Bio-inspired method for generating self-trapped beams in the nonlinear Schrödinger equation

Daniel Torres-Valladares, Felipe J. Villaseñor-Cavazos, and Servando Lopez-Aguayo*

Tecnologico de Monterrey, Escuela de Ingeniería y Ciencias, Avenida Eugenio Garza Sada 2501, Monterrey, NL 64849, Mexico

(Received 3 April 2021; revised 5 July 2021; accepted 1 September 2021; published 27 September 2021)

We introduce a stochastic optimization technique to obtain self-trapped beams in the generalized nonlinear Schrödinger equation. The method is based on combining a variational approach with a bio-inspired method: the cuckoo search algorithm that relies on Lévy flights. The proposed technique can be easily adapted to generate diverse self-trapped structures in a plethora of nonlinear media. Unlike the standard variational technique and some of the numerical algorithms previously reported, this algorithm allows for the optimization of different families of self-trapped beams concurrently.

DOI: 10.1103/PhysRevA.104.033524

I. INTRODUCTION

A soliton is a stable self-trapped wave in nonlinear media with particlelike properties [1]. Solitons have been objects of study in several different areas, such as optics for waveguides [2–4], random lasers [5], high-energy and particle physics [6], dark matter studies [7,8], and condensed-matter physics [9,10], to name a few. Optical spatial solitons are self-trapped optical beams that are generated by a balance between diffraction and self-focusing effects due to a nonlinear term [11].

Methods to obtain and study solitary waves include the inverse scattering theory [12], Darboux-Backlund transformation [13], self-similarity technique [14], and F-expansion technique [15,16]. These methods are generally used for the case of so-called integrable systems, such as the Kerr medium. For nonintegrable systems a general approach is not known. Numerical methods are frequently needed to solve these systems.

The numerical algorithms for optimization can be classified as either deterministic or stochastic algorithms [17,18]. Deterministic algorithms arise naturally in the classical mechanics of a single particle based on gradient computation, such as the Newton-Raphson method and the gradient descent technique. Stochastic algorithms arise more naturally from statistical physics, such as the case of the simulated annealing algorithm used in thermodynamics [19] or the Metropolis-Hastings algorithm used to study the behavior of neutrons in fissile material [20]. Metaheuristic algorithms based on stochastic optimization have been gaining popularity in the recent years. They have the advantage of not needing the gradient of the objective function, or any other information besides the objective function itself. Although there are a wide variety of metaheuristics, most of them are characterized by two phases: exploitation and exploration. In exploitation, or local search, new solutions are generated by a combination of previous solutions. In exploration, or global search, solutions are generated stochastically [21].

The generation of structured localized light in nonlinear media is a challenging problem due to the lack of standard analytical and computational tools to obtain the corresponding stationary modes. Therefore, in addition to the numerical procedures, semianalytical methods such as variational techniques are used.

In this paper, we propose a variational approach [22] combined with a bio-inspired metaheuristic algorithm, the cuckoo search (CS) [23], as a robust numerical approximation for generating a plethora of structured light on demand that can be used for either integrable or not integrable nonlinear media. We demonstrate that this technique can overcome a key disadvantage of standard variational method, which is that the optimization parameters must be part of a tractable ansatz.

Though solitons are solitary waves with particle behavior in nonlinear media, it is important to mention that from a mathematical point of view, the term "soliton" must fulfill several rigorous requirements. Though in optics the term "soliton" is generally interchangeable with the term "solitary wave," in this paper we chose to use the term "solitary wave" to avoid any confusion.

The structure of our paper is as follows. In Sec. II we explain the theoretical model as well as the stochastic method used to obtain solitary-wave-like profiles. In Sec. III we compare the numerical results obtained with the analytical ones for the Kerr medium. We also report the numerical results for media where there are no known analytical results and we corroborate that the algorithm is useful in generating self-trapped beams by performing numerical propagation. Finally, in Sec. IV we discuss the behavior of the algorithm as we increase the dimensionality of the problem and the number of iterations.

II. METHODOLOGY

A. Theoretical model

We start our analysis by solving for Ψ, the complex scalar field modeled by the generalized $(2+1)$ nonlinear Schrödinger equation (GNLSE):

$$i\frac{\partial \Psi}{\partial z} + \nabla_{\perp}^{2}\Psi + \mathcal{F}(|\Psi|^{2})\Psi = 0, \tag{1}$$

*servando@tec.mx

TABLE I. Nonlinear terms in. Eq. (4). ρ is the radial distance in the plane transverse to the propagation.

Medium	$N\mathcal{L}(\|U\|^2)$
Kerr	$-\frac{1}{2}\|U\|^4$
Saturation	$[\ln(s\|U\|^2+1)-s\|U\|^2]/s^2$
nth-order Bessel	$-p[J_n(k\rho)]^2\|U\|^2$

where $\mathcal{F}(|\Psi|^2)$ is the corresponding function that describes the nonlinear effects and $\nabla_\perp^2$ defines the transverse Laplacian. Note that the generalized nonlinear Schrödinger equation is in dimensionless units. In the context of propagation of optical scalar fields in nonlinear media, Ψ is the dimensionless complex amplitude of the optical field; x, y, and z are the spatial coordinates where x and y are transverse coordinates normalized to the beam width w_0; and z is the longitudinal coordinate normalized to the diffraction length $L_d = \eta_0\omega w_0^2/c$. Here η_0 defines the unperturbed refractive index, ω models the carrying frequency, and c is the speed of light [24].

In this paper we restrict the analysis and results reported for the case of the two-dimensional GNLSE, but a similar analysis can be used to study cases from one to three dimensions. Thus, we set $\nabla_\perp^2\Psi = \frac{\partial\Psi^2}{\partial x^2} + \frac{\partial\Psi^2}{\partial y^2}$. In general, the variational approach used to solve Eq. (1) consists of changing a nonlinear evolution problem to a variational problem corresponding to a Lagrangian density $\mathcal{L}(\Psi, \frac{\partial\Psi}{\partial z}, \nabla_\perp^2\Psi, \Psi^*, \frac{\partial\Psi^*}{\partial z}, \nabla_\perp^2\Psi^*)$ [22]. The equivalent Lagrangian density of the evolution Eq. (1) is given by [25,26]

$$\mathcal{L} = \frac{i}{2}\left(\Psi\frac{\partial\Psi^*}{\partial z} - \Psi^*\frac{\partial\Psi}{\partial z}\right) + |\nabla_\perp\Psi|^2 + N\mathcal{L}(|\Psi|^2). \quad (2)$$

We are looking for approximations for localized wave entities that can propagate in nonlinear media while maintaining a constant size and shape. The transverse optical field amplitude distribution of the beam is represented with the same function through the entire propagation [11]. Therefore we set

$$\Psi(\mathbf{r}_\perp, z) = U(\mathbf{r}_\perp)e^{i\lambda z}, \quad (3)$$

where λ is a real propagation constant. This assumption reduces Eq. (2) to

$$\mathcal{L} = \lambda|U(\mathbf{r}_\perp)|^2 + |\nabla_\perp U(\mathbf{r}_\perp)|^2 + N\mathcal{L}[|U(\mathbf{r}_\perp)|^2], \quad (4)$$

where $\mathbf{r}_\perp$ is the system of coordinates perpendicular to the propagation direction. Possible forms of the anharmonic term $\mathrm{N}\mathcal{L}(|\Psi|)$ in the Lagrangian density [Eq. (4)] are shown in Table I. We assume media with focusing nonlinearity and we report results for three media: the so-called integrable case for the Kerr medium, a saturation medium which can be useful to study photorefractive materials such as $LiNbO_3$ [27], and the final case which consists of a combination of a saturation medium and a nth-order optical Bessel lattice [28].

To find approximate solutions, a trial function or ansatz is proposed for $U(\mathbf{r}_\perp)$ with some variational parameters. The parameters are chosen in such a way that the function minimizes the integral of the Lagrangian in Eq. (4) with respect to $\mathbf{r}_\perp$. Therefore, a k set of equations is obtained:

$$\left|\nabla_{a,b,c,\ldots}\int L d\mathbf{r}_\perp\right| = 0 \approx |\vec{\mu}| \quad (5)$$

where $\nabla_{a,b,c,\ldots}$ is the gradient operator acting over the k-variational parameters $a, b, c, \ldots,$ which is a finite set of parameters to be tuned, and $|\vec{\mu}|$ is the norm of the vector $\vec{\mu}$ given by $|\vec{\mu}| = \sqrt{\sum\mu_k^2}$. Therefore, in order to find approximate solitary wave solutions, we look to minimize $|\vec{\mu}|$.

Theoretically, Eq. (5) should be equated to zero. However, since we are performing a numerical analysis and using trial functions instead of the actual form of the solution, a residual will be obtained. Hence, obtaining approximate solutions to the differential equation in Eq. (4) is equivalent to solving the nonlinear system of equations stated by Eq. (5). This problem could result in a highly nontrivial task, even for the case $n = 2$. Therefore, in this paper we explore the use of CS, a stochastic numerical method, to find approximate solitary wave solutions. We show that CS is a valuable and robust technique to generate solitary waves in many nonlinear media.

B. Cuckoo search method

The CS algorithm was first introduced in optimization problems by Yang and Deb [29]; it is a bio-inspired metaheuristic population search algorithm. The CS uses random processes to search in a bounded space to find approximate solutions to a given target function. The target function could be either maximized or minimized, according to the problem. For this paper the target function is Eq. (5), so the closer $|\vec{\mu}(a_k, b_k, \ldots)| = |\vec{\mu}|_k$ is to zero, the better the solution set $P_k^{(l)} = [a_k, b_k, \ldots]^{(l)}$ $k \in \{1, \ldots, \kappa\}$ (κ being the total number of sets proposed). In this section subindices are used to represent set members and superindices are used to represent the step or iteration of the algorithm.

The algorithm is based on the behavior of cuckoo birds. Some of these birds practice breeding parasitism, meaning they let other species of birds take care of their own eggs without the other bird's knowledge. Cuckoo species try to camouflage the appearance of their eggs with those of the host bird to reduce the probability that the invading eggs will be discovered [30].

In CS, each set of solutions P_k is called a nest (κ in total) and each individual solution is an egg $a, b, c, \ldots$ (n in total). Initially all eggs or solutions are generated randomly. The quality of the camouflage is understood to be how good a particular solution is. As stated previously, the quality of a set P_k is measured with the target function [Eq. (5)].

A very crucial question for any stochastic algorithm is: what is the most convenient statistical strategy to be used in an optimization based on a random search? Though this is still an open problem, it has been suggested that for a space where there are sparse optimum values, a Lévy flight motion is the most adequate strategy in some scenarios [31–33].

A crucial step in this algorithm is to generate novel solutions with a stochastic optimization based in Lévy flights. A Lévy flight is a random walk with a step size that follows a

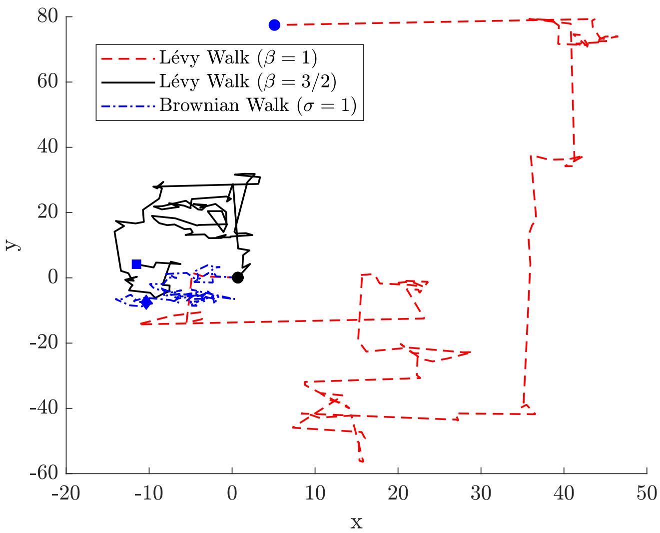


FIG. 1. First 100 steps of Lévy flights for random walks for $\beta = 3/2$ (solid line) and $\beta = 1$ (dashed line) and a Brownian walk (dotted-dashed line) for $\sigma = 1$. All start at the black circle located at $x = 0$ and $y = 0$.

Lévy distribution defined as follows:

$$L(s) = \frac{1}{\pi}\int_0^\infty \cos(\tau s)e^{-\gamma\tau^\beta}d\tau, \quad (6)$$

where γ and β are some parameters of the probability distribution and s is a stochastic variable. For most applications $\gamma = 1$ for simplicity [34]. Since Eq. (6) does not have a general analytical solution, usually approximations are used to reproduce Lévy distributions [35].

Lévy flights allow the algorithm to find new solutions in the search space. As shown in Fig. 1, a difference between a Lévy flight (solid and dashed line) and a Brownian walk (dotted-dashed line) in a two-dimensional space is that Lévy flights cover longer distances within a search space which helps to avoid converging to local minima in the algorithm. The variance of a Lévy flight grows as $\sigma^2_{\text{Lévy}} \sim l^{3-\beta}$. This contrasts the variance of a Brownian walk which grows as $\sigma^2_{\text{Brownian}} \sim l$, l being the current step number and $\beta \in [1, 2]$ being of special interest for mimicking an anomalous diffusion process. From a physical point of view, the parameter β has been suggested to model diverse physical effects, such as quantum tunneling for $0 < \beta \leqslant 1$ or superdiffusion, including turbulence for $1 < \beta < 2$ and diffusion processes for $\beta = 2$ [36]. For purposes of this paper, we set $\beta = 3/2$ since it has been reported as a standard [23].

In our algorithm, new solutions $P_k^{(l+1)}$ are generated by a Lévy flight. This process can be expressed by

$$P_k^{(l+1)} = P_k^{(l)} + [\alpha_k \times L_k(s)], \quad (7)$$

where the vector at the current step l is $P_k^{(l)} = [a_k, b_k]^{(l)}$, and the parameters a_k and b_k are within some chosen range. L_k is defined as

$$L_k(s) = \frac{u}{|v|^{1/\beta}}, \quad (8)$$

where $u \sim N(0, \sigma_u^2 = 1)$ and $v \sim N(0, \sigma_v^2)$ are two random numbers generated from a normal distribution centered at zero with σ_u and σ_v being the standard deviations. The latter is described by

$$\sigma_v = \left(\frac{\Gamma(1+\beta)\sin\left(\pi\frac{\beta}{2}\right)}{\Gamma\left(\frac{1+\beta}{2}\right)2^{(\beta-1)/2}\beta}\right)^{1/\beta}, \quad (9)$$

where $\Gamma(\eta)$ stands for the *Gamma function*, $\Gamma(\eta) = \int_0^\infty t^{\eta-1}e^{-t}dt$. The standard deviations, σ_u and σ_v, are constructed in such a way that the quotient $u/|v|^{1/\beta}$ obeys a symmetric Lévy distribution according to Mantegna's algorithm [35,37]. Mantegna's algorithm is one of the most efficient ways of producing Lévy flights [35]. For other procedures of computing Lévy distributions see [38]. Finally, α_k in Eq. (7) is a scale factor that is calculated as

$$\alpha_k = 0.01(P_{\text{best}} - P_k). \quad (10)$$

Here P_{best} is the current best solution [i.e., the $P_k = [a_k, b_k]$ that makes $|\vec{\mu}|$ in Eq. (5) closest to zero] and 0.01 is a factor to prevent Lévy flights from becoming too aggressive [39]. Note that in this way the best solution is not modified since $\alpha_{\text{best}} = 0$.

Corresponding tails of the Lévy distributions do not fall as fast as the tails of the Gaussian distributions. Hence, a particularly important characteristic of a Lévy flight is that it combines small steps with longer steps within as shown in Fig. 1 causing the variance of the distribution to be infinite [40].

Lévy steps do not have a characteristic length scale which makes them especially good for chaos, field fractal related problems [41], and optimization procedures [42,43] since it is less likely to fall into local minima or maxima. Lévy distributions have also been found naturally in other areas such as astrophysics [44], nuclear physics [45], and plasma physics [46]. Remarkably, Lévy flights have been reported to be more efficient than Brownian random walks for exploring large scale space with sparse optimal values. This can be explained because the variance of Lévy flights increases much more rapidly than that of Brownian random walks [35].

In this paper, we use Lévy flights and Lévy walks interchangeably; however, the former is strictly used for discontinuous trajectories and infinite propagation velocity, while the latter is a stochastic process oriented for continuous trajectories and finite propagation velocity [47].

After the Lévy flight is done, the next step of the CS algorithm will discard solutions with a probability of $\Omega \in [0, 1]$. For this paper $\Omega = 0.25$. This mimics the natural process of the host bird finding a parasite egg and abandoning it. The discarded solutions are subjected to an exploitation process where they combine with each other. This is analogous to improving the camouflage of eggs in nature. In this step of CS, a local search occurs to improve the quality of the solutions, and hence the algorithm ensures that the best eggs (high-quality solutions) will reach the next generations. Figure 2 shows a flowchart of the CS algorithm implemented to obtain approximate solitary wave solutions. We will show later in this paper that even though the diagram is presented for just two variational parameters, the algorithm could be used to tune more parameters in a straightforward way.

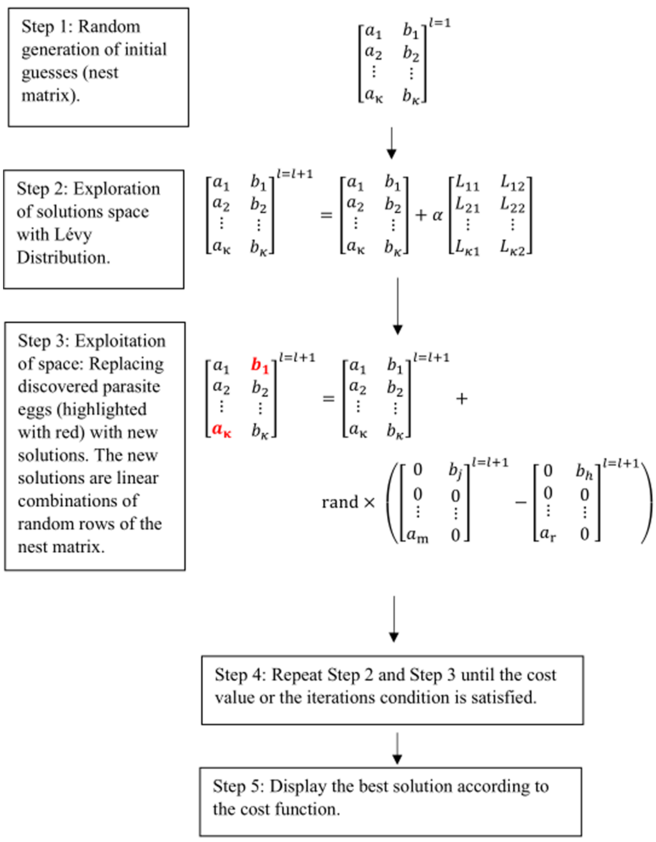


FIG. 2. Diagram of the CS algorithm implemented to approximate solitary waves solutions with two variational parameters. Here each row is composed of the two elements a_k and b_k ($k = 1, 2, \dots, \kappa$) which stand for the amplitude and width parameters, respectively. The superindex, l, is used to show the number of iterations of the algorithm.

In summary, the main idea of the CS is the use of random walks based on Lévy flights to make every set $P_k^{(l)}$ converge into the set $\mathcal{S}$ as $l \to \infty$, where the set $\mathcal{S}$ satisfies $|\vec{\mu}|_{\mathcal{S}} \equiv 0$.

III. RESULTS OF CUCKOO SEARCH TO APPROXIMATE SOLITARY WAVES

It is necessary to propose trial functions in order to use the variational approach to find a solitary wave solution to Eq. (4). Three trial functions that have been used in literature are [48]

$$U(r_\perp, \theta) = a \exp{(-r_\perp^2/b^2)} \cos{(m\theta)} r_\perp^m, \tag{11}$$

$$U(r_\perp, \theta) = a \exp{(-r_\perp^2/b^2)} \exp{(im\theta)} r_\perp^m, \tag{12}$$

$$U(r_\perp, \theta) = a\, \mathrm{sech}(-r_\perp/b), \tag{13}$$

where θ is the angular coordinate in the transverse plane, $r_\perp = \sqrt{x^2 + y^2}$ with x and y being the Cartesian coordinates, m is an integer denoting the $m + 1$ humps for the solitary wave solution, i is the complex unit, and a and b are variational parameters to be tuned. In this paper, a and b physically correspond to the amplitude and width of the optical beam, respectively.

TABLE II. Analytical solutions for the variational parameters a and b when $m = 0$ for Gaussian and hyperbolic secant envelopes described in the Kerr medium by Eqs. (12) and (13), respectively.

Trial function	$a(\lambda)$	$b(\lambda)$
Gaussian [Eq. (12)]	$2\sqrt{\lambda}$	$\sqrt{2/\lambda}$
Hyperbolic [Eq. (13)]	$2.17\sqrt{\lambda}$	$0.76/\sqrt{\lambda}$

The three proposed trial functions come from physical insights. Equation (11) is emulating first-order Hermite-Gauss beams that constitute solutions to the linear paraxial wave equation [49]. Equation (12) resembles the Laguerre-Gauss optical beams. Last, Eq. (13) is indeed the true soliton solution but for the one-dimensional case of the nonlinear Schrödinger equation in a pure Kerr medium. Note that these two first trial functions are mimicking solutions for $N\mathcal{L}(|\Psi|^2) = 0$.

For the integrable case of a fundamental bright soliton solution ($m = 0$ in a Kerr medium) for Eqs. (12) and (13), analytical solutions for the parameters a and b are already known [48]. They are shown in Table II as a function of the propagation constant λ [see Eq. (3)].

An integral of the trial function must be calculated to obtain these values through the variational approach. The main limitation of the traditional variational approach is that a complex trial function is not useful because it cannot be integrated analytically. Therefore, it is not possible to tune analytically the variational parameters in this case. On the other hand, a very simple trial function could not generate valuable optical beams to analyze. The significance of the CS algorithm is that it is able to numerically tune the variational parameters regardless of the complexity of the trial function.

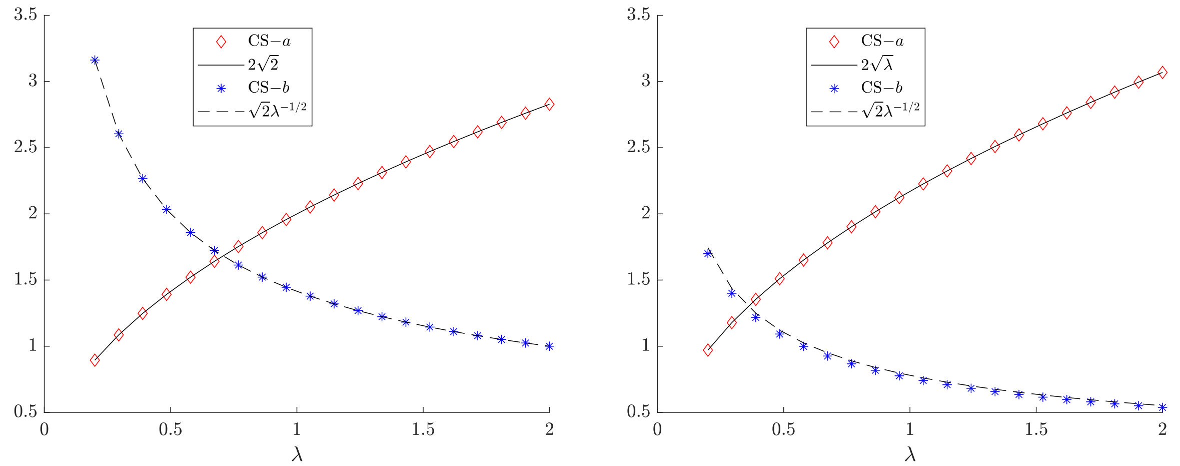
A. Testing algorithm performance: Kerr medium

As a first test of the CS method, we look for approximate solitary waves in a Kerr medium using the corresponding nonlinear term described in Table I. We use Eqs. (12) and (13) as trial functions. We choose these functions because the analytical values are known for these scenarios, therefore we can compare numerical and analytical results. The average absolute percentage deviation between these results is calculated to have a quantitative analysis of the error of the form

$$\epsilon_A = \frac{1}{N} \sum_{k=1}^{N} \left| \frac{A(\lambda_k) - Q(\lambda_k)}{A(\lambda_k)} \right|, \tag{14}$$

where $A(\lambda_k)$ are the analytical results for either $a(\lambda_k)$ or $b(\lambda_k)$ for the Kerr medium shown in Table II, $Q(\lambda_k)$ represents the average of M numerical results obtained by CS for $a(\lambda_k)$ and $b(\lambda_k)$, and N is the total number of λ_k's computed. A standard deviation $\sigma_i = \sqrt{\frac{1}{M-1} \sum_{p=1}^{M} [A(\lambda) - Q(\lambda_k)]^2}$ is calculated for each λ_k and the total standard deviation is $\sigma = \frac{1}{N}\sqrt{\sum_i^M \sigma_i^2}$. For the fundamental mode, i.e., $m = 0$, we compute a and b for $N = 20$ different values of $\lambda \in [0.2, 2]$. This was performed ten times for each λ value ($M = 10$) and a visual assessment of the results is shown in Fig. 3.

CS is able to compute the amplitude a and width b parameters necessary to approximate the Gaussian [see Fig. 3(a)] and



(a) (b)

FIG. 3. Analytical and numerical solutions for parameters a and b for Eqs. (12) (a) and (13) (b) (Gaussian and sech beams, respectively) for the fundamental mode ($m = 0$) in the Kerr medium.

hyperbolic secant [see Fig. 3(b)] functions on optical solitary waves in a satisfactory way. In order to quantify the approximation, the corresponding ϵ_A and σ are shown in Table III. CS is a robust algorithm that has a high accuracy due to a low ϵ_A and high precision due to a low variance, as shown in Table III. The low variance describes the high stability of convergence of the CS algorithm while the ϵ_A indicates that the code delivers a satisfactory answer. These results constitute a positive criterion for the performance of the CS algorithm and motivate its use to explore more complex nonlinear media. Since the CS is a stochastic algorithm, it is convenient to make several runs of the computational algorithm to observe that the algorithm is converging to an optimum value.

B. Lévy flights vs Brownian walk

In the previous section we showed that the CS algorithm is able to obtain solitary-wave-like profiles. We conjecture that an important part of the CS algorithm's efficiency lies in the Lévy flights' lack of a characteristic length scale [35]. In order to study this proposition, we proceed to compare the performance of CS using a Lévy walk (as originally formulated) with CS using a Brownian walk. For a Lévy walk we use parameter $\beta = 3/2$ as stated in Sec. II B. By changing the parameter to $\beta = 2$ we produce a Brownian walk [35]. The results of the comparison between a Brownian walk ($\beta = 2$) and a Lévy walk ($\beta = 3/2$) for the variational parameter a and trial function Eq. (11) when $m = 0$ in a Kerr medium are shown in Table IV.

As seen in Table IV when a Lévy walk begins it is rather unstable. There is a larger error and variance in the approximation of the variational values. This behavior is due to the aggressive steps of the Lévy flight. Comparatively, the Brownian walk is characterized by more subtle steps, as it is shown in Fig. 1. However, as the algorithm continues to iterate, the Lévy walks converge to more accurate results than the Brownian walks. Lévy walks are able to avoid local minima by exploring the search space more efficiently. Consequently they obtain better results than their Brownian counterpart.

TABLE III. Statistical results for the variational parameters a and b for $m = 0$ to Eqs. (12) and (13) in the Kerr medium. Error ϵ_A is measured according to Eq. (14). The standard deviation σ is also reported.

Trial function		$a \times 10^{-4}$	$b \times 10^{-4}$
Gaussian	ϵ_A	2.6	6.8
	σ	3.1	13.5
Hyperbolic	ϵ_A	13.9	4.50
	σ	2.4	5.9

C. Saturation medium

We proceed to compute solitary wave solutions in a more realistic medium in which saturation effects are considerable. We set the nonlinear term in Eq. (4) to the saturation term described in Table I. To the best of our knowledge, there is no previous report of analytical results for the variational approach.

In Fig. 4 we show the results by using a CS for the parameters a and b for Eq. (11) when $m = 1$ and saturation values of $s = 0.05$ and 0.2. Note that this scenario corresponds to

TABLE IV. Statistical results for the optimization of the variational parameter a of the Gaussian trial function [Eq. (11)] when $m = 0$ in the Kerr medium for $\lambda = 1$ using two different random walks: a Brownian and a Lévy walk. The results are shown for different numbers of iterations (I) and each were compared to the analytic solutions (shown in Table II). The average error is calculated according to Eq. (14) ($N = 10$) and the standard deviation σ is also reported.

Type of walk		25 I	75 I
Lévy	ϵ_A	0.312	0.009
	σ	0.555	0.008
Brownian	ϵ_A	0.061	0.012
	σ	0.053	0.013

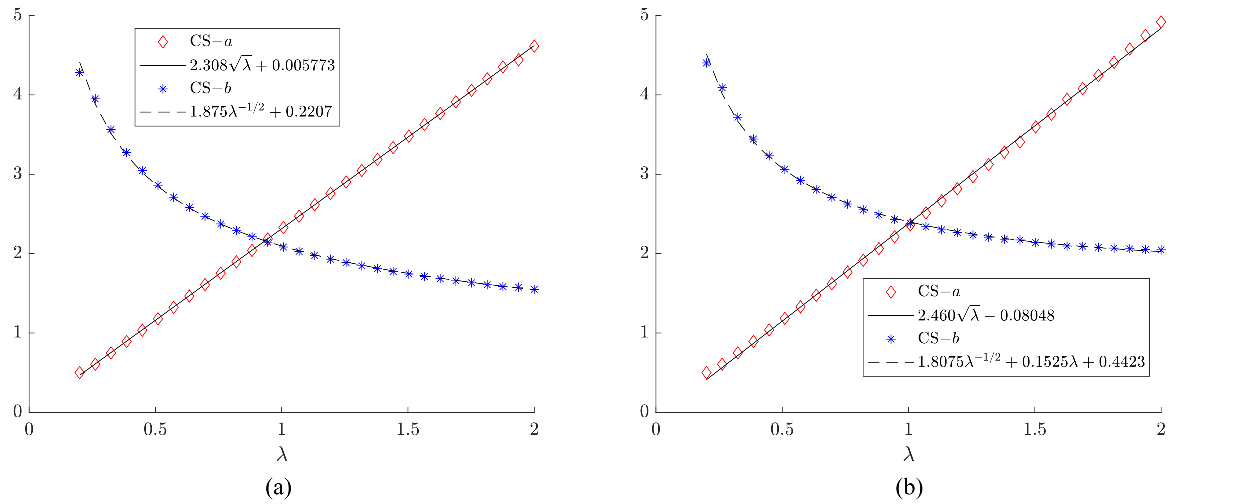


FIG. 4. Numerical solutions for optical dipole solitary waves with parameters a and b for Eq. (11) when $m = 1$ in the saturation medium (see Table I) for $s = 0.05$ (a) and $s = 0.2$ (b). A fitted curve for each parameter is also plotted. $R^2 = 0.9999$ for a and $R^2 = 0.9982$ for b in (a) and $R^2 = 0.9992$ for a and $R^2 = 0.9977$ for b in (b).

optimizing the parameters for a dipole optical solitary wave. Additionally, we fit a function for the solutions previously shown in Table II. This proposed form of the solution is based on the actual solutions for the Gaussian trial function in a Kerr medium for the dipole solitary wave, i.e., $m = 1$ [26]. The general form of the proposed function is

$$A(\lambda) = C_1\lambda^{-1/2} + C_2\lambda + C_3, \tag{15}$$

where C_1, C_2, and C_3 are the coefficients found with a standard function fitting algorithm. These curves were fitted with the set of a's and b's previously found by CS for 30 equidistant λ's $\in [0.2, 2]$. In this case, C_1 was set to zero for the a's since for the Kerr solutions they are known to be linear with respect to λ.

We compute the variational parameters for two saturation values and we find that for the amplitude parameter a, C_2 is 2.308 and 2.460, while C_3 is 0.0058 and 0.0805 for saturation values of $s = 0.05$ and 0.2, respectively. For the width parameter b with a saturation of $s = 0.05$, the effect of the linear term C_2 is found to be negligible, while C_1 and C_3 are 1.875 and 1.8075, respectively. However, for a saturation of $s = 0.2$, the coefficient C_2 is no longer negligible and C_1, C_2, and C_3 are 1.8075, 0.1525, and 0.4423, respectively. Therefore, this procedure shows that CS can also be useful to compute analytical expressions for solitary wave approximations like the one explored in Eq. (15).

To quantify the solitary wave approximation, we compute the coefficient of determination, which is given by

$$R^2 = 1 - \frac{\sum_{k=1}^{n}(\eta_k - \phi_k)^2}{\sum_{k=1}^{n}(\eta_k - \bar{\eta})^2}, \tag{16}$$

where η_k are the numerical results of the variational parameters numerically obtained, ϕ_k is the fitted data [with the form of Eq. (15)], $\bar{\eta}$ is the average of the set $\{\eta_1, \eta_2, \ldots, \eta_n\}$, and n is the sample size. The R^2 values are 0.9999 and 0.9982 for a and b, respectively, for $s = 0.05$. Similarly, the R^2 values are 0.9992 and 0.9977 for a and b, respectively, for $s = 0.2$. Thus, $R^2 \to 1$ shows that the function described by Eq. (15) closely fits the data. For low saturation [see Fig. 4(a)] there is no linear term for λ in b, while for higher saturation value [see Fig. 4(b)] there is one. Therefore, a saturated medium can be studied as a perturbation of the Kerr medium.

Since there are no known analytical solutions to compare our results, we use the generated optical profiles with CS as initial conditions to study their optical evolution as self-trapped structures. Therefore, we corroborate the approximate solitary wave solutions obtained by propagating them using a standard pseudospectral technique and we show the corresponding propagation dynamics in Figs. 5(a)–5(c) for

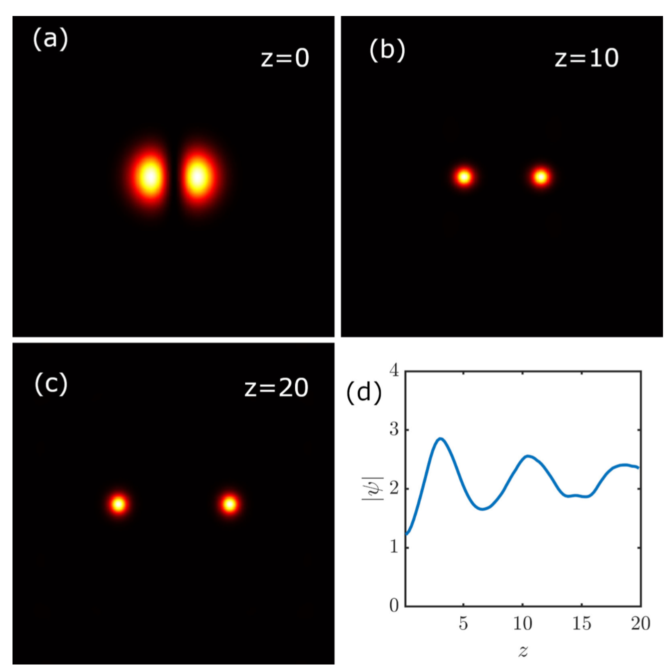


FIG. 5. (a)–(c) Propagation dynamics for an optical dipole solitary wave with parameters $a = 0.69$ and $b = 4.09$ for Eq. (11) when $m = 1$ and $\lambda = 0.25$ in a saturation medium (see Table I) for a transverse display area of 40. (d) Evolution of the peak amplitude.

a saturation medium with $s = 0.2$. We report that despite showing oscillations in the peak amplitude [see Fig. 5(d)], initially the solitary wave remains self-trapped. Also, because the intensity-dependent refractive index between the two initial humps has a minimum value resulting from the destructive interference of the initial optical field, both humps feel a repulsive force caused by the nonlinear term. Therefore, this initial dipole soliton eventually decays into two fundamental solitary waves as is expected in any realistic local media.

A problem beyond the scope of the present paper that should be explored is the stabilization of the generated structures. Physical mechanisms such as nonlocal effects, competing nonlinearities, or particular boundary conditions should be studied in order to achieve stability. Remarkably, the initial light structure profiles generated by the CS algorithm show propagation dynamics where the initial profiles generated remain self-trapped for several diffraction lengths. This is contrary to the most common scenarios where either the initial profile is completely distorted as a result of the diffraction or the initial profile experiences catastrophic collapse. Thus, the numerical results reported show that CS can be useful to generate approximate solitary wave solutions in nonlinear systems.

D. More complex media

In this section, we report results of the CS algorithm for a more complex medium generated by a combination of saturation effects and an optical lattice described by a nth-order Bessel function (see Table I) that modulates the refractive index. A varying refractive index in the medium drastically changes the propagation, allowing new dynamics and localizations of optical beams. This modulated refractive index can be experimentally induced by imprinting an optical pattern onto a photorefractive crystal (see, for example, [28]). The parameters were chosen as follows: lattice depth $p = 15$, saturation parameter $s = 0.2$, transverse scale of the Bessel beam $k = 3$, and the corresponding Bessel order $n = 5$. Note that J_n in Table I stands for a nth-order Bessel function of the first kind.

From a physical point of view, Bessel beams are interesting because of their nondiffracting behavior. This is due to the superposition of plane waves with wave number k that have the same inclination angle with respect to the propagation axis. The nondiffracting properties make the Bessel beams quite useful in applications such as the manipulation of biological and colloidal material or soliton routing and steering [50,51].

For the simulation, we use the trial function described by Eq. (12) that corresponds to optical solitary wave vortices. For these optical structures, the energy flow rotates around the vortex core at the center, and the velocity of energy flow is infinite at this point and thus the optical intensity must vanish. The results for different λ's and different topological charges (i.e., different m values) are shown in Table V. Note that the topological charge, or dislocation strength, can also be defined by the circulation of the phase gradient around the singularity located at the center of the optical field.

Similar to the case of the saturation medium, here there are no analytical solutions to compare these CS results with, therefore we perform a numerical propagation to study the evolution of the initial profile. The results of the propagation are shown in Fig. 6. In Figs. 6(a)–6(c) it can be observed that the initial profile is partially conserved after 50 propagation units. In Fig. 6(d) we observe that the peak amplitude of the propagated beam oscillates, but the initial distribution of the energy of the beam remains self-preserved during the propagation. This confirms again that CS is correctly generating approximate solitary wave solutions in complex nonintegrable media. Contrary to the dipole soliton case, here there is nonzero angular momentum associated with the spatial structure of the beam, resulting in the rotation of these self-trapped structures during propagation.

Similarly, we compute vortex solitary waves with topological charge of $m = 2$ and parameters $p = 15$, $s = 0.2$, $k = 3$, $n = 5$ with the trial function Eq. (12) for different λ's. We show some values obtained in Table VI. The initial profile is shown in Fig. 7(a). In comparison to the single topological charge shown in Fig. 6, here the optical vortex solitary wave also oscillates during propagation; however, the initial circular shape is more deformed [see Figs. 7(b) and 7(c)]. Eventually the initial ring-shaped vortex decays into fundamental modes

TABLE V. Numerical results for a combination of media (saturation and nth-order Bessel) for the trial function 12 with $m = 1$ and parameters $p = 15$, $s = 0.2$, $k = 3$, and $n = 5$.

λ	2	3	4
a	10.52	6.17	7.01
b	1.68	2.98	4.62

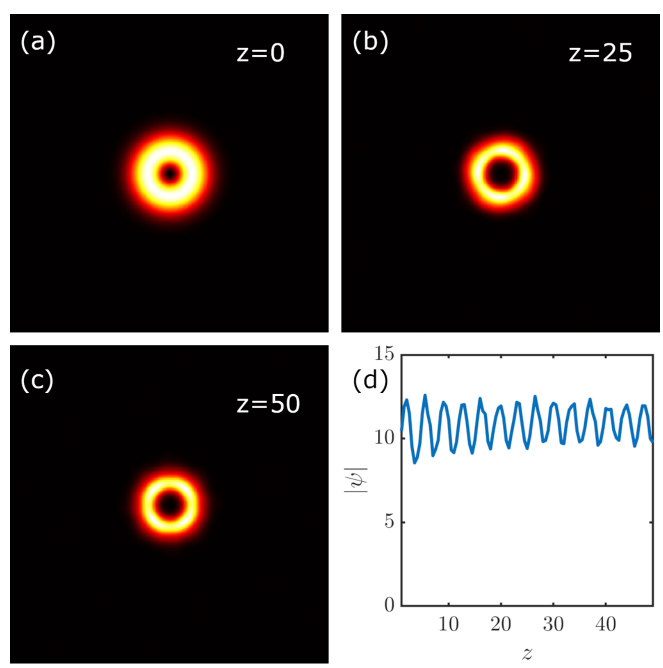


FIG. 6. (a)–(c) Propagation dynamics of a single vortex solitary wave done in a combination of a saturation medium with $s = 0.2$ and an optical Bessel lattice with parameters $p = 15$, $n = 5$, and $k = 3$ (see Table I) in a transverse display area of 30. The initial state generated with CS obtained parameters $a = 6.17$ and $b = 2.98$. (d) Evolution of the peak amplitude.

TABLE VI. Numerical results for a combination of media saturation and nth-order Bessel (see Table I) for the trial function 12 with $m = 2$ and parameters $p = 15$, $s = 0.2$, $k = 3$, and $n = 5$.

λ	2	3	4
a	1.04	1.59	1.54
b	3.18	3.34	4.48

because of transverse instabilities. These modes fly off the initial configuration along tangential trajectories due to the conservation of angular momentum. Therefore, the oscillation in the peak amplitude is maintained for up to 80 propagation units and the vortex solitary wave decays into modes, each with zero angular momentum, as expected.

The results of the CS algorithm for self-trapped structures depend strongly on the trial function proposed. For most cases, the initial optical profile launched will experience breathinglike dynamics, contrary to what would result from more deterministic algorithms such as Newton-Raphson, where a closer profile to the exact invariant optical structure might be obtained. CS allows us to study more general shapes of self-trapped optical structures and avoid limitations from deterministic approaches when looking for optical profiles with certain symmetries.

E. Scaling the CS algorithm

We increase the dimensionality of the parameters to show the robustness of the CS algorithm by optimizing three beams,

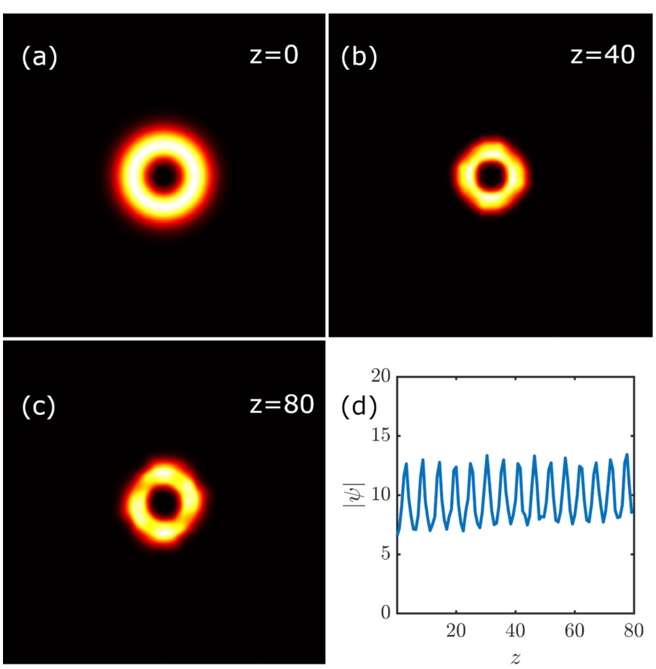


FIG. 7. (a)–(c) Propagation dynamics of a solitary wave with topological charge $m = 2$ done in a combination of a saturation medium with $s = 0.2$ and an optical Bessel lattice with parameters $p = 15$, $n = 5$, and $k = 3$ in a transverse display area of 35. (a) Initial state generated with CS obtaining parameters $a = 1.59$ and $b = 3.34$. (d) Evolution of the peak amplitude.

TABLE VII. Numerical results for the solitonic system made by three trial functions [Eq. (12) for $m = 0, 1, 2$, $\lambda = 2$] in the Kerr medium. The corresponding six parameters are obtained at the same time with a single run of CS.

	$m = 0$	$m = 1$	$m = 2$
a	2.83	3.97	2.17
Error (%)	0.01	0.73	0.47
b	0.99	1.42	1.73
Error (%)	0.03	0.45	0.17

and therefore six parameters, simultaneously. Although the problem is still nonlinear, the superposition principle can be used since each beam is far enough away to not interfere with the domain of the others. We can take this interference to be null. Therefore, we expect that solving these three beams simultaneously will closely match with the results of solving for each beam individually. Table VII shows the CS results and their respective error according to the analytical results [26] for optimizing Eq. (12), $m = 0, 1, 2$ and $\lambda = 2$ in the Kerr medium. This result demonstrates the potential of the algorithm to optimize more complex test functions with multiple variational parameters. Further results using the CS algorithm for a large number of variational parameters could be analyzed by using cloud computing, graphics processing units, or parallel computing techniques that are naturally supported by CS. However, the primary objective of this paper is to introduce the CS algorithm and to demonstrate its basic and general use.

We perform a similar procedure for optimizing six parameters, but now for a saturation medium with $s = 0.05$ and $\lambda = 1$, and we use the combination of three spatially modulated vortex solitary waves or azimuthons described by [52] as a trial function:

$$U = \sum_{j=1}^{3} a^{(j)} \exp\left[r_{\perp,j}^2/(b^{(j)})^2\right] r_{\perp,j}^2 [\cos(\Theta_j)\cos(2\theta_j) + \sin(\Theta_j)\sin(2\theta_j)], \tag{17}$$

where $r_{\perp,j}$ and θ_j are the same as in Eqs. (11)–(13), but displaced from the origin to avoid overlap between the beams. $\Theta_{1,2,3}$ are 20°, 30°, and 40°, respectively, and show different modulations in the azimuthons. To obtain these results (Table VIII) an important adjustment is needed to make the algorithm converge. The algorithm searches for the optimal values of the variational parameters within a given interval (this interval could be chosen individually for each parameter); however, if the lower limit for the parameter modulating the amplitude for a is not big enough, the algorithm could

TABLE VIII. Numerical results for azimuthons constructed with three trial functions [Eq. (17), $\lambda = 1$] in the saturation medium for $s = 0.05$. The corresponding six parameters are obtained at the same time with a single run of CS.

	$j = 1$	$j = 2$	$j = 3$
a^j	0.91	0.96	0.96
b^j	2.56	2.59	2.61

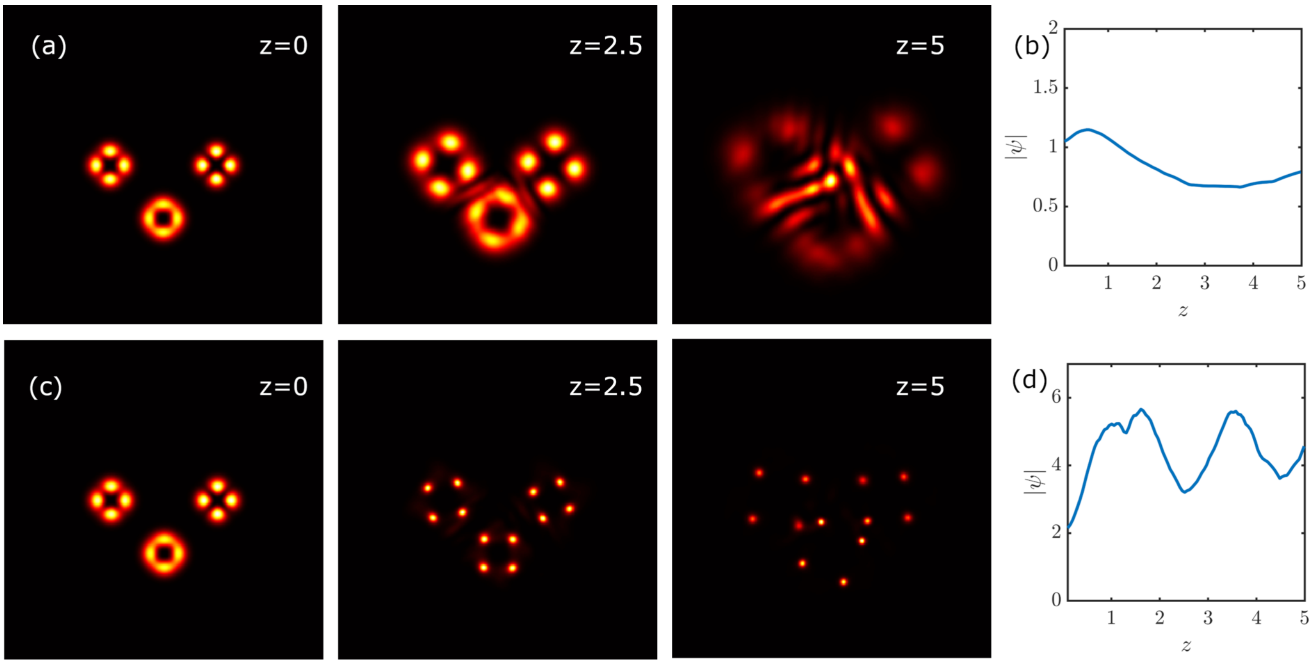


FIG. 8. Propagation dynamics of three azimuthons with charge $m = 2$. (a) Propagation dynamics without optimized parameters. Here the corresponding amplitude values are only half of the corresponding value obtained by CS. (b) The peak amplitude evolution for the three interacting beams. (c) Propagation dynamics with optimized parameters obtained by a single CS. (d) The peak amplitude evolution for the three interacting beams. The propagations were done in a transverse display area $L = 60$ and in a saturation medium with $s = 0.05$.

converge to the trivial solution $U = 0$. For most of the results obtained here, the lower limit is chosen to be 0.65, though a lower limit of 0.9 can be chosen for the $a^{(j)}$ parameters to avoid the trivial solution, as shown in Table VIII.

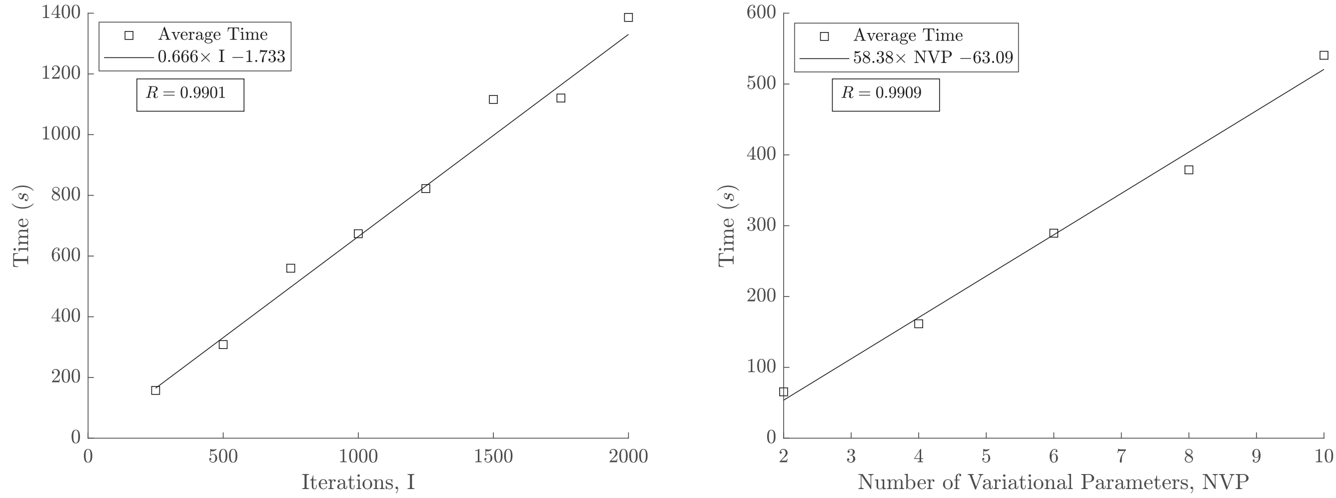
In order to verify the self-trapped nature of these solutions, we propagate the approximate solitonic system by using a standard pseudospectral technique as before. For a general combination of the parameters a and b, the azimuthons experience stronger deformations depicted in Fig. 8(a), where after only five units of propagation, the initial configuration is completely lost, despite the peak amplitude not changing drastically [see Fig. 8(b)]. However, using the parameters obtained by CS in the propagation makes the configuration more robust, as depicted in Fig. 8(c), though the peak amplitude oscillates at a higher percentage than for the past scenario [see Fig. 8(d)]. Each of the three azimuthons decays finally into four fundamental modes due to instability issues, as expected. Similar studies that use different stabilization mechanisms such as nonlocal media [53] could improve the propagation of this kind of complex solitonic system. Note that these azimuthons can be characterized by two integer numbers: the topological charge m and the number of the intensity peaks N_p. For this scenario we set $m = 2$ and $N_p = 4$, fulfilling the condition $N_p \geqslant 2m$ that is necessary for observing rotation azimuthons in saturated media [52]. From a physical point of view, the corresponding rotation of these optical structures has two different components. The first contribution comes from their particular phase distribution and has a wavelike origin, while the second contribution is due to the azimuthal modulation of the intensity of the optical field which can be associated more with particlelike behavior.

IV. GENERAL ALGORITHM BEHAVIOR

The CS algorithm is tested to analyze the order of growth of running time with respect to the total iterations. In Fig. 9(a) the results for $m = 0$ and $\lambda = 1$ are shown. The algorithm shows an apparent linear relationship between time and the total number of iterations, and is confirmed with $R^2 = 0.9901$. Evidently, the fitted equation shown in Fig. 10(a) can change according to the computational power, but estimations of the total amount of time for N iterations could be done just by running a few iterations and extrapolating these results.

In Fig. 9(b) we show the dependence of time vs the total number of parameters to adjust (the dimension of the problem). There is a quasilinear behavior with $R^2 = 0.9909$. A crucial remark is that these results are obtained by leaving the number of iterations constant. However, if the dimensions of the problem are increased, more iterations could be needed to converge to an accurate result. Therefore, as we increase the total number of parameters, the total order of growth of running time will be nonlinear.

We depict the cost function vs iteration behavior for three different trial functions [Eq. (11) with $m = 0, 1, 2$] in Fig. 10. Figures 10(a) and 10(b) have a x-log and y-log axis, respectively. Note that for the case of $m = 0$ the line stops before the iteration 1500 because the cost function reaches zero, which has no representation in a log scale. As m in Eq. (11) increases, the trial function has a more complicated shape, making the cost function converge slower. Therefore, it is natural to expect different values for the cost functions corresponding to different trial functions. In general, the more complex the trial



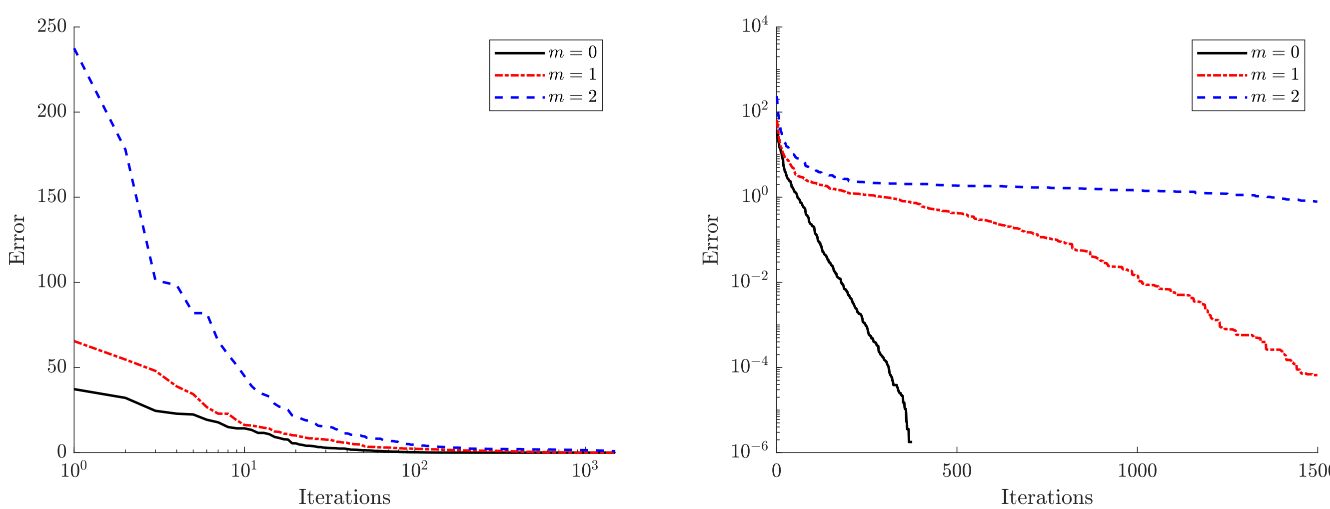
(a) (b)

FIG. 9. (a) Iterations vs time. The plot shows a linear dependence with $R^2 = 0.9901$. (b) Number of variational parameters vs time. The plot shows a linear dependence with $R^2 = 0.9909$. In this case the number of iterations is fixed to 5.

function, the slower the algorithm will converge to the lowest bound. Furthermore, this limit value could vary depending on the trial function used.

V. CONCLUSIONS

In this paper, we report a robust stochastic method to find approximate optical spatial solitary waves in nonlinear media based on Lévy flights. We have developed a metaheuristic algorithm, based on the bio-inspired cuckoo search algorithm, that is able to overcome the cumbersome mathematical treatment of a trial function required in the variational approach for generating approximate solitary wave solutions. The algorithm is scalable in the sense that it grows linearly in time with increased iterations and dimensions. However, increasing the number of variables will require more iterations, making the computational time not linear. The CS algorithm successfully obtained solitary wave solutions for different saturation and media, namely, Kerr and more complex media, including a combination of the saturation effects and optical Bessel lattices. For the Kerr medium the obtained solutions were compared with the analytical results. The absolute errors for this medium were obtained and the variances for ten runs were on the order of $\approx 10^{-4}$. The results show that the algorithm has a highly converging accuracy and precision. The algorithm has been tested for even more complex media. Due to the lack of previous analytical solutions to compare these results with, numerical propagations were performed to study the evolution of the initial profiles. Remarkably, the diffraction and self-focusing effects are balanced, showing that the CS algorithm can successfully be used to obtain approximate solitary wave solutions even for more complex nonlinear systems. Furthermore, the algorithm has been able to optimize solitary wave systems with several dimensional variational



(a) (b)

FIG. 10. Average cost function iteration evolution for Eq. (11) with $m = 0, 1, 2$ and $N = 20$ in the x-log axis (a) and in the y-log axis (b).

parameters. This shows the relevance of applying the CS method to more complex trial functions and using it as an efficient optimization tool for designing optical systems based on the exploration of more complex self-trapped optical structures. Finally, note that modifying the cost value function described in this paper by Eq. (5) could result in even more specific and efficient stochastic algorithms for solving particular optimization problems that arise from many other nonlinear physical systems.

ACKNOWLEDGMENT

We would like to thank Zerrin Vural for her helpful comments while writing this paper.

[1] N. J. Zabusky and M. D. Kruskal, Phys. Rev. Lett. **15**, 240 (1965).
[2] A. Katti, R. Yadav, and A. Prasad, Wave Motion **77**, 64 (2018).
[3] S. He, Z. Zhang, H. Liu, S. Akhmadaliev, S. Zhou, X. Wang, and P. Wu, Appl. Phys. Express **12**, 076502 (2019).
[4] S. Konar and V. A. Trofimov, Pramana **85**, 975 (2015).
[5] S. Perumbilavil, A. Piccardi, R. Barboza, O. Buchnev, M. Kauranen, G. Strangi, and G. Assanto, Nat. Commun. **9**, 3863 (2018).
[6] C.-W. Lee, P. Kurzyński, and H. Nha, Phys. Rev. A **92**, 052336 (2015).
[7] M. Li, T. Xu, and D. Meng, J. Phys. Soc. Jpn. **85**, 124001 (2016).
[8] A. Paredes, D. N. Olivieri, and H. Michinel, Physica D **403**, 132301 (2020).
[9] E. J. Meier, F. A. An, and B. Gadway, Nat. Commun. **7**, 13986 (2016).
[10] U. A. Laudyn, M. Kwaśny, F. A. Sala, M. A. Karpierz, N. F. Smyth, and G. Assanto, Sci. Rep. **7**, 12385 (2017).
[11] Z. Chen, M. Segev, and D. N. Christodoulides, Rep. Prog. Phys. **75**, 086401 (2012).
[12] M. J. Ablowitz and H. Segur, *Solitons and the Inverse Scattering Transform* (SIAM, Philadelphia, 1981).
[13] X. Geng and H.-W. Tam, J. Phys. Soc. Jpn. **68**, 1508 (1999).
[14] V. I. Kruglov, A. C. Peacock, and J. D. Harvey, Phys. Rev. Lett. **90**, 113902 (2003).
[15] J.-F. Zhang, C.-Q. Dai, Q. Yang, and J.-M. Zhu, Opt. Commun. **252**, 408 (2005).
[16] N. Z. Petrović, M. Belić, W.-P. Zhong, R.-H. Xie, and G. Chen, Opt. Lett. **34**, 1609 (2009).
[17] H.-G. Beyer and B. Sendhoff, Comput. Methods Appl. Mech. Eng. **196**, 3190 (2007).
[18] Q. P. Zheng, J. Wang, and A. L. Liu, IEEE Trans. Power Syst. **30**, 1913 (2014).
[19] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, Science **220**, 671 (1983).
[20] D. B. Hitchcock, Am. Stat. **57**, 254 (2003).
[21] S. J. Nanda and G. Panda, Swarm Evolut. Computat. **16**, 1 (2014).
[22] D. Anderson, M. Lisak, and A. Berntson, Pramana **57**, 917 (2001).
[23] X.-S. Yang, Cuckoo search (cs) algorithm, https://la.mathworks.com/matlabcentral/fileexchange/29809-cuckoo-search-cs-algorithm (2013).
[24] Y. V. Kartashov, V. A. Vysloukh, and L. Torner, Prog. Opt. **52**, 63 (2009).
[25] Q. Guo, B. Luo, and S. Chi, Opt. Commun. **259**, 336 (2006).
[26] E. I. Duque, S. Lopez-Aguayo, and B. A. Malomed, Opt. Express **26**, 7451 (2018).
[27] G. C. Valley, M. Segev, B. Crosignani, A. Yariv, M. M. Fejer, and M. C. Bashaw, Phys. Rev. A **50**, R4457 (1994).
[28] R. Fischer, D. N. Neshev, S. Lopez-Aguayo, A. S. Desyatnikov, A. A. Sukhorukov, W. Krolikowski, and Y. S. Kivshar, Opt. Express **14**, 2825 (2006).
[29] X.-S. Yang and S. Deb, in *Proceedings of the 2009 World Congress on Nature and Biologically Inspired Computing (NaBIC)* (IEEE, New York, 2009), pp. 210–214.
[30] R. B. Payne and M. D. Sorensen, *The Cuckoos* (Oxford University Press, Oxford, UK, 2005), Vol. 15.
[31] G. M. Viswanathan, S. V. Buldyrev, S. Havlin, M. Da Luz, E. Raposo, and H. E. Stanley, Nature (London) **401**, 911 (1999).
[32] N. Levernier, J. Textor, O. Bénichou, and R. Voituriez, Phys. Rev. Lett. **124**, 080601 (2020).
[33] B. Guinard and A. Korman, Sci. Adv. **7**, eabe8211 (2021).
[34] R. N. Mantegna and H. E. Stanley, Phys. Rev. Lett. **73**, 2946 (1994).
[35] X.-S. Yang, *Nature-Inspired Metaheuristic Algorithms*, 2nd ed. (Luniver Press, Frome, 2010).
[36] M. Gutowski, arXiv:math-ph/0106003.
[37] R. N. Mantegna, Phys. Rev. E **49**, 4677 (1994).
[38] M. Leccardi, in *Proceedings of the Fifth EUROMECH Nonlinear Dynamics Conference* (Citeseer, Italy, 2005).
[39] X.-S. Yang and S. Deb, arXiv:1005.2908.
[40] A. F. Kamaruzaman, A. M. Zain, S. M. Yusuf, and A. Udin, in *Information Technology for Manufacturing Systems IV*, Applied Mechanics and Materials Vol. 421 (Trans Tech Publications, Switzerland, 2013), pp. 496–501.
[41] J. Bouchaud, in *Lévy Flights and Related Topics in Physics* (Springer, New York, 1995), pp. 237–250.
[42] Y. Ling, Y. Zhou, and Q. Luo, IEEE Access **5**, 6168 (2017).
[43] H. Haklı and H. Uğuz, Appl. Soft Comput. **23**, 333 (2014).
[44] I. I. Shevchenko, Phys. Rev. E **81**, 066216 (2010).
[45] S. Singh, Y. Chu, M. Lukin, and S. Yelin, in *Advances in Atomic, Molecular, and Optical Physics* (Elsevier, Amsterdam, 2015), Vol. 64, pp. 273–327.
[46] S. Bulanov, T. Z. Esirkepov, J. Koga, S. Bulanov, Z. Gong, X. Yan, and M. Kando, J. Plasma Phys. **83**, 905830202 (2017).
[47] B. Dybiec, E. Gudowska-Nowak, E. Barkai, and A. A. Dubkov, Phys. Rev. E **95**, 052102 (2017).
[48] D. Anderson, M. Bonnedal, and M. Lisak, Phys. Fluids **22**, 1838 (1979).
[49] D. Buccoliero, A. S. Desyatnikov, W. Krolikowski, and Y. S. Kivshar, Phys. Rev. Lett. **98**, 053901 (2007).
[50] J. Durnin, J. Opt. Soc. Am. A **4**, 651 (1987).
[51] D. McGloin and K. Dholakia, Contemp. Phys. **46**, 15 (2005).
[52] A. S. Desyatnikov, A. A. Sukhorukov, and Y. S. Kivshar, Phys. Rev. Lett. **95**, 203904 (2005).
[53] E. Garza, S. Lopez-Aguayo, and J. C. Gutiérrez-Vega, Phys. Rev. A **99**, 033804 (2019).