Adaptive system optimization using (simultaneous) random directions stochastic approximation

Prashanth L A* and Shalabh Bhatnagar†

Department of Computer Science and Automation, Indian Institute of Science, INDIA

Abstract

We present the first adaptive random directions Newton algorithm under i.i.d., symmetric, uniformly distributed perturbations for a general problem of optimization under noisy observations. We also present a simple gradient search scheme under the aforementioned perturbation random variates. Our Newton algorithm requires generating $N$ perturbation variates and three simulations at each iteration unlike the well studied simultaneous perturbation Newton search algorithm of Spall [2000] that requires $2N$ iterates and four simulations. We prove the convergence of our algorithms to a local minimum and also present rate of convergence results. Our asymptotic mean square errors (AMSE) analysis indicates that our gradient algorithm requires 60% less number of simulations to achieve a given accuracy as compared to a similar algorithm in Kushner and Clark [1978], Chin [1997] that incorporates Gaussian perturbations. Moreover, our adaptive Newton search algorithm results in an AMSE that is on par and sometimes even better than the Newton algorithm of Spall [2000]. Our experiments are seen to validate the theoretical observations. In particular, our experiments show that our Newton algorithm 2RDSA requires only 75% of the total number of loss function measurements as required by the Newton algorithm of Spall [2000] while providing the same accuracy levels as the latter.

Keywords: Stochastic optimization, stochastic approximation, random directions, SPSA.

1 Introduction

Problems of optimization under uncertainty arise in several areas of engineering and science, such as signal processing, operations research, computer networks, and manufacturing systems. The problems themselves may involve system identification, model fitting, optimal control, or performance evaluation based on observed data. We shall be concerned in this paper with the problem of minimizing a function $f$, given only noise-corrupted measurements.

Algorithms based on gradient search in general require objective function gradients to be estimated from loss function measurements. The finite difference Kiefer-Wolfowitz (KW), see Kiefer and Wolfowitz [1952], estimates are the earliest known and require $2N$ system simulations for a gradient estimate where $N$ is the parameter dimension. This makes the scheme, also called finite difference stochastic approximation (FDSA), disadvantageous for large parameter dimensions. The random directions stochastic approximation (RDSA) approach, see Kushner and Clark [1978] (pp. 58-60), alleviates this problem by requiring two system simulations regardless of the parameter dimension. This it does by randomly perturbing all the $N$ parameter component directions using independent random vectors that are uniformly distributed over the surface of the $N$-dimensional unit sphere. Generating these random vectors in practice (particularly for large $N$) is however computationally difficult because of the inherent dependence between component perturbations. It has been observed in Chin [1997], see also Rubinstein [1981], Bhatnagar [2007], Bhatnagar et al. [2013] that RDSA also works in the case when the component perturbations are independent Gaussian or Cauchy distributed.

*prashanth@csa.iisc.ernet.in
†shalabh@csa.iisc.ernet.in
RDSA with Gaussian perturbations has also been independently derived in [Katkovnik and Kulchitsky, 1972] by approximating the gradient of expected performance by its convolution with a multivariate Gaussian that is then seen (via an integration-by-parts argument) as a convolution of the objective function with a scaled Gaussian. This procedure as such requires only one simulation (regardless of the parameter dimension) with a perturbed parameter (vector) whose component directions are perturbed using independent \(N(0, 1)\) samples. A two-simulation finite difference version that has lower bias than the above SF scheme is studied in [Styblinski and Tang, 1990], [Chin, 1997], [Bhatnagar, 2007].

Amongst all gradient based random perturbation approaches, the simultaneous perturbation stochastic approximation (SPSA) of [Spall, 1992] has been the most popular and widely studied in applications, largely due to its ease of implementation and observed numerical performance when compared with other approaches. Here each component direction of the parameter is perturbed using independent, zero mean symmetric Bernoulli distributed random variables. In [Chin, 1997], performance comparisons between RDSA, SPSA, and FDSA have been studied through both analytical and numerical comparisons of the mean square error metric, and it is observed that SPSA outperforms both RDSA under Gaussian perturbations and FDSA.

Within the class of simulation based search methods, there are also methods that estimate the Hessian in addition to the gradient. Such methods are often more accurate than simple gradient search schemes. In [Fabian, 1971], the Hessian is estimated using \(O(N^2)\) samples of the cost objective at each iterate while in [Ruppert, 1985], the Hessian is estimated assuming knowledge of objective function gradients.

During the course of the last fifteen years, there has been considerable research activity aimed at developing adaptive Newton based random search algorithms for stochastic optimization. The first algorithm in this direction, proposed in [Spall, 2000], is based on the simultaneous perturbation approach and involves the generation of \(2N\) independent symmetric Bernoulli distributed random variables at each update epoch unlike (gradient) SPSA that requires only \(N\) such random variables per iterate. The Hessian estimator in this algorithm requires four parallel simulations with different perturbed parameters at each update epoch. Two of these simulations are also used for gradient estimation. The Hessian estimator is projected to the space of positive definite and symmetric matrices at each iterate for the algorithm to progress along a descent direction. In [Bhatnagar, 2005], three other simultaneous perturbation estimators of the Hessian that require three, two, and one simulation(s) respectively have been proposed in the context of long-run average cost objectives. The resulting algorithms incorporate two-timescale stochastic approximation, see Chapter 6 of [Borkar, 2008]. Certain three-simulation balanced simultaneous perturbation Hessian estimates have been proposed in [Bhatnagar and Prashanth, 2015]. In addition, certain Hessian inversion procedures that require lower computational effort have also been proposed, see also [Bhatnagar et al., 2013]. In [Zhu and Spall, 2002], a similar algorithm as in [Spall, 2000] is considered except that for computational simplicity, the geometric mean of the eigen-values (projected to the positive half line) is used in place of the Hessian inverse in the parameter update step. In [Spall, 2009], certain enhancements of the four-simulation Hessian estimates of [Spall, 2000] using some feedback and weighting mechanisms have been proposed. In [Bhatnagar, 2007], Newton based smoothed functional algorithms based on Gaussian perturbations have been proposed. A text book account of random search approaches (both Gradient and Newton based) is available in [Bhatnagar et al., 2013].

We present in this paper novel RDSA algorithms based on uniformly distributed perturbations. These include both first order gradient as well as second order Newton methods. Our contributions can be summarized as follows:

1. We propose an adaptive Newton based RDSA algorithm. The benefits of using our procedure are two-fold. First, it only requires generating \(N\) perturbation variates at each iteration even for the Newton scheme (\(N\) being the parameter dimension), unlike the simultaneous perturbation Newton algorithms of [Spall, 2000], [Bhatnagar, 2005], [Spall, 2009], [Bhatnagar and Prashanth, 2015], [Zhu and Spall, 2002] that require generation of \(2N\) perturbation variates. Second, the number of system simulations required per iteration in our procedure is three, while the same in [Spall, 2000] is four. This results in significant savings in computational effort and resources.

2. A significant novelty of our procedure is that the component perturbation variates are only required to be independent and uniformly distributed over a symmetric interval around zero, unlike [Kushner and Clark, 1978], [Chin, 1997] where these are assumed uniformly distributed over the surface of an \(N\)-dimensional sphere. Generating the latter perturbations is computationally tedious.
We also present for the first time a gradient RDSA scheme with our proposed i.i.d., uniform perturbation variates. Our gradient estimates require two parallel simulations with balanced (randomly) perturbed parameters while the Hessian estimates in the Newton procedure are also balanced and require three simulations with an additional simulation based on the nominal (running) parameter update. The overall procedure thus requires three simulations for the Newton scheme and two simulations for the gradient algorithm.

We show the unbiasedness of our gradient and Hessian estimators and prove the almost sure convergence of our algorithms to a local minimum of the objective function.

We also analyze the rate of convergence of our algorithms and show comparisons with FDSA, SPSA, and RDSA with Gaussian distributed perturbations. From the asymptotic mean square errors (AMSE) analysis, we observe that

(a) the gradient RDSA scheme with uniform perturbation variates that we propose requires 60% less number of simulations to achieve a given accuracy in comparison to an RDSA scheme that uses Gaussian perturbations; and

(b) the adaptive RDSA scheme with the Hessian estimate that we propose results in an AMSE that is on par and sometimes even better than the Newton algorithm (2SPSA) of Spall [2000]. A significant advantage of our adaptive RDSA scheme is that it uses only three simulations per iteration, while 2SPSA requires four, to achieve the same accuracy.

Finally, our numerical results show that our Newton algorithm 2RDSA requires only 75% of the number of function evaluations as required by 2SPSA (Spall [2000]) while providing the same accuracy.

2 First order random directions SA (1RDSA)

We want to minimize a function $f(\cdot)$, given only noise-corrupted measurements. A first-order scheme for this problem has the following form:

$$x_{n+1} = x_n - a_n \nabla f(x_n).$$

where $\nabla f(x_n)$ is an estimate of $\nabla f(x_n)$. RDSA is a well-known random search scheme that estimates the gradient $\nabla f$ using only noisy measurements of $f$. As illustrated in Fig. 1 the idea is to obtain noisy measurements of $f$ at parameters $x_n + \delta_n d_n$ and $x_n - \delta_n d_n$. Let us denote these values by $y^+_n$ and $y^-_n$, i.e., $y^+_n = f(x_n + \delta_n d_n) + \xi^+_n$, $y^-_n = f(x_n - \delta_n d_n) + \xi^-_n$. Here $\xi^+_n, \xi^-_n$ represent $\mathbb{R}$-valued noise terms that satisfy $\mathbb{E}[\xi^+_n | \mathcal{F}_n] = \mathbb{E}[\xi^-_n | \mathcal{F}_n] = 0$, where $\mathcal{F}_n = \sigma(x_m, m \leq n; d_m, m < n)$ denotes a sequence of sigma-fields.

Our RDSA estimate of the gradient is given by

$$\nabla f(x_n) = \frac{3}{\eta^2} d_n \left[ \frac{y^+_n - y^-_n}{2\delta_n} \right].$$

Figure 1: Overall flow of 1-RDSA algorithm.
In the above, \( d_n = (d_{n1}, \ldots, d_{nN})^T \), where \( d_{ni}, i = 1, \ldots, N \) are independent random variables with distribution \( U[-\eta, \eta] \) for some \( \eta > 0 \).

**Remark 1. (Computational Efficiency)** As described previously, we are the first to study RDSA under i.i.d., symmetric, uniformly distributed perturbations whereas previous studies, see [Kushner and Clark, 1978; Section 2.3.5] and [Chin, 1997]), assumed the perturbation vector \( \eta \) to be uniformly distributed over the surface of the unit sphere. Generating samples of such distributions is far more computationally tedious than the i.i.d. uniform samples that we require for our algorithm.

### 2.2 Main results

We shall make the following assumptions:

(A1) \( f : \mathbb{R}^N \rightarrow \mathbb{R} \) is a continuously differentiable function with bounded second derivative.

(A2) The step-sizes \( a_n, \delta_n > 0, \forall n \) with \( a_n, \delta_n \rightarrow 0 \) as \( n \rightarrow \infty \), \( \sum_n a_n = \infty \), \( \sum_n \left( \frac{\delta_n}{n} \right)^2 < \infty \).

(A3) \( \sup_n \|x_n\| < \infty \) w.p. 1.

(A4) For some \( \alpha_0, \alpha_1 > 0 \) and for all \( n \), \( \mathbb{E} \xi_n^{1+2} \leq \alpha_0 \) and \( \mathbb{E} (x_n + \delta_n d_n)^2 \leq \alpha_1 \).

The above assumptions are standard in the analysis of simultaneous perturbation methods, cf. [Spall, 1992]. In particular, (A1) is required to ensure the underlying ODE is well-posed, (A2) comprises of standard stochastic approximation conditions on step-sizes and perturbation parameters, (A3) is a stability assumption required to ensure that (1) converges and finally, (A4) is another standard stochastic approximation assumption that is required in order to ensure that the associated martingale sequence in the analysis of (1) is almost surely convergent.

We next present three results: First, Lemma 1 establishes that the bias in the gradient estimate (2) is of the order \( O(\delta_n^2) \). Second, Theorem 1 proves that the iterate \( x_n \) governed by (1) converges almost surely and finally, Theorem 2 provides a central limit theorem type result. The reader is referred to Section 5 for a sketch of the proofs, with detailed proofs being provided in Appendix A.

**Lemma 1. (Bias in the gradient estimate)** Under (A1)-(A4), we have almost surely

\[
\mathbb{E} \left[ \nabla f(x_n) \bigg| x_n \right] - \nabla f(x_n) = O(\delta_n^2). \tag{3}
\]

**Remark 2.** The above lemma gives the bias in the gradient estimate of 1RDSA and previous results (see [Kushner and Clark, 1978; Chin, 1997]) indicate that this bias is of the same order as of 1RDSA with perturbations \( d_n \) that are either Gaussian or distributed on the surface of the unit sphere as well as SPSA. Computationally, the cost of generating perturbations in our scheme is comparable to that of SPSA, while the other perturbation choices for RDSA are expensive.

**Theorem 1. (Strong Convergence)** Let \( x^* \) be an (unique) asymptotically stable solution of the following ordinary differential equation (ODE): \( \dot{x}_t = -\nabla f(x_t) \), with domain of attraction \( D(x^*) \), i.e., \( D(x^*) := \{ x_0 \, | \, \lim_{t \rightarrow \infty} x(t \mid x_0) = x^* \} \), where \( x(t \mid x_0) \) is the solution to the aforementioned ODE with initial condition \( x_0 \). Assume (A1)-(A4). Then, if there exists a compact subset \( X \) of \( D(x^*) \) such that \( x_n \in X \) infinitely often, we have

\[
x_n \rightarrow x^* \text{ as } n \rightarrow \infty \text{ for almost all sample points.}
\]

We require the following variant of (A4) in order to establish a CLT for 1RDSA:

(A4′) For some \( \zeta, \alpha_0, \alpha_1 > 0 \) and for all \( n \), \( \mathbb{E} \xi_n^{1+2+\zeta} \leq \alpha_0 \) and \( \mathbb{E} f(x_n + \delta_n d_n)^{1+2+\zeta} \leq \alpha_1 \).

Let \( a_n = a_0 / n^\alpha \) and \( \delta_n = \delta_0 / n^\gamma \), where \( a_0, \delta_0 > 0, \alpha \in (0, 1] \) and \( \gamma \geq 1/6 \) such that (A2) holds. Let \( \nabla^2 f(x) \) denote the Hessian of \( f \) and \( \mathbb{E} (\xi_n^2 - \xi_n)^2 \rightarrow \sigma^2 \) as \( n \rightarrow \infty \).
Theorem 2. (Asymptotic Normality) Consider (A1)-(A3) and (A4'). Let $\beta = \alpha - 2\gamma > 0$ and $P$ be an orthogonal matrix with $P\nabla^2 f(x)P^T = \frac{1}{a_0} \text{diag}(\lambda_1, \ldots, \lambda_N)$. Then,

$$n^{3/2}(x_n - x^*) \to_{n \to \infty} N(\mu, PMP^T),$$

where $M = \frac{a_0^2 \sigma^2}{4\delta_0^2} \text{diag}((2\lambda_1 - \beta_+)^{-1}, \ldots, (2\lambda_N - \beta_+)^{-1})$ and $\mu = 1.8(a_0 \nabla^2 f(x^*) - \beta^+ P\nabla I/2)\nabla^T$ if $\gamma = \alpha/6$ and $0$ if $\gamma > \alpha/6$. Here $I$ is the identity matrix of size $N \times N$, $\beta_+ = \beta$ if $\alpha = 1$ and $0$ if $\alpha < 1$ and $T$ is a $N$-vector with $l$th element given by

$$\frac{1}{6} a_0 \delta_0^2 \left[ \nabla_{ii}^2 f(x^*) + 3 \sum_{i=1,i\neq l}^{N} \nabla_{ij}^2 f(x^*) \right].$$

3 Second order random directions SA (2RDSA)

3.1 The basic algorithm

A second order algorithm has the following form:

$$x_{n+1} = x_n - a_n \tilde{\Upsilon}(\tilde{\Pi}_n)^{-1}\tilde{\nabla} f(x_n),$$

$$\tilde{\Pi}_n = \frac{n}{n+1} \tilde{\Pi}_{n-1} + \frac{1}{n+1} \tilde{H}_n,$$

where $\tilde{\nabla} f(x_n)$ is the estimate of $\nabla f(x_n)$ defined in (3) and $\tilde{\Pi}_n$ is an estimate of the Hessian $\nabla^2 f(\cdot)$. Further, $\tilde{\Upsilon}$ is an operator that projects a matrix onto the set of positive definite matrices. In order to ensure that the recursion (5) moves along a descent direction, it is required that the eigenvalues of the Hessian estimate are positive and the operator $\tilde{\Upsilon}$ ensures that this happens. The basic algorithm above is similar to the adaptive scheme analyzed by Spall [2000]. However, the gradient and Hessian estimates in our case use random directions SA, while Spall [2000] employs SPSA based estimates for the same, see Remark 3.

As in the case of 1RDSA, we are given only noisy measurements of $f$ and we use three measurements per iteration in (5) to estimate both the gradient as well as the Hessian of the objective $f$. These measurements correspond to parameter values $x_n$, $x_n + \delta_n d_n$, and $x_n - \delta_n d_n$, respectively. Let us denote these values by $y_n$, $y_n^+$ and $y_n^-$, i.e.,

$$y_n = f(x_n) + \xi_n, \quad y_n^+ = f(x_n + \delta_n d_n) + \xi_n^+, \quad y_n^- = f(x_n - \delta_n d_n) + \xi_n^-.$$ 

Here $\xi_n, \xi_n^+, \xi_n^-$ denote the noise terms that satisfy $E[\xi_n^+ | F_n] = E[\xi_n^- | F_n] = E[\xi_n | F_n] = 0$.

Using the three measurements and the random directions obtained from $d_n$, we form the Hessian estimate $\tilde{H}_n$ as follows:

$$\tilde{H}_n = \frac{9}{2\eta^2} M_n \left( \frac{y_n^+ + y_n^- - 2y_n}{\delta_n^2} \right),$$

where $M_n := \left[ \begin{array}{cccc} \frac{5}{2} \left( (d_n^2)^2 - \frac{\sigma^2}{3} \right) & d_n^2 d_n^1 & \cdots & d_n^2 d_n^N \\ d_n^2 d_n^1 & \frac{5}{2} \left( (d_n^1)^2 - \frac{\sigma^2}{3} \right) & \cdots & d_n^1 d_n^N \\ \vdots & \vdots & \ddots & \vdots \\ d_n^N d_n^1 & d_n^N d_n^2 & \cdots & \frac{5}{2} \left( (d_n^N)^2 - \frac{\sigma^2}{3} \right) \end{array} \right].$

Henceforth, we shall refer to algorithm (5)–(6) with Hessian estimate (7) as 2RDSA.

Remark 3. In comparison to the second order SPSA algorithm (2SPSA) from Spall [2000], we would like to point out that the computational effort per-iteration of 2RDSA is significantly low. The former algorithm requires 4 measurements per iteration and requires the generation of $2N$ perturbation random variables, while 2RDSA uses 3 function measurements and requires only $N$ perturbation random variables at each iteration epoch. Our experiments also show that 2RDSA gives the same accuracy as 2SPSA while requiring only 75% of the loss function measurements.
3.2 Main results

Recall that $\mathcal{F}_n = \sigma(x_m, m \leq n; d_m, m < n)$ denotes the underlying sequence of sigma-fields. We make the following assumptions that are similar to [Spall 2000].

(C1) $E[\xi^+_n | \mathcal{F}_n] = E[\xi^-_n | \mathcal{F}_n] = E[\xi_n | \mathcal{F}_n] = 0$ a.s. for all $n$.

(C2) The step-sizes $a_n, \delta_n > 0, \forall n$ with $a_n, \delta_n \to 0$ as $n \to \infty$, $\sum_n a_n = \infty$ and $\sum_n \left(\frac{a_n}{\delta_n}\right)^2 < \infty$.

(C3) For some $\rho > 0$ and almost all $x_n$, the function $f$ is thrice continuously differentiable with a uniformly (in $y$) bounded fourth derivative for all $x$ such that $\|y - x\| \leq \rho$.

(C4) For each $n$ and all $x$, there exists a $\rho > 0$ not dependent on $n$ and $x$, such that $(x - x^*)^T \bar{f}_n(x) \geq \rho \|x_n - x\|$, where $f_n(x) = \Upsilon(\mathcal{P}_n)^{-1} \nabla f(x)$.

(C5) For each $i = 1, \ldots, N$ and any $\rho > 0$,

$$P(\{\bar{f}_{ni}(x_n) \geq 0 \text{ i.o.} \} \cap \{\bar{f}_{ni}(x_n) < 0 \text{ i.o.} \} \cap \{|x_{ni} - x_i^*| \geq \rho \ \forall i\}) = 0.$$

(C6) $\Upsilon(H_n)^{-1}$ exists a.s for all $n$ and $\delta_n^2 \Upsilon(H_n)^{-1} \to 0$ a.s. Further, for some $\zeta, \rho > 0$,

$$E(\|\Upsilon(H_n)^{-1}\|^{2+\zeta}) \leq \rho.$$

(C7) For any $\tau > 0$ and nonempty $S \subseteq \{1, \ldots, N\}$, there exists a $\rho'(\tau, S) > \tau$ such that

$$\limsup_{n \to \infty} \left| \frac{\sum_{i \notin S} |x_{ni} - x_i^*| \bar{f}_{ni}(x)}{\sum_{i \in S} |x_{ni} - x_i^*| \bar{f}_{ni}(x)} \right| < 1 \text{ a.s.}$$

for all $|(x - x^*)_i| < \tau$ when $i \notin S$ and $|(x - x^*)_i| \geq \rho'(\tau, S)$ when $i \in S$.

Lemma 2. (Bias in Hessian estimate) Under (C1)-(C7), we have almost surely

$$E \left[ \tilde{H}_n | x_n \right] - \nabla^2 f(x_n) = O(\delta_n^2). \quad (8)$$

Remark 4. (Comparison with 2SPSA) From the above lemma, it is evident that the bias in the Hessian estimate in 2RDSA is of the same order as that in 2SPSA of [Spall 2000]. An advantage with our scheme is that it requires three system simulations per iteration, while the corresponding number is four for 2SPSA.

Theorem 3. (Strong Convergence - parameter) Assume (C1)-(C7). Then $x_n, n \geq 1$ obtained from (5) converges almost surely to $x^*$ as $n \to \infty$.

For ensuring convergence of the Hessian recursion, we require the following additional assumptions:

(C8) In addition to (C2), assume that $\sum_n \frac{1}{(n+1)\delta_n} < \infty$.

(C9) For some $\alpha_0, \alpha_1 > 0$ and for all $n$, $E\xi_n^2 \leq \alpha_0, E\xi_n^{2+\zeta} \leq \alpha_0, Ef(x_n)^2 \leq \alpha_1$ and $Ef(x_n + \delta_n d_n)^2 \leq \alpha_1$.

Theorem 4. (Strong Convergence - Hessian) Assume (C1)-(C9). Then $\mathcal{P}_n$ governed by (6) converges almost surely to $\nabla^2 f(x^*)$.

We next present a CLT for 2RDSA, which requires the following additional assumptions:

(C10) For some $\zeta, \alpha_0, \alpha_1 > 0$ and for all $n$, $E\xi_n^{2+\zeta} \leq \alpha_0, E\xi_n^{2+2\zeta} \leq \alpha_0, Ef(x_n)^{2+\zeta} \leq \alpha_1$ and $Ef(x_n + \delta_n d_n)^{2+\zeta} \leq \alpha_1$.

(C11) The operator $\Upsilon$ is chosen such that $\Upsilon(\mathcal{P}_n) - \mathcal{P}_n \to 0$ as $n \to \infty$.

Let $a_n = a_0/n^\alpha$ and $\delta_n = \delta_0/n^\beta$, where $a_0, \delta_0 > 0, \alpha \in (0, 1]$ and $\gamma \geq 1/6$. Let $\beta = \alpha - 2\gamma$. Let $E(\xi_n^+ - \xi_n^-)^2 \to \sigma^2$ as $n \to \infty$. The main result is as follows:
Theorem 5. (Asymptotic Normality) Under (C1)-(C9) and assuming $\nabla^2 f(x^*)^{-1}$ exists, we have

$$n^{3/2}(x_n - x^*) \xrightarrow{\text{dist}}_{n \to \infty} N(\mu, \Omega),$$

where

$$\Omega = \frac{a_0^2 \sigma^2}{4 \delta_{\mu}^2 (8a_0 - 4\beta_+)} \nabla^2 f(x^*)^{-2}$$

and

$$\mu = (a_0 - \beta_+/2)^{-1} \nabla^2 f(x^*)^{-1} T$$

if $\gamma = \alpha/6$ and $0$ if $\gamma > \alpha/6$, with $\beta_+ = \beta$ if $\alpha = 1$ and $0$ if $\alpha < 1$. Further, $T$ is as defined in Theorem 3.

The reader is referred to Section 5 for a sketch of the proofs, with proof details in Appendix B.

4 Convergence rates using AMSE

The result in Theorem 2 shows that $n^{3/2}(x_n - x^*)$ asymptotically normal for 1RDSA. The related second moment for this distribution, denoted by $E_{\text{1RDSA}}(a, c)$, is given by

$$E_{\text{1RDSA}}(a_0, \delta_0) := \mu^T \mu + \text{trace}(PMP^T),$$

where $a_0$ is the step-size constant, $\delta_0$ is the constant in the perturbation sequence $\delta_k$ and $\mu$, $P$ and $M$ are as defined in Theorem 2. Under certain assumptions (cf. Gerencsér [1999]), it can be shown that $E_{\text{1RDSA}}(a, c)$ coincides with $n^3 \text{E} \|x_n - x^*\|^2_2$, where $\beta = 2/3$.

With step-size $a_n = a_0/n$, setting $a_0$ optimally requires knowledge of the minimum eigenvalue $\lambda_0$ of the Hessian $\nabla^2 f(x^*)$, i.e., $a_0 > \beta/2\lambda_0$. Under this choice, we obtain

$$E_{\text{1RDSA}}(a_0, \delta_0) = (3.6\delta_{\alpha}^2 a_0 \| (2a_0 \nabla^2 f(x^*) - \beta)^{-1} T \|_2^2 + \delta_0^{-2} \text{trace} ((2a_0 \nabla^2 f(x^*) - \beta)^{-1} S),$$

where $T$ is as defined in Theorem 2 and $S = \frac{\gamma}{2} I$. Since $\lambda_0$ is unknown, obtaining the above rate is problematic and one can get rid of the dependency of $a_0$ on $\lambda_0$ either by averaging of iterates or employing an adaptive (second order) scheme. The former would employ step-size $a_0 = a_0/n^\alpha$, with $\alpha \in (1/2, 1)$ and couple this choice with averaging of iterates as $\bar{x}_n = 1/n \sum_{m=1}^{n-1} x_m$. The latter adaptive scheme would correspond to 2RDSA, which performs a Newton step to update $x_n$ in (3). The AMSE for the adaptive and iterate averaged variants are denoted by $E_{\text{2RDSA}}(a_0, \delta_0)$ and $E_{\text{1RDSA-avg}}(\delta_0)$, respectively and can be derived using Theorems 2 and 3 as follows:

$$E_{\text{2RDSA}}(a_0, \delta_0) = \left(3.6\delta_{\alpha}^2 a_0 \| \frac{\beta}{2a_0 - \beta} \frac{\nabla^2 f(x^*)^{-1} T}{2} \|_2^2 + \frac{a_0^2}{\delta_{\alpha}^2 (2a_0 - \beta)} \text{trace} \left( \frac{\nabla^2 f(x^*)^{-1}}{2} S \frac{\nabla^2 f(x^*)^{-1}}{2} T \right) \right),$$

$$E_{\text{1RDSA-avg}}(\delta_0) = \left(3.6\delta_{\alpha}^2 a_0 \| \frac{\beta}{2a_0 - \beta} \frac{\nabla^2 f(x^*)^{-1} T}{2} \|_2^2 + \frac{1}{\delta_{\alpha}^2 (2a_0 - \beta)} \text{trace} \left( \frac{\nabla^2 f(x^*)^{-1}}{2} S \frac{\nabla^2 f(x^*)^{-1}}{2} T \right) \right).$$

Setting $a_0 = 1$ for 2RDSA, we see that one obtains the same rate as that with iterate averaging. Moreover, both these schemes do not have the dependency on $\lambda_0$. Moreover, using the reasoning similar to that in Dippon and Renz [1997] (see expressions (5.2) and (5.3) there), we obtain

$$\forall \delta_0, E_{\text{2RDSA}}(1, \delta_0) < 2 \min_{a_0 > \beta/(2\lambda_0)} E_{\text{1RDSA}}(a_0, \delta_0).$$

Note that the above bound holds for any choice of $\delta_0$. Thus, 2RDSA is a robust scheme as a wrong choice for $a_0$ would adversely affect the bound for 1RDSA, while 2RDSA has no such dependency on $a_0$.

Remark 5. [Chin [1997]] simplifies the AMSE for 1RDSA by solving $E_{\text{1RDSA}}(a_0, \delta_0)$ for $\delta_0$ after setting $a_0$ optimally using $\lambda_0$. Using $N(0, 1)$ for $d_n$ and comparing the resulting AMSE of 1RDSA to that of first order SPSA with symmetric Bernoulli distributed perturbations, they report a ratio of 3 : 1 for the number of measurements to achieve a given accuracy. Here 3 is a result of the fact that for $N(0, 1)$ distributed $d_n$, $\mathbb{E}d_n^4 = 3$, while 1 for SPSA comes from a bound on the second and inverse second moments, both of which are 1 for the Bernoulli case.

Using $U[-\eta, \eta]$ distributed $d_n$ in 1RDSA would bring down this ratio to 1.8 : 1. However, this result comes with a huge caveat - that $a_0$ and $\delta_0$ are set optimally. Setting these quantities require knowledge of the objective, especially, $\nabla^2 f(x^*)$ and the vector $T$ and these are not known except perhaps in some superficial scenarios.
Comparing 1RDSA vs 1SPSA.

Taking the ratio of AMSE of 1RDSA to that of 1SPSA with symmetric Bernoulli \( \pm 1 \)-valued perturbations, we obtain:

\[
\frac{E_{1RDSA}(a_0, \delta_0)}{E_{1SPSA}(a_0, \delta_0)} = \frac{(2\delta_0 a_0 \| (2a_0 \nabla^2 f(x^*) - \beta)^{-1} T \|_2^2) + \delta_0^{-2} \text{trace} \left( (2a_0 \nabla^2 f(x^*) - \beta)^{-1} S \right)}{(2\delta_0 a_0 \| (2a_0 \nabla^2 f(x^*) - \beta)^{-1} T \|_2^2 + \delta_0^{-2} \text{trace} \left( (2a_0 \nabla^2 f(x^*) - \beta)^{-1} S \right)}
\]

\[
= 1 + \frac{0.24}{1 + (\delta_0^{-2} \text{trace} \left( (2a_0 \nabla^2 f(x^*) - \beta)^{-1} S \right) / (2\delta_0 a_0 \| (2a_0 \nabla^2 f(x^*) - \beta)^{-1} T \|_2^2)}^2
\]

From the above, we observe that 1SPSA has a better AMSE in comparison to 1RDSA, but it is not clear if the difference is ‘large’. This is because the ratio in the denominator above depends on the objective function (via \( \nabla^2 f(x^*) \) and \( T \)) and a high ratio value would make the difference between 1RDSA and 1SPSA negligible. Contrast this with the 1.8 : 1 ratio obtained if one knows the underlying objective function (see remark above).

Comparing 2RDSA vs 2SPSA.

Let \((A) := \frac{2\delta_0^2 \| (\nabla^2 f(x^*)^{-1} T \|_2^2}{2\delta_0 a_0 \| (2a_0 \nabla^2 f(x^*) - \beta)^{-1} S \nabla^2 f(x^*)^{-1}}\) and \((B) := \frac{1}{\eta_0 (2-\beta)} \text{trace} \left( \nabla^2 f(x^*)^{-1} S \nabla^2 f(x^*)^{-1} \right)\). Taking the ratio of AMSE of 2RDSA to that of 2SPSA, we obtain:

\[
\frac{E_{2RDSA}(1, \delta_0)}{E_{2SPSA}(1, \delta_0)} = \frac{3.24(A) + (B)}{(A) + (B)}
\]

However, 2SPSA uses four system simulations per iteration, while 2RDSA uses only three. So, in order to achieve a given accuracy, the ratio of the number of simulations needed for 2RDSA (denoted by \( n_{2RDSA} \)) to that for 2SPSA (denoted by \( n_{2SPSA} \)) is

\[
\frac{n_{2RDSA}}{n_{2SPSA}} = \frac{3}{4} \cdot \frac{E_{2RDSA}(1, 0)}{E_{2SPSA}(1, 0)} = \frac{3}{4} \cdot \frac{3.24(A) + (B)}{(A) + (B)} = 1 + \frac{5.72(A) - (B)}{4(A) + 4(B)}
\]

Thus, if \(5.72(A) - (B) < 0\), 2RDSA’s AMSE is better than 2SPSA. On the other hand, if \(5.72(A) - (B) \approx 0\), 2RDSA is comparable to 2SPSA and finally, in the case where \(5.72(A) - (B) > 0\), 2SPSA is better, but the difference may be minor unless \(5.72(A) >> (B)\), as we have \(4(A) + 4(B)\) is the denominator above. Note that the quantities \((A)\) and \((B)\) are problem-dependent, as they require \(\nabla^2 f(x^*)\) and \(T\) that relate to the objective function.

Unlike the first order algorithms, one cannot conclude that 2SPSA is better an 2RDSA even when \(\nabla^2 f(x^*)\) and \(T\) are known. Given the fact that 2RDSA uses less number of simulations per iteration may tilt the balance in favor of 2RDSA (a fact also confirmed by our experiments). This is because the resulting AMSE for 2RDSA may turn out to be better than that for 2SPSA.

5 Analysis

5.1 Proofs for 1RDSA

Proof. (Lemma[1]) Let \( F_n = \sigma(x_m, m \leq n; d_m, m < n), n \geq 1 \). By Taylor’s expansions, we obtain

\[ f(x_n \pm \delta_n d_n) = f(x_n) \pm \delta_n d_n^T \nabla f(x_n) + \frac{\delta_n^2}{2} d_n^T \nabla^2 f(x_n) d_n + O(\delta_n^3). \]

Hence, \( \mathbb{E} \left[ d_n \left( f(x_n + \delta_n d_n) - f(x_n - \delta_n d_n) \right) \right] = \mathbb{E} \left[ d_n d_n^T \nabla f(x_n) x_n \right] + O(\delta_n^2) \)

\[ = \mathbb{E} \left[ d_n d_n^T \nabla f(x_n) + O(\delta_n^2) \right] = \frac{\eta^2}{3} \nabla f(x_n) + O(\delta_n^2), \]

where the last equality follows from \( \mathbb{E}[(d_n^2)] = \frac{\eta^2}{3} \) and \( \mathbb{E}[d_n d_n^T] = 0 \). The claim follows. \( \square \)

Proof. (Theorems[1] and [2]) These follow from an application of the results from Kushner and Clark [1978] and Fabian [1968]. The reader is referred to Appendix[A] for details. \( \square \)
5.2 Proofs for 2RDSA

Proof. (Lemma 2) We start with suitable Taylor’s expansions of \( f \) to obtain
\[
\frac{f(x_n + \delta_n d_n) + f(x_n - \delta_n d_n) - 2f(x_n)}{\delta_n^2} = \sum_{i=1}^N (d^i_n)^2 \nabla^2_{ii} f(x_n) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N d^i_n d^j_n \nabla^2_{ij} f(x_n) + O(\delta_n^2).
\]
Now, taking the conditional expectation of the Hessian estimate \( \hat{H}_n \) and observing that \( \mathbb{E}[\xi_n^+ + \xi_n^- - 2\xi_n \mid \mathcal{F}_n] = 0 \), we obtain the following:
\[
\mathbb{E}[\hat{H}_n \mid \mathcal{F}_n] = \frac{9}{2\eta^4} \mathbb{E} \left[ M_n \times \left( \sum_{i=1}^{N-1} (d^i_n)^2 \nabla^2_{ii} f(x_n) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N d^i_n d^j_n \nabla^2_{ij} f(x_n) + O(\delta_n^2) \right) \mid \mathcal{F}_n \right]. \tag{10}
\]
A calculation using the distribution of \( d_n \) shows that the \( l \)th diagonal term inside the conditional expectation in (10) can be simplified as:
\[
\frac{45}{4\eta^4} \mathbb{E} \left[ (d^l_n)^2 - \frac{\eta^2}{3} \left( \sum_{i=1}^{N-1} (d^i_n)^2 \nabla^2_{ii} f(x_n) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N d^i_n d^j_n \nabla^2_{ij} f(x_n) \right) \mid \mathcal{F}_n \right] = \nabla^2_{ll} f(x_n). \tag{11}
\]
Along similar lines, the \((k, l)\)th term in (10) can be simplified as:
\[
\frac{9}{2\eta^4} \mathbb{E} \left[ d^k_n d^l_n \left( \sum_{i=1}^{N-1} (d^i_n)^2 \nabla^2_{ii} f(x_n) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N d^i_n d^j_n \nabla^2_{ij} f(x_n) \right) \mid \mathcal{F}_n \right] = \nabla^2_{kl} f(x_n). \tag{12}
\]
The claim follows. The reader is referred to Appendix B for a detailed proof.

Proof. (Theorems 3 and 4) The proof of Theorem 3 follows in a manner similar to Theorem 1a of Spall [2000], once we observe that the estimate \( \nabla f(x_n) \) in (5) satisfies: \( \mathbb{E}\left[ \nabla f(x_n) \mid x_n \right] = \nabla f(x_n) + \beta_n \), where the bias term \( \beta_n \) is such that \( \beta_n^2 \| \beta_n \| \) is uniformly bounded for sufficiently large \( n \).

For proving Theorem 4 we first use a martingale convergence result to show that \( \frac{1}{n+1} \sum_{m=0}^n \nabla f(x) \to 0 \) a.s. Next, using Proposition 2 we can conclude that \( \frac{1}{n+1} \sum_{m=0}^n \mathbb{E}\left[ \hat{H}_m \mid x_n \right] \to \nabla^2 f(x^*) \) a.s. and the claim follows since \( \nabla f(x_n) = \frac{1}{n+1} \sum_{m=0}^n \hat{H}_m \).

Proof. (Theorem 5) As in the case of 1RDSA, we verify conditions (2.2.1)-(2.2.3) of Fabian [1968] to establish the result and the details are provided in Appendix B.

6 Numerical Experiments

Setting. We compare the performance of 1RDSA and 2RDSA along with their SPSA counterparts 1SPSA and 2SPSA, using the following loss function in \( N = 10 \) dimensions:
\[
f(x) = x^T Ax + b^T x,
\]
where \( A \) is such that \( NA \) is an upper triangular matrix with each entry one, \( b \) is the \( N \)-dimensional vector of ones. The noise structure is similar to that used in Spall [2000]. For any \( x \), the noise is \([x', 1]^T \), where \( z \approx \mathcal{N}(0, \sigma^2 I_{11} \times 11) \). We set \( \sigma = 0.001 \) in our experiment\(^1\). The optimum \( x^* \) for \( f \) is such that each coordinate of \( x^* \) is \(-0.9091 \), with \( f(x^*) = -4.55 \).

For 1RDSA and 1SPSA, we set \( \delta_n = 1.9/n^{0.101} \) and \( a_n = 1/n \). For 2RDSA and 2SPSA, we set \( \delta_n = 3.8/n^{0.101} \) and \( a_n = 1/n^{0.6} \). SPSA algorithms use Bernoulli \( \pm 1 \)-valued perturbations, while RDSA variants use \( U[-1, 1] \) distributed perturbations. For all the algorithms, the initial point \( x_0 \) is the \( N \)-dimensional vector of ones.

\(^1\)The reader is referred to Appendix C for additional experiments with \( \sigma = 0.01 \) and \( \sigma = 0 \).
Table 1: Normalized MSE averaged over 50 replications for $N = 10$ dimensional problem

| No. of function measurements | 1SPSA  | 1RDSA  | 2SPSA  | 2RDSA  |
|------------------------------|--------|--------|--------|--------|
| 1000                         | 0.0407 | 0.0453 | 0.0062 | 0.0026 |
| 2000                         | 0.0325 | 0.0365 | $3.3285 \times 10^{-4}$ | $3.9851 \times 10^{-5}$ |

**Results.** Table 1 presents the normalized mean square error (MSE) for all the algorithms. Normalized MSE is defined as the ratio $\|x_{n_{\text{end}}} - x^*\|_2^2 / \|x_0 - x^*\|_2^2$, where $x_{n_{\text{end}}}$ is the algorithm iterate at the end of the simulation. Note, $n_{\text{end}}$ is algorithm-specific and a function of the number of measurements. For instance, with 2000 measurements, $n = 1000$ for both 1SPSA and 1RDSA, as they use two measurements per iteration. For both 2SPSA and 2RDSA, an initial 25% of the number of measurements was used up by 1SPSA/1RDSA and the resulting iterate was used to initialize 2SPSA/2RDSA. Thus, with 2000 measurements available, the initial 500 measurements are used for 1SPSA/1RDSA and the remaining 1500 are used up by 2SPSA/2RDSA. This results in $n_{\text{end}}$ of $1500/4 = 375$ for 2SPSA and $1500/3 = 500$ for 2RDSA. Note that the difference here is due to the fact the 2RDSA uses 3 simulations per iteration, while 2SPSA needs 4.

From Table 1, it is evident that second-order algorithms outperform the first-order counterparts and this is consistent with earlier results, for instance Spall [2000]. Moreover, 2RDSA results in the best normalized MSE. In fact, running 2RDSA for 375 iterations, which was the number used for 2SPSA with 2000 measurements available (see Table 1), the resulting normalized MSE was found to be $3.3436 \times 10^{-4}$ and this is comparable to the corresponding result of $3.3285 \times 10^{-4}$ for 2SPSA. However, 2RDSA used up only 75% number of simulations in comparison to 2SPSA, in order to achieve this result.

7 Conclusions

We considered a general problem of optimization under noisy observations and presented the first adaptive random directions Newton algorithm based on i.i.d., symmetric, uniformly distributed perturbations. In addition, we also presented a simple gradient search scheme using the aforementioned perturbations. While our gradient search scheme requires the same number of perturbations and system simulations per iteration as the simultaneous perturbation gradient scheme of Spall [1992], our Newton scheme only requires half the number of perturbations and three-fourths the number of simulations as compared to the well known simultaneous perturbation Newton algorithm of Spall [2000]. We proved the convergence of our algorithms and analyzed their rates of convergence. We observed from asymptotic mean square analysis and numerical comparisons that our algorithms are computationally efficient. In particular, our Newton algorithm results in AMSE that is close to or better in some cases than the algorithm in Spall [2000]. Further, our numerical experiments show that our Newton algorithm requires only 75% of the number of function evaluations as required by the Newton algorithm of while providing the same accuracy levels as the latter algorithm.

As future work, it would be of interest to extend our algorithms to scenarios where the noise random variables form a parameterized Markov process and to develop multiscale algorithms in this setting for long-run average or infinite horizon discounted costs. Such algorithms will be of relevance in the context of reinforcement learning, for instance, as actor-critic algorithms.
References

S. Bhatnagar. Adaptive multivariate three-timescale stochastic approximation algorithms for simulation based optimization. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 15(1):74–107, 2005.

S. Bhatnagar. Adaptive Newton-based smoothed functional algorithms for simulation optimization. *ACM Transactions on Modeling and Computer Simulation*, 18(1):2:1–2:35, 2007.

S. Bhatnagar and L. A. Prashanth. Simultaneous perturbation Newton algorithms for simulation optimization. *Journal of Optimization Theory and Applications*, 164(2):621–643, 2015.

S. Bhatnagar, Prasad H.L., and Prashanth L.A. *Stochastic Recursive Algorithms for Optimization: Simultaneous Perturbation Methods (Lecture Notes in Control and Information Sciences)*, volume 434. Springer, 2013.

V. S. Borkar. *Stochastic Approximation: A Dynamical Systems Viewpoint*. Cambridge University Press and Hindustan Book Agency (Jointly Published), Cambridge University Press, U. K. and Hindustan Book Agency, New Delhi (jointly published), 2008.

D. C. Chin. Comparative study of stochastic algorithms for system optimization based on gradient approximations. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 27(2):244–249, 1997.

J. Dippon and J. Renz. Weighted means in stochastic approximation of minima. *SIAM Journal on Control and Optimization*, 35(5):1811–1827, 1997.

V. Fabian. On asymptotic normality in stochastic approximation. *The Annals of Mathematical Statistics*, pages 1327–1332, 1968.

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

L. Gerencsér. Convergence rate of moments in stochastic approximation with simultaneous perturbation gradient approximation and resetting. *IEEE Transactions on Automatic Control*, 44(5):894–905, 1999.

V. Ya Katkovnik and Yu Kulchitsky. Convergence of a class of random search algorithms. *Automation Remote Control*, 8:1321–1326, 1972.

E. Kiefer and J. Wolfowitz. Stochastic estimation of the maximum of a regression function. *Ann. Math. Statist.*, 23:462–466, 1952.

H. J. Kushner and D. S. Clark. *Stochastic Approximation Methods for Constrained and Unconstrained Systems*. Springer Verlag, New York, 1978.

R. G. Laha and V. K. Rohatgi. *Probability Theory*. Wiley, New York, 1979.

R. Y. Rubinstein. *Simulation and the Monte Carlo Method*. Wiley, New York, 1981.

D. Ruppert. A Newton-Raphson version of the multivariate Robbins-Monro procedure. *Annals of Statistics*, 13:236–245, 1985.

J. C. Spall. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Trans. Auto. Cont.*, 37(3):332–341, 1992.

J. C. Spall. Adaptive stochastic approximation by the simultaneous perturbation method. *IEEE Trans. Autom. Contr.*, 45:1839–1853, 2000.

J. C. Spall. Feedback and weighting mechanisms for improving Jacobian estimates in the adaptive simultaneous perturbation algorithm. *IEEE Transactions on Automatic Control*, 54(6):1216–1229, 2009.
M. A. Styblinski and T.-S. Tang. Experiments in nonconvex optimization: stochastic approximation with function smoothing and simulated annealing. *Neural Networks*, 3:467–483, 1990.

X. Zhu and J. C. Spall. A modified second-order SPSA optimization algorithm for finite samples. *Int. J. Adapt. Control Signal Process.*, 16:397–409, 2002.
Appendix

A Proofs for 1RDSA

Proof of Lemma 1

Proof. Recall that $\mathcal{F}_n = \sigma(x_m, m \leq n; d_m, m < n)$, $n \geq 1$.

By Taylor’s expansions, we obtain

$$f(x_n + \delta_n d_n) = f(x_n) + \delta_n d_n \nabla f(x_n) + \frac{\delta_n^2}{2} d_n^T \nabla^2 f(x_n) d_n + O(\delta_n^3),$$

$$f(x_n - \delta_n d_n) = f(x_n) - \delta_n d_n \nabla f(x_n) + \frac{\delta_n^2}{2} d_n^T \nabla^2 f(x_n) d_n + O(\delta_n^3),$$

respectively. Hence,

$$\frac{f(x_n + \delta_n d_n) - f(x_n - \delta_n d_n)}{2\delta_n} = d_n^T \nabla f(x_n) + O(\delta_n^2).$$

From the foregoing,

$$\mathbb{E} \left[ d_n \left( \frac{f(x_n + \delta_n d_n) - f(x_n - \delta_n d_n)}{2\delta_n} \right) \bigg| \mathcal{F}_n \right]$$

$$= \mathbb{E} [d_n d_n^T \nabla f(x_n)] + O(\delta_n^2)$$

$$= \mathbb{E} [d_n d_n^T \nabla f(x_n)] + O(\delta_n^2)$$

$$= \mathbb{E} \left[ \begin{array}{cccc} (d_n^2)^2 & d_n^2 d_n^2 & \cdots & d_n^2 d_N^2 \\ d_n^2 d_n^2 & (d_n^2)^2 & \cdots & (d_n^2)^2 \\ \vdots & \vdots & \ddots & \vdots \\ d_n^2 d_n^2 & d_n^2 d_n^2 & \cdots & (d_N^2)^2 \end{array} \right] \nabla f(x_n) + O(\delta_n^2)$$

$$= \frac{n^2}{3} \nabla f(x_n) + O(\delta_n^2).$$

In the above, for the last equality, we have used the facts that $\mathbb{E}[(d_n^2)^2] = \frac{1}{2n} \int_{-n}^{n} x^2 dx = \frac{n^2}{3}$ and $\mathbb{E} [d_n^2 d_n^2] = \mathbb{E} [d_n^2] \mathbb{E} [d_n^2] = 0$. The claim follows.

Proof of Theorem 1

Proof. We first rewrite the update rule (1) as follows:

$$x_{n+1} = x_n - a_n (\nabla f(x_n) + \eta_n + \beta_n),$$

where $\eta_n = \nabla f(x_n) - \mathbb{E}(\nabla f(x_n) | \mathcal{F}_n)$ is a martingale error term and $\beta_n = \mathbb{E}(\nabla f(x_n) | \mathcal{F}_n) - \nabla f(x_n)$ is the bias in the gradient estimate. Convergence of (13) can be inferred from Theorem 2.3.1 on pp. 29 of Kushner and Clark [1978], provided we verify that the assumptions A2.2.1 to A2.2.3 and A2.2.4” of Kushner and Clark [1978] are satisfied for $x_n$ governed by (1). We mention these assumptions as (B1)-(B4) below.

(B1) $\nabla f$ is a continuous $\mathbb{R}^N$-valued function.

(B2) The sequence $\beta_n, n \geq 0$ is a bounded random sequence with $\beta_n \to 0$ almost surely as $n \to \infty$.

(B3) The step-sizes $a_n, n \geq 0$ satisfy

$$a(n) \to 0 \text{ as } n \to \infty \text{ and } \sum_n a_n = \infty.$$
\( \eta_n, n \geq 0 \) is a sequence such that for any \( \epsilon > 0 \),
\[
\lim_{n \to \infty} P \left( \sup_{m \geq n} \left\| \sum_{i=n}^{m} a_i \eta_i \right\| \geq \epsilon \right) = 0.
\]

The above assumptions can be verified for (13) as follows:

- (A1) implies (B1).
- Lemma 1 above establishes that the bias \( \eta_n \) is \( O(\delta_n^2) \) and since \( \delta_n \to 0 \) as \( n \to \infty \) (see (A2)), it is easy to see that (B2) is satisfied.
- (A2) implies (B3) is satisfied.
- We now verify (B4) using arguments similar to those used by Spall [1992]: We first recall a martingale inequality attributed to Doob (also given as (2.1.7) on pp. 27 of Kushner and Clark [1978]):
\[
P \left( \sup_{m \geq n} \left\| W_m \right\| \geq \epsilon \right) \leq \frac{1}{\epsilon^2} \lim_{m \to \infty} E \left\| W_m \right\|^2
\]
(14)

We apply the above inequality in our setting to the martingale sequence \( \{W_n\} \), where \( W_n := \sum_{i=0}^{n-1} a_i \eta_i \), \( n \geq 1 \), to obtain
\[
P \left( \sup_{m \geq n} \left\| \sum_{i=n}^{m} a_i \eta_i \right\| \geq \epsilon \right) \leq \frac{1}{\epsilon^2} \lim_{n \to \infty} \sum_{i=n}^{\infty} a_i^2 E \left\| \eta_i \right\|^2
\]
(15)

We now bound \( E \left\| \eta_i \right\|^2 \) as follows:
\[
E \hat{\nabla}_i f(x_n)^2 = \frac{3}{\eta^2} (Ed_n)^2 \left\| y_n^+ - y_n^- \right\|_2^2
\]
\[
= \frac{1}{4\delta_n^2} E \left[ y_n^+ - y_n^- \right]^2
\]
\[
\leq \frac{(\alpha_1 + \alpha_0)}{\delta_n^2}.
\]

The second equality above follows owing to the fact that \( (Ed_n)^2 = \frac{2}{\eta^2} \), while the last inequality follows from (A4). Thus, \( E \left\| \eta_i \right\|^2 \leq \frac{N(\alpha_1 + \alpha_0)}{\delta_n^2} \). Plugging this in (15), we obtain
\[
\lim_{n \to \infty} P \left( \sup_{m \geq n} \left\| \sum_{i=n}^{m} a_i \eta_i \right\| \geq \epsilon \right) \leq \frac{N(\alpha_1 + \alpha_0)}{\epsilon^2} \lim_{n \to \infty} \sum_{i=n}^{\infty} a_i^2 E \left\| \eta_i \right\|^2 = 0.
\]

For the equality above, we have used the fact that \( \sum_{n} \left( \frac{\alpha_1}{\delta_n^2} \right)^2 < \infty \) from (A2).

The claim follows from Theorem 2.3.1 on pp. 29 of Kushner and Clark [1978].

**Proof of Theorem 2**

_Proof_. Follows from Proposition 1 of Chin [1997] after observing that \( \frac{3}{\eta^2} Ed_n = I \) and \( \frac{9}{\eta^4} Ed_n^4 = 1.8 \) for any \( i = 1, \ldots, N \).
B Proofs for 2RDSA

Proof of Lemma 2

Proof. By Taylor’s expansions, we obtain

\[
\frac{f(x_n + \delta_n d_n) - f(x_n - \delta_n d_n) - 2f(x_n)}{\delta_n^2} = d_n^T \nabla^2 f(x_n) d_n + O(\delta_n^2)
\]

\[
= \sum_{i=1}^{N} \sum_{j=1}^{N} d_i^T d_j \nabla_{ij}^2 f(x_n) + O(\delta_n^2)
\]

\[
= \sum_{i=1}^{N} (d_i^T)^2 \nabla_{ii}^2 f(x_n) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} d_i^T d_j \nabla_{ij}^2 f(x_n) + O(\delta_n^2)
\]

Now, taking the conditional expectation of the Hessian estimate \( \hat{H}_n \) and observing that \( \mathbb{E}[\xi_n^+ + \xi_n^- - 2 \xi_n | F_n] = 0 \), we obtain the following:

\[
\mathbb{E}[\hat{H}_n | F_n] = \frac{9}{2 \eta^2} \mathbb{E} \left[ M_n \times \left( \sum_{i=1}^{N-1} (d_i^T)^2 \nabla_{ii}^2 f(x_n) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} d_i^T d_j \nabla_{ij}^2 f(x_n) + O(\delta_n^2) \right) | F_n \right] \quad (16)
\]

Note that the \( O(\delta_n^2) \) term inside the conditional expectation above remains \( O(\delta_n^2) \) even after the multiplication with \( M_n \). We analyse the diagonal and non-diagonal terms in the multiplication of the matrix \( M_n \) with the scalar above, ignoring the \( O(\delta_n^2) \) term.

Diagonal terms in \( (16) \).

Consider the \( l \)th diagonal term inside the conditional expectation in \( (16) \):

\[
\frac{45}{4 \eta^2} \left( (d_i^T)^2 - \frac{\eta^2}{3} \right) \left( \sum_{i=1}^{N} (d_i^T)^2 \nabla_{ii}^2 f(x_n) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} d_i^T d_j \nabla_{ij}^2 f(x_n) \right)
\]

\[
= \frac{45}{4 \eta^2} (d_i^T)^2 \sum_{i=1}^{N} (d_i^T)^2 \nabla_{ii}^2 f(x_n) + \frac{45}{2 \eta^2} (d_i^T)^2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} d_i^T d_j \nabla_{ij}^2 f(x_n)
\]

\[
- \frac{15}{4 \eta^2} \sum_{i=1}^{N} (d_i^T)^2 \nabla_{ii}^2 f(x_n) - \frac{15}{2 \eta^2} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} d_i^T d_j \nabla_{ij}^2 f(x_n) \quad (17)
\]

From the distribution of \( d_i^T, d_j^T \) and the fact that \( d_i^T \) is independent of \( d_j^T \) for \( i < j \), it is easy to see that

\[
\mathbb{E} \left( (d_i^T)^2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} d_i^T d_j \nabla_{ij}^2 f(x_n) \left| F_n \right. \right) = 0 \quad \text{and} \quad \mathbb{E} \left( \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} d_i^T d_j \nabla_{ij}^2 f(x_n) \left| F_n \right. \right) = 0,
\]

respectively. Thus, the conditional expectation of the second and fourth terms on the RHS of \( (17) \) equals zero.
The first term on the RHS of (17) with the conditional expectation can be simplified as follows:

\[
\frac{45}{4\eta^2} \mathbb{E} \left( (d_n^i)^2 \sum_{i=1}^N (d_n^i)^2 \nabla_{ii}^2 f(x_n) \mid \mathcal{F}_n \right) = \frac{45}{4\eta^2} \mathbb{E} \left( (d_n^i)^4 \nabla_{ii}^2 f(x_n) + \sum_{i=1, i \neq l}^N (d_n^i)^2 (d_n^l)^2 \nabla_{il}^2 f(x_n) \right)
\]

\[
= \frac{45}{4\eta^2} \left( \frac{\eta^4}{9} \nabla_{ii}^2 f(x_n) + \frac{\eta^4}{9} \sum_{i=1, i \neq l}^N \nabla_{il}^2 f(x_n) \right).
\]

For the second inequality above, we have used the fact that \( \mathbb{E}[(d_n^i)^4] = \frac{1}{2\eta} \int_0^\eta x^4 dx = \frac{9}{5} \) and \( \mathbb{E}[(d_n^i)^2(d_n^l)^2] = \mathbb{E}[(d_n^i)^2] \mathbb{E}[(d_n^l)^2] = \eta^4 / 9 \), \( \forall l \neq i \).

The third term in (17) with the conditional expectation and without the negative sign can be simplified as follows:

\[
\frac{15}{4\eta^2} \mathbb{E} \left( \sum_{i=1}^N (d_n^i)^2 \nabla_{ii}^2 f(x_n) \mid \mathcal{F}_n \right) = \frac{15}{4\eta^2} \sum_{i=1}^N \mathbb{E} [(d_n^i)^2] \nabla_{ii}^2 f(x_n)
\]

\[
= \frac{5}{4} \sum_{i=1}^N \nabla_{ii}^2 f(x_n).
\]

Combining the above, we obtain

\[
\frac{45}{4\eta^2} \mathbb{E} \left[ (d_n^i)^2 - \frac{\eta^2}{3} \right] \left( \sum_{i=1}^N (d_n^i)^2 \nabla_{ii}^2 f(x_n) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N d_n^i d_n^j \nabla_{ij}^2 f(x_n) \right) \mid \mathcal{F}_n \right] = \nabla_{ii}^2 f(x_n).
\]

(18)

**Off-diagonal terms in (16).**

We now consider the \((k, l)\)th term in (16): Assume w.l.o.g that \(k < l\). Then,

\[
\frac{9}{2\eta^2} \mathbb{E} \left[ d_n^k d_n^l \left( \sum_{i=1}^N (d_n^i)^2 \nabla_{ii}^2 f(x_n) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N d_n^i d_n^j \nabla_{ij}^2 f(x_n) \right) \mid \mathcal{F}_n \right]
\]

\[
= \frac{9}{2\eta^2} \sum_{i=1}^N \mathbb{E} \left( (d_n^k d_n^l d_n^i)^2 \nabla_{ii}^2 f(x_n) + \frac{9}{\eta^4} \sum_{i=1}^{N-1} \sum_{j=i+1}^N \mathbb{E} \left( (d_n^k d_n^l d_n^i d_n^j)^2 \nabla_{ij}^2 f(x_n) \right) \right)
\]

\[
= \nabla_{kl}^2 f(x_n).
\]

(19)

The last equality follows from the fact that the first term in (19) is 0 since \(k \neq l\), while the second term in (19) can be seen to be equal to \(\frac{9}{\eta^4} \mathbb{E} \left( (d_n^k d_n^l)^2 (d_n^i)^2 \nabla_{kl}^2 f(x_n) = \nabla_{kl}^2 f(x_n) \right) \). The claim follows.

**Proof of Theorem 3**

**Proof.** From Lemma 1 we observe that the gradient estimate \(\hat{\nabla} f(x_n)\) in (5) satisfies:

\[
\mathbb{E} \left[ \hat{\nabla} f(x_n) \mid x_n \right] = \nabla f(x_n) + \beta_n,
\]

where the bias term \(\beta_n\) is such that \(\delta_n^{-2} \|\beta_n\|\) is uniformly bounded for sufficiently large \(n\).

The rest of the proof follows in a manner similar to Theorem 1a of Spall [2000]. In particular, note that this proof does not assume any form of the Hessian estimate and only requires assumptions (C1)-(C7), which are
similar to those in Spall [2000]. The only variation in our case, in comparison to Spall [2000], is the gradient estimation uses a RDSA scheme while Spall [2000] use first order SPSA. Thus, only the first step of the proof varies and in our case, Lemma 1 controls the bias with same order as that of SPSA, leading to the final result.

**Proof of Theorem 5**

**Proof.** We use the proof technique from Spall [2000]. The proof proceeds along the following steps:

Let \( W_m = \hat{H}_m - \mathbb{E} \left[ \hat{H}_m \mid x_m \right] \). Then, we know that \( \mathbb{E} W_m = 0 \). In addition, we have \( \sum_{m} \frac{\| W_m \|^2}{n^2} < \infty \).

The latter follows by first observing that \( \mathbb{E} \left[ \delta_m^2 \| \hat{H}_m \|^2 \right] < \infty, \forall n \) uniformly as a consequence of (C9) and then coupling this fact with (C8). Now, applying a martingale convergence result from pp. 397 of Laha and Rohatgi [1979] to \( W_m \), we obtain

\[
\frac{1}{n+1} \sum_{m=0}^{n} \hat{H}_m - \mathbb{E} \left[ \hat{H}_m \mid x_m \right] \to 0 \text{ a.s.}
\]  

(21)

From Proposition 2 we know that \( \mathbb{E} \left[ \hat{H}_m \mid x_m \right] = \nabla^2 f(x_n) + O(\delta_n^2). \)

\[
\frac{1}{n+1} \sum_{m=0}^{n} \mathbb{E} \left[ \hat{H}_m \mid x_m \right] = \frac{1}{n+1} \sum_{m=0}^{n} \nabla^2 f(x_m) + O(\delta_n^2) \to \nabla^2 f(x^*) \text{ a.s.}
\]

The final step above follows from the fact that the Hessian is continuous near \( x_n \) and Theorem 3 which implies \( x_n \) converges almost surely to \( x^* \). Thus, we obtain

\[
\frac{1}{n+1} \sum_{m=0}^{n} \hat{H}_m \to \nabla^2 f(x^*) \text{ a.s.}
\]

and the claim follows by observing that \( \bar{H}_m = \frac{1}{n+1} \sum_{m=0}^{n} \hat{H}_m \).

**Proof of Theorem 5**

**Proof.** We use the well-known result for establishing asymptotic normality of stochastic approximation schemes from Fabian [1968]. As in the case of SPSA based algorithms (cf. Spall [1992], Spall [2000]) as well as 1RDSA, it can be shown that, for sufficiently large \( n \), there is a \( \bar{x}_n \) on the line segment that connects \( x_n \) and \( x^* \), such that the following holds:

\[
\mathbb{E}[\tilde{\nabla} f(x_n) \mid x_n] = H(\bar{x}_n)(x_n - x^*) + \beta_n,
\]

where \( \beta_n = \mathbb{E}(\tilde{\nabla} f(x_n) \mid x_n) - \nabla f(x_n) \) is the bias in the gradient estimate. Next, we write the estimation error \( x_{n+1} - x^* \), in a form that is amenable for applying the result from Fabian [1968], as follows:

\[
x_{n+1} - x^* = (I - n^{-\alpha} \Gamma_n)(x_n - x^*) + n^{-(\alpha+\beta)/2} \Phi_n V_n + n^{(\alpha-\beta)/2} \Upsilon(T_n)^{-1} T_n,
\]

(22)

where \( \Gamma_n = a_0 \Gamma(T_n)^{-1} H(\bar{x}_n), \Phi_n = -a_0 \Gamma(T_n)^{-1}, V_n = n^{-1} (\tilde{\nabla} f(x_n) - \mathbb{E}(\tilde{\nabla} f(x_n) \mid x_n)) \) and \( T_n = -a_0 n^{\beta/2} \beta_n \). The above recursion is similar to that for 2SPSA of Spall [2000], except that we estimate the Hessian and gradients using RDSA and not SPSA.

For establishing the main claim, one needs to verify conditions (2.2.1) to (2.2.3) in Theorem 2.2 of Fabian [1968]. This can be done as follows:
From the results in Theorems 3 and 4, we know that $x_n, \nabla^2 f(x_n)$ converge to $x^*$ and $\nabla^2 f(x^*)$, respectively. Thus, $\Gamma_n \to a_0, \Phi_n \to -a_0 \nabla^2 f(x^*)^{-1}$. Moreover, $T_n$ is identical to that in 1RDSA and hence, $T_n \to 0$ if $\gamma > \alpha/6$ and if $\gamma = \alpha/6$, then the limit of $T_n$ is to a vector $T$ as defined in Theorem 2. These observations together imply that condition (2.2.1) of Fabian [1968] is satisfied.

$V_n$ is also identical to that in 1RDSA and hence, $E(V_n V_n^T | x_n) \to \frac{1}{4} \delta^{-1} \sigma^2 I$. This implies condition (2.2.2) of Fabian [1968] is satisfied.

Condition (2.2.3) can be verified using arguments that are the same as those in Spall [1992] for first order SPSA.

Now, applying Theorem 2.2 of Fabian [1968], it is straightforward to obtain the expressions for the mean $\mu$ and covariance matrix $\Gamma$ of the limiting Normal distribution.

\section{Additional Experiments}

\begin{table}[h]
\centering
\caption{Normalized MSE averaged over 50 replications with $\sigma = 0.01$}
\begin{tabular}{|c|c|c|c|}
\hline
No. of function measurements & 1SPSA & 1RDSA & 2SPSA & 2RDSA \\
\hline
1000 & 0.0408 & 0.0453 & 0.0063 & 0.0027 \\
2000 & 0.0325 & 0.0404 & $4.6916 \times 10^{-4}$ & $3.316 \times 10^{-4}$ \\
\hline
\end{tabular}
\end{table}

\begin{table}[h]
\centering
\caption{Normalized MSE averaged over 50 replications with $\sigma = 0$}
\begin{tabular}{|c|c|c|c|}
\hline
No. of function measurements & 1SPSA & 1RDSA & 2SPSA & 2RDSA \\
\hline
1000 & 0.0407 & 0.0453 & 0.0062 & 0.0026 \\
2000 & 0.0325 & 0.0403 & $3.2907 \times 10^{-4}$ & $3.8122 \times 10^{-5}$ \\
\hline
\end{tabular}
\end{table}