Proceedings of the 2016 Winter Simulation Conference
T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, eds.

A RANDOMIZED ALGORITHM FOR CONTINUOUS OPTIMIZATION

Ajin George Joseph
Shalabh Bhatnagar

Dept. of Computer Science and Automation
Indian Institute of Science
Bangalore, 560 094, INDIA

ABSTRACT

The cross entropy (CE) method is a model based search method to solve optimization problems where the objective function has minimal structure. The Monte-Carlo version of the CE method employs the naive sample averaging technique which is inefficient, both computationally and space wise. We provide a novel stochastic approximation version of the CE method, where the sample averaging is replaced with bootstrapping. In our approach, we reuse the previous samples based on discounted averaging, and hence it can save the overall computational and storage cost. Our algorithm is incremental in nature and possesses attractive features such as computational and storage efficiency, accuracy and stability. We provide conditions required for the algorithm to converge to the global optimum. We evaluated the algorithm on a variety of global optimization benchmark problems and the results obtained corroborate our theoretical findings.

1 INTRODUCTION

In the paper, we consider the following optimization problem:

$$Find \quad x^* \in \arg\max_{x \in \mathscr{X} \subset \mathbb{R}^m} \mathscr{H}(x). \tag{1}$$

Here $\mathscr{H} : \mathbb{R}^m \to \mathbb{R}$ is a deterministic, multi-modal, bounded real-valued continuous function and the solution space $\mathscr{X}$ is a compact subset of $\mathbb{R}^m$. We assume that x^* is unique and $x^* \in \text{interior}(\mathscr{X})$. The continuity of $\mathscr{H}$ implies that $\mathscr{H}(x^*)$ is not an isolated point.

The problem is made more challenging by considering a "black-box" scenario, *i.e.*, a closed form expression of the objective function is unavailable, however for a given $x \in \mathscr{X}$, the value of the objective function $\mathscr{H}(x)$ is available. A few predominant algorithms which solve problems of this kind are, simultaneous perturbation stochastic approximation (SPSA) (Spall 1992), model reference adaptive search (MRAS) (Hu, Fu, and Marcus 2007), cross entropy (CE) method (Rubinstein and Kroese 2013),(Kroese, Porotsky, and Rubinstein 2006), estimation of distribution algorithms (EDA) (Zhang and Mühlenbein 2004) and gradient-based adaptive stochastic search (GASS) (Zhou and Hu 2014). SPSA is a randomized finite difference method, while the rest of the above methods belong to a broader class of methods called the *model based search methods*. The model based search methods are zero-order or gradient-free techniques, *i.e.*, do not need knowledge of the gradient of the objective function. Hence the algorithm can be applied in any setting, where the function does not possess smooth differentiable structure. The goal of this method is to find a "*model*" or *probability distribution* which concentrates on the global maximum x^*. The search is therefore performed on a parametrized family of distributions $\mathscr{F} = \{f_\theta(\cdot) | \theta \in \Theta\}$, where f_θ is a probability density function on the solution space $\mathscr{X}$. It follows an iterative procedure where at each iteration k, a model over the space $\mathscr{X}$ is developed and as k goes to infinity, the model sequence better represents the promising region (the neighbourhood around x^*).

Exponential family of distributions: The common choice for $\mathscr{F}$ is the exponential family of distributions: $\mathscr{C} \triangleq \{f_\theta(x) = h(x)e^{\theta^\top \Gamma(x) - K(\theta)} \mid \theta \in \Theta \subset \mathbb{R}^d\}$, where $h : \mathbb{R}^m \longrightarrow \mathbb{R}$, $\Gamma : \mathbb{R}^m \longrightarrow \mathbb{R}^d$ and $K : \mathbb{R}^d \longrightarrow \mathbb{R}$ and

$K(\theta) = \log \int h(x) e^{\theta^\top \Gamma(x)} dx$. The Gaussian distribution with mean vector $\mu \in \mathbb{R}^m$ and the covariance matrix $\Sigma \in \mathbb{R}^{m \times m}$ belongs to $\mathscr{C}$. In this case,

$$f_\theta(x) = ((2\pi)^m |\Sigma|)^{-1/2} \exp\{-(x-\mu)^\top \Sigma^{-1}(x-\mu)/2\}, \tag{2}$$

and so one may let $h(x) = (2\pi)^{-m/2}, \Gamma(x) = (x, xx^\top)^\top$, $\theta = (\Sigma^{-1}\mu, -\frac{1}{2}\Sigma^{-1})^\top$.

In this paper, we consider the well known cross entropy (CE) method. The cross entropy method is inspired from the algorithm proposed in (Rubinstein 1997) to estimate the probability of rare events in stochastic networks. Later, an adaptive scheme based on this algorithm found significant inroads in combinatorial optimization (Rubinstein 1999). Solutions to various NP-hard problems were obtained using the CE method. The cross entropy was also applied to continuous optimization problems (Kroese, Porotsky, and Rubinstein 2006). In this paper, we study the properties of the CE method, understand its limitations and propose a modified approach to resolve them. We also propose the conditions required for CE method to converge to the global maximum of the objective function.

Notation: We use $\mathbf{x}$ to denote a random variable and x for deterministic variable. For $A \subset \mathbb{R}^m$, I_A represents the indicator function of A, *i.e.*, $I_A(x) = 1$ if $x \in A$ and 0 otherwise. Let $f_\theta(\cdot)$ denote the *probability density function* parametrized by θ and $\mathbb{E}_\theta[\cdot]$ be the *expectation w.r.t.* f_θ. For $\rho \in (0,1)$, let $\gamma_\rho^{\mathscr{H}}(\theta)$ denote the $(1-\rho)$-quantile of $\mathscr{H}(\mathbf{x})$ *w.r.t.* the f_θ, *i.e.*, $\gamma_\rho^{\mathscr{H}}(\theta) \triangleq \sup\{l : P_\theta(\mathscr{H}(\mathbf{x}) \geq l) \geq \rho$. Let $supp(f) \triangleq \overline{\{x | f(x) \neq 0\}}$ denote the support of f and interior(A) be the *interior* of set A. Also $\lceil a \rceil$ denote the smallest integer greater than a. For $x \in \mathbb{R}^m$, let $\|x\|_\infty$ represent the sup-norm, *i.e.*, $\|x\|_\infty = \max_i |x_i|$.

1.1 Cross Entropy (Ideal Version)

The CE method aims to find a sequence of model parameters $\{\theta_k \in \Theta\}_{k \in \mathbb{Z}^+}$ and an increasing sequence of thresholds $\{\gamma_k \in \mathbb{R}\}_{k \in \mathbb{Z}^+}$ with the property that the support of the model f_{θ_k} satisfies $supp(f_{\theta_k}) \subseteq \{x | \mathscr{H}(x) \geq \gamma_k\}$. By assigning greater weight to the higher values of $\mathscr{H}$ at each iteration, the expected behaviour of the probability distribution sequence should improve. This is achieved by solving at each instant $k+1$, the following optimization problem:

$$\theta_{k+1} = \arg\max_{\theta \in \Theta} \Phi_k(\theta, \gamma_{k+1}), \tag{3}$$

where $\Phi_k(\theta, \gamma) \triangleq \mathbb{E}_{\theta_k}\left[\varphi(\mathscr{H}(\mathbf{x})) I_{\{\mathscr{H}(\mathbf{x}) \geq \gamma\}} \log f_\theta(\mathbf{x})\right]$ and $\varphi : \mathbb{R} \to \mathbb{R}_+$ is a monotone strictly increasing positive function. Note that when f_θ belongs to the exponential family of distributions, Φ_k is concave in θ and hence the equality in (3) is well-defined. The most common choice for γ_{k+1} is $\gamma_\rho(\theta_k)$: the $(1-\rho)$-quantile of $\mathscr{H}(\mathbf{x})$ *w.r.t.* f_{θ_k}, where $\rho \in (0,1)$ is set *a priori*. (We drop the superscript $\mathscr{H}$, since $\mathscr{H}$ is fixed.)

In this paper, we take the Gaussian distribution as the preferred choice for f_θ. The model is parametrized as $\theta = (\mu, \Sigma)^\top$, where $\mu \in \mathbb{R}^m$ is the mean vector and $\Sigma \in \mathbb{R}^{m \times m}$ is the covariance matrix. Hence the distribution space $\mathscr{F} = \{f_\theta | \theta = (\mu, \Sigma)^\top \in \Theta \subset \mathbb{R}^{m(m+1)}\}$. It is easy to verify that the above parametrization has one-to-one correspondence with the parametrization $(\Sigma^{-1}\mu, -\frac{1}{2}\Sigma^{-1})^\top$ given in (2). For brevity we denote by $\vartheta(\theta) = (\vartheta_1, \vartheta_2)^\top \triangleq (\Sigma^{-1}\mu, -\frac{1}{2}\Sigma^{-1})^\top$. We further assume that the model parameter space Θ is a compact subset of $\mathbb{R}^{m(m+1)}$ and is large enough so that the solution to (3) is contained in interior(Θ).

We obtain a closed-form expression for θ_{k+1} by equating $\nabla \Phi_k$ to 0 and using (2) for $f_\theta(\cdot)$ as follows:

$$\nabla_{\vartheta_1} \Phi_k(\theta, \gamma) = 0 \Rightarrow \mu = \frac{\mathbb{E}_{\theta_k}\left[\mathbf{g_1}\big(\mathscr{H}(\mathbf{x}), \mathbf{x}, \gamma\big)\right]}{\mathbb{E}_{\theta_k}\left[\mathbf{g_0}(\mathscr{H}(\mathbf{x}), \gamma)\right]} \triangleq \Upsilon_1(\theta_k, \gamma), \tag{4}$$

$$\nabla_{\vartheta_2} \Phi_k(\theta, \gamma) = 0 \Rightarrow \Sigma = \frac{\mathbb{E}_{\theta_k}\left[\mathbf{g_2}\big(\mathscr{H}(\mathbf{x}), \mathbf{x}, \gamma, \mu\big)\right]}{\mathbb{E}_{\theta_k}\left[\mathbf{g_0}\big(\mathscr{H}(\mathbf{x}), \gamma\big)\right]} \triangleq \Upsilon_2(\theta_k, \gamma), \tag{5}$$

where $\mathbf{g_0}\big(\mathscr{H}(x), \gamma\big) \triangleq \varphi(\mathscr{H}(x)) I_{\{\mathscr{H}(x) \geq \gamma\}}$, $\quad \mathbf{g_1}\big(\mathscr{H}(x), x, \gamma\big) \triangleq \varphi(\mathscr{H}(x)) I_{\{\mathscr{H}(x) \geq \gamma\}} x$ and $\mathbf{g_2}\big(\mathscr{H}(x), x, \gamma, \mu\big) \triangleq \varphi(\mathscr{H}(x)) I_{\{\mathscr{H}(x) \geq \gamma\}} (x-\mu)(x-\mu)^\top$. It is easy to verify that Υ_1 and Υ_2 are well-defined.

1.2 Cross Entropy (Monte Carlo Version)

The operator $\mathbb{E}_\theta[\cdot]$ and the quantile $\gamma_\rho(\theta)$ used in (4) and (5) are hard to compute in general. Hence their stochastic counterparts are employed. The stochastic versions of Υ_1 and Υ_2 are as follows:

$$\bar{\Upsilon}_1(\theta,\gamma,\Lambda) \triangleq \frac{\frac{1}{N}\sum_{i=1}^{N} \mathbf{g_1}\big(\mathscr{H}(\mathbf{x}_i),\mathbf{x}_i,\gamma\big)}{\frac{1}{N}\sum_{i=1}^{N} \mathbf{g_0}\big(\mathscr{H}(\mathbf{x}_i),\gamma\big)}, \qquad \bar{\Upsilon}_2(\theta,\gamma,\Lambda) \triangleq \frac{\frac{1}{N}\sum_{i=1}^{N} \mathbf{g_2}\big(\mathscr{H}(\mathbf{x}_i),\mathbf{x}_i,\gamma,\mu\big)}{\frac{1}{N}\sum_{i=1}^{N} \mathbf{g_0}\big(\mathscr{H}(\mathbf{x}_i),\gamma\big)}. \tag{6}$$

where $N = |\Lambda|$ and $\Lambda = \{\mathbf{x}_1,\ldots,\mathbf{x}_N\} \sim f_\theta$.

A naive approach is used to estimate $\gamma_\rho(\theta)$ as follows: $\bar{\gamma}_\rho(\theta) \triangleq \mathscr{H}_{(\lceil(1-\rho)N\rceil)}$, (7)

where $\mathscr{H}_{(i)}$ is the ith-order statistic of $\{\mathscr{H}(\mathbf{x}_i)\}_{i=1}^N$, $N = |\Lambda|$ and $\Lambda = \{\mathbf{x}_1,\ldots,\mathbf{x}_N\} \sim f_\theta$.

A user configured *observation allocation rule* $\{N_k \in \mathbb{Z}_+\}_{k\in\mathbb{Z}_+}$ is used to decide the sample size required for each iteration, where $N_k \uparrow \infty$. The Monte Carlo CE method is given in Algorithm 2.

Algorithm 1 Monte Carlo CE Method

Step 0: Choose an initial *p.d.f.* $f_{\bar{\theta}_0}(\cdot)$ on $\mathscr{X}$ and $\varepsilon > 0$.

Step 1: [Sampling Candidate Solutions] Sample N_k *i.i.d.* solutions $\Lambda_k = \{\mathbf{x}_1,\ldots,\mathbf{x}_{N_k}\}$ using $f_{\bar{\theta}_k}(\cdot)$.

Step 2: [Threshold Evaluation] Calculate the sample $(1-\rho)$-quantile.
$\bar{\gamma}_{k+1} = \bar{\gamma}_\rho(\bar{\theta}_k)$.

Step 3: [Threshold Comparison]
if $\bar{\gamma}_{k+1} \geq \bar{\gamma}^*_k + \varepsilon$ **then**
$\bar{\gamma}^*_{k+1} = \bar{\gamma}_{k+1}$.
else
$\bar{\gamma}^*_{k+1} = \bar{\gamma}^*_k$.
end if

Step 4: [Model Parameter Update]
$\bar{\theta}_{k+1} = \big(\bar{\Upsilon}_1(\bar{\theta}_k,\bar{\gamma}^*_{k+1},\Lambda_k),\bar{\Upsilon}_1(\bar{\theta}_k,\bar{\gamma}^*_{k+1},\Lambda_k)\big)^\top$.

Step 5: If the stopping rule is satisfied, then return $\bar{\theta}_{k+1}$and terminate, else set $k := k+1$ and go Step 1.

1.3 Motivation of the paper

Even though the Monte Carlo CE tracks the ideal CE, it has a few significant drawbacks: (i) The naive approach of the Monte-Carlo CE does not utilize prior information efficiently. Note that Monte-Carlo CE possesses a stateless behaviour. At each iteration k, a completely new collection of samples are drawn using the distribution f_{θ_k}. The samples are used to derive the estimates $\bar{\gamma}_{k+1}$, $\bar{\mu}_{k+1}$ and $\bar{\Sigma}_{k+1}$. The algorithm does not utilize the estimates generated prior to k (ii) The second drawback is the poor computational and space complexity. The performance of the Monte Carlo version depends heavily on the sample size N_k. In most practical cases, the best value of N_k can only be obtained by trying the same for various values in a brute force manner. The estimate $\bar{\gamma}_k$ requires the order statistic $\mathscr{H}_{(i)}$ which is obtained by sorting the list $\{\mathscr{H}(\mathbf{x}_i)\}_{i=1}^{N_k}$. The summation operation in (6) requires $\mathscr{O}(N_k)$ time, while the sort operation required for the order statistic $\mathscr{H}_{(i)}$ in (7) requires $\mathscr{O}(N_k \log N_k)$. Note that N_k diverges and hence this super-linear relationship is computationally expensive and algorithm becomes well nigh intractable. Also note that the space complexity of the Monte Carlo CE is $\mathscr{O}(N_k)$ which is mainly attributed to the space occupied by the samples Λ_k. This is a heavy requirement too. All these are further worsened by the direct relationship of the sample size N_k with the dimension m of the solution space $\mathscr{X}$, *i.e.*, higher the dimension, more the required number of samples. Variants of the CE method such as gradient based CE method (Hu, Hu, and Chang 2012) and modified CE method (Wang and Enright 2013) also suffer similar drawbacks.

1.4 Our Contribution

The above mentioned drawbacks on the inefficient information utilization and the heavy cost on the space and computational requirements are primarily attributed to the non-incremental and stateless nature of the algorithm. In this paper, we resolve these shortcomings of the CE algorithm by remodelling the same using the stochastic approximation framework. We *replace the sample averaging with a bootstrapping approach,* i.e.*, deriving new estimates using the current estimates*. The algorithm possesses various features which we find desirable: (1) *Stability* (2) *Limited restriction on the objective function,* i.e. *without imposing heavy structural restrictions on the objective function* (3) *Incremental in nature,* i.e.*, evolves at each time instant according to the data (the function value $\mathscr{H}(\cdot)$) available at that particular instant. In other words the solution is built incrementally.* (4) *Efficient use of prior information,* i.e.*, the algorithm adopts an adaptive nature where the function values $\mathscr{H}(\cdot)$ are requested only when required. The boostrapping nature of the algorithm guarantees a continuous evolution (in contrast to the stateless nature of the Monte-Carlo version) and hence no data or prior information is under-utilized*. A recent study (Hu, Hu, and Chang 2012),(Hu, Fu, and Marcus 2007) shows that CE method is only a local improvement (local optimization) algorithm. In (Hu, Hu, and Chang 2012), a few counter examples are also provided. But in many practical cases, CE method exhibits good global convergence behaviour. In this paper, we explore this dichotomy and propose conditions which facilitate the convergence of the CE method to the global optimum.

2 PROPOSED ALGORITHM: CE2-ND

In this paper, we propose a new algorithm CE2-ND which stands for Cross Entropy 2-Normal Distribution. The idea is to track the ideal CE method using stochastic approximation. We provide a stochastic approximation algorithm whose asymptotic behaviour is equivalent to that of the ideal CE algorithm.

Stochastic approximation algorithms (Borkar 2008, Kushner and Clark 2012, Robbins and Monro 1951) are a natural way of encoding prior information and are expressed as recursive equations of the form:

$$\mathbf{z}_{k+1} = \mathbf{z}_k + \alpha_{k+1}\Delta\mathbf{z}(\mathbf{z}_k, b_k, D_{k+1})), \tag{8}$$

where $\Delta\mathbf{z}(z,b,D) = h(z)+b+D$ is the *increment term*, b_k is the *bias term* with $b_k \to 0$, D_k is a *random noise with zero-mean* and $h(\cdot)$ a *Lipschitz continuous function*. The *learning rate* α_k satisfies $\Sigma\alpha_k = \infty, \Sigma\alpha_k^2 < \infty$.

In the Monte Carlo version of CE, naive approaches are employed to estimate $\gamma_\rho(\cdot)$, Υ_1 and Υ_2. At each iteration k, a completely new collection of samples Λ_k is used to derive the estimates. This is a stateless approach where any structural relationship between the distribution parameter θ and $\gamma_\rho(\theta)$ or between θ and Υ_1, Υ_2 are completely discarded. The continuity relationship that holds between them can in fact be exploited to accelerate the whole procedure. Bootstrapping which is inherent in the stochastic approximation techniques can be employed to achieve this.

We found the following lemma from (Homem-de Mello 2007) to be beneficial in deriving a stochastic recursion to track the $(1-\rho)$-quantile of $\mathscr{H}(\cdot)$ *w.r.t.* a given probability distribution.

Lemma 1. *The $(1-\rho)$-quantile of a bounded real-valued function $H(\cdot)$ $\big(H(x) \in [H_1, H_2]\big)$ w.r.t the probability distribution $f_\theta(\cdot)$ is reformulated as the optimization problem:*

$$\textit{Find } \gamma_\rho^H(\theta) = \underset{u\in[H_1,H_2]}{\arg\min}\, \mathbb{E}_\theta\left[\psi(H(\mathbf{x}),u)\right], \tag{9}$$

where $\psi(H(x),u) = (1-\rho)(H(x)-u)I_{\{H(x)\geq u\}} + \rho(u-H(x))I_{\{H(x)\leq u\}}$.

We use this lemma to develop a stochastic gradient recursion which solves the optimization problem in (9). The increment term for the recursion is the sub-differential of ψ *w.r.t.* u, and is given by

$$\nabla_u\psi(H(x),u) = -(1-\rho)I_{\{H(x)\geq u\}} + \rho I_{\{H(x)\leq u\}}. \tag{10}$$

For the model parameter update, we track $\Upsilon_1(\theta_k,\gamma)$ and $\Upsilon_2(\theta_k,\gamma)$ (from (4) and (5)). For this we introduce two new variables, η and ξ, whose stochastic recursions track Υ_1 and Υ_2 respectively. The increment functions for the respective stochastic recursions are defined as follows:

$$\Delta\eta(x,\gamma) = \mathbf{g_1}(\mathscr{H}(x),x,\gamma) - \eta\,\mathbf{g_0}(\mathscr{H}(x),\gamma), \tag{11}$$

$$\Delta\xi(x,\mu,\gamma) = \mathbf{g_2}(\mathscr{H}(x),x,\gamma,\mu) - \xi\,\mathbf{g_0}(\mathscr{H}(x),\gamma). \tag{12}$$

The CE2-ND algorithm is given in Algorithm 2.

Algorithm 2 CE2-ND

Data: $\varepsilon_1 \in (0,1), \alpha_k, \lambda_k \in (0,1)$, $\theta_0 = (\mu_0, \Sigma_0)^\top$.
Init: $\gamma_0 = 0$, $\eta_0 = 0_{m\times 1}$, $\xi_0 = 0_{m\times m}$, $T_0 = 0$, $\gamma_0^* = -\infty$, $\lambda = \lambda_0$.
while stopping criteria not satisfied **do**

$\mathbf{x}_{k+1} \sim \widetilde{f}_{\theta_k}(\cdot)$ where $\widetilde{f}_{\theta_k} = (1-\lambda) f_{\theta_k} + \lambda f_{\theta_0}$;

- $\big[$**Tracking $(1-\rho)$-quantile of $\mathscr{H}(\cdot)$ *w.r.t.* $\widetilde{f}_{\theta_k}$**$\big]$

$$\gamma_{k+1} = \gamma_k - \alpha_{k+1}\nabla_u \psi(\mathscr{H}(\mathbf{x}_{k+1}),\gamma_k); \tag{13}$$

- $\big[$**Tracking μ_{k+1} of equation (4)**$\big]$

$$\eta_{k+1} = \eta_k + \alpha_{k+1}\Delta\eta(\mathbf{x}_{k+1},\gamma_k); \tag{14}$$

- $\big[$**Tracking Σ_{k+1} of equation (5)**$\big]$

$$\xi_{k+1} = \xi_k + \alpha_{k+1}\Delta\xi(\mathbf{x}_{k+1},\eta_k,\gamma_k); \tag{15}$$

- $\big[$**Threshold comparison**$\big]$

$$T_{k+1} = T_k + \lambda\left(I_{\{\gamma_{k+1}>\gamma_k^*\}} - I_{\{\gamma_{k+1}\le\gamma_k^*\}} - T_k\right); \tag{16}$$

- $\big[$**Model parameter update**$\big]$

if $T_{k+1} > \varepsilon_1$ **then**

$$\theta_{k+1} = \theta_k + \alpha_{k+1}\left((\eta_k,\xi_k)^\top - \theta_k\right); \tag{17}$$

$$\gamma_{k+1}^* = \gamma_k; \quad T_k = 0; \quad \lambda = \lambda_k; \tag{18}$$

else

$$\gamma_{k+1}^* = \gamma_k^*; \qquad \theta_{k+1} = \theta_k;$$

end if
$k := k+1$;
end while

The algorithm uses only a single sample $\mathbf{x}_{k+1}$ per iteration. The computational cost per iteration is proportional to m^2, where m is the dimension of the solution space $\mathscr{X}$. This is a significant improvement in terms of computational and space requirements.

Note that in the algorithm, a mixture distribution $\widetilde{f}_{\theta_k}$ is used to generate the sample $\mathbf{x}_{k+1}$, where $\widetilde{f}_{\theta_k} = (1-\lambda) f_{\theta_k} + \lambda f_{\theta_0}$ with λ the mixing weight. λ takes its values from a pre-defined decaying sequence $\{\lambda_k\}_{k\in\mathbb{Z}_+}$, with assignment happening in (18) during the model parameter update. The initial distribution parameter θ_0 is chosen such that the density function f_{θ_0} has strictly positive values for every point in the solution space $\mathscr{X}$, *i.e.*, $f_{\theta_0}(x) > 0, \forall x \in \mathscr{X}$. The mixture approach facilitates exploration of the solution space and prevents the iterates from getting locked in suboptimal solutions.

Assumption (A2): The learning rate α_k and the mixing weight λ_k are deterministic, non-increasing and

satisfy the following:

$$\lambda_k \in (0,1], \alpha_k \in (0,1], \quad \lim_{k\to\infty} \lambda_k = 0, \quad \sum_{k=1}^{\infty} \alpha_k = \infty, \ \sum_{k=1}^{\infty} \alpha_k^2 < \infty. \tag{19}$$

The step size α_k controls the rate at which the algorithm accrues the information and it is sensitive to the objective function and hence requires proper tuning. A very common choice of step-size used in stochastic approximation algorithms is the constant step-size, *i.e.*, $\alpha_k = \alpha \in (0,1) \forall k$. In this case, the convergence can only be claimed with high probability (Borkar 2008). This implies that there are rare events where the algorithm diverges.

The threshold comparison is achieved using the recursion (16) of the random variable T_k. Note that the model parameter θ_k is not updated at each epoch k. Rather it is updated whenever T_k arches over ε_1, where $\varepsilon_1 \in (0,1)$ is a constant. So the update of θ_k only happens along a subsequence $\{k_{(n)}\}_{n\in\mathbb{Z}_+}$ of $\{k\}_{k\in\mathbb{Z}_+}$. Between $k = k_{(n)}$ and $k = k_{(n+1)}$, the variable γ_k tracks $\gamma_\rho(\widetilde{\theta}_{k_{(n)}})$: the $(1-\rho)$-quantile of $\mathscr{H}$ *w.r.t.* $\widetilde{f}_{\theta_{k_{(n)}}}$. The threshold γ_k^* is also updated in (18) during the ε_1 crossover. Thus $\gamma^*_{k_{(n)}}$ is the estimate of $\gamma_\rho(\widetilde{\theta}_{k_{(n-1)}})$: the $(1-\rho)$-quantile of $\mathscr{H}$ *w.r.t.* $\widetilde{f}_{\theta_{k_{(n-1)}}}$. Put succinctly, T_k tracks the evolution of γ_k and tries to deduce a reasonable comparison between $\gamma_\rho(\widetilde{\theta}_{k_{(n)}})$ and $\gamma_\rho(\widetilde{\theta}_{k_{(n-1)}})$.

Proposition 1: T_k belongs to $(-1,1)$, $\forall k > 0$.

Proof: By rearranging terms in (16) we get $T_{k+1} = (1-\lambda)T_k + \lambda(I_{\{\gamma_{k+1}\geq\gamma^*_{k+1}\}} - I_{\{\gamma_{k+1}<\gamma^*_{k+1}\}})$,

where $\lambda \in (0,1)$. In the worst case, either $I_{\{\gamma_{k+1}\geq\gamma^*_{k+1}\}} = 1, \forall k$ or $I_{\{\gamma_{k+1}<\gamma^*_{k+1}\}} = 1, \forall k$. Since the two events are mutually exclusive, we will only consider the former event $\{I_{\{\gamma_{k+1}\geq\gamma^*_{k+1}\}} = 1, \forall k\}$. In this case

$$\lim_{k\to\infty} T_k = \lim_{k\to\infty} \left(\lambda + \lambda(1-\lambda) + \cdots + \lambda(1-\lambda)^{k-1}\right) = 1.$$

Similarly for the event $\{I_{\{\gamma_{k+1}<\gamma^*_{k+1}\}} = 1, \forall k\}$, we have $\lim_{k\to\infty} T_k = -1$. ■

3 CONVERGENCE ANALYSIS

Assumption (A1): The sequence$\{\gamma_k\}_{k\in\mathbb{Z}_+}$ in equation (13) satisfy $\sup_k |\gamma_k| < \infty$ *a.s.*.

Remark 3: *The assumption (A1) is a technical requirement to prove convergence of the algorithm. A commonly used procedure to ensure almost sure boundedness of iterates in a stochastic iterative scheme is to project these after each update to an a priori chosen (large enough) compact and convex set. In this case, the bound on the compact set can be derived from the bound on* $\mathscr{H}(\cdot)$.

As mentioned above, θ_k is updated only along a subsequence $\{k_{(n)}\}_{n\in\mathbb{Z}_+}$ of $\{k\}_{k\in\mathbb{Z}_+}$. Between $k = k_{(n)}$ and $k = k_{(n+1)}$, the model parameters θ_k remain constant. So we can analyse the limiting behaviour of γ_k, ξ_k and η_k by keeping θ_k fixed. We now have the following result for recursion (13):

Lemma 2. *Assume* $\theta_k \equiv \theta, \forall k$. *Let assumption (A1) hold. Then* $\gamma_k \to \gamma_\rho(\widetilde{\theta})$ *as* $k \to \infty$ *w.p. 1., where* $\widetilde{f}_\theta = (1-\lambda)f_\theta + \lambda f_{\theta_0}$.

Interpretation of Lemma 2: *Lemma 2 claims that if the model parameter is held constant, i.e.,* $\theta_k \equiv \theta, \forall k$, *then* γ_k *successfully tracks* $\gamma_\rho(\widetilde{\theta})$*: the* $(1-\rho)$*-quantile of* $\mathscr{H}$ *w.r.t.* $\widetilde{f}_\theta$.

Now we analyse the asymptotic behaviour of the sequences $\{\eta_k\}_{k\in\mathbb{Z}_+}$ and $\{\xi_k\}_{k\in\mathbb{Z}_+}$. We keep the model parameter θ_k constant during the analysis.

Lemma 3. *Assume $\theta_k \equiv \theta, \forall k$ and $\gamma_k^* \equiv \gamma^*, \forall k$, then a.s.*

$$\text{(i)} \lim_{k\to\infty} \eta_k = \eta_* = \frac{\mathbb{E}_{\widetilde{\theta}}\left[\mathbf{g_1}\big(\mathscr{H}(\mathbf{x}), \mathbf{x}, \gamma_\rho(\widetilde{\theta})\big)\right]}{\mathbb{E}_{\widetilde{\theta}}\left[\mathbf{g_0}\big(\mathscr{H}(\mathbf{x}), \gamma_\rho(\widetilde{\theta})\big)\right]},$$

$$\text{(ii)} \lim_{k\to\infty} \xi_k = \xi_* = \frac{\mathbb{E}_{\widetilde{\theta}}\left[\mathbf{g_2}\big(\mathscr{H}(\mathbf{x}), \mathbf{x}, \gamma_\rho(\widetilde{\theta}), \eta^*\big)\right]}{\mathbb{E}_{\widetilde{\theta}}\left[\mathbf{g_0}\big(\mathscr{H}(\mathbf{x}), \gamma_\rho(\widetilde{\theta})\big)\right]}.$$

(iii) *Further if $\gamma_\rho(\widetilde{\theta}) > \gamma^*$, then $\lim_{k\to\infty} T_k = 1$ a.s., where $\widetilde{f}_\theta = (1-\lambda) f_\theta + \lambda f_{\theta_0}$.*

Interpretation of Lemma 3: *Lemma 3 claims that if the model parameter is held constant, i.e., $\theta_k \equiv \theta, \forall k$, then η_k and ξ_k successfully track $\Upsilon_1(\widetilde{\theta}, \gamma_\rho(\widetilde{\theta}))$ and $\Upsilon_2(\widetilde{\theta}, \gamma_\rho(\widetilde{\theta}))$ respectively. Part (iii) of Lemma 3 claims that T_k gives a credible comparison of the thresholds. In practical cases, we choose ε_1 close to 1.*

Notation: For the subsequence $\{k_{(n)}\}_{n>0}$ of $\{k\}_{k\in\mathbb{Z}_+}$, we denote $k_{(n)}^- \triangleq k_{(n)} - 1$ for $n > 0$.

Along the subsequence $\{k_{(n)}\}_{n\geq 0}$ with $k_0 = 0$ the update of θ_k can be expressed as follows:

$$\theta_{k_{(n+1)}} = \theta_{k_{(n)}} + \alpha_{k_{(n+1)}} \Delta\theta_{k_{(n+1)}}, \tag{20}$$

where $\Delta\theta_{k_{(n+1)}} = (\eta_{k_{(n+1)}^-}, \xi_{k_{(n+1)}^-})^\top - \theta_{k_{(n)}}$. We show now that the increment term $\Delta\theta_{k_{(n+1)}}$ in equation (20) is indeed an estimate of $\nabla_{\vartheta(\theta)} \Psi(\theta)\big|_{\theta=\widetilde{\theta}_{k_{(n)}}}$, where

$$\boxed{\Psi(\theta) = \log \mathbb{E}_\theta\left[\mathbf{g_0}\left(\mathscr{H}(\mathbf{x}), \gamma_\rho(\theta)\right)\right]} \tag{21}$$

with $\theta = (\mu, \Sigma)^\top$ and $\vartheta(\theta) = (\Sigma^{-1}\mu, -\frac{1}{2}\Sigma^{-1})^\top$.

We now state a key lemma about the gradient of Ψ.

Lemma 4. *For the given function $\mathscr{H}(\cdot) \in \mathbb{R}$, $\theta = (\mu, \Sigma)^\top$ and $\vartheta = (\vartheta_1, \vartheta_2)^\top = (\Sigma^{-1}\mu, -\frac{1}{2}\Sigma^{-1})^\top$, we have*

$$\nabla_{\vartheta_1} \Psi(\theta) = \frac{\mathbb{E}_\theta\left[\mathbf{g_1}\big(\mathscr{H}(\mathbf{x}), \mathbf{x}, \gamma_\rho(\theta)\big)\right]}{\mathbb{E}_\theta\left[\mathbf{g_0}\big(\mathscr{H}(\mathbf{x}), \gamma_\rho(\theta)\big)\right]} - \mu.$$

$$\nabla_{\vartheta_2} \Psi(\theta) = \frac{\mathbb{E}_\theta\left[\mathbf{g_2}\big(\mathscr{H}(\mathbf{x}), \mathbf{x}, \gamma_\rho(\theta), \mu\big)\right]}{\mathbb{E}_\theta\left[\mathbf{g_0}\big(\mathscr{H}(\mathbf{x}), \gamma_\rho(\theta)\big)\right]} - \Sigma.$$

We now present our main result. The following theorem shows that the convergence of the model sequence $\{\theta_k\}_{k\in\mathbb{Z}_+}$ generated by CE2-ND is indeed guaranteed and provides a characterization of the limit points. Additionally, by imposing certain structural restrictions on the objective function $\mathscr{H}$, the convergence of the algorithm to the degenerate distribution concentrated on the global maximum x^* is ensured.

Theorem 5. *(***Convergence to global maximum***) Let $\varphi(x) = exp(rx), r \in \mathbb{R}_+$. Assume that the objective function $\mathscr{H}$ satisfies the following two conditions: (i) $\nabla^2\mathscr{H}$ exists and (ii) $\frac{\partial^2\mathscr{H}}{\partial x_i \partial x_j}$ is continuous for $1 \leq \forall i, \forall j \leq m$. Let the learning rate α_k, $k \in \mathbb{Z}_+$ satisfy (19). Let $\{\theta_k = (\mu_k, \Sigma_k)^\top\}_{k\in\mathbb{Z}_+}$ be the sequence generated by CE2-ND and assume $\theta_k \in$ interior(Θ), $\forall k \in \mathbb{Z}_+$. Also, let the assumptions (A1) and (A2) hold. Then*

$$\lim_{k\to\infty} \theta_k = \lim_{k\to\infty} (\mu_k, \Sigma_k)^\top = (x^*, 0_{m\times m})^\top, \quad w.p.1$$

where x^* is defined in (1).

Proof: Rewriting the equation (17) along the subsequence $\{k_{(n)}\}_{n\in\mathbb{Z}_+}$, we have for $n\in\mathbb{Z}_+$,

$$\theta_{k_{(n+1)}} = \theta_{k_{(n)}} + \alpha_{k_{(n+1)}}\big((\eta_{k^-_{(n+1)}}, \xi_{k^-_{(n+1)}})^\top - \theta_{k_{(n)}}\big). \tag{22}$$

Also $\sup_n \|\theta_{k_{(n)}}\| < \infty$ *a.s.* Rearranging the equation (22) we get, for $n\in\mathbb{Z}_+$,

$$\theta_{k_{(n+1)}} = \theta_{k_{(n)}} + \alpha_{k_{(n+1)}}\Big(\mathbb{E}\Big[\nabla_{\vartheta(\theta)}\Psi(\theta_{k_{(n)}})\Big|\theta_{k_{(n)}}\Big] + o(1)\Big), \tag{23}$$

where the $o(1)$ term corresponds to errors in the estimation of η_k and ξ_k and each decay to zero *a.s.*

Now consider the gradient flow ODE

$$\frac{d\theta(t)}{dt} = \nabla_{\vartheta(\theta)}\Psi(\theta(t)), \qquad t\in\mathbb{R}_+. \tag{24}$$

By appealing to Theorem 2, Chapter 2 of (Borkar 2008), the asymptotic equivalence between the equations (23) and (24) can be easily established. Therefore the recursion (17) reduces to a stochastic gradient ascent which optimizes the objective function $\Psi(\cdot)$. Hence the limiting behaviour of the model sequence $\{\theta_k\}_{k\in\mathbb{Z}_+}$ can be obtained by analysing the same of the above ODE. The equilibrium points of the ODE (24) can be obtained by equating $\nabla\Psi$ to 0.

$$\text{Equating } \nabla_{\vartheta_1}\Psi(\theta) \text{ to } 0_{m\times 1}, \text{ we get } \mu = \frac{\mathbb{E}_\theta\left[\mathbf{g_1}\big(\mathscr{H}(\mathbf{x}),\mathbf{x},\gamma_\rho(\theta)\big)\right]}{\mathbb{E}_\theta\left[\mathbf{g_0}\big(\mathscr{H}(\mathbf{x}),\gamma_\rho(\theta)\big)\right]}. \tag{25}$$

$$\text{Equating } \nabla_{\vartheta_2}\Psi(\theta) \text{ to } \mathbb{O}(=0_{m\times m}), \text{ we get } \frac{\mathbb{E}_\theta\left[\mathbf{g_2}\big(\mathscr{H}(\mathbf{x}),\mathbf{x},\gamma_\rho(\theta),\mu\big)\right]}{\mathbb{E}_\theta\left[\mathbf{g_0}\big(\mathscr{H}(\mathbf{x}),\gamma_\rho(\theta)\big)\right]} - \Sigma = \mathbb{O}. \tag{26}$$

For brevity let $L(\theta) = \mathbb{E}_\theta\left[\mathbf{g_0}\big(\mathscr{H}(\mathbf{x}),\gamma_\rho(\theta)\big)\right]$ and $\hat{\mathbf{g}}_\mathbf{0}(\mathbf{x},\theta) \triangleq \mathbf{g_0}\big(\mathscr{H}(\mathbf{x}),\gamma_\rho(\theta)\big)$.
Substituting the expression for μ from (25) in (26) and after further simplification we get,

$$(1/L(\theta))\,\mathbb{E}_\theta\left[\hat{\mathbf{g}}_\mathbf{0}(\mathbf{x},\theta)\mathbf{x}\mathbf{x}^\top\right] - \mu\mu^\top - \Sigma = \mathbb{O}. \tag{27}$$

Since $\Sigma = \mathbb{E}_\theta\left[\mathbf{x}\mathbf{x}^\top\right] - \mu\mu^\top$, the equation (27) implies

$$\begin{aligned}(1/L(\theta))\mathbb{E}_\theta\left[\hat{\mathbf{g}}_\mathbf{0}(\mathbf{x},\theta)\mathbf{x}\mathbf{x}^\top\right] - \mathbb{E}_\theta\left[\mathbf{x}\mathbf{x}^\top\right] = \mathbb{O} \quad &\Longrightarrow_1 \quad (1/L(\theta))\mathbb{E}_\theta\left[\left(\hat{\mathbf{g}}_\mathbf{0}(\mathbf{x},\theta) - L(\theta)\right)\mathbf{x}\mathbf{x}^\top\right] = \mathbb{O} \\ \Longrightarrow_2 \quad \Sigma\Sigma\mathbb{E}_\theta\left[\nabla_x^2\mathbf{g_0}(\mathbf{x},\theta)\right] = \mathbb{O} \quad &\Longrightarrow_3 \quad \Sigma^2\mathbb{E}_\theta\left[\varphi(\mathscr{H}(\mathbf{x}))G(\mathbf{x})I_{\{\mathscr{H}(\mathbf{x})\geq\gamma_\rho(\theta)\}}\right] = \mathbb{O},\end{aligned} \tag{28}$$

where $G(x) \triangleq r^2\nabla\mathscr{H}(x)\nabla\mathscr{H}(x)^\top + r\nabla^2\mathscr{H}(x)$. Note that $\Longrightarrow_2$ follows from "integration by parts" rule for multivariate Gaussian and $\Longrightarrow_3$ follows from the assumption $\varphi(x) = exp(rx)$. Note that for each $x\in\mathscr{X}$, $G(x)$ is a $m\times m$ square matrix. Since $(\nabla_i\mathscr{H})^2 \geq 0$, we can find an $r\in\mathbb{R}_+$ and $1\leq i\leq m$ *s.t.* $G_{ii}(x) > 0$, $\forall x\in\mathscr{X}$. This further implies that $\mathbb{E}_\theta[\varphi(\mathscr{H}(\mathbf{x}))G(\mathbf{x})I_{\{\mathscr{H}(\mathbf{x})\geq\gamma_\rho(\theta)\}}] \neq \mathbb{O}$, $\forall\theta\in\Theta$. Hence, from (28) we get $\Sigma = \mathbb{O}$. This proves that for any $x\in\mathbb{R}^m$, the degenerate distribution concentrated on x given by $\theta_x = (x, 0_{m\times m})^\top$ is an equilibrium point of the ODE (24). Also note that the ODE (24) is asymptotically stable at all local maxima of $\Psi(\cdot)$. The existence of the Lyapunov function $V_x : U_x \to \mathbb{R}_+$ on the open neighbourhood U_x of θ_x, defined as $V_x(\theta) \triangleq \Psi(\theta_x) - \Psi(\theta)$ is enough to prove the local asymptotic stability.

To prove that the limit is indeed θ^*, the degenerate distribution concentrated at x^*, we use *proof by contradiction* technique. So assume to the contrary, *i.e.*, $\theta_k \to \widehat{\theta} = (\widehat{x}, 0_{m\times m})^\top$, where $\widehat{x}\neq x^*$. Note that

$\mathbb{E}_\theta\left[\mathbf{g_1}\big(\mathscr{H}(\mathbf{x}),\mathbf{x},\gamma_\rho(\theta)\big)\right]$, $\mathbb{E}_{\theta_0}\left[\mathbf{g_1}\big(\mathscr{H}(\mathbf{x}),\mathbf{x},\gamma_\rho(\theta)\big)\right]$, $\mathbb{E}_\theta\left[\mathbf{g_0}\big(\mathscr{H}(\mathbf{x}),\gamma_\rho(\theta)\big)\right]$ and $\mathbb{E}_{\theta_0}\left[\mathbf{g_0}\big(\mathscr{H}(\mathbf{x}),\gamma_\rho(\theta)\big)\right]$ are all continuous on θ. This implies that we can find scalars $\varepsilon_2>0,\delta_2>0$ and $k\in\mathbb{Z}_+$ *s.t.*

$$\begin{aligned}
&\left\|\theta_k-\widehat{\theta}\right\|_\infty<\delta_2,\\
&\left\|\mathbb{E}_{\widehat{\theta}}\left[\mathbf{g_1}\big(\mathscr{H}(\mathbf{x}),\mathbf{x},\gamma_\rho(\widehat{\theta})\big)\right]-\mathbb{E}_{\theta_k}\left[\mathbf{g_1}\big(\mathscr{H}(\mathbf{x}),\mathbf{x},\gamma_\rho(\theta_k)\big)\right]\right\|_\infty<\varepsilon_2,\\
&\left\|\mathbb{E}_{\theta_0}\left[\mathbf{g_1}\big(\mathscr{H}(\mathbf{x}),\mathbf{x},\gamma_\rho(\widehat{\theta})\big)\right]-\mathbb{E}_{\theta_0}\left[\mathbf{g_1}\big(\mathscr{H}(\mathbf{x}),\mathbf{x},\gamma_\rho(\theta_k)\big)\right]\right\|_\infty<\varepsilon_2,\\
&\left\|\mathbb{E}_{\widehat{\theta}}\left[\mathbf{g_0}\big(\mathscr{H}(\mathbf{x}),\gamma_\rho(\widehat{\theta})\big)\right]-\mathbb{E}_{\theta_k}\left[\mathbf{g_0}\big(\mathscr{H}(\mathbf{x}),\gamma_\rho(\theta_k)\big)\right]\right\|_\infty<\varepsilon_2,\\
&\left\|\mathbb{E}_{\theta_0}\left[\mathbf{g_0}\big(\mathscr{H}(\mathbf{x}),\gamma_\rho(\widehat{\theta})\big)\right]-\mathbb{E}_{\theta_0}\left[\mathbf{g_0}\big(\mathscr{H}(\mathbf{x}),\gamma_\rho(\theta_k)\big)\right]\right\|_\infty<\varepsilon_2.
\end{aligned}\tag{29}$$

$$\text{Now consider}\qquad \nabla_{\vartheta_1}\Psi(\theta)\big|_{\theta=\widetilde{\theta}_k}=(1/L(\theta))\,\mathbb{E}_\theta\left[\mathbf{g_1}\big(\mathscr{H}(\mathbf{x}),\mathbf{x},\gamma_\rho(\theta)\big)\right]-\mu\Big|_{\theta=\widetilde{\theta}_k}\tag{30}$$

We denote by $e=(1,\ldots,1)^\top\in\mathbb{R}^m$. Applying sup norm on either side of (30) and using (29) we get,

$$\begin{aligned}
\left\|\nabla_{\vartheta_1}\Psi(\theta)\big|_{\theta=\widetilde{\theta}_k}\right\|_\infty &\geq \left\|\frac{(1-\lambda)\mathbb{E}_{\widehat{\theta}}\left[\mathbf{g_1}\big(\mathscr{H}(\mathbf{x}),\mathbf{x},\gamma_\rho(\widehat{\theta})\big)\right]+\lambda\mathbb{E}_{\theta_0}\left[\mathbf{g_1}\big(\mathscr{H}(\mathbf{x}),\mathbf{x},\gamma_\rho(\widehat{\theta})\big)\right]-\varepsilon_2 e}{(1-\lambda)\mathbb{E}_{\widehat{\theta}}\left[\mathbf{g_0}\big(\mathscr{H}(\mathbf{x}),\gamma_\rho(\widehat{\theta})\big)\right]+\lambda\mathbb{E}_{\theta_0}\left[\mathbf{g_0}\big(\mathscr{H}(\mathbf{x}),\gamma_\rho(\widehat{\theta})\big)\right]+\varepsilon_2}-\widehat{x}-\delta_2 e\right\|_\infty\\
&\geq \left\|\frac{(1-\lambda)\widehat{x}\varphi(\mathscr{H}(\widehat{x}))+\lambda\mathbb{E}_{\theta_0}\left[\mathbf{g_1}\big(\mathscr{H}(\mathbf{x}),\mathbf{x},\mathscr{H}(\widehat{x})\big)\right]-\varepsilon_2 e}{(1-\lambda)\varphi(\mathscr{H}(x^*))+\lambda\mathbb{E}_{\theta_0}\left[\mathbf{g_0}\big(\mathscr{H}(\mathbf{x}),\mathscr{H}(\widehat{x})\big)\right]+\varepsilon_2}-\widehat{x}-\delta_2 e\right\|_\infty\\
&\geq \left\|(K_1(\widehat{x},\varepsilon_2)-1)\widehat{x}+K_2(\widehat{x},\varepsilon_2)\frac{\mathbb{E}_{\theta_0}\left[\mathbf{g_1}\big(\mathscr{H}(\mathbf{x}),\mathbf{x},\gamma_\rho(\widehat{\theta})\big)\right]}{\mathbb{E}_{\theta_0}\left[\mathbf{g_0}\big(\mathscr{H}(\mathbf{x}),\gamma_\rho(\widehat{\theta})\big)\right]}-(\varepsilon_2+\delta_2)e\right\|_\infty>K_3>0,
\end{aligned}$$

where $K_2(\cdot,\cdot)>0$ and $0<K_1(\cdot,\cdot)<1$ with $K_1(x_1,x_2)\to 1$ as $x_1\to x^*$ and $x_2\to 0$. This is a contradiction since $\Psi(\theta)$ is continuously differentiable (easily verifiable). However for θ^*, we have $\mathbb{E}_{\theta_0}\left[\mathbf{g_1}\big(\mathscr{H}(\mathbf{x}),\mathbf{x},\gamma_\rho(\theta^*)\big)\right]=0_{m\times 1}$ and $\mathbb{E}_{\theta_0}\left[\mathbf{g_0}\big(\mathscr{H}(\mathbf{x}),\gamma_\rho(\theta^*)\big)\right]=0$. We also have $K_1(x^*,\varepsilon_2)\to 1$ as $\varepsilon_2\to 0$. Hence a lower positive bound for $\left\|\nabla_{\vartheta_1}\Psi(\theta)\big|_{\theta=\widetilde{\theta}_k}\right\|_\infty$ cannot be obtained for θ^*. This eliminates the possibility of the model sequence $\{\theta_k\}$ converging to any degenerate distribution but to θ^*. This concludes the proof. ■

4 EXPERIMENTAL ILLUSTRATIONS

We tested CE2-ND on several global optimization benchmark functions from (Jamil and Yang 2013). To evaluate the algorithm, we compare it against the Monte-Carlo CE (MCCE) and the state-of-the-art gradient based Monte-Carlo CE (GMCCE) (Hu, Hu, and Chang 2012), which is a modified version of the Monte-Carlo CE. Here, we consider $\varphi=\exp(rx), r>0$. In each of the plots shown in this section, the solid graph represents the trajectory of $\mathscr{H}(\mu_k)$, while the dotted horizontal line is the global maximum $\mathscr{H}^*$ of the objective function $\mathscr{H}$. The x-axis represents the real time relative to the start of the algorithm. All the algorithms use the same initial distribution θ_0, which helps to compare them without any initial bias. The results shown are averages over 10 independent simulations obtained with the same initial distribution θ_0.

We consider the following benchmark functions for performance evaluation:

1. **Griewank function** [$m=200$][Continuous, Differentiable, Non-Separable, Scalable, Multimodal]

$$\mathscr{H}_1(x)=-1-\frac{1}{4000}\sum_{i=1}^{m}x_i^2+\prod_{i=1}^{m}\cos\left(x_i/\sqrt{i}\right).\tag{31}$$

2. **Levy function** [$m = 50$][Continuous, Differentiable, Multimodal]

$$\mathscr{H}_2(x) = -1 - \sin^2(\pi y_1) - (y_m - 1)^2(1 + \sin^2(2\pi y_m)) - \sum_{i=1}^{m}[(y_i - 1)^2(1 + 10\sin^2(\pi y_i + 1))],$$

$$\text{where} \quad y_i = 1 + \frac{x_i - 1}{4}.$$

3. **Trigonometric function** [$m = 30$][Continuous, Differentiable, Non-Separable, Scalable, Multimodal]

$$\mathscr{H}_3(x) = -1 - \sum_{i=1}^{m}[8\sin^2(7(x_i - 0.9)^2) + 6\sin^2(14(x_i - 0.9)^2) - (x_i - 0.9)^2].$$

4. **Rastrigin function** [m = 30][Continuous, Differentiable, Scalable, Multimodal]

$$\mathscr{H}_4(x) = -\sum_{i=1}^{m}(x_i^2 - 10\cos(2\pi x_i)) - 10m.$$

5. **Qing function** [m = 30][Continuous, Differentiable, Separable, Scalable, Multimodal]

$$\mathscr{H}_{10}(x) = -\sum_{i=1}^{m}(x_i^2 - i)^2.$$

6. **Bukin function** [$m = 2$][Multimodal, Continuous, Non-Differentiable, Non-Separable, Non-Scalable]

$$\mathscr{H}_9(x) = -100\sqrt{x_2 - 0.01x_1^2} - 0.01|x_1 + 10| - 20.0.$$

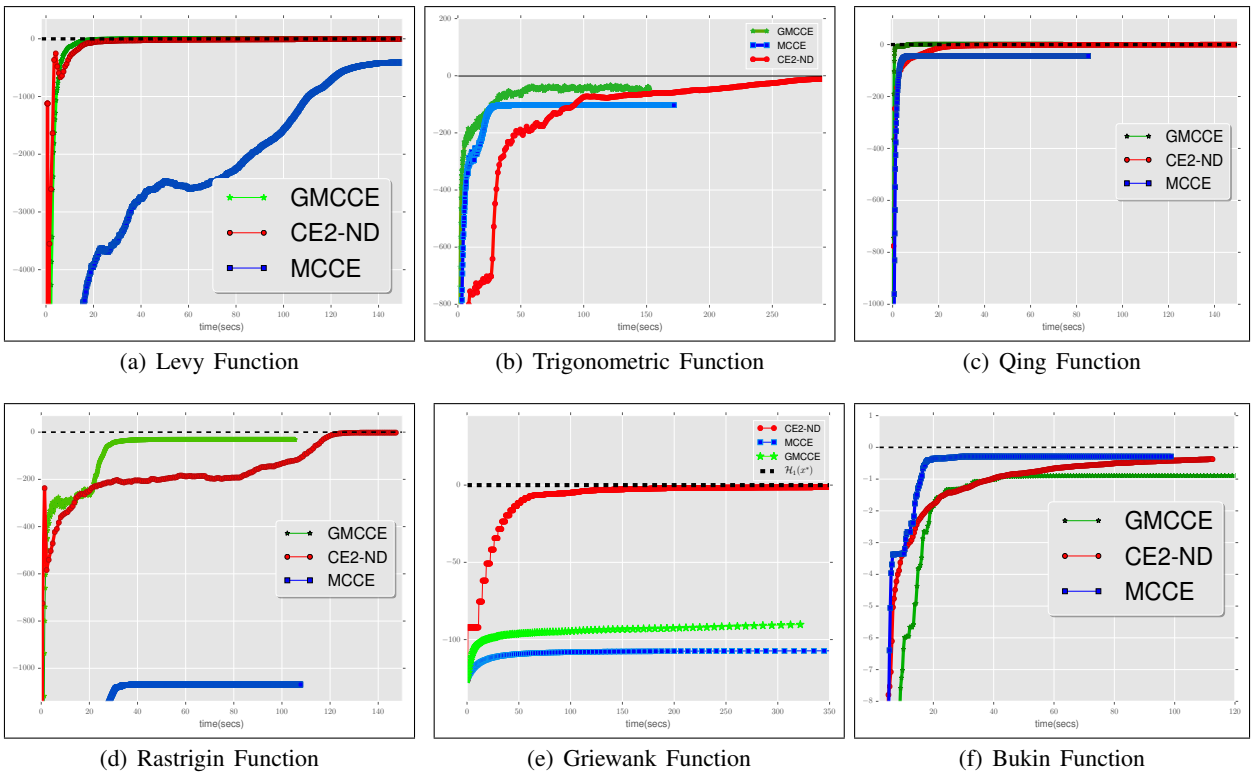
The results of the numerical experiments are shown in Figure 1. The various parameter values used in the experiments are shown in Table 1 and Table 2. To illustrate the advantage of CE2-ND in terms of memory requirements, we present in Figure 2, the real time memory usage of CE2-ND and GMCCE.

Table 1: The parameter values used in the experiments.

CE2-ND						**GMCCE**			
$\mathscr{H}(\cdot)$	r	α_k	λ_k	ε_1	ρ	r	α_k	ρ	N_k
$\mathscr{H}_1$	1.0	$k^{-0.52}$	$k_{(n)}^{-3.0}$	0.9	0.001	0.1	0.1	0.001	$N_{k+1} = 1.03 * N_k$, $N_0 = 700$
$\mathscr{H}_2$	0.001	0.1	$k_{(n)}^{-3.0}$	0.9	0.1	0.001	0.1	0.1	$N_{k+1} = 1.001 * N_k$, $N_0 = 700$
$\mathscr{H}_3$	0.001	0.03	$k_{(n)}^{-3.0}$	0.9	0.001	0.001	0.001	0.1	$N_{k+1} = 1.001 * N_k$, $N_0 = 700$
$\mathscr{H}_4$	0.01	0.2	$k_{(n)}^{-3.0}$	0.9	0.1	0.001	0.2	0.01	$N_{k+1} = 1.001 * N_k$, $N_0 = 800$
$\mathscr{H}_5$	0.00001	0.05	$k_{(n)}^{-3.0}$	0.9	0.01	0.001	0.2	0.01	$N_{k+1} = 1.001 * N_k$, $N_0 = 1000$
$\mathscr{H}_6$	0.1	$k_{(n)}^{-0.52}$	$k_{(n)}^{-3.0}$	0.9	0.01	0.1	0.1	0.01	$N_{k+1} = 1.001 * N_k$, $N_0 = 2000$

From the experiments, we have made the following observations:

- The algorithm CE2-ND shows good performance compared to GMCCE and MCCE. The algorithm CE2-ND also exhibits good global convergence behaviour when applied to the functions $\mathscr{H}_1$ to $\mathscr{H}_5$. However, when applied to the Bukin function, all the three algorithms fail to converge to the global optimum. It is easy to verify that Bukin function is non-differentiable and hence does not satisfy the criteria proposed in Theorem 5. This particular illustration corroborates the findings of Theorem 5. The algorithm exhibits robustness with respect to the initial distribution θ_0. An initial distribution which weighs the solution space reasonably well, seems to be sufficient. The values of the parameters λ_k and ε_1 are same for all test cases. This implies that these parameters require minimal tuning in most cases. As with any stochastic



(a) Levy Function (b) Trigonometric Function (c) Qing Function

(d) Rastrigin Function (e) Griewank Function (f) Bukin Function

Figure 1: The performance comparison of CE2-ND against GMCCE and MCCE. Here y-axis is $\mathscr{H}(\mu_k)$ and x-axis is the time in secs relative to the start of the algorithm.

approximation algorithm, the choice of the learning rate α_k is vital. The algorithm seems to be dependent on the quantile factor ρ which needs further investigation.

- The computational and storage requirements of the algorithm CE2-ND are minimal. This is attributed to the streamlined and incremental nature of the algorithm. This attribute makes the algorithm suitable in settings where the computational and storage resources are scarce.

5 CONCLUSIONS

We developed, in this paper, a stochastic approximation version of the cross entropy (CE) method. Our technique generalises the Monte-Carlo cross entropy method as it requires only one sample (as opposed to N_k) at each (kth) update epoch. The proposed algorithm is incremental in nature and possesses attractive features like robustness, stability as well as computational and space efficiency. We showed the almost sure convergence of our algorithm and proposed conditions required to achieve the convergence to the global maximum of the objective function. Numerical experiments over diverse benchmark functions are shown to support the theoretical findings.

REFERENCES

Borkar, V. S. 2008. "Stochastic approximation: A dynamical systems viewpoint". *Cambridge Univ. Press*.

Homem-de Mello, T. 2007. "A study on the cross-entropy method for rare-event probability estimation". *INFORMS Journal on Computing* 19 (3): 381–394.

Hu, J., M. C. Fu, and S. I. Marcus. 2007. "A model reference adaptive search method for global optimization". *Operations Research* 55 (3): 549–568.

Table 2: The initial distribution θ_0 used in the various cases and the global maximum $\mathscr{H}^*$ of the respective functions.

$\mathscr{H}(\cdot)$	θ_0	$\mathscr{H}^*$
$\mathscr{H}_1$	$(50.0, 50.0, \ldots, 50.0)^\top$, $100 * I_{200\times 200}$	0
$\mathscr{H}_2$	$(30.0, 30.0, \ldots, 30.0)^\top$, $250 * I_{50\times 50}$	-1
$\mathscr{H}_3$	$(10.0, 10.0, \ldots, 10.0)^\top$, $100 * I_{30\times 30}$	-1
$\mathscr{H}_4$	$(25.0, 25.0, \ldots, 25.0)^\top$, $100 * I_{30\times 30}$	0
$\mathscr{H}_5$	$(20.0, 20.0, \ldots, 20.0)^\top$, $200 * I_{30\times 30}$	0
$\mathscr{H}_6$	$(30.0, 30.0)^\top$, $250 * I_{2\times 2}$	0

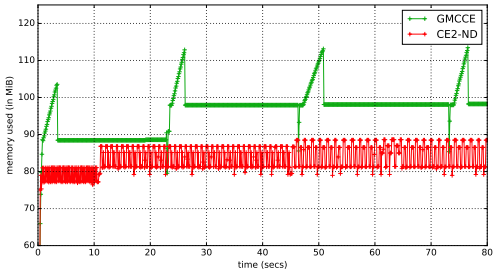


Figure 2: Memory usage comparison: CE2-ND uses less memory compared to GMCCE. The spikes in GMCCE is due to the sample generation.

Hu, J., P. Hu, and H. S. Chang. 2012. "A stochastic approximation framework for a class of randomized optimization algorithms". *Automatic Control, IEEE Transactions on* 57 (1): 165–178.

Jamil, M., and X. Yang. 2013. "A literature survey of benchmark functions for global optimisation problems". *International Journal of Mathematical Modelling and Numerical Optimisation* 4 (2): 150–194.

Kroese, D. P., S. Porotsky, and R. Y. Rubinstein. 2006. "The cross-entropy method for continuous multi-extremal optimization". *Methodology and Computing in Applied Probability* 8 (3): 383–407.

Kushner, H. J., and D. S. Clark. 2012. *Stochastic approximation methods for constrained and unconstrained systems*, Volume 26. Springer Science & Business Media.

Robbins, H., and S. Monro. 1951. "A stochastic approximation method". *The Annals of Mathematical Statistics*:400–407.

Rubinstein, R. 1999. "The cross-entropy method for combinatorial and continuous optimization". *Methodology and computing in applied probability* 1 (2): 127–190.

Rubinstein, R. Y. 1997. "Optimization of computer simulation models with rare events". *European Journal of Operational Research* 99 (1): 89–112.

Rubinstein, R. Y., and D. P. Kroese. 2013. *The cross-entropy method: a unified approach to combinatorial optimization, Monte-Carlo simulation and machine learning*. Springer Science & Business Media.

Spall, J. C. 1992. "Multivariate stochastic approximation using a simultaneous perturbation gradient approximation". *Automatic Control, IEEE Transactions on* 37 (3): 332–341.

Wang, B., and W. Enright. 2013. "Parameter estimation for ODEs using a cross-entropy approach". *SIAM Journal on Scientific Computing* 35 (6): A2718–A2737.

Zhang, Q., and H. Mühlenbein. 2004. "On the convergence of a class of estimation of distribution algorithms". *Evolutionary Computation, IEEE Transactions on* 8 (2): 127–136.

Zhou, E., and J. Hu. 2014. "Gradient-based adaptive stochastic search for non-differentiable optimization". *Automatic Control, IEEE Transactions on* 59 (7): 1818–1832.

AUTHOR BIOGRAPHIES

AJIN GEORGE JOSEPH is a Ph.D. student in the Dept. of Computer Science and Automation, Indian Institute of Science, Bangalore. His research interests are stochastic approximation and reinforcement learning. His email address is ajin@csa.iisc.ernet.in.

SHALABH BHATNAGAR is a Professor in the Dept. of Computer Science and Automation, Indian Institute of Science, Bangalore. His research interests include stochastic control, stochastic approximation and simulation optimization. His email address is shalabh@csa.iisc.ernet.in.