



**HAL**  
open science

# Enhancing the performance of a bistable energy harvesting device via the cross-entropy method

Americo Cunha Jr

► **To cite this version:**

Americo Cunha Jr. Enhancing the performance of a bistable energy harvesting device via the cross-entropy method. 2020. hal-01531845v2

**HAL Id: hal-01531845**

**<https://hal.science/hal-01531845v2>**

Preprint submitted on 31 Mar 2020 (v2), last revised 30 Nov 2020 (v4)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

# Enhancing the performance of a bistable energy harvesting device via the cross-entropy method

Americo Cunha Jr

Received: date / Accepted: date

**Abstract** This work deals with the solution of a non-convex optimization problem to enhance the performance of an energy harvesting device, which involves a nonlinear objective function and a discontinuous constraint. This optimization problem, which seeks to find a suitable configuration of parameters that maximize the electrical power recovered by a bistable energy harvesting system, is formulated in terms of the dynamical system response and a binary classifier obtained from 0-1 test for chaos. A stochastic solution strategy that combines penalization and the cross-entropy method is proposed and numerically tested. Computational experiments are conducted to address the performance of the proposed optimization approach by comparison with a reference solution, obtained via an exhaustive search in a refined numerical mesh. The obtained results illustrate the effectiveness and robustness of the cross-entropy optimization strategy, showing that the proposed framework may be a very useful and powerful tool to solve optimization problems involving nonlinear energy harvesting dynamical systems.

**Keywords** energy harvesting · nonlinear dynamics · cross-entropy method · evolutionary algorithm · non-convex optimization · global optimization

## 1 Introduction

Energy harvesting is a process of energy conversion in which a certain amount of energy is pumped from

an abundant source (e.g. the sun, wind, ocean, environmental vibrations, electromagnetic radiation, etc.) into another system that stores and/or use this energy for its self-operation [24, 47, 57]. This type of technology has a broad field of applicability, that includes: (i) large-scale devices such as ocean thermal energy converters [38], magnetic levitators [51], etc.; (ii) middle size applications like electromechanical systems [26, 37], sensors/actuators [5, 23, 58], living trees [63] etc.; (iii) small or very small systems, namely medical implants [46], micro/nano electro-mechanical systems (MEMS/NEMS) [43, 56], graphene structures [41], bio cells [8], etc.

Enhance the amount of energy collected by an energy harvester is a key problem in the operation of these kind of devices, being the object of interest of several research works among the last decade, with efforts divided essentially in two fronts: (i) approaches with great physical appeal, which seek to explore complex geometric arrangements and nonlinearities in a smart way to improve the performance of the system [1, 3, 18, 21, 32, 33, 49, 59, 65]; (ii) works that look at this task from a more mathematical perspective, conducting theoretical analysis in mathematical models [50], using control theory to optimize the underlying dynamics [2, 20, 62] or formulating and solving non-convex optimization problems to find the optimal system design [27, 39, 40, 45, 64].

The approaches that seek to optimize energy harvesters exploring their physics, in general, try to take advantage of a key (and very particular) characteristic of the dynamics, often neglected in a naive look, in favor of better performance. Although they can produce very efficient solutions for certain types of systems, they lack generality, which makes them, in a sense, an “art”, as

---

A. Cunha Jr  
Rio de Janeiro State University, Institute of Mathematics and Statistics, Rio de Janeiro, Brazil  
ORCID: 0000-0002-8342-0363  
E-mail: americ@ime.uerj.br

it requires a high level of knowledge about the system behavior by the designer's part.

On the other hand, approaches based on numerical optimization have a good level of generality, however, they come up against theoretical and computational difficulties inherent to the solution of non-convex optimization problems, these being defined by objective functions that are constructed from observations of the nonlinear dynamical systems.

These difficulties are faced not only when dealing with energy harvesters, but they can also be seen in the optimization of several types of nonlinear dynamical systems, generally being circumvented with the use of computational intelligence-based algorithms [42, 48], since traditional gradient-based methods have no guarantee of finding global extremes in the absence of convexity [6, 7, 44].

Among the computational intelligence algorithms available in the literature, the most frequently used for optimization in nonlinear dynamics include genetic algorithms (and their variants), particle swarm optimization, differential evolution, artificial neural networks (and other machine learning methods), etc [48]. As they are global search methods, they are often successful in overcoming the (local) difficulties faced by gradient-based methods, at the price of losing computational efficiency. But the loss of efficiency is not the only weakness of these global methods, most of them need to have the underlying control parameters tuned for proper functioning. This task is usually done manually, in a trial and error fashion, which is not at all practical, as the meaning of many of these parameters is often not intuitive. This peculiarity practically "condemns" these tools to be used in the black-box format, without much control by the user, which may induce significant losses in performance and accuracy.

However, the computational intelligence literature has at least one global search algorithm that is robust and simple, for which the control parameters are very intuitive, it is known as the *cross-entropy (CE) method*, proposed by R. Rubinstein in 1997 [52, 53, 54] for rare events simulation. Soon after it was realized that it could be a very appealing global search technique for challenging combinatorial and continuous optimization problems [36, 55]. It is a sampling technique, from the family of Monte Carlo methods, which iteratively attacks the problem, refining the candidates to be the solution according to a certain optimality criterion.

Surprisingly, the nonlinear dynamics literature has been neglecting the CE method, although it is a relatively known technique in the combinatorial optimization community. To the best of the author's knowledge, there are few studies in the open literature applying

the CE method for numerical optimization problems involving nonlinear mechanical systems [15, 16, 17, 25, 34, 60], which leaves space for new contributions in this line. In particular, the development of a simplistic and robust optimization framework, easily customizable, which can be used by naive users without major performance losses.

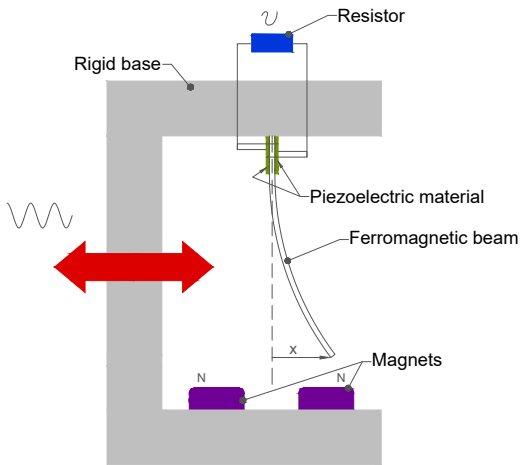
In this context, this work deals with the numerical optimization of nonlinear energy harvesting systems, trying to find a strategy to increase the performance of a bistable energy harvesting device subjected to a periodic excitation. For this purpose, a non-convex optimization problem, with a nonlinear objective function and discontinuous constraint, is formulated and a stochastic strategy of solution that combines penalization and the CE method is proposed and numerically tested. As this method has a theoretical guarantee of convergence and evidence in the literature proving its effectiveness [36, 55], the proposed cross-entropy framework for optimization of dynamical systems has the potential to be a very general and robust numerical methodology.

The rest of this manuscript is organized as follows. Section 2 presents the energy harvesting device of interest and the underlying dynamical system. An optimization problem associated with this nonlinear system is defined in section 3, and the CE method strategy of solution is presented in section 4. Numerical experiments, conducted to test the effectiveness and robustness of the stochastic solution strategy, are shown in section 5. Finally, in section 6, final remarks are set out.

## 2 Nonlinear dynamical system

### 2.1 Physical system

The energy harvesting system of interest in this work is the (sinusoidal excited) piezo-magneto-elastic beam presented by Erturk et al. [22], which is based on the (stochastically excited) inverted pendulum energy harvester proposed by Cottone et al. [9]. An illustration of this bistable energy harvesting system can be seen in Figure 1, where it is possible to see that it consists of a vertical fixed-free beam made of ferromagnetic material, a rigid base and a pair of magnets. In the beam upper part, there is a pair of piezoelectric laminae coupled to a resistive circuit. The rigid base is periodically excited by a harmonic force, which, together with the magnetic force generated by magnets, induces large amplitude vibrations. The piezoelectric laminae convert the energy of movement into electrical power, which is dissipated in the resistor.



**Figure 1** Illustration of the bistable piezo-magneto-elastic energy harvesting system.

Although this system dissipates energy, instead of using it to supply some secondary system, it is a typical prototype of a piezoelectric energy harvesting system. In an application of interest, the resistor is replaced by a more complex electrical circuit, which stores (and possibly manipulates) the voltage delivered by the mechanical system.

## 2.2 Initial value problem

The dynamic behavior of the system of interest is described by the following initial value problem

$$\ddot{x} + 2\xi\dot{x} - \frac{1}{2}x(1-x^2) - \chi v = f \cos(\Omega t), \quad (1)$$

$$\dot{v} + \lambda v + \kappa \dot{x} = 0, \quad (2)$$

$$x(0) = x_0, \quad \dot{x}(0) = \dot{x}_0, \quad v(0) = v_0, \quad (3)$$

where  $\xi$  is the damping ratio;  $\chi$  is the piezoelectric coupling term in mechanical equation;  $\lambda$  is a reciprocal time constant;  $\kappa$  is the piezoelectric coupling term in electrical equation;  $f$  is the external excitation amplitude;  $\Omega$  is the external excitation frequency. The initial conditions are  $x_0$ ,  $\dot{x}_0$  and  $v_0$ , which respectively represent, the beam edge initial position, initial velocity and the initial voltage over the resistor. Also,  $t$  denotes the time, so that the beam edge displacement at time  $t$  is given by  $x(t)$ , and the resistor voltage at  $t$  is represented by  $v(t)$ . The upper dot is an abbreviation for time-derivative, and all of these parameters are dimensionless.

## 2.3 Mean output power

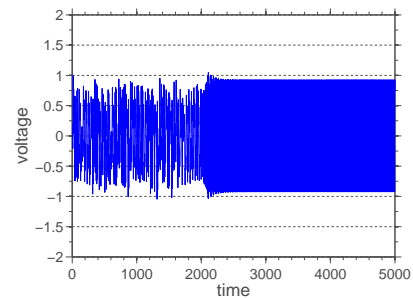
Since the objective in this work is to maximize the amount of energy recovered by an energy harvesting process, the main quantity of interest (QoI) associated with the nonlinear dynamical system under analysis is the mean output power

$$P = \frac{1}{T_f - T_0} \int_{\tau=T_0}^{T_f} \lambda v^2(\tau) d\tau, \quad (4)$$

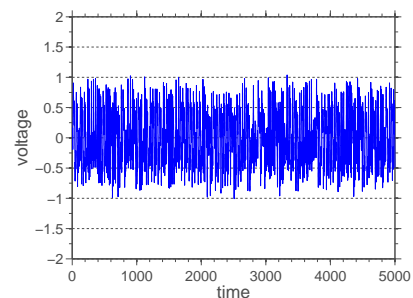
which is defined as the temporal average of the instantaneous power  $\lambda v^2$  over a given time interval  $[T_0, T_f]$ . This QoI plays the role of objective function in the optimization problem defined in section 3.

## 2.4 Dynamic classifier

Due to dynamical system nonlinearity, the steady-state dynamical response (over a given time interval) of the energy harvesting device may be chaotic or regular (non-chaotic), such as illustrated in Figure 2, which shows typical voltage time-series for this kind of bistable oscillator.



(a) regular dynamics



(b) chaotic dynamics

**Figure 2** Two typical voltage time-series for the bistable energy harvesting system. The time-series in (a) has a regular steady-state dynamics, while in (b) a chaotic steady-state dynamics is observed.

To distinguish between the chaotic and regular dynamic regimes, the 0-1 test for chaos by Gottwald and

Melbourne [28, 29, 30, 31] is employed. This test, which is based on an extension of the dynamical system to a two-dimensional Euclidean group [4], uses a binary classifier  $K$  to identify the dynamic regime of operation underlying the system of interest. This classifier is constructed from a dynamical system observation (time-series)  $\phi(t)$ . In fact, it is sufficient to use  $\Phi = (\phi(t_1), \phi(t_2), \dots, \phi(t_N))$ , a discrete version of the observable  $\phi(t)$  that is obtained through a numerical integration (sampling) process.

Receiving the discrete observation (time-series) as an input, the analytical procedure of the 0-1 test consists of the following steps:

1. A real parameter  $c \in [0, 2\pi)$  is chosen and, for  $n = 1, 2, \dots, N$ , the discrete version of  $\phi(t)$  is used to define the translation variables

$$p_n(c) = \sum_{j=1}^n \phi(t_j) \cos(jc), \quad (5)$$

$$q_n(c) = \sum_{j=1}^n \phi(t_j) \sin(jc). \quad (6)$$

2. Then, for  $n = 1, 2, \dots, N$ , the time-averaged mean square displacement of the dynamics trajectory in  $(p_c, q_c)$  space is computed by

$$M_n(c) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N \tilde{M}_n(c), \quad (7)$$

where

$$\tilde{M}_n(c) = (p_{j+n}(c) - p_j(c))^2 + (q_{j+n}(c) - q_j(c))^2. \quad (8)$$

3. Finally, defining the mean-square vector

$$\mathbf{M}_n = (M_1, M_2, \dots, M_n), \quad (9)$$

and the temporal mesh vector

$$\mathbf{t}_n = (t_1, t_2, \dots, t_n), \quad (10)$$

the dynamic classifier is constructed through the correlation

$$K_c = \lim_{n \rightarrow \infty} \frac{\text{cov}(\mathbf{t}_n, \mathbf{M}_n)}{\sqrt{\text{var}(\mathbf{t}_n) \text{var}(\mathbf{M}_n)}}, \quad (11)$$

where  $\text{cov}(\cdot, \cdot)$  and  $\text{var}(\cdot)$  respectively denote the covariance and variance statistical operators.

It can be proved that  $K_c \in \{0, 1\}$ , being  $K_c = 0$  for regular dynamics and  $K_c = 1$  for chaotic dynamics [30, 4]. For numerical implementation purposes, the above procedure is performed several times, for several randomly chosen values of  $c \in [0, 2\pi)$ , and considering the voltage time-series as the system observation, i.e.,  $\phi(t) = v(t)$ . The discrete version of  $\phi(t)$  is constructed through temporal integration via the standard 4th order Runge-Kuta method. The limit processes in Eqs.(7) and (11) are replaced by the condition  $n \ll N$ , and the classifier  $K$  is calculated as the median of  $K_c$  realizations, i.e.,  $K = \text{median}(K_c)$ . Indeed, for a careful done numerical simulation one has  $K \approx 0$  or  $K \approx 1$ . See [4, 31] for further details.

### 3 Non-convex optimization problem

#### 3.1 Problem definition

The objective of this work is to find a set of parameters that maximize the mean output power dissipated in the resistor. For this purpose, the electromechanical system properties ( $\xi$ ,  $\chi$ ,  $\lambda$  and  $\kappa$ ) sound as natural choices for the design variables, once they are intrinsic to the energy harvester embedded physical characteristics, something a designer might want to modify for better performance. However, the present work uses the dynamical system excitation parameters, namely the amplitude  $f$  and the frequency  $\Omega$ , as design variables. Although this choice sounds quite artificial at first since, in general, the designer does not have control of the level (or frequency) of vibration in which the energy harvester is operating, it is also of interest in the optimization of these systems to know in which vibration settings the device can recover more energy. As the numerical methodology developed here is completely general, it can be easily applied in the cases where the other parameters are used as the design variables, but by using  $f$  and  $\Omega$  to test the proposed methodology, an advantage is gained in terms of physical intuition about underlying physics, as the literature presents better knowledge about the effect of these two parameters on the behavior of this bistable energy harvesting system.

Many combinations of  $(f, \Omega)$  lead the dynamical system to operate in the chaotic regime, an undesirable condition for electric power usage in principle, since the process of rectifying the chaotic electrical signal, which is necessary to enable its use in applications, can consume a significant part of the available energy. Although it is possible to explore the chaotic dynamics in favor of greater efficiency of the system [2, 18], as done by the author and collaborators in [20], through the use of chaos control techniques, this is not the approach

followed in the present work, which seeks to develop a very general numerical framework.

Thus, as not every pair  $(f, \Omega)$  is an acceptable choice for the optimal design, it is necessary to impose a constraint that ensures the regularity (non-chaoticity) of the system dynamics. Taking advantage of the 0-1 test for chaos, described in section 2.4, the constraint to ensure a regular (non-chaotic) dynamic regime can be formulated as  $K = 0$ . Note that the optimization problem is extremely nontrivial, once the constraint presents jump-type discontinuities since  $K \in \{0, 1\}$ .

### 3.2 Constrained problem formulation

In an abstract way, one can formulate the constrained nonlinear optimization problem described above as find a feasible vector of design variables  $\mathbf{x}^*$  that maximize a certain objective function  $\mathbf{x} \in \mathcal{D} \mapsto \mathcal{S}(\mathbf{x}) \in \mathbb{R}$ , i.e.,

$$\mathbf{x}^* = \arg \max_{\mathbf{x} \in \mathcal{D}} \mathcal{S}(\mathbf{x}), \quad (12)$$

where the set of admissible (feasible) parameters is defined by the bounded region

$$\mathcal{D} = \{\mathbf{x} \mid \mathbf{x}_{min} \leq \mathbf{x} \leq \mathbf{x}_{max} \text{ and } \mathcal{G}(\mathbf{x}) = 0\}. \quad (13)$$

In this context, it is straightforward to see that the design variables vector is  $\mathbf{x} = (f, \Omega)$ , the objective function is  $\mathcal{S}(\mathbf{x}) = P(f, \Omega)$ , the binary constraint is  $\mathcal{G}(\mathbf{x}) = K(f, \Omega)$ , and the limits of the design variables are  $\mathbf{x}_{min} = (f_{min}, \Omega_{min})$  and  $\mathbf{x}_{max} = (f_{max}, \Omega_{max})$ .

In the case of a problem with more variables, or with different objective function and/or constraints, the adaptations to be made are straightforward.

### 3.3 Penalized problem formulation

A penalized version of this constrained nonlinear optimization problem is introduced here in order to facilitate the computational implementation of the solution algorithm. In this formulation the constraint  $\mathcal{G}(\mathbf{x}) = 0$  is replaced by the weaker condition  $\mathcal{G}(\mathbf{x}) \leq \varepsilon \ll 1$ , once in practice the best one has is  $K \approx 0$ . In this way, the problem defined by (12) is replaced by the penalized problem which seeks a pair  $\mathbf{x}^*$  such that

$$\mathbf{x}^* = \arg \max_{\mathbf{x} \in \mathcal{D}} \tilde{\mathcal{S}}(\mathbf{x}), \quad (14)$$

where the set of feasible parameters is now defined by

$$\mathcal{D} = \{\mathbf{x} \mid \mathbf{x}_{min} \leq \mathbf{x} \leq \mathbf{x}_{max}\}, \quad (15)$$

and the penalized objective function is given by

$$\tilde{\mathcal{S}}(\mathbf{x}) = \mathcal{S}(\mathbf{x}) - \alpha \max\{0, \mathcal{G}(\mathbf{x}) - \varepsilon\}. \quad (16)$$

The penalty parameter is heuristically chosen, being  $\alpha = 10$  the value used in all simulations reported in this work. In the same way,  $\varepsilon = 1/10$  is adopted.

## 4 The cross-entropy method

### 4.1 The general idea of CE method

The key idea of the CE method is to transform the given non-convex optimization problem into an “equivalent” rare event estimation problem, that can be efficiently treated by a Monte Carlo like algorithm. The only requirement is that the problem has a single solution, i.e., the global extreme is unique.

In this framework, the feasible region is sampled according to a given probability distribution chosen by the user, and the low-order statistics (mean and standard deviation) of these samples are used to update the optimum point estimation, and the stopping criteria metric (respectively). It is an iterative process consisting of two steps:

1. *Sampling*: The feasible region is sampled according to a given probability distribution, and a special subset of these points, dubbed the *elite sample set*, is defined (it is used to construct an optimal solution candidate in the next step);
2. *Learning*: The parameters of the probability distribution are updated using statistics obtained from the elite sample set, modifying the given distribution in a sense that try to make it as close as possible to a Dirac delta centered on the global optimum. The distribution mean provides the approximation for the global optimum.

Throughout this process, the update of the mean is done in order to move the distribution center towards the optimization problem optimum, while the standard deviation tendency is to be reduced, thus “shrinking” the distribution around its central value. The composition of these effects of translation and “closing” characterizes the process in which the distribution is “shaped” towards a point mass Dirac distribution centered on the optimal.

### 4.2 Theoretical framework

Without loss of generality, suppose that the problem of interest is to maximize an objective function  $\mathbf{x} \in \mathcal{D} \mapsto \mathcal{S}(\mathbf{x}) \in \mathbb{R}$ , as stated in Eqs.(12). If the penalized formulation from Eq.(14) is adopted, just consider  $\mathbf{x} \in \mathcal{D} \mapsto \tilde{\mathcal{S}}(\mathbf{x}) \in \mathbb{R}$  instead of  $\mathcal{S}(\mathbf{x})$ .

Denote the global optimal by  $\mathbf{x}^*$  and the corresponding maximum value by  $\gamma^* = \mathcal{S}(\mathbf{x}^*)$ , i.e.,  $\gamma^* = \max_{\mathbf{x} \in \mathcal{D}} \mathcal{S}(\mathbf{x})$  for all  $\mathbf{x} \in \mathcal{D}$ .

Using this global maximum as a reference value, and considering a randomized version of the vector  $\mathbf{x}$ , denoted by  $\mathbf{X}$ , it is possible to construct the random event  $\mathcal{S}(\mathbf{X}) \geq \gamma^*$ , which represents the scenarios where the random variable  $\mathcal{S}(\mathbf{X})$  is bigger or equal to the (deterministic) scalar value  $\gamma^*$ . In other words, this random event considers the possibility of choosing, randomly, points in the feasible region  $\mathcal{D}$  that produce values greater than or equal to the global maximum. Note that this random event is related to the optimization problem defined in Eq.(12), it can be thought of as its randomized version.

Since  $\gamma^*$  is the maximum value of  $\mathcal{S}(\mathbf{x})$ , no realization of  $\mathbf{X}$  can produce  $\mathcal{S}(\mathbf{X}) > \gamma^*$  and only  $\mathbf{X} = \mathbf{x}^*$  can make  $\mathcal{S}(\mathbf{X}) = \gamma^*$ . Therefore, the probability of this random event is zero, i.e.,  $\mathcal{P}\{\mathcal{S}(\mathbf{X}) \geq \gamma^*\} = 0$ .

Relaxing the reference value for a scalar  $\gamma < \gamma^*$ , one has the random event  $\mathcal{S}(\mathbf{X}) \geq \gamma$ , which defines a rare event if  $\gamma \approx \gamma^*$ , for which  $\mathcal{P}\{\mathcal{S}(\mathbf{X}) \geq \gamma\} \approx 0$ .

In his 1997 seminal work [52], R. Rubinstein conceived the CE method as a tool to efficiently estimate such type of probabilities, associated with events located at the distribution tail. Sometime later [53], due to the connection between the random event and the optimization problem described above, he noticed that this rare event estimation process could be used as a global search method to optimize the given objective function.

Notice that, with the aid of the expected value operator  $\mathbb{E}\{\cdot\}$ , and the indicator function

$$\mathbb{1}_{\mathcal{A}}(\mathbf{x}) = \begin{cases} 1, & \text{if } \mathbf{x} \in \mathcal{A} \\ 0, & \text{if } \mathbf{x} \notin \mathcal{A} \end{cases} \quad (17)$$

the probability of the random event  $\mathcal{S}(\mathbf{X}) \geq \gamma$  can be written as

$$\mathcal{P}\{\mathcal{S}(\mathbf{X}) \geq \gamma\} = \mathbb{E}\left\{\mathbb{1}_{\mathcal{S}(\mathbf{X}) \geq \gamma}\right\}, \quad (18)$$

which provides a practical way of estimating its value. Once the right side of Eq.(18) is the mean value of the random event  $\mathbb{1}_{\mathcal{S}(\mathbf{X}) \geq \gamma}$ , it can be easily computed through the sample mean of  $N_s$  samples of  $\mathbf{X}$ , i.e.,

$$\mathcal{P}\{\mathcal{S}(\mathbf{X}) \geq \gamma\} \approx \frac{1}{N_s} \sum_{k=0}^{N_s} \mathbb{1}_{\mathcal{S}(\mathbf{X}_k) \geq \gamma}, \quad (19)$$

where the realizations  $\mathbf{X}_k$  ( $k = 1, \dots, N_s$ ) are drawn according to  $g(\mathbf{x}; \mathbf{v})$ , the probability density function of  $\mathbf{X}$ , parametrized by the hyper-parameters vector  $\mathbf{v}$ . Very often,  $\mathbf{v} = (\boldsymbol{\mu}, \boldsymbol{\sigma})$ , where  $\boldsymbol{\mu}$  and  $\boldsymbol{\sigma}$  are the mean and the standard deviation vectors of  $\mathbf{X}$ , respectively.

The idea of the CE method is to approximate the solution of the underlying optimization problem by solving the rare event probability estimation from Eq.(18). For this purpose, it employs a multilevel approach, that generates an optimal sequence of statistical estimators for the pair  $(\gamma, \mathbf{v})$ , denoted by  $(\hat{\gamma}_\ell, \hat{\mathbf{v}}_\ell)$ , such that

$$\hat{\gamma}_\ell \xrightarrow{a.s.} \gamma^* \quad \text{and} \quad g(\mathbf{x}, \hat{\mathbf{v}}_\ell) \xrightarrow{a.s.} \delta(\mathbf{x} - \mathbf{x}^*), \quad (20)$$

i.e., the reference level  $\gamma$  tends with probability 1 to the maximum value, and the family of distributions  $g(\cdot, \mathbf{v})$  goes (almost sure) towards a point mass distribution, centered on the optimization problem global optimum.

This sequence of estimators is optimal in the sense that it minimizes the Kullback-Leibner divergence between  $\mathbb{1}_{\mathcal{S}(\mathbf{X}) \geq \gamma}$  and  $g(\cdot, \mathbf{v})$  [55].

More concretely, the feasible region  $\mathcal{D}$  is sampled with  $N_s$  independent and identically distributed (iid) realizations of the random vector  $\mathbf{X}$ , drawn from its density  $g(\mathbf{x}; \mathbf{v})$ . For each of these samples, the objective function is evaluated, generating the sequence of values  $\mathcal{S}(\mathbf{X}_1), \mathcal{S}(\mathbf{X}_2), \dots, \mathcal{S}(\mathbf{X}_{N_s})$ .

Then, an elite sample  $\mathcal{E}_t = \{\mathbf{X}_k : \mathcal{S}(\mathbf{X}_k) \geq \hat{\gamma}_t\}$  is defined by lumping the  $N_e < N_s$  points that better performed, i.e., those which produced the highest values for  $\mathcal{S}(\mathbf{x})$ . Note that this elite set is defined in terms of the maximum value statistical estimator, given by

$$\hat{\gamma}_t = \mathcal{S}_{(N_s - N_e + 1)}. \quad (21)$$

The hyper-parameters vector  $\mathbf{v}$  is updated using the maximum likelihood estimator so that, with the aid of the elite set  $\mathcal{E}_t$ , it is written as

$$\hat{\mathbf{v}}_t = \arg \max_{\mathbf{v}} \sum_{\mathbf{X}_k \in \mathcal{E}_t} \ln(g(\mathbf{X}_k; \mathbf{v})). \quad (22)$$

Depending on the distribution chosen for  $\mathbf{X}$ , the stochastic program from Eq.(22) needs to be solved numerically. However, for the distributions of the exponential family, which includes the Gaussian and its truncated version, this estimator can be calculated in an analytic way, with each component of mean and deviation given by the formulas

$$\hat{\boldsymbol{\mu}}_t = \frac{\sum_{\mathbf{X}_k \in \mathcal{E}_t} X_k}{N_e}, \quad (23)$$

and

$$\hat{\boldsymbol{\sigma}}_t = \sqrt{\frac{\sum_{\mathbf{X}_k \in \mathcal{E}_t} (X_k - \hat{\boldsymbol{\mu}}_t)^2}{N_e}}, \quad (24)$$

respectively.

Although this process has a theoretical guarantee of converging to a point mass distribution centered on the global optimum, in computational practice it is common to see the updated distribution numerically degenerating before it “reaches the target”. Sometimes the standard deviation decreases very quickly, causing the distribution to “shrink” in a region far from the global optimum. In this scenario, only samples far from the optimum point are drawn, producing poor estimates for the optimization problem solution.

At the theoretical level, where it is possible to sample an infinite number of times, this pathological situation is bypassed, because at some point a sample on the tail is drawn, moving the distribution center out of the “frozen region”, as this outlier has a lot of weight in the mean estimate. However, in computational terms, as any robust implementation requires a maximum number of iterations (levels), the algorithm may stop with the distribution center within a “frozen region”.

This problem may be solved through the use of a smooth updating scheme for the hyper-parameters

$$\hat{\boldsymbol{\mu}}_t := \alpha \hat{\boldsymbol{\mu}}_t + (1 - \alpha) \hat{\boldsymbol{\mu}}_{t-1}, \quad (25)$$

$$\hat{\boldsymbol{\sigma}}_t := \beta_t \hat{\boldsymbol{\sigma}}_t + (1 - \beta_t) \hat{\boldsymbol{\sigma}}_{t-1}, \quad (26)$$

$$\beta_t = \beta - \beta \left(1 - \frac{1}{t}\right)^q, \quad (27)$$

where the smooth parameters are such that  $0 < \alpha \leq 1$ ,  $0.8 \leq \beta \leq 0.99$  and  $5 \leq q \leq 10$ , and the estimations at  $t$  and  $t - 1$  are obtained by solving the Eq.(22).

### 4.3 The numerical algorithm

The geometric idea of the CE method is described at the beginning of this section, being complemented by the theoretical formalism presented in the previous subsection. The compilation of these ideas in the form of an easy to implement computational algorithm is presented below:

1. Define the number of samples  $N_s$ , the number of elite samples  $N_e < N_s$ , a convergence tolerance `tol`, the maximum of iteration levels  $\ell_{max}$ , a family of probability distributions  $g(\cdot, \mathbf{v})$ , an initial vector of hyper-parameters  $\hat{\mathbf{v}}_0$  for  $g$ , and set the level counter  $\ell = 0$ ;
2. Update the level counter  $\ell = \ell + 1$ ;
3. Generate a total of  $N_s$  independent and identically distributed (iid) samples from  $g(\cdot, \hat{\mathbf{v}}_{\ell-1})$ , denoted by  $\mathbf{X}_1, \dots, \mathbf{X}_{N_s}$ ;
4. Evaluate the objective function  $\mathcal{S}(\mathbf{X}_n)$  at the samples  $\mathbf{X}_1, \dots, \mathbf{X}_{N_s}$  and sort the results  $\mathcal{S}_{(1)} \leq \dots \leq \mathcal{S}_{(N_s)}$  to select the elite sample (the  $N_e$  points which better performed);

5. Update the estimators  $\hat{\gamma}_\ell$  and  $\hat{\mathbf{v}}_\ell$  with aid of the elite sample, using the maximum likelihood method;
6. Repeat the steps (2) up to (5) while a (standard deviation dependent) stopping criterion is not met.

A schematic representation of this algorithm, illustrating all the stages of the sampling and learning phases, can be seen in Figure 3.

### 4.4 Remarks about CE method

Among the several characteristics that make this relatively novel technique interesting, the following can be highlighted:

- *Simplicity* – very intuitive algorithm with few control parameters ( $N_s, N_e, \ell_{max}$  and `tol`), each of them with a very clear interpretation;
- *Robustness* – theoretical results ensure that, under typical conditions, the method is guaranteed to converge if the problem has a single global extreme;
- *Efficiency* – the method typically presents fast convergence in comparison with classic global search meta-heuristics, such as genetic algorithms;
- *Generality* – the method does not require any regularity of the objective function and can be applied to almost any type of non-convex optimization problem (even non-differentiable or discontinuous);
- *Generalizable* – the theory is general, in principle it can be applied to problems of any finite dimension, the computational cost being the limiting factor in practice;
- *Easy implementation* – the algorithm can be implemented with a few lines of code in a high level programming language.

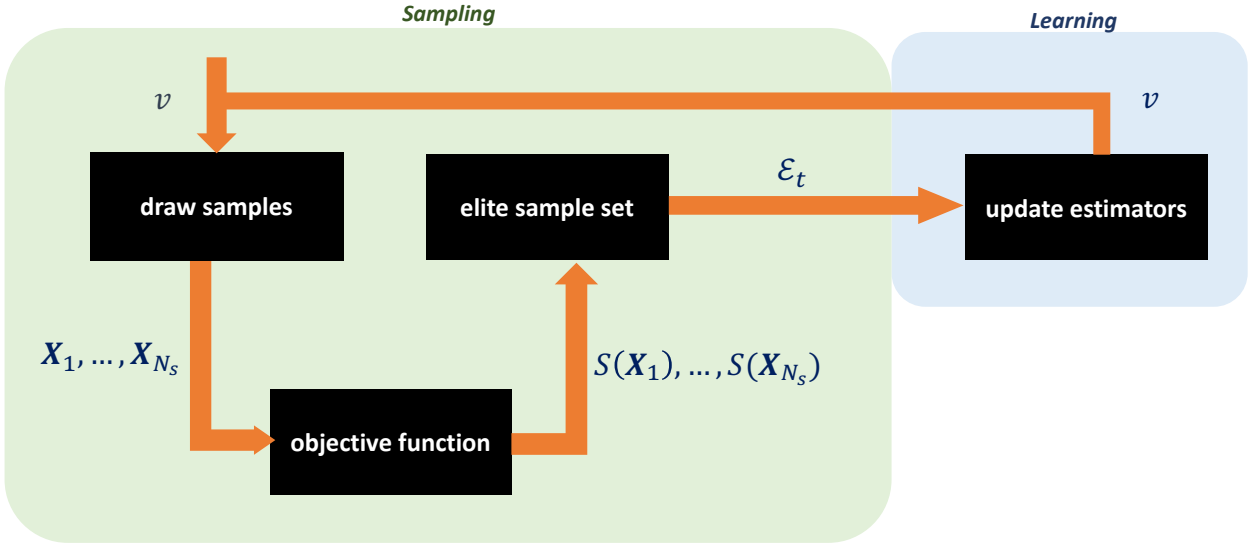
The mathematical development associated with the formalism described above is relatively non-trivial, but it has been well established throughout the first decade of the method, including theorems that strictly establish the conditions where the method has guaranteed convergence. These details are suppressed from this paper as they are outside the scope of the journal. But for the interested reader, the following references are recommended [19, 35, 54, 55].

The computational experiments in the next section illustrate the efficiency and robustness of this framework in solving a non-trivial problem of optimizing an energy harvesting device.

## 5 Numerical experiments

For the numerical experiments conducted here, the following numerical values are adopted for the dynamical





**Figure 3** Schematic representation of the CE algorithm for optimization.

system parameters:  $\xi = 0.01$ ,  $\chi = 0.05$ ,  $\kappa = 0.5$  and  $\lambda = 0.05$ . The initial condition is defined by  $x_0 = 1$ ,  $\dot{x}_0 = 0$  and  $v = 0$ . The dynamics is integrated over the time interval  $[0, T] = [0, 2500]$ , and the mean output power is computed over the last 50% of this time-series.

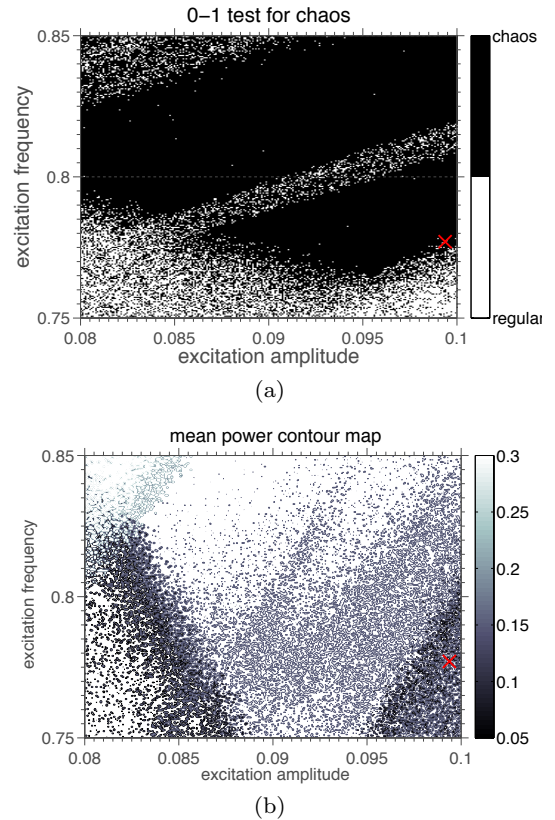
### 5.1 Reference solution

In order to analyze the effectiveness and robustness of the stochastic solution strategy proposed here, a reference solution is computed by a standard exhaustive search on a fine grid over the domain

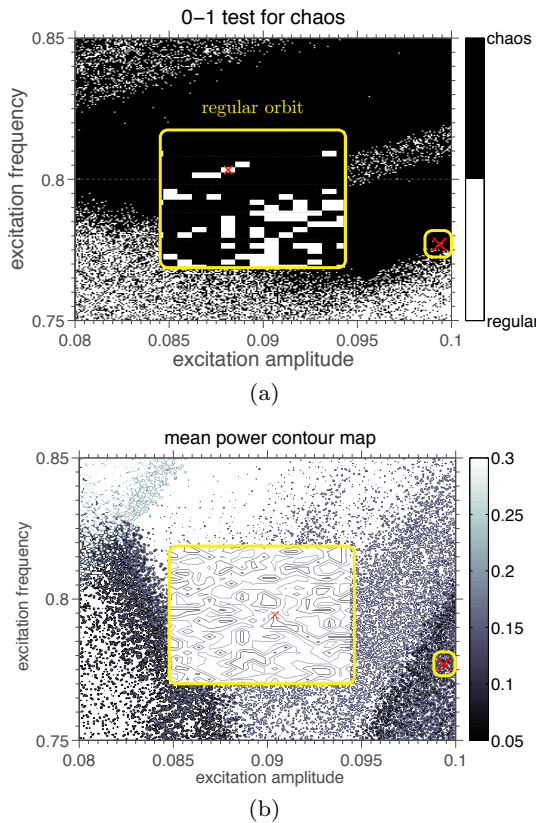
$$\mathcal{D} = \{(f, \Omega) \mid 0.08 \leq f \leq 0.1, 0.75 \leq \Omega \leq 0.85\}.$$

In this standard approach, a structured  $256 \times 256$  uniform numerical grid is used to discretize  $\mathcal{D}$ . The system dynamics is then integrated for each grid point, with the optimization constraint being evaluated next. The objective function is evaluated at all feasible points, and the extreme value is updated at each step of the grid scanning process. Two contour maps, associated with the reference solution, are shown in Figure 4: (a) constraint function, and (b) objective function. The pair  $(f, \Omega) = (0.0994, 0.7771)$  that corresponds to the global maximum is indicated in both contour maps by a red cross, being associated with a mean output power  $P = 0.3438$ . Figure 5 shows a magnification of the global maximum neighborhood, where it is possible to better appreciate the contour levels shape, and verify that it corresponds to a regular dynamic regime configuration. Note also that this result is compatible with the literature, which points to a better performance of piezoelectric harvesters for low frequencies and high

amplitudes of excitation [59, 40]. For sake of reference, this solution was obtained in 14 hours in a MacBook Pro Core i7 2.2 GHz 16GB 1333 MHz DDR3.



**Figure 4** Contour maps of the: (a) constraint function defined by 0-1 test for chaos, and (b) objective function defined by mean output power.



**Figure 5** Magnification of the contour maps, around the global maximum, of the: (a) constraint function, and (b) objective function.

## 5.2 Cross-entropy solution

In the approach based on the CE method, the domain is randomly sampled using for this  $N^s = 50$  points. This sampling is done according to a truncated Gaussian distribution, parametrized by mean vector  $\boldsymbol{\mu} = (\mu_f, \mu_\Omega)$  and standard deviation vector  $\boldsymbol{\sigma} = (\sigma_f, \sigma_\Omega)$ . The number of elite samples is chosen as  $N^e = \text{round}(N^s/10)$ , maximum number of levels is set  $t_{max} = 100$ , while the convergence criterion is adopted as  $\max\{\sigma_f, \sigma_\Omega\} < \text{tol}$ , for a tolerance  $\text{tol} = 1 \times 10^{-3}$ . The smoothing parameters are  $\alpha = 0.7$ ,  $\beta = 0.8$  and  $q = 5$ .

A visual illustration of the CE method is presented in Figure 6, which shows the domain sampling at different levels (iterations) of the algorithm. An animation of the algorithm in action is available in [11]. The reader can also appreciate the evolution of this random algorithm in Table 1, where each line displays the level index, the value of the means<sup>1</sup> and standard deviations of  $f$  and  $\Omega$  in addition to the optimal value obtained for the objective function.

<sup>1</sup> Remember that the approximation for the optimum point is given by the vector of means, i.e.,  $(\mu_f^*, \mu_\Omega^*) \approx (\mu_f, \mu_\Omega)$ .

**Table 1** Evolution of CE algorithm using  $N^s = 50$  samples.

$\ell$	P	$\mu_f$	$\mu_\Omega$	$\sigma_f$	$\sigma_\Omega$
01	+0.00729	0.08752	0.79356	0.02516	0.12591
02	+0.00634	0.09100	0.77788	0.01006	0.04820
03	-0.97244	0.09241	0.76584	0.00714	0.02012
04	+0.00779	0.09515	0.75804	0.00411	0.01003
05	+0.33424	0.09686	0.76026	0.00281	0.00576
06	+0.33530	0.09706	0.76248	0.00192	0.00425
07	+0.33726	0.09752	0.76526	0.00175	0.00338
08	+0.33792	0.09820	0.76673	0.00168	0.00243
09	+0.33854	0.09869	0.76805	0.00145	0.00240
10	+0.33990	0.09855	0.76977	0.00112	0.00195
11	+0.34022	0.09877	0.77115	0.00097	0.00167
12	+0.34083	0.09858	0.77247	0.00095	0.00149
13	+0.34155	0.09875	0.77361	0.00082	0.00130
14	+0.34126	0.09936	0.77329	0.00069	0.00112
15	+0.34143	0.09942	0.77342	0.00062	0.00116
16	+0.34115	0.09954	0.77319	0.00060	0.00110
17	+0.34191	0.09979	0.77392	0.00049	0.00096

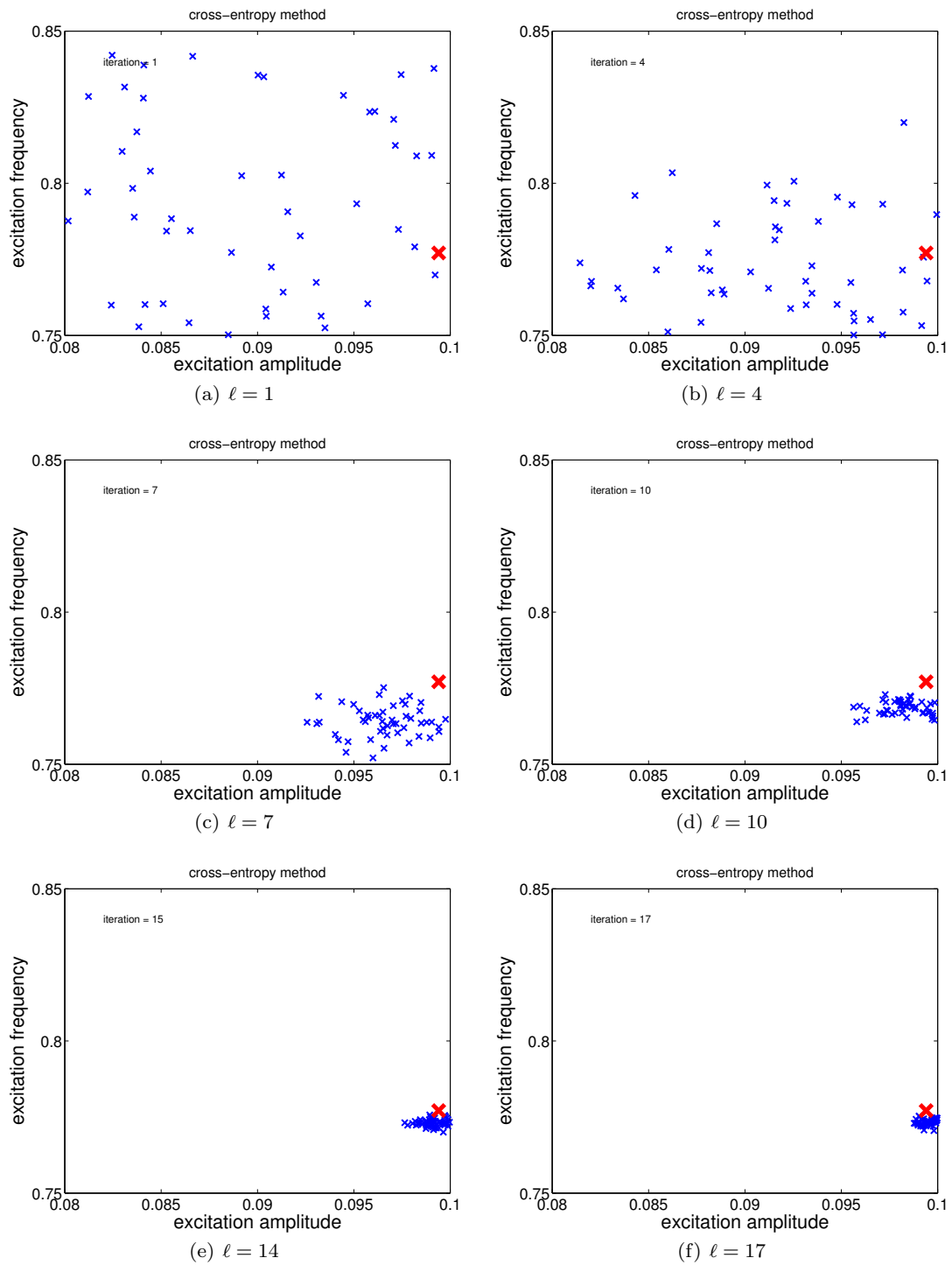
**Table 2** Performance of CE algorithm for different number of samples.

samples	levels	CPU time* (seconds)	speed-up	function evaluation
reference	—	50 632	—	65536
25	15	304	167	375
50	17	696	73	850
75	29	1736	29	2175
100	42	3286	15	4200

\*MacBook Pro Core i7 2.2 GHz 16GB 1333 MHz DDR3

Note that the approximation obtained is very close to the reference value of section 5.1, obtained after 17 iterative steps only, which corresponds to a speed-up of more than 70, when compared with the exhaustive search (see in Table 2 the CPU time spent). The accuracy can still be slightly improved as shown in Table 3, which presents the CE results with  $N^s = 75$ . The corresponding animation can be seen in [12]. The price to be paid for this additional gain of accuracy is a loss of performance, which makes the speed-up fall from more than 70 to 29 (see Table 2). An experiment with only  $N^s = 25$  samples is also conducted (see the video in [10]), obtaining a result with no significant loss of accuracy and more than doubling the speed-up to an impressive value of 167. A numerical experiment with  $N^s = 100$  is also reported in Table 2, although no figure or table from this simulation is shown in the text.

The speed-up shown in Table 2 is defined as the ratio between the calculation time between direct search (reference) and CE solution. Although this metric provides a good measure of efficiency for large values of CPU times, it is extremely dependent on the machine used. Thus, to provide a machine-independent performance measure, this table also lists the number of eval-



**Figure 6** Illustration of CE method sampling of the domain at different levels (iterations) of the algorithm. The reference solution is indicated with a red cross.

**Table 3** Evolution of CE algorithm using  $N^s = 75$  samples.

$\ell$	P	$\mu_f$	$\mu_\Omega$	$\sigma_f$	$\sigma_\Omega$
01	+0.00676	0.09039	0.78364	0.02378	0.11257
02	+0.00694	0.09134	0.78555	0.01052	0.05551
03	+0.00677	0.09146	0.77791	0.00755	0.02934
04	-7.87800	0.09159	0.77091	0.00587	0.02273
05	+0.00585	0.09168	0.76281	0.00506	0.01298
06	+0.18206	0.09420	0.76179	0.00511	0.01140
07	+0.33294	0.09703	0.76184	0.00337	0.00908
08	+0.33606	0.09834	0.76411	0.00261	0.00652
09	+0.33771	0.09880	0.76678	0.00205	0.00474
10	+0.33819	0.09906	0.76827	0.00157	0.00400
11	+0.33835	0.09897	0.76836	0.00133	0.00315
12	+0.33876	0.09881	0.76952	0.00119	0.00275
13	+0.34010	0.09925	0.77121	0.00097	0.00263
14	+0.34006	0.09939	0.77147	0.00081	0.00237
15	+0.34082	0.09953	0.77219	0.00073	0.00208
16	+0.34101	0.09946	0.77254	0.00067	0.00180
17	+0.34122	0.09941	0.77329	0.00062	0.00174
18	+0.34140	0.09931	0.77316	0.00055	0.00158
19	+0.34166	0.09930	0.77362	0.00050	0.00155
20	+0.34114	0.09947	0.77355	0.00044	0.00149
21	+0.34193	0.09964	0.77385	0.00040	0.00140
22	+0.34034	0.09962	0.77375	0.00038	0.00148
23	+0.34093	0.09957	0.77368	0.00036	0.00142
24	+0.34154	0.09950	0.77389	0.00035	0.00132
25	+0.34193	0.09957	0.77427	0.00033	0.00119
26	+0.34227	0.09959	0.77479	0.00032	0.00110
27	+0.34203	0.09963	0.77490	0.00031	0.00104
28	+0.34220	0.09973	0.77502	0.00029	0.00100
29	+0.34203	0.09972	0.77487	0.00027	0.00095

**Table 4** Evolution of CE algorithm using  $N^s = 25$  samples.

$\ell$	P	$\mu_f$	$\mu_\Omega$	$\sigma_f$	$\sigma_\Omega$
01	-0.01735	0.09252	0.78304	0.02400	0.10592
02	-1.71710	0.08654	0.77635	0.00951	0.03866
03	+0.00538	0.08894	0.76449	0.00849	0.01564
04	+0.00781	0.09366	0.75846	0.00491	0.00903
05	+0.33423	0.09698	0.76137	0.00351	0.00671
06	+0.33689	0.09742	0.76432	0.00227	0.00450
07	+0.33610	0.09765	0.76447	0.00210	0.00310
08	+0.33809	0.09810	0.76732	0.00194	0.00244
09	+0.33903	0.09850	0.76834	0.00141	0.00170
10	+0.33874	0.09857	0.76807	0.00137	0.00129
11	+0.33924	0.09907	0.76836	0.00117	0.00110
12	+0.33892	0.09892	0.76867	0.00110	0.00104
13	+0.33937	0.09894	0.76930	0.00110	0.00105
14	+0.33924	0.09927	0.76928	0.00103	0.00086
15	+0.34030	0.09935	0.77019	0.00086	0.00067

uations of the objective function, since this is the most expensive operation to be performed by the optimization algorithm. Note that in this new metric, the ratio between the objective function evaluations in the reference and the CE solution produces values very close to the speed-ups obtained previously, confirming the computational efficiency of the proposed approach in a machine independent fashion.

The above results allow us to conclude that the optimization approach based on CE method is very robust (able to find the global optimal) and efficient (computationally feasible) to address this nonlinear non-convex optimization problem, which has non-trivial numerical solution since the existence of jump-like discontinuities in the constraint prevents gradient-based algorithms from being used.

Last but not least, it is worth mentioning that the CE method has a great advantage in terms of simplicity when compared to most of the meta-heuristics used in non-differentiable optimization since the algorithm involves only four control parameters, all of which have a very intuitive meaning, which is a considerable advantage compared to genetic algorithms for example.

## 6 Concluding remarks

This work presented the formulation of a nonlinear non-convex optimization problem that seeks to maximize the efficiency of a bistable energy harvesting system driven by a sinusoidal excitation and constrained to operate in non-chaotic dynamic regimes only. Since the problem has jump-type discontinuities, which prevents the use of gradient-based methods, a stochastic approach based on the cross entropy method was proposed to construct a numerical approximation for the optimal solution. Tests to verify the efficiency and accuracy of this cross-entropy approach were conducted, showing that the proposed strategy of solution is quite robust and effective, constituting a very appealing tool to deal with the typical (non-convex) problems related to the optimization of energy harvesting systems.

The impressive results reported here can still be improved in terms of performance, through the use of parallelization strategies. In particular, it would be interesting to test the cloud computing parallelization strategy proposed by [13]. In addition, the optimization framework is extremely versatile, allowing extension to problems involving robust optimization such as those reported by [14, 61], in which the global optimum is not sought in the strict sense, but in a sense in which the system performance is maximized in average terms, respecting probabilistic constraints.

## Acknowledgements

The author is very grateful to Dr. Welington de Oliveira (MINES ParisTech), for the fruitful discussions about the optimization problem addressed in this work.

## Funding

This research received financial support from the Brazilian agencies Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001, and the Carlos Chagas Filho Research Foundation of Rio de Janeiro State (FAPERJ) under the following grants: 211.304/2015, 210.021/2018, 210.167/2019, and 211.037/2019.

## Compliance with ethical standards

### Conflict of Interest

The author declares that he has no conflict of interest.

### Code availability

To facilitate the reproduction of the results of this paper, as well as to popularize the use of the CE method in the dynamical systems community, the code used in the simulations is available in the following repository: <https://github.com/amicocunhajr/HarvesterOpt>

## References

1. Abdelkefi A, Nayfeh AH, Hajj MR (2012) Enhancement of power harvesting from piezoaeroelastic systems. *Nonlinear Dynamics* 68:531–541, DOI <https://doi.org/10.1007/s11071-011-0234-9>
2. Barbosa WOV, De Paula AS, Savi MA, Inman DJ (2015) Chaos control applied to piezoelectric vibration-based energy harvesting systems. *The European Physical Journal Special Topics* 224:2787–2801, DOI <https://doi.org/10.1140/epjst/e2015-02589-1>
3. Benacchio S, Malher A, Boisson J, Touzé C (2016) Design of a magnetic vibration absorber with tunable stiffnesses. *Nonlinear Dynamics* 85:893–911, DOI <https://doi.org/10.1007/s11071-016-2731-3>
4. Bernardini D, Litak G (2016) An overview of 0–1 test for chaos. *Journal of the Brazilian Society of Mechanical Sciences and Engineering* 38:1433–1450, DOI <https://doi.org/10.1007/s40430-015-0453-y>
5. Bhatti NA, Alizai MH, Syed AA, L M (2016) Energy harvesting and wireless transfer in sensor network applications: Concepts and experiences. *ACM Transactions on Sensor Networks* 12:24:1–24:40, DOI <https://doi.org/10.1145/2915918>
6. Bonnans J, Gilbert JC, Lemarechal C, Sagastizábal CA (2009) *Numerical Optimization: Theoretical and Practical Aspects*, 2nd edn. Springer
7. Boyd S, Vandenberghe L (2004) *Convex Optimization*. Cambridge University Press
8. Catacuzzeno L, Orfei F, Michele AD, Sforna L, Franciolini F, Gammaitoni L (2019) Energy harvesting from a bio cell. *Nano Energy* 56:823–827
9. Cottone F, Vocca H, Gammaitoni L (2009) Nonlinear energy harvesting. *Phys Rev Lett* 102:080601, DOI <https://doi.org/10.1103/PhysRevLett.102.080601>
10. Cunha Jr A (2020) Cross-entropy optimization of bistable energy harvesting system (25 samples). <https://youtu.be/0EvzdVX1PqA>, (Accessed 28 March 2020)
11. Cunha Jr A (2020) Cross-entropy optimization of bistable energy harvesting system (50 samples). <https://youtu.be/-JB3eniIdDY>, (Accessed 28 March 2020)
12. Cunha Jr A (2020) Cross-entropy optimization of bistable energy harvesting system (75 samples). <https://youtu.be/uIZM4SjCbrw>, (Accessed 28 March 2020)
13. Cunha Jr A, Nasser R, Sampaio R, Lopes H, Britman K (2014) Uncertainty quantification through Monte Carlo method in a cloud computing setting. *Computer Physics Communications* 185:1355–1363, DOI <https://doi.org/10.1016/j.cpc.2014.01.006>
14. Cunha Jr A, Soize C, Sampaio R (2015) Computational modeling of the nonlinear stochastic dynamics of horizontal drillstrings. *Computational Mechanics* 56:849–878, DOI <https://doi.org/10.1007/s00466-015-1206-6>
15. Dantas E (2019) A cross-entropy strategy for parameters identification problems. Monograph, Universidade do Estado do Rio de Janeiro, <https://dx.doi.org/10.13140/RG.2.2.18045.51688>
16. Dantas E, Cunha Jr A, Silva TAN (2019) A numerical procedure based on cross-entropy method for stiffness identification. In: 5th International Conference on Structural Engineering Dynamics (ICE-Dyn 2019), Viana do Castelo, Portugal
17. Dantas E, Cunha Jr A, Soeiro FJCP, Cayres BC, Weber HI (2019) An inverse problem via cross-entropy method for calibration of a drill string torsional dynamic model. In: 25th ABCM International Congress of Mechanical Engineering (COBEM 2019), Uberlândia, Brazil, DOI <http://dx.doi.org/10.26678/abcm.cobem2019.cob2019-2216>

18. Daqaq MF, Crespo RS, Ha S (2020) On the efficacy of charging a battery using a chaotic energy harvester. *Nonlinear Dynamics* 99:1525–1537, DOI <https://doi.org/10.1007/s11071-019-05372-0>
19. De Boer P, Kroese DP, Mannor S, Rubinstein RY (2005) A tutorial on the cross-entropy method. *Annals of Operations Research* 134:19–67, DOI <https://doi.org/10.1007/s10479-005-5724-z>
20. de la Roca L, Peterson JVLL, Pereira MC, Cunha Jr A (2019) Control of chaos via OGY method on a bistable energy harvester. In: 25th ABCM International Congress of Mechanical Engineering (COBEM 2019), Uberlândia, Brazil, DOI <http://dx.doi.org/10.26678/abcm.cobem2019.cob2019-1970>
21. Dekemele K, Van Torre P, Loccufier M (2019) Performance and tuning of a chaotic bi-stable meso to mitigate transient vibrations. *Nonlinear Dynamics* 98:1831–1851, DOI <https://doi.org/10.1007/s11071-019-05291-0>
22. Erturk A, Hoffmann J, Inman DJ (2009) A piezomagnetoelastic structure for broadband vibration energy harvesting. *Applied Physics Letters* 94:254102, DOI <https://doi.org/10.1063/1.3159815>
23. Gallo CA, Tofoli FL, Rade DA, Steffen, Jr V (2012) Piezoelectric actuators applied to neutralize mechanical vibrations. *Journal of Vibration and Control* 18:1650–1660, DOI <https://doi.org/10.1177/1077546311422549>
24. Gammaitone L (2012) There's plenty of energy at the bottom (micro and nano scale nonlinear noise harvesting). *Contemporary Physics* 53:119–135, DOI <https://doi.org/10.1080/00107514.2011.647793>
25. Ghidry H (2015) Reliability-based design optimization with cross-entropy method. Master's thesis, Norwegian University of Science and Technology, Trondheim
26. Ghouli Z, Hamdi M, Belhaq M (2017) Energy harvesting from quasi-periodic vibrations using electromagnetic coupling with delay. *Nonlinear Dynamics* 89:1625–1636, DOI <https://doi.org/10.1007/s11071-017-3539-5>
27. Godoy TC, Trindade MA (2012) Effect of parametric uncertainties on the performance of a piezoelectric energy harvesting device. *Journal of the Brazilian Society of Mechanical Sciences and Engineering* 34:552–560, DOI <https://dx.doi.org/10.1590/S1678-58782012000600003>
28. Gottwald GA, Melbourne I (2004) A new test for chaos in deterministic systems. *Proceedings of the Royal Society of London Series A* 460:603–611, DOI <https://doi.org/10.1098/rspa.2003.1183>
29. Gottwald GA, Melbourne I (2009) On the implementation of the 0-1 test for chaos. *SIAM Journal on Applied Dynamical Systems* 8:129–145, DOI <https://doi.org/10.1137/080718851>
30. Gottwald GA, Melbourne I (2009) On the validity of the 0-1 test for chaos. *Nonlinearity* 22:1367–1382, DOI <https://doi.org/10.1088/0951-7715/22/6/006>
31. Gottwald GA, Melbourne I (2016) The 0-1 Test for Chaos: A review, vol 915, Springer. DOI <https://doi.org/10.1007/978-3-662-48410-4>
32. Harne RL (2012) Theoretical investigations of energy harvesting efficiency from structural vibrations using piezoelectric and electromagnetic oscillators. *The Journal of the Acoustical Society of America* 132:162–172, DOI <https://doi.org/10.1121/1.4725765>
33. Ibrahim A, Towfighian S, Younis M, Su Q (2016) Magnetoelastic beam with extended polymer for low frequency vibration energy harvesting. In: Meyendorf NG, Matikas TE, Peters KJ (eds) *Smart Materials and Nondestructive Evaluation for Energy Systems 2016*, International Society for Optics and Photonics, SPIE, vol 9806, pp 71 – 85, DOI <https://doi.org/10.1117/12.2219276>
34. Issa MVS, Cunha Jr A, Soeiro FJCP, Pereira A (2018) Structural optimization using the cross-entropy method. In: XXXVIII Congresso Nacional de Matemática Aplicada e Computacional (CN-MAC 2018), Campinas, Brazil, DOI <https://dx.doi.org/10.5540/03.2018.006.02.0443>
35. Kroese DP, Taimre T, Botev ZI (2011) *Handbook of Monte Carlo Methods*. Wiley
36. Kroese DP, Rubinstein RY, Cohen I, Porotsky S, Taimre T (2013) *Cross-Entropy Method*, Springer, pp 326–333. DOI [https://doi.org/10.1007/978-1-4419-1153-7\\_131](https://doi.org/10.1007/978-1-4419-1153-7_131)
37. Leadenham S, Erturk A (2020) Mechanically and electrically nonlinear non-ideal piezoelectric energy harvesting framework with experimental validations. *Nonlinear Dynamics* 99:625–641, DOI <https://doi.org/10.1007/s11071-019-05091-6>
38. Liu W, Xu X, Chen F, Liu Y, Li S, Liu L, Chen Y (2019) A review of research on the closed thermodynamic cycles of ocean thermal energy conversion. *Renewable and Sustainable Energy Reviews* p 109581, DOI <https://doi.org/10.1016/j.rser.2019.109581>
39. Lopes VG, Peterson JVLL, Cunha Jr A (2017) Numerical study of parameters influence over the dynamics of a piezo-magneto-elastic energy harvesting device. In: XXXVII Congresso Nacional de Matemática Aplicada e Computacional (CN-MAC 2017), São José dos Campos, Brazil, DOI

- <http://dx.doi.org/10.5540/03.2018.006.01.0407>
40. Lopes VG, Peterson JVLL, Cunha Jr A (2019) Non-linear Characterization of a Bistable Energy Harvester Dynamical System. In: Belhaq M (ed) *Topics in Nonlinear Mechanics and Physics: Selected Papers from CSNDD 2018* (Springer Proceedings in Physics), Springer, Singapore, pp 71–88, DOI [https://dx.doi.org/10.1007/978-981-13-9463-8\\_3](https://dx.doi.org/10.1007/978-981-13-9463-8_3)
  41. López-Suárez M, Rurali R, Gammaitoni L, Abadal G (2011) Nanostructured graphene for energy harvesting. *Physical Review B* 84:161401, DOI <https://doi.org/10.1103/PhysRevB.84.161401>
  42. Mangla C, Ahmad M, Uddin M (2020) Optimization of complex nonlinear systems using genetic algorithm. *International Journal of Information Technology* DOI 10.1007/s41870-020-00421-z
  43. Nabavi S, Zhang L (2016) MEMS piezoelectric energy harvester design and optimization based on Genetic Algorithm. In: 2016 IEEE International Ultrasonics Symposium (IUS), pp 1–4, DOI <https://doi.org/10.1109/ULTSYM.2016.7728786>
  44. Nocedal J, Wright S (2006) *Numerical Optimization*, 2nd edn. Springer
  45. Peterson JVLL, Lopes VG, Cunha Jr A (2016) Maximization of the electrical power generated by a piezo-magneto-elastic energy harvesting device. In: XXXVI Congresso Nacional de Matemática Aplicada e Computacional (CNMAC 2016), Gramado, Brazil, DOI <http://dx.doi.org/10.5540/03.2017.005.01.0200>
  46. Pfenniger A, Stahel A, Koch VM, Obrist D, Vogel R (2014) Energy harvesting through arterial wall deformation: A FEM approach to fluid–structure interactions and magneto-hydrodynamics. *Applied Mathematical Modelling* 38:3325–3338, DOI <https://doi.org/10.1016/j.apm.2013.11.051>
  47. Priya S, Inman DJ (2009) *Energy Harvesting Technologies*. Springer
  48. Quaranta G, Lacarbonara W, Masri SF (2020) A review on computational intelligence for identification of nonlinear dynamical systems. *Nonlinear Dynamics* 99:1709–1761, DOI 10.1007/s11071-019-05430-7
  49. Ramlan R, Brennan MJ, Mace BR, Kovacic I (2010) Potential benefits of a non-linear stiffness in an energy harvesting device. *Nonlinear Dynamics* 59:545–558, DOI <https://doi.org/10.1007/s11071-009-9561-5>
  50. Rechenbach B, Willatzen M, Lassen B (2016) Theoretical study of the electromechanical efficiency of a loaded tubular dielectric elastomer actuator. *Applied Mathematical Modelling* 40:1232–1246, DOI <https://doi.org/10.1016/j.apm.2015.06.029>
  51. Rocha RT, Balthazar JM, Tuset AM, de Souza SLT, Janzen FC, Arbex HC (2019) On a non-ideal magnetic levitation system: nonlinear dynamical behavior and energy harvesting analyses. *Nonlinear Dynamics* 95:3423–3438, DOI <https://doi.org/10.1007/s11071-019-04765-5>
  52. Rubinstein RY (1997) Optimization of computer simulation models with rare events. *European Journal of Operations Research* 99:89–112, DOI [https://doi.org/10.1016/S0377-2217\(96\)00385-2](https://doi.org/10.1016/S0377-2217(96)00385-2)
  53. Rubinstein RY (1999) The cross-entropy method for combinatorial and continuous optimization. *Methodology and Computing in Applied Probability* 2:127–190, DOI <https://doi.org/10.1023/A:1010091220143>
  54. Rubinstein RY, Kroese DP (2004) *The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation and Machine Learning*. Information Science and Statistics, Springer-Verlag
  55. Rubinstein RY, Kroese DP (2016) *Simulation and the Monte Carlo Method*, 3rd edn. Wiley Series in Probability and Statistics, Wiley
  56. Selvan KV, Ali MSM (2016) Micro-scale energy harvesting devices: Review of methodological performances in the last decade. *Renewable and Sustainable Energy Reviews* 54:1035–1047, DOI <https://doi.org/10.1016/j.rser.2015.10.046>
  57. Spies P, Pollak M, Mateu L (2015) *Handbook of Energy Harvesting Power Supplies and Applications*. Pan Stanford
  58. Trindade MA (2016) *Passive and Active Structural Vibration Control*, Springer International Publishing, pp 65–92. DOI [https://doi.org/10.1007/978-3-319-29982-2\\_4](https://doi.org/10.1007/978-3-319-29982-2_4)
  59. Vocca H, Neri I, Travasso F, Gammaitoni L (2012) Kinetic energy harvesting with bistable oscillators. *Applied Energy* 97:771–776, DOI <https://doi.org/10.1016/j.apenergy.2011.12.087>
  60. Wang B (2012) Parameter estimation for ODEs using a cross-entropy approach. Master’s thesis, University of Toronto, Toronto
  61. Wolszczak P, Lonkwic P, Cunha Jr A, Litak G, Molski S (2019) Robust optimization and uncertainty quantification in the nonlinear mechanics of an elevator brake system. *Meccanica* 54:1057–1069, DOI <https://doi.org/10.1007/s11012-019-00992-7>
  62. Yang T, Cao Q (2019) Time delay improves beneficial performance of a novel hybrid energy harvester. *Nonlinear Dynamics* 96:1511–1530, DOI <https://doi.org/10.1007/s11071-019-04868-z>
  63. Ying Q, Yuan W, Hu N (2015) Improving the efficiency of harvesting electricity from living trees.

- Journal of Renewable and Sustainable Energy 7, DOI <https://doi.org/10.1063/1.4935577>
64. Zheng B, Chang CJ, Gea HC (2009) Topology optimization of energy harvesting devices using piezoelectric materials. *Structural and Multidisciplinary Optimization* 38:17–23, DOI <https://doi.org/10.1007/s00158-008-0265-0>
  65. Zhou S, Cao J, Lin J (2016) Theoretical analysis and experimental verification for improving energy harvesting performance of nonlinear monostable energy harvesters. *Nonlinear Dynamics* 86:1599–1611, DOI <https://doi.org/10.1007/s11071-016-2979-7>