Pergamon

PII: S0005-1098(96)00149-5

Automatica, Vol. 33, No. 1, pp. 109–112, 1997

0005-1098/97 $17.00 + 0.00

Technical Communique

A One-measurement Form of Simultaneous Perturbation Stochastic Approximation*

JAMES C. SPALL†

Key Words—Optimization; gradient estimation; simultaneous perturbation; SPSA.

Abstract—The simultaneous perturbation stochastic approximation (SPSA) algorithm has proven very effective for difficult multivariate optimization problems where it is not possible to obtain direct gradient information. As discussed to date, SPSA is based on a highly efficient gradient approximation requiring only two measurements of the loss function independent of the number of parameters being estimated. This note presents a form of SPSA that requires only one function measurement (for any dimension). Theory is presented that identifies the class of problems for which this one-measurement form will be asymptotically superior to the standard two-measurement form. © 1997 Elsevier Science Ltd. All rights reserved.

1. *Introduction*

The simultaneous perturbation stochastic approximation (SPSA) algorithm has recently attracted considerable attention for multivariate optimization problems where it is difficult or impossible to obtain a gradient of the objective function with respect to the parameters being optimized (see e.g. Alessandri and Parasini, 1995; Hill and Fu, 1995; Smith and Chin, 1995; Maeda *et al.*, 1995; Rezayat, 1995). SPSA is for problems in the multivariate Kiefer–Wolfowitz SA setting, where only (possibly noisy) measurements of the objective (say, loss) function are assumed available. The essential feature of SPSA is its highly efficient gradient approximation. As described in Spall (1987, 1992), the gradient approximation requires only two loss function measurements, regardless of the problem dimension. This contrasts with the $2p$ function measurements required in conventional finite-difference methods, where p is the problem dimension (the number of parameters being optimized). The central theoretical result in Spall (1992) is that in many practical problems this p-fold savings in function evaluations per gradient approximation translates directly into a p-fold savings in function evaluations to solve the optimization problem (i.e., the algorithms take the same number of iterations to achieve a given level of mean-square accuracy in the optimization parameters, but SPSA takes p times fewer measurements per iteration).

This note introduces a variant of SPSA that requires only *one* function evaluation to construct the gradient approximation. Theory is presented that provides guidelines on when it is advantageous to use this one-measurement form versus the 'standard' two-measurement form mentioned above. The implications of the theory are illustrated in a numerical study for a small-scale problem involving a multivariate polynomial loss function. For convenience, we shall refer to the one- and two-measurement forms of SPSA as SPSA1 and SPSA2 respectively.

* Received 1 April 1996; received in final form 5 August 1996. This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Editor Peter Dorato. Corresponding author James C. Spall. Tel. +1 301 953 5000; E-mail james.spall@jhuapl.edu.

† The Johns Hopkins University, Applied Physics Laboratory, Laurel, MD 20723-6099, U.S.A.

2. *Algorithm*

Our fundamental aim is to minimize a loss function $L(\theta)$, where θ is some p-dimensional vector of adjustable parameters. Consistent with the usual framework for nonlinear continuous optimization, we seek a minimum θ^* such that $g(\theta) \equiv \partial L(\theta)/\partial\theta = 0$. (This paper deals with the local unconstrained optimization problem only; for SPSA2, an extension to the constrained optimization problem is given in Sadegh (1996) and an extension to the global problem of multiple roots to $g(\theta) = 0$ is given in Chin (1994).) It is assumed that only measurements of $L(\theta)$ (typically with additive noise) are available and that no direct measurements (with or without noise) of $g(\theta)$ are available. This is identical to the well-known framework of Kiefer–Wolfowitz SA (Ruppert, 1983).

SPSA has the standard iterative form

$$\hat{\theta}_{k+1} = \hat{\theta}_k - a_k \hat{g}_k(\hat{\theta}_k), \tag{1}$$

where a_k represents a scalar gain coefficient and $\hat{g}_k(\cdot)$ represents the SP approximation to the unknown gradient $g(\cdot)$. In Spall (1987, 1992), $\hat{g}_k(\cdot)$ is represented by a two-measurement approximation. For this paper, we propose the following gradient approximation based on one measurement of the loss function:

$$\hat{g}_k(\hat{\theta}_k) = \frac{y_k}{c_k}\begin{bmatrix}\Delta_{k1}^{-1}\\ \Delta_{k2}^{-1}\\ \vdots \\ \Delta_{kp}^{-1}\end{bmatrix}, \tag{2}$$

where $y_k = L(\hat{\theta}_k + c_k\Delta_k) + \epsilon_k$, c_k is a positive scalar, $\Delta_k = (\Delta_{k1}, \Delta_{k2}, \ldots, \Delta_{kp})^{\mathrm{T}}$ is a vector of zero-mean independent random variables (typically generated by Monte Carlo), and ϵ_k is the measurement noise. Note that ϵ_k is not necessarily independent of $\hat{\theta}_k$ or Δ_k, but it does satisfy a weaker Martingale-type condition, as given in Section 3 below.

To see intuitively why $\hat{g}_k(\cdot)$ above is reasonable, note, by a Taylor expansion of the lth element about an arbitrary parameter value θ, that

$$\hat{g}_{kl}(\theta) = \frac{L(\theta + c_k\Delta_k) + \epsilon_k}{c_k\Delta_{kl}} = \frac{L(\theta)}{c_k\Delta_{kl}} + \frac{c_k g(\theta)^{\mathrm{T}}\Delta_k}{c_k\Delta_{kl}} + \frac{c_k^2\Delta_k^{\mathrm{T}}H(\theta)\Delta_k}{2c_k\Delta_{kl}} + \frac{c_k^3 L'''(\bar{\theta}_k)\Delta_k\otimes\Delta_k\otimes\Delta_k}{6c_k\Delta_{kl}} + \frac{\epsilon_k}{c_k\Delta_{kl}}, \tag{3}$$

where $H(\theta)$ is the Hessian matrix of $L(\theta)$ and $L'''(\bar{\theta}_k)$ is the third derivative of L with respect to θ^{T} evaluated at some point between θ and $\theta + c_k\Delta_k$ (as in Spall, 1992, Lemma 1). Then, under the same assumptions on the distribution of the $\{\Delta_{ki}\}$ given in Spall (1992) (particularly independence, symmetry and finite inverse moments), we have from (2.3)

$$E\hat{g}_{kl}(\theta) = g_l(\theta) + O(c_k^2) \quad \forall l = 1, 2, \ldots, p,$$

i.e. $\hat{g}_k(\cdot)$ is an unbiased estimator of the true (unknown) gradient to within an $O(c_k^2)$ bias. A fundamental difference in the two-measurement and one-measurement forms of SPSA is that for the two-measurement form, contributions due to

$L(\theta)$ and $H(\theta)$ (the first and third terms on the right-hand side of (3)) are *identically* 0 (versus simply *mean* 0) as a result of cancellation effects. This plays a significant role in the efficiency analysis and suggests why, despite taking twice the number of measurements, SPSA2 is generally the better algorithm to use in most practical applications. However, there are situations where SPSA1 will be the more efficient algorithm, and the theory in Section 3 considers this.

3. *Efficiency*

This section analyzes the relative efficiency of SPSA1 and SPSA2. The arguments here closely follow those of Spall (1992), which were based on an asymptotic distribution theory in Fabian (1968) (the author is unaware of any formal distribution theory for the *finite-sample* case). We shall find that one cannot make a universal claim that either SPSA1 or SPSA2 is the more efficient algorithm but that it is possible to identify classes of problems for which SPSA1 will be (asymptotically) more efficient than SPSA2. The fundamental measure of efficiency here is the number of loss function measurements (not number of iterations *per se*), since it is loss function evaluations that represent the primary cost in the optimization process.

For the most part, the regularity conditions here are identical to those in Spall (1992), and hence will not be repeated.† The only difference is that, because of the different gradient approximations, two of the conditions on the measurement noise are altered slightly. In particular, we assume

$$E(\epsilon_k \mid \hat{\theta}_k, \Delta_k) = 0 \quad \forall k$$

$$\text{var}(\epsilon_k) \to \sigma_\epsilon^2 \quad \text{as } k \to \infty \quad \text{for some } \sigma_\epsilon^2$$

$$(\text{frequently } \text{var}(\epsilon_k) = \sigma_\epsilon^2 \ \forall k)$$

(these contrasts with the corresponding noise conditions on pp. 333 and 335 respectively) of Spall, 1992. Then, following the arguments of Proposition 1 in Spall (1992), we have the following convergence result:

$$\hat{\theta}_k \to \theta \quad \text{a.s.}, \quad k \to \infty$$

For the asymptotic distribution result that is critical to the efficiency analysis, we consider gain sequences of the standard form $a_k = a/(k+1+A)^\alpha$ and $c_k = c/(k+1)^\gamma$, a, c, α, $\gamma > 0$, $A \geq 0$, and let $\beta = \alpha - 2\gamma$. Then, following the assumptions and proof of SPSA2,

$$k^{\beta/2}(\hat{\theta}_k - \hat{\theta}^*) \xrightarrow{\text{dist}} N(\mu, PM_1P^{\mathrm{T}}), \quad k \to \infty, \tag{4}$$

where μ and P are as in Spall (1992, Proposition 2) (the detailed definitions are not critical to the analysis here) and M_1 (the subscript 1 is used to contrast with the corresponding expression M_2 used below for SPSA2) is given by

$$M_1 = a^2c^{-2}\rho^2\{\sigma_\epsilon^2 \,\text{diag}\,[(2\lambda_1 - \beta_+)^{-1}, \ldots, (2\lambda_p - \beta_+)^{-1}] + L(\theta^*)^2 I\}, \tag{5}$$

with $E\Delta_{ki}^{-2} \to \rho^2$ as $k \to \infty \ \forall l$, λ_i is the ith eigenvalue of the matrix $aH(\theta^*)$, and $\beta_+ = \{\beta \text{ if } \alpha = 1; 0 \text{ if } \alpha < 1\}$.

Since β, μ and P are identical in SPSA1 and SPSA2, the difference in efficiency centers on the difference between M_1 and M_2, where

$$M_2 = \tfrac{1}{4}a^2c^{-2}\rho^2\sigma^2 \,\text{diag}\,[(2\lambda_1 - \beta_+)^{-1}, \ldots, (2\lambda_p - \beta_+)^{-1}], \tag{6}$$

with σ^2 representing the variance of the sum of the two noise terms entering the SPSA2 gradient approximation (if the noise terms are uncorrelated then $\sigma^2 = 2\sigma_\epsilon^2$). Note from (5) and (6) that M_1 will tend to be larger (in the matrix sense) than M_2 owing to the presence of the $L(\theta^*)^2 I$ contribution (i.e. $\sigma_\epsilon^2 > \frac{1}{4}\sigma^2$). However, since SPSA1 uses only half the measurements of SPSA2 per iteration, efficiency gains are still possible, as discussed below.

We aim to address the following question: To achieve the *same level* of mean-square accuracy in estimating θ, what is the relative number of loss function measurements needed by SPSA1 and SPSA2? The asymptotic distribution results in (4) and Spall (1992, Proposition 2) provide the machinery for answering this question, provided that the absolute number of measurements in both algorithms is sufficiently large. Let k_1 and k_2 denote the number of iterations and n_1 and n_2 the number of loss function evaluations in SPSA1 and SPSA2 respectively. Then from (4) and Spall (1992, Proposition 2), we seek the ratio

$$\frac{n_1}{n_2} : k_1^{-\beta}(\text{tr}\, PM_1P^{\mathrm{T}} + \mu^{\mathrm{T}}\mu) = k_2^{-\beta}(\text{tr}\, PM_2P^{\mathrm{T}} + \mu^{\mathrm{T}}\mu),$$

since the terms being equated represent the asymptotically based approximation to $E\,\|\hat{\theta}_k - \theta^*\|^2$ for SPSA1 and SPSA2 (this, of course, is under the standard conditions—e.g. uniform integrability—that the second moments of the asymptotic distributions correspond to the second moments of the corresponding random process). From the fact that $k_1 = n_1$, $k_2 = \frac{1}{2}n_2$,

$$\frac{n_1}{n_2} \to \frac{1}{2}\left(\frac{\text{tr}\, PM_1P^{\mathrm{T}} + \mu^{\mathrm{T}}\mu}{\text{tr}\, PM_2P^{\mathrm{T}} + \mu^{\mathrm{T}}\mu}\right)^{1/\beta} \tag{7}$$

as the number of measurements for both procedures becomes large.

The expression (7) provides a powerful means for analyzing the relative asymptotic efficiency (analogous to (4.2) in Spall (1992) for analyzing the efficiency of SPSA and the classical Kiefer–Wolfowitz finite-difference method). In particular, by specifying values of terms that are used in the algorithm (a, c etc.) and that represent properties of the measurements (σ_ϵ^2, $H(\theta^*)$ etc.), one can determine if $n_1/n_2 \leq$ or > 1 as a guide to when SPSA1 or SPSA2 is the more efficient algorithm for a given scenario.

To make (7) more concrete, let us consider an important special case. Suppose $\sigma^2 = 2\sigma_\epsilon^2$ (as mentioned earlier) and that $L(\theta^*) = 0$ (this may be true in certain mean-square minimization or tracking problems or in cases where an equivalent loss function can be formed by subtracting a known value of the original loss function at θ^*). Further, suppose that both algorithms run with the same gain coefficients a and c and that $\beta = \frac{2}{3}$ (corresponding to the asymptotically optimal $\alpha = 1$ and $\gamma = \frac{1}{6}$; see e.g. Fabian (1971) or Chin (1997)). Then, from (5) and (6), $M_1 = 2M_2$ (since M_2 is now identical to M_1 except that σ^2 is replaced by $2\sigma_\epsilon^2$). Under these conditions, (7) reduces to

$$\frac{n_1}{n_2} \to \frac{1}{2}\left(\frac{2\,\text{tr}\, PM_2P^{\mathrm{T}} + \mu^{\mathrm{T}}\mu}{\text{tr}\, PM_2P^{\mathrm{T}} + \mu^{\mathrm{T}}\mu}\right)^{3/2}. \tag{8}$$

From (8), we find that as $\mu^{\mathrm{T}}\mu/\text{tr}\, PM_2P^{\mathrm{T}}$ ranges from 0 to ∞, n_1/n_2 ranges from $\sqrt{2}$ to $\frac{1}{2}$. In particular, $\mu^{\mathrm{T}}\mu/\text{tr}\, PM_2P^{\mathrm{T}} = 0.7024$ is the point under or over which n_1/n_2 is greater than or less than one. Qualitatively, this makes sense, since the μ contribution for both SPSA1 and SPSA2 is identical for a given number of iterations (see (4)); hence, as the μ part dominates, one would expect SPSA1 to be the more efficient algorithm.

Obviously, the analysis associated with (8) is no longer valid if $L(\theta^*) \neq 0$. Then one should appeal to the more general form in (7). In such settings, the range of problems for which n_1/n_2 is asymptotically less than one is smaller.

4. *Numerical example*

This section provides an illustration of SPSA1 to a simple quartic function with $p = 5$. Let

$$L(\theta) = \theta^{\mathrm{T}}\theta + 0.1\sum_{i=1}^{5}\theta_i^3 + 0.01\sum_{i=1}^{5}\theta_i^4,$$

which has $\theta^* = 0$. The measurement noise is independently normally distributed with mean 0 and $\sigma_\epsilon = 0.01$, and

† Let us take this opportunity to clarify two slightly ambiguous conditions in Spall (1992). In Lemma 1, one of the statements should read as follows: '... suppose that $\forall\theta$ in an open ***ball about*** $\hat{\theta}_k$ ***whose radius*** is not a function of k or ω, $L^{(3)}(\theta) \equiv \partial^3 L/\partial\theta^{\mathrm{T}}\,\partial\theta^{\mathrm{T}}\,\partial\theta^{\mathrm{T}}$ exists continuously...' (***bold italics*** highlight change). In Proposition 1, Condition A3 should read: $\sup_k \|\hat{\theta}_k\| < \infty$ a.s.

$\hat{\theta}_0 = (0.1, 0.1, 0.1, 0.1, 0.1)^T$. All simulations were conducted with MATLAB on a UNIX-based workstation.

We shall consider the setting where (8) applies, and shall present numerical results for two settings: one where (asymptotically) $n_1/n_2 > 1$ and one where $n_1/n_2 < 1$. Consistent with practical guidelines given in Spall (1996), we take c of magnitude similar to σ_ϵ (we take $c = 0.06$ for both SPSA1 and SPSA2, somewhat larger than the $c \approx \sigma_\epsilon$ suggested in Spall (1996) to accommodate the greater variability inherent in a one-measurement form of the gradient approximation and the greater rate of decay implies by the associated γ coefficient ($\gamma = 0.16667$ here versus $\gamma = 0.101$ suggested in Spall, 1996)). The Δ_{ki} values were independently Bernoulli ± 1 distribued for all k and i. Also, for both algorithms, we include the additive constant A in the denominator of a_k to enhance numerical stability in the early iterations (i.e. we take $a_k = a/(20 + k)$), which, of course, has no effect on the asymptotic theory.

With $a = 0.17$, (8) implies that the asymptotic ratio $n_1/n_2 = 0.7606$, while with $a = 0.27$, the asymptotic ratio implies the opposite in terms of efficiency, namely $n_1/n_2 = 1.414$. We ran numerical studies for both of these cases. For $a = 0.17$, we took $n_1 = 3042$ and $n_2 = 4000$ (the round number for n_2 was selected, and n_1 was then derived based on the ratio value 0.7606). The value of a was chosen to be slightly greater than the lower bound to a for this loss function of $\frac{1}{6}$, as derived from the conditions in Fabian (1968) or Spall (1992). The theory associated with (8) suggests that the mean-square errors (MSEs) associated with these two sample sizes will be equal, and, in fact, that is nearly what happened. The ratio of observed MSEs (SPSA1/SPSA2) was 0.98, based on an average of the terminal observed MSEs over 500 realizations for each algorithm. For $a = 0.27$, we take $n_1 = 4000$ and $n_2 = 2828$, consistent with the ratio 1.414 mentioned above. The ratio of observed MSEs here was 0.91 (so SPSA1 performed slightly better than asymptotic theory would predict).

There were some other observations associated with the relative behavior of SPSA1 and SPSA2. Although we found that SPSA1 could outperform SPSA2 in certain cases (as above), we also found that it was more sensitive to small changes in the underlying data-generating process and to the choice of initial conditions. In fact, for certain initial conditions not close to the solution, SPSA1 diverged while SPSA2 converged, including cases where, based on asymptotic thery, SPSA1 is the more efficient algorithm. The reason for this relative lack of robustness is apparent by comparing the Taylor expansion in (3) with the analogous expansion for SPSA2: for SPSA2, the terms involving $L(\theta)$ and $H(\theta)$ disappear versus merely having mean 0 in SPSA1. When θ is far from θ^*, these terms may significantly degrade the gradient estimate (hence the greater sensitivity to initial conditions). These numerical results suggest that although SPSA1 can be more efficient than SPSA2 in some implementations, one should exercise caution in its implementation, since nonasymptotic effects may have a greater detrimental impact.

5. *Concluding remarks and extensions*

This paper has presented an extension of the SPSA algorithm of Spall (1987, 1992). The primary contribution is to reduce from two to one the number of loss function measurements needed per iteration (to approximate the gradient of the loss). Theory has been presented that identifies the class of problems for which this twofold savings in measurements per iteration translates into a measurement savings over the course of the complete optimization process. In particular, asymptotic results have been presented that identify settings where the one- and two-measurement forms of SPSA yield the same level of mean-square accuracy but where SPSA1 takes less than twice the number of iterations of SPSA2 to achieve this accuracy (and hence uses fewer loss function measurements).

One of the areas for which SPSA1 seems most appropriate is in feedback control problems. As described in, say, Spall and Cristion (1994, 1995) or Rezayat (1995), one can use SPSA to build controllers without the need to build a model of the process. Although this can be very effective for the control of complex (nonlinear stochastic) systems, there is the potential for difficulty if the process dynamics change dramatically in the course of collecting the two measurements for the SPSA2 gradient approximation. SPSA1 has obvious potential advantages in such settings, since it approximates the gradient based on one instantaneous measurement. Other topics of potential interest are to explore the connection of SPSA1 and SPSA2 to the one- and two-measurement algorithms of Polyak and Tsybakov (1992) and Maeda (1996), to evaluate whether the iterate averaging idea of Ruppert (1988, 1991) and Polyak and Juditsky (1992) might enhance convergence properties, and to determine how constrained optimization might be implemented (Sadegh (1996) treats the SPSA2 case).

In summary, for most problems, the two-measurement form of SPSA previously introduced will be the preferred algorithm. Asymptotic theory suggests that it will generally be the more efficient algorithm (in terms of total function evaluations required), and empirical evidence suggests that it is also a more robust algorithm (to changes in initial conditions, gain coefficients, noise levels etc.). However, there is a class of problems for which the one-measurement form of SPSA is (asymptotically) more efficient, and it is for this class that the algorithm of this paper should be considered. This is especially the case in feedback control applications, where nonstationarities may make the instantaneous aspect of the one-measurement gradient approximation especially appealing.

Acknowledgements—This work was supported by the JHU/APL IRAD Program and U.S. Navy Contract N00039-95-C-0002.

References

Allessandri, A. and T. Parisini (1995). Nonlinear modelling and state estimation in a real power plant using neural networks and stochastic approximation. In *Proc. American Control Conf.*, Seattle, WA, pp. 1561–1567.

Chin, D. C. (1994). A more efficient global optimization algorithm based on Styblinski and Tang. *Neural Networks,* **7,** 573–574.

Chin, D. C. (1997). Comparative study of stochastic algorithms for system optimization based on gradient approximations. *IEEE Trans. Syst. Man Cybernetics,* **SMC-27,** in press.

Fabian, V. (1968). On asymptotic normality in stochastic approximation. *Ann. Math. Statist.,* **39,** 1327–1332.

Fabian, V. (1971). Stochastic approximation. In J. J. Rustagi (Ed.), *Optimizing Methods in Statistics,* pp. 439–470, Academic Press, New York.

Hill, S. D. and M. C. Fu (1995). Transfer optimization via simultaneous perturbation stochastic approximation. In C. Alexopoulos, K. Kang, W. R. Lilegdon and D. Goldsman (Eds), *Proc. Winter Simulation Conf.,* Washington, DC, pp. 242–249.

Maeda, Y. (1996). Time difference simultaneous perturbation method. *Electron. Lett.,* **32,** 1016–1018.

Maeda, Y., H. Hirano and Y. Kanata (1995). A learning rule of neural networks via simultaneous perturbation and its hardware implementation. *Neural Networks,* **8,** 251–259.

Polyak, B. T. and A. B. Juditsky (1992). Acceleration of stochastic approximation by averaging. *SIAM J. Control Optim.,* **30,** 838–855.

Polyak, B. T. and A. B. Tsybakov (1992). On stochastic approximation with arbitrary noise (the K-W case). *Adv. Soviet Math.,* **12,** 107–113.

Rezayat, F. (1995). On the use of an SPSA-based model-free controller in quality improvement. *Automatica,* **31,** 913–915.

Ruppert, D. (1983). Kiefer–Wolfowitz procedure. In N. L. Johnson and S. Kotz (Eds), *Encyclopedia of Statistical Sciences,* Vol. 4, pp. 379–381. Wiley, New York.

Ruppert, D. (1988). Efficient estimators from a slowly convergent Robbins–Monro process. Technical Report 781, School of Operations Research and Industrial Engineering, Cornell University.

Ruppert, D. (1991). Stochastic approximation. In B. K. Ghosh and P. K. Sen (Eds), *Handbook of Sequential Analysis*, pp. 503–529. Marcel Dekker, New York.

Sadegh, P. (1996). Constrained optimization via stochastic approximation with a simultaneous perturbation gradient approximation. *Automatica*, to appear (also Technical Report 3/96, Institute of Mathematical Modelling, Technical University of Denmark, Lyngby).

Smith, R. H. and D. C. Chin (1995). Evaluation of an adaptive traffic control technique with underlying system changes. In C. Alexopoulos, K. Kang, W. R. Lilegdon and D. Goldsman (Eds), *Proc. Winter Simulation Conf.*, Washington, DC, pp. 1124–1130.

Spall, J. C. (1987). A stochastic approximation technique for generating maximum likelihood parameter estimates. In *Proc. American Control Conf.*, Minneapolis, MN, pp. 1161–1167.

Spall, J. C. (1992). Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Trans. Autom. Control*, **AC-37**, 332–341.

Spall, J. C. (1996). Implementation of the simultaneous perturbation algorithm for stochastic optimization. JHU/APL Technical Report.

Spall, J. C. and J. A. Cristion (1994). Nonlinear adaptive control using neural networks: estimation based on a smoothed form of simultaneous perturbation gradient approximation. *Statistica Sinica*, **4**, 1–27.

Spall, J. C. and J. A. Cristion (1995). Model-free control of nonlinear stochastic systems in discrete time. In *Proc. 34th IEEE Conf. on Decision and Control*, New Orleans, LA, pp. 2199–2204.