population-based methods such as genetic algorithms [11] are inappropriate here since they require *multiple* systems to form the population from which the cross-over and other operations in the algorithm are performed. SPSA works with

only a single solution, corresponding to a single system for which the sensors must be configured, and hence is feasible in practical configuration settings.

The rest of the paper is organized as follows. Section 2 discusses the choice of optimality criterion. Section 3 offers a brief review of the SPSA algorithm which is central to the experimental methodology of the paper. Section 4 concerns optimal sensor location for signal detection in structures. This includes both computer simulation and real experimental results for an aluminum plate and a steel I-beam respectively. The study has application in the nondestructive evaluation of structures. Finally, Section 5 offers concluding remarks.

2. Criterion for Sensor Configuration

The direct optimization approach using the SPSA algorithm offers a uniquely efficient alternative for sensor configuration problems in complex systems. The realizability and the relevance of the technique, on the other hand, is conditioned on (and only on) the existence of an appropriate criterion which is computable from response recording. In this section, we discuss a suitable selection for the criterion. It should be noted, however, that the choice of a criterion to be used within the SPSA based optimization approach is not unique and may be influenced by particular design objectives and further available information.

The formal setting to be considered here is as follows. Let $\{X_\theta(t)\}$, $t = 1, 2, \dots$, denote a set of responses received by the sensors with configuration θ under the operating condition of interest. Each element of the sequence, $X_\theta(t)$, is a N -dimensional vector where N is the number of sensors. The sequences $\{X_\theta(t)\}$ are realizations of some unknown random process, and a realization of the process can be generated by experimentation with the system. We seek a configuration θ that provides good measurements of all possible realizations of the process. We introduce the following criterion:

$$E\left\{\left(\det\left\{\sum_t X_\theta(t)X_\theta(t)'\right\}\right)^{1/N}\right\} \quad (2.1)$$

where the expectation $E\{\cdot\}$ is with respect to the process generating $\{X_\theta(t)\}$, and the summation is taken over a time window of interest.

The rationale for selecting the criterion is as follows.

1. Defining the overall response of a sensor as the sum of squares of the sensor responses over the time window of interest, it is evident that the diagonal elements of $\sum_t X_\theta(t)X_\theta(t)'$ account for the overall response of the sensors. The determinant of a positive semi-definite matrix increases with its diagonal elements.
2. The off-diagonal elements of $\sum_t X_\theta(t)X_\theta(t)'$ account for the correlation among the sensors responses. The determinant of a positive semi-definite matrix decreases with large-in-magnitude off-diagonal elements.
3. The $1/N$ exponent scales the units properly such that the criterion is measured in the same physical units as the overall response.
4. The quantity $(\det\{\sum_t X_\theta(t)X_\theta(t)'\})^{1/N}$ is readily computable given that a response realization $\{X_\theta(t)\}$ is available. This computed value obviously gives an unbiased estimate of (2.1).

5. The criterion can be related to detection of the signal $X_\theta(t)$ in presence of additive mean zero measurement noise with unknown covariance, which is uncorrelated with the process generating $X_\theta(t)$, see [12].

We wish to emphasize again that the direct approach using the SPSA algorithm *is not* uniquely related to the criterion presented in this section. Any criterion that is a suitable performance measure, and is computable from response recordings can be used within the experimental methodology of the paper as follows in the next section.

3. Overview of the SPSA algorithm

Since the SPSA algorithm plays a central role in the experimental methodology, this section provides a brief overview of the generic technique. A step by step implementation of the algorithm for the sensor configuration problem is presented in [12].

Consider the problem of determining a parameter $\theta \in \mathbb{R}^p$ that maximizes a differentiable objective function $J(\theta)$, where the explicit dependence of the loss function upon θ is unknown, but where for each θ a noisy value for the objective function can be obtained (i.e., no gradient information is directly available). The SPSA algorithm has recently attracted considerable attention for challenging optimization problems of this type in application areas such as adaptive control, pattern recognition, discrete event systems, neural network training, and model parameter estimation, see e.g., [13], [14], [15], [16], [10], and [17]. In this section, we briefly present the SPSA algorithm. The reader is referred to [6] for a detailed treatment.

Let $\hat{\theta}_k$ denote the estimate for the parameter θ at the k th iteration and $g(\theta)$ denote $\partial J/\partial \theta$. The SPSA algorithm has the form $\hat{\theta}_{k+1} = \hat{\theta}_k + a_k \hat{g}_k(\hat{\theta}_k)$, where the gain sequence $\{a_k\}$ satisfies certain conditions (as follows) and $\hat{g}_k(\hat{\theta}_k)$ is a simultaneous perturbation approximation to $g(\hat{\theta}_k)$ at iteration k . We define the simultaneous perturbation estimate for the gradient as follows. Let $\Delta_k \in \mathbb{R}^p$ be a vector of p mutually independent mean zero random variables $\{\Delta_{k1}, \Delta_{k2}, \dots, \Delta_{kp}\}$ satisfying certain conditions (see Lemma 1 and assumption A_2 of [6]; note especially the important finite inverse moments condition that *precludes* normal and uniform distributions but admits, e.g., the symmetric Bernoulli ± 1 distribution). Consistent with the usual framework of stochastic approximation, we have noisy observations of the objective function. In particular, at the k th iteration, consider the two observations:

$$\begin{aligned} j_k^{(+)} &= J(\hat{\theta}_k + c_k \Delta_k) + \epsilon_k^{(+)} \\ j_k^{(-)} &= J(\hat{\theta}_k - c_k \Delta_k) + \epsilon_k^{(-)} \end{aligned}$$

where $\{c_k\}$ is a gain sequence, and $\epsilon_k^{(+)}$ and $\epsilon_k^{(-)}$ represent noise terms that satisfy $E\{\epsilon_k^{(+)} - \epsilon_k^{(-)} | \hat{\theta}_k, \Delta_k\} = 0$. The gain sequences $\{a_k\}$ and $\{c_k\}$ are positive for all k and tend to zero as $k \rightarrow \infty$. Moreover, $\sum_{k=0}^{\infty} a_k = \infty$, $\sum_{k=0}^{\infty} (a_k/c_k)^2 < \infty$.

The basic simultaneous perturbation form for the estimate of $g(\cdot)$ at the k th iteration is then

$$\hat{g}_k(\hat{\theta}_k) = \left[\frac{j_k^{(+)} - j_k^{(-)}}{2c_k \Delta_{k1}} \dots \frac{j_k^{(+)} - j_k^{(-)}}{2c_k \Delta_{kp}} \right]^T. \quad (3.1)$$

Note that at each iteration, only *two* observations are needed to form the estimate regardless of the number of parameters.

In [6], it is shown that under fairly general conditions, the SPSA iterates converge almost surely to the true optimum. The same reference derives the asymptotic distribution of the iterate, which under reasonably general conditions is shown to be Gaussian. This result is important for quantifying the accuracy of the estimate. The reference [6] also compares the asymptotic behavior of SPSA and FDSA, both theoretically and through numerical studies. The results indicate that under fairly general conditions, SPSA requires p times fewer number of measurements in order to achieve a specified level of accuracy (in terms of the asymptotic mean square error of the estimate). Finally, [18] treats constrained optimization via SPSA. The constrained algorithm is of particular interest since possible sensor configurations are often subject to e.g. geometric restrictions.

4. Application: Signal Detection in Complex Structures

In this section, we consider an application of the procedure to sensor placement for signal detection in complex structures. A very important problem in this connection is the nondestructive evaluation of structures by acoustic signal sensing. The use of acoustic emission (AE) signal sensing has recently received considerable attention in problems related to the nondestructive evaluation, see e.g. [1]. As a result of crack formation in a structure, acoustic emission signals are generated and propagated throughout the structure. It is then possible to detect a forming crack by doing acoustic signal measurements using a number of acoustic sensors placed on the surface of the structure. As a result of structure complexity and material and geometric irregularities in many practically important objects such as highway bridges, it is often not possible to develop models for response determination solely based on physical and geometrical considerations. The lack of reliable models and the convenient availability of data (including laser induced simulations of AE cracking events) point to the experimental methodology of the paper as a promising tool for solving this challenging sensor placement problem (see also the discussions in Section 1). In this section, we present an application of the methodology of the previous sections to a simulated aluminum plate (Subsection 4.1) and to an actual steel I-beam (Subsection 4.2).

4.1. Simulation Results for Sensor Placement on a Plate

Here, we present a computer simulation-based optimization of the sensor location for detection of impulse pressure inputs to a plate. We apply the developed procedure and find the optimal Cartesian coordinates of 10 sensors placed on the surface of a $1m \times 1m$ plate (hence the number of optimization parameters is equal to 20). Notice that in a real application, the experimental procedure typically involves *real* experiments and not computer simulations. In this spirit, the optimization procedure will have no knowledge of the plate equations implemented within the simulation. For the simulation purpose, we use the following partial differential equation (see [19], page 233)

$$\begin{aligned} (Eh^3/12(1-\nu^2))\nabla^4 w(x,y,t) + \\ \rho h \frac{\partial^2 w(x,y,t)}{\partial t^2} = q(x,y,t), \end{aligned} \quad (4.1)$$

where x and y denote respectively the horizontal coordinate and the vertical coordinate, $\nabla^4 = (\frac{\partial^4}{\partial x^4} + 2\frac{\partial^4}{\partial x^2\partial y^2} + \frac{\partial^4}{\partial y^4})$ is the biharmonic operator, $w(x, y, t)$ denotes the deflection, $q(x, y, t)$ denotes the input pressure, E is the Young module, ρ is the density of the plate, h is the thickness of the plate, and finally $0 < \nu < 1$ is the Poisson ratio (relating the orthogonal strain-stress couples). The following constants are used in the simulations: $E = 69.0 \times 10^9 N/m^2$, $\rho = 2.7 \times 10^3 kg/m^3$ (these values of E and ρ are typical for an aluminum plate; see any relevant standard table), $\nu = 0.01$, $h = 2.0 \times 10^{-5} m$. The plate is subject to clamped edge boundary conditions, i.e. for all t , $w(x, y, t) = 0$ and $\partial w(x, y, t)/\partial t = 0$ along $x = 0$, $x = 1$, $y = 0$, and $y = 1$.

The input pressure is taken to be

$$q(x, y, t) = \begin{cases} 10^4 \delta(t) N/m^2 & x = x_0, y = y_0 \\ 0 & \text{Otherwise} \end{cases}$$

where $\delta(t)$ is the Dirac delta function. The prior knowledge of $(x_0, y_0)'$ is embedded in a distribution obtained by truncating a $N((0.5, 0.5)', 0.025^2 I)$ distribution (where I as usual denotes the identity matrix) such that $Pr(0.1 \leq x, y \leq 0.9) = 1$. Finally, each sensor records the deflection of the structure at the location point. The sampling time of data is 0.05 seconds for all the sensors.

The tuning of SPSA constants has been previously studied in the literature. Following the guidelines of [20] which are obtained empirically, we select $a_k = 0.1/k^{0.602}$, $c_k = 0.1/k^{0.101}$ for the SPSA algorithm. The reference [20] also recommends addition of a so called stability constant to k in a_k , where the constant is typically around 10 percent of the total number of iterations (although the stability constant has not been considered here, it can usually improve the performance of SPSA significantly). Random perturbations are sampled from a Bernoulli (± 1) distribution which is asymptotically optimal (see [21]). The initial placement, $\hat{\theta}_0$, for implementing the algorithm is obtained by randomly distributing the sensors over the square $0.1 \leq x, y \leq 0.9$.

For each placement of the sensors, a function evaluation includes the following. We randomly draw an input location according to the probability distribution for the possible input locations, solve the partial differential equation (4.1) numerically to find the sensor response for each sensor, and form the response (deflection) sequence $\{X_\theta(t)\}$, $t = 0.05, \dots, 50 \times 0.05$, where the element i of $X_\theta(t)$ is equal to the response of sensor i at time t . A noisy realization of $J(\theta)$ is then obtained by computing $(\det\{\sum_t X_\theta(t)X_\theta(t)'\})^{1/N}$.

We apply the SPSA algorithm with a moderate number of function computations, say 500, corresponding to $500/2=250$ iterations and denote the resulting sensor placement by $\hat{\theta}_{250}$. The initial and final placement of the sensors are plotted in Figure 1 for comparison. Note the shift of the sensors towards the mean value for (x_0, y_0) (signal generation center). A similar convergence pattern as the one illustrated in Figure 1 is observed for varying initial conditions of the SPSA algorithm.

To evaluate the relative objective function values at the initial and final solutions, we fix the sensors at the locations $\hat{\theta}_{250}$ and $\hat{\theta}_0$ respectively, do 100 objective function evaluations for each one of the sensor placements, and average over the 100 computed values (recall that the direct analytical value of

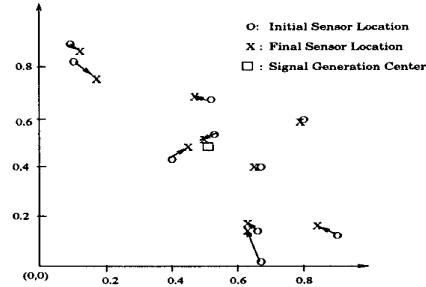


Figure 1: Initial and final placement of the sensors on the plate.

$J(\theta)$ is unavailable). The obtained average values for $J(\theta)$ are equal to 0.0517 and 0.0309 respectively. Notice that a considerable increase in the value of the criterion is obtained using the SPSA algorithm with relatively small number of function evaluations. We repeat the same procedure for two randomly selected sensor placements, θ_1 and θ_2 , in order to investigate sensor placements where the sensors are close to the signal generation center but provide redundant information due to being densely located. The placement θ_1 is obtained by randomization over $0.4 \leq x, y \leq 0.6$, and θ_2 is obtained by randomization over $0.2 \leq x, y \leq 0.8$. The redundant information is especially apparent in the case for θ_1 . The average function value obtained for θ_1 is in the scale of 10^{-4} and the average function value for θ_2 is 0.0138, both much worse than the SPSA solution. For a more detailed comparison of the above sensor configurations for signal detection in the presence of Gaussian measurement noise, see [12]. The same reference also compares the SPSA solution with solutions obtained by other random search procedures.

4.2. Small Scale I-beam Experiments

Here, we present results obtained by real experimentations on a steel I-beam (a beam with I-shaped cross section). Because of practical laboratory constraints, our experimentation is limited to locating 3 acoustic sensors on the center line of the I-beam, i.e. the number of optimization parameters is equal to 3. The I-beam has a length of approximately 120cm and a height of approximately 15cm. The acoustic sensors transform a mechanical deflection to an electrical voltage.

To simulate AE cracking events, we use a high energy laser with an energy varying within approximately 10% of the tuned energy level. A more detailed description of the experimental apparatus and use of AE in nondestructive evaluation is given in [22]. We consider a situation where AE events occur with equal likelihood within an approximately 5cm long line piece along the center line and around a point 50cm from one end of the I-beam (i.e., a uniform distribution for the location of AE events is used). Our data collection/processing equipments consist of an oscilloscope where the outputs of the transducers are recorded (in mv), and a computer where the oscilloscope data are down-loaded and SPSA iterations are performed using MATLAB[®] software.

We apply the experimental procedure using 20 iterations of SPSA and obtain the placement $\hat{\theta}_{20}$. The SPSA constants are selected as $q = 1$, $a_k = 0.1/k^{0.602} (cm/mv)^2$, $c_k = 0.01/k^{0.101} (cm)$. Figure 2 shows the initial and final

placements of the sensors. It is interesting to note that the final placement is further away from the mean AE source but the sensors are also further apart, reducing redundant information. This can be justified considering the fact that steel is known to have small signal attenuation coefficient relative to aluminum as considered in Subsection 4.1 (see [23]). Therefore, it is more important to reduce the redundant information rather than to locate the sensors close to the mean AE source.

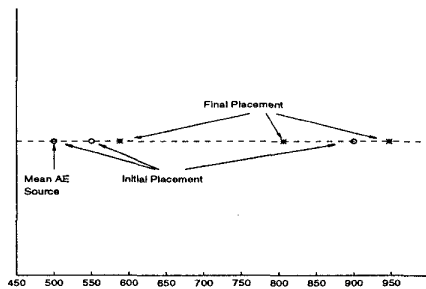


Figure 2: Initial and final placement of the sensors on the steel I-beam.

5. Conclusion

The paper studies the problem of optimal configuration of a number of sensors for a system. The presented approach is direct and uses experimental data without an intermediate step of modeling the sensor response. The approach is easily implementable and does not suffer from modeling inaccuracies. The SPSA algorithm provides a uniquely powerful tool and plays a central role within the experimental methodology of sensor configuration design. We have demonstrated the approach on the problem of sensor placement for signal detection in complex structures, using both computer simulations and real experiments. The approach here represents a special case of a more general SPSA-based experimental design approach.

References

- [1] R. K. Miller and P. McIntire. Acoustic emission testing. In *Nondestructive Testing Handbook*. American Society for Nondestructive Testing- Columbus- Ohio, 1987.
- [2] I. Grabec, W. Sachse, and E. Govekar. Solving AE problems by neural networks. In W. Sache, J. Roget, and K. Yamaguchi, editors, *Acoustic Emission: Current Practice and Future Directions*, pages 165–182. American Society for Testing and Materials, Philadelphia, 1991. ASTM STP 1077.
- [3] R. H. Myers, A. I. Khuri, and W. H. Carter. Response surface methodology:1966-1988. *Technometrics*, 31:137–157, 1989.
- [4] J. Kiefer and J. Wolfowitz. Stochastic estimation of a regression function. *Ann. Math. Stat.*, 23:462–466, 1952.
- [5] H. J. Kushner and D. S. Clark. *Stochastic Approximation for Constrained and Unconstrained Systems*. Springer-Verlag, Berlin, 1978.
- [6] J. C. Spall. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Transactions on Automatic Control*, 37(3):332–341, 1992.
- [7] D. C. Chin. Comparative study of stochastic algorithms for system optimization based on gradient approximations. *IEEE Transactions on Systems, Man, and Cybernetics*, 27:244–249, 1997.
- [8] S. P. Brooks and B. J. T. Morgan. Optimization using simulated annealing. *The Statistician*, 44:241–257, 1995.
- [9] M. A. Styblinski and T. S. Tang. Experiments in non-convex optimization: Stochastic approximation with function smoothing and simulated annealing. *Neural Networks*, 3:467–483, 1990.
- [10] D. C. Chin. A more efficient global optimization based on Styblinski and Tang. *Neural Nets.*, 7:573–574, 1994.
- [11] L. Davis. *Handbook of Genetic Algorithms*. International Thomson Computer Press, 1996.
- [12] P. Sadegh and J. C. Spall. Optimal sensor configuration for complex systems. Submitted to *Journal of American Statistical Association* (a copy can be provided upon request), 1998.
- [13] F. Rezaat. On the use of SPSA-based model free controller in quality improvement. *Automatica*, 31:913–915, 1995.
- [14] Y. Maeda, H. Hirano, and Y. Kanata. A learning rule of neural networks via simultaneous perturbation and its hardware implementation. *Neural Nets.*, 8:251–259, 1995.
- [15] S. D. Hill and M. C. Fu. Transfer optimization via simulation perturbation stochastic approximation. In *Proc. Winter Simulation Conference*, pages 242–249, 1995.
- [16] G. Cauwenberghs. *Analog VLSI Autonomous Systems for Learning and Optimization*. PhD thesis, Dept of Electrical Engineering, California Institute of Technology, 1994.
- [17] T. Parisini and A. Alessandri. Non-linear modeling and state estimation in a real power plant using neural networks and stochastic approximation. In *Proc. American Control Conference*, pages 1561–1567, 1995.
- [18] P. Sadegh. Constrained optimization via stochastic approximation with a simultaneous perturbation gradient approximation. *Automatica*, 33(5):889–892, 1997.
- [19] K. F. Graff. *Wave Motion in Elastic Solids*. Oxford Engineering Science Series. Oxford, Clarendon, 1975.
- [20] J. C. Spall. Implementation of simultaneous perturbation algorithm for stochastic optimization. *IEEE Transactions on Aerospace and Electronic Systems*, 34, 1998. In press, tentatively issue no. 3.
- [21] P. Sadegh and J. C. Spall. Optimal random perturbations for stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Transactions on Automatic Control*, in press, 1998.
- [22] J. C. Spall, J. L. Maryak, and M. S. Asher. Neural network approach to locating acoustic emission sources in nondestructive evaluation. *J. of Sound and Vibration*, 211:133–143, 1998.
- [23] ASM. Nondestructive evaluation and quality control. In *Metals Handbook*, volume 17. American Society for Metals, 1989.