



**HAL**  
open science

# Testing the Instanton Approach to the Large Amplification Limit of a Diffraction-Amplification Problem

Philippe Mounaix

► **To cite this version:**

Philippe Mounaix. Testing the Instanton Approach to the Large Amplification Limit of a Diffraction-Amplification Problem. 2024. hal-04737082

**HAL Id: hal-04737082**

**<https://hal.science/hal-04737082v1>**

Preprint submitted on 15 Oct 2024

**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.

# Testing the Instanton Approach to the Large Amplification Limit of a Diffraction-Amplification Problem

Philippe Mounaix<sup>1,\*</sup>

<sup>1</sup>*CPHT, CNRS, École polytechnique,  
Institut Polytechnique de Paris, 91120 Palaiseau, France.*

(Dated: October 15, 2024)

## Abstract

The validity of the instanton analysis approach is tested numerically in the case of the diffraction-amplification problem  $\partial_z \psi - \frac{i}{2m} \partial_{x^2}^2 \psi = g|S|^2 \psi$  for  $\ln U \gg 1$ , where  $U = |\psi(0, L)|^2$ . Here,  $S(x, z)$  is a complex Gaussian random field,  $z$  and  $x$  respectively are the axial and transverse coordinates, with  $0 \leq z \leq L$ , and both  $m \neq 0$  and  $g > 0$  are real parameters. We consider a class of  $S$ , called the ‘one-max class’, for which we devise a specific biased sampling procedure. As an application,  $p(U)$ , the probability distribution of  $U$ , is obtained down to values less than  $10^{-2270}$  in the far right tail. We find that the agreement of our numerical results with the instanton analysis predictions in Mounaix (2023 *J. Phys. A: Math. Theor.* **56** 305001) is remarkable. Both the predicted algebraic tail of  $p(U)$  and concentration of the realizations of  $S$  onto the leading instanton are clearly confirmed, which validates the instanton analysis numerically in the large  $\ln U$  limit for  $S$  in the one-max class.

Keywords: stochastic partial differential equations, instanton analysis, extreme event statistics

---

\*Electronic address: [philippe.mounaix@polytechnique.edu](mailto:philippe.mounaix@polytechnique.edu)

## I. INTRODUCTION

In a recent work [1], the instanton analysis approach [2–4] was used to determine the tail of  $p(U)$ , the probability distribution of  $U = |\psi(0, L)|^2$ , for  $\ln U \gg 1$ ,  $\psi(x, z)$  being the solution to the diffraction-amplification problem<sup>1</sup>:

$$\begin{cases} \partial_z \psi(x, z) - \frac{i}{2m} \partial_x^2 \psi(x, z) = g |S(x, z)|^2 \psi(x, z), \\ 0 \leq z \leq L, \quad x \in \Lambda \subset \mathbb{R}, \quad \text{and } \psi(x, 0) = 1. \end{cases} \quad (1)$$

Here,  $z$  and  $x$  respectively denote the axial and transverse coordinates in a domain of length  $L$  and (one-dimensional) cross-section  $\Lambda$ . For technical convenience, we will take for  $\Lambda$  the circle of length  $\ell$ . The boundary condition at  $z = 0$  is taken to be a constant for simplicity. Both  $m \neq 0$  and  $g > 0$  are real parameters and  $S(x, z)$  is a transversally homogeneous complex Gaussian noise with zero mean and normalization  $L^{-1} \int_0^L \langle |S(x, z)|^2 \rangle dz = 1$ .

From the instanton analysis of the corresponding Martin-Siggia-Rose action, it was found in [1] that  $S$  concentrates onto long filamentary instantons,  $S_{\text{inst}}$ , as  $\ln U \rightarrow +\infty$ . These filamentary instantons run along specific non-random paths, denoted by  $x_{\text{inst}}(\cdot)$ , that maximize the largest eigenvalue  $\mu_1[x(\cdot)]$  of the covariance operator  $T_{x(\cdot)}$  defined by

$$(T_{x(\cdot)} f)(z) = \int_0^L C(x(z) - x(z'), z, z') f(z') dz', \quad f(z) \in L^2([0, L]), \quad (2)$$

where  $C(x - x', z, z') = \langle S(x, z) S(x', z')^* \rangle$ . In equation (2),  $x(\cdot)$  is a continuous path in  $\Lambda$  satisfying  $x(L) = 0$ , and for the class of  $S$  considered in [1], every maximizing path  $x_{\text{inst}}(\cdot)$  is also continuous with  $x_{\text{inst}}(L) = 0$  (see [1] for details). In the most common case of ‘single-filament instantons’ for which there is only one maximizing path, and assuming a non-degenerate  $\mu_1[x_{\text{inst}}(\cdot)]$ , one has  $S(x, z) \sim S_{\text{inst}}(x, z)$  ( $\ln U \rightarrow +\infty$ ) with

$$S_{\text{inst}}(x, z) = \frac{c_1}{\mu_{\text{max}}} \int_0^L C(x - x_{\text{inst}}(z'), z, z') \phi_1(z') dz', \quad (3)$$

where  $\mu_{\text{max}}$  is short for  $\mu_1[x_{\text{inst}}(\cdot)]$ ,  $\phi_1$  is the fundamental eigenfunction of  $T_{x_{\text{inst}}(\cdot)}$ , and  $c_1$  is a complex Gaussian random variable with  $\langle c_1 \rangle = \langle c_1^2 \rangle = 0$  and  $\langle |c_1|^2 \rangle = \mu_{\text{max}}$ . The tail of  $p(U)$  for  $\ln U \gg 1$  can then be deduced from the statistics of  $S_{\text{inst}}$  as given in equation (3).

---

<sup>1</sup> This problem is of interest in, e.g., laser-plasma interaction and nonlinear optics in which  $\psi$  is the complex time-envelope of the scattered light electric field,  $g$  and  $S$  being proportional to the average laser intensity and the complex time-envelope of the laser electric field, respectively [5].

The result is a leading algebraic tail  $\propto U^{-\zeta}$  with  $\zeta = (1 + 1/2\mu_{\max}g)$ , modulated by a slow varying amplitude (slower than algebraic) [1].

In the absence of a mathematically rigorous proof, the need for checking the validity of these analytical results numerically cannot be overlooked. To this end, it is essential to have a good sampling of the realizations of  $S$  for which  $\ln U \gg 1$ . Unfortunately, such realizations are extremely rare events, far beyond the reach of any direct sampling with a reasonable sample size. For instance, for the same Gaussian field  $S$  and parameters as in section 5 of [1] and in the simple diffraction-free limit,  $m^{-1} = 0$ , in which  $U$  can be computed exactly, it can be checked that  $p(\ln U \geq 10^3) = O(10^{-100})$ . It is thus clearly unrealistic to expect that the asymptotic analytical results can be tested by direct numerical simulations. To gain access to the asymptotic regime and check the validity of the instanton analysis we need a specific approach. One possible strategy is to bias the underlying distribution of  $S$  towards the outcomes of interest. In the case of nonlinear equations with additive noise, this has been successfully achieved by means of the ‘importance sampling algorithm’ [6] frequently used in rare event physics (see e.g. [7–9] and references therein). For the diffraction-amplification problem (1) with  $S$  admitting a single and non-degenerate instanton, like in equation (3) (see the appendix), it turns out that a different, somewhat simpler, method can be devised, based on the existence of a nonlinear fit to numerical data giving a highly accurate approximation of  $U$  as a function of  $S$ , when  $\ln U$  is large. It is then possible to determine the extreme upper tail of  $p(U)$  and the corresponding realizations of  $S$  from numerical simulations. Development of the biased sampling procedure and its application to check the validity of the instanton analysis in the case of problem (1) is the subject of the present work.

Before entering the details of the calculations, it is useful to summarize our main results.

- Let  $T_C$  denote the covariance operator defined by

$$(T_C f)(x, z) = \int_0^L \int_{\Lambda} C(x - x', z, z') f(x', z') dx' dz', \quad f(x, z) \in L^2(\Lambda \times [0, L]). \quad (4)$$

Write  $r_{\parallel}$ , and  $r_{\perp}$  the  $L^2$ -norms of the components of  $T_C^{-1/2}S$  parallel and perpendicular to any given direction in  $L^2$  close enough to the one maximizing  $U$  at fixed  $\|T_C^{-1/2}S\|_2$ , which is unique for the class of  $S$  we consider (see Section III A). By analyzing a large number of numerical solutions to equation (1), we identify the existence of an implicit equation relating  $U$ ,  $r_{\parallel}$ , and  $r_{\perp}$  when  $\ln U$  is large and all the other quantities

characterizing  $T_C^{-1/2}S$  are fixed. The reason why  $T_C^{-1/2}S$  appears rather than  $S$  will be made clear at the end of section II and above equation (16). More specifically, writing  $r_{\parallel} = \sqrt{\eta} \cos \theta$  and  $r_{\perp} = \sqrt{\eta} \sin \theta$ , with  $0 \leq \theta \leq \pi/2$  and  $\eta > 0$ , we check that our numerical data satisfy

$$\cos \theta - \left( \frac{\log_{10} U}{a\eta - b} \right)^{\alpha} = 0, \quad (5)$$

with very good accuracy for  $1820 \leq \log_{10} U \leq 2030$  and  $1000 \leq \eta \leq 1100$ , where  $a = 1.86428$ ,  $b = 25.7163$ , and  $\alpha \sim 0.5$  is a random exponent depending on the quantities characterizing  $T_C^{-1/2}S$  other than  $\theta$  and  $\eta$ .

- From this result, we derive the expression of the conditional probability distribution  $p(U, \theta | \Sigma_{\text{orv}})$  valid for  $\log_{10} U \gg 1$ , where  $\Sigma_{\text{orv}}$  stands for all the random variables characterizing  $T_C^{-1/2}S$  other than  $\theta$  and  $\eta$  ('orv' is short for 'other random variables'). For a given  $U$  with  $\log_{10} U \gg 1$ , we then obtain (i)  $p(U | \Sigma_{\text{orv}})$  by integrating  $p(U, \theta | \Sigma_{\text{orv}})$  numerically over  $0 \leq \theta \leq \pi/2$ , and (ii)  $p(\theta | U, \Sigma_{\text{orv}})$  as  $p(U, \theta | \Sigma_{\text{orv}}) / p(U | \Sigma_{\text{orv}})$ .
- We draw the realizations of  $S$  given  $U$  with  $\log_{10} U \gg 1$  according to the procedure defined by: (a) draw  $\Sigma_{\text{orv}}$ ; (b) derive  $p(\theta | U, \Sigma_{\text{orv}})$ ; (c) draw  $\theta$  from  $p(\theta | U, \Sigma_{\text{orv}})$  instead of from its unconditional probability distribution; and (d) use the resulting  $T_C^{-1/2}S$  in  $S = T_C^{1/2}(T_C^{-1/2}S)$  to get  $S$ .
- We estimate  $p(U)$  as the sample mean of  $p(U | \Sigma_{\text{orv}})$  over realizations of  $\Sigma_{\text{orv}}$  for different values of  $U$  with  $\log_{10} U \gg 1$  and we compare the results with the predictions of the instanton analysis. Let  $\mathfrak{D}$  denote the  $L^2$ -distance between  $S / \|S\|_2$  and  $S_{\text{inst}} / \|S_{\text{inst}}\|_2$ . For a given  $U$  with  $\log_{10} U \gg 1$ , we estimate  $p(\mathfrak{D} | U)$  as a sample mean over realizations of  $\Sigma_{\text{orv}}$  and  $\theta$  (see equation (23) and below). Finally, we plot  $|S|^2 / \|S\|_2^2$  for different realizations of  $S$  with particular values of  $\mathfrak{D}$  as well as the sample mean of  $|S|^2 / \|S\|_2^2$  over realizations of  $S$  with close values of  $\mathfrak{D}$ , and we compare the results with the theoretical instanton profile.

The outline of the paper is as follows. The Gaussian field  $S$  that we consider is specified in section II, where we also recall some results of [1] needed in the sequel. Section III gives preliminary numerical results used in section IV. Finally, in section IV we define our biased sampling procedure and use it to test the instanton analysis of problem (1) numerically, in the limit of a large  $\ln U$  and for the class of  $S$  considered.

## II. MODEL AND DEFINITIONS

Since the present work is the continuation of the numerical study initiated in [1], section 5, we consider the same Gaussian field  $S(x, z)$ . Namely, we take

$$S(x, z) = \sum_{n=-50}^{\prime 50} s_n \sqrt{\varsigma_n} \exp i \left[ \frac{2\pi n}{\ell} x + \left( \frac{2\pi n}{\ell} \right)^2 \frac{z}{2} \right], \quad (6)$$

where  $\sum_n'$  means  $\sum_{n \neq 0}$ . Here, the  $s_n$ s are independent and identically distributed (i.i.d.) complex Gaussian random variables with  $\langle s_n \rangle = \langle s_n s_m \rangle = 0$  and  $\langle s_n s_m^* \rangle = \delta_{nm}$ . The spectral density  $\varsigma_n$  normalized to  $\sum_{n=-50}^{\prime 50} \varsigma_n = 1$  is given by the Gaussian spectrum

$$\varsigma_n \propto \exp \left[ - \left( \frac{\pi n}{\ell} \right)^2 \right]. \quad (7)$$

Equation (6) is reminiscent of models of spatially smoothed laser beams [10], where  $S$  is a solution to the paraxial wave equation

$$\partial_z S(x, z) + \frac{i}{2} \partial_{x^2}^2 S(x, z) = 0, \quad (8)$$

with boundary condition  $S(x, 0) = \sum_{n=-50}^{\prime 50} s_n \sqrt{\varsigma_n} \exp(2i\pi n x / \ell)$ . The  $n = 0$  mode is excluded from the Fourier representation (6) to ensure that the space average  $\ell^{-1} \int_{\Lambda} S(x, z) dx$  is zero for all  $z$  and every realization of  $S$ , as expected for the electric field of a smoothed laser beam.

For each realization of  $S$  on a cylinder of length  $L = 10$  and circumference  $\ell = 20$ , we solve equation (1) by using a symmetrized  $z$ -split method [11] which propagates the diffraction term,  $(i/2m) \partial_{x^2}^2 \psi(x, z)$ , in Fourier space and the amplification term,  $g|S(x, z)|^2 \psi(x, z)$ , in real space. We take  $m = 0.7$ , and  $g = 0.5$ . For  $S$  in equation (6) with Gaussian spectrum (7) and the given values of  $L$  and  $\ell$ , it is shown in [1] that:

(i) there is a single instanton,  $S_{\text{inst}}$ , which is a single-filament instanton of the form (3) in which  $x_{\text{inst}}(\cdot) \equiv 0$ ,  $\mu_{\text{max}} = 4.34984$ , and

$$C(x, z, z') = C(x, z - z') = \sum_{n=-50}^{\prime 50} \varsigma_n \exp i \left[ \frac{2\pi n}{\ell} x + \left( \frac{2\pi n}{\ell} \right)^2 \frac{z - z'}{2} \right]; \quad (9)$$

(ii) the convolution representation of  $S_{\text{inst}}$  in equation (3) with  $x_{\text{inst}}(\cdot) \equiv 0$  and  $C$  in equation (9) is equivalent to the Fourier representation

$$S_{\text{inst}}(x, z) = \sum_{n=-50}^{\prime 50} s_n \sqrt{\varsigma_n} \exp i \left[ \frac{2\pi n}{\ell} x + \left( \frac{2\pi n}{\ell} \right)^2 \frac{z}{2} \right], \quad (10)$$

where  $\mathbf{s}$  (with components  $\mathbf{s}_n$ ) is an eigenvector of the  $100 \times 100$  positive definite Hermitian matrix

$$M_{nm} = \sqrt{\varsigma_n \varsigma_m} \int_0^L \exp i \left[ \left( \frac{2\pi}{\ell} \right)^2 (m^2 - n^2) \frac{z}{2} \right] dz, \quad (n, m \neq 0), \quad (11)$$

associated with the eigenvalue  $\mu_{\max}$ . (Note that  $M$  and  $T_{x_{\text{inst}}(\cdot) \equiv 0}$  have the same eigenvalues with the same degeneracies [12]). The  $\mathbf{s}_n$ s in equation (10) are correlated complex Gaussian random variables with  $\langle \mathbf{s}_n \rangle = \langle \mathbf{s}_n \mathbf{s}_m \rangle = 0$  and  $\langle \mathbf{s}_n \mathbf{s}_m^* \rangle = \mathbf{e}_n^{(1)} \mathbf{e}_m^{(1)*}$ , where  $\mathbf{e}^{(1)}$  (with components  $\mathbf{e}_n^{(1)}$ ) is the normalized fundamental eigenvector of  $M$  (see [1], section 4.1, for details).

Figure 1 shows the contour plots of  $|S_{\text{inst}}|^2$  and the autocorrelation profile  $|C|^2$  (also referred to as ‘hot spot profile’ in the laser-matter interaction literature [10]).

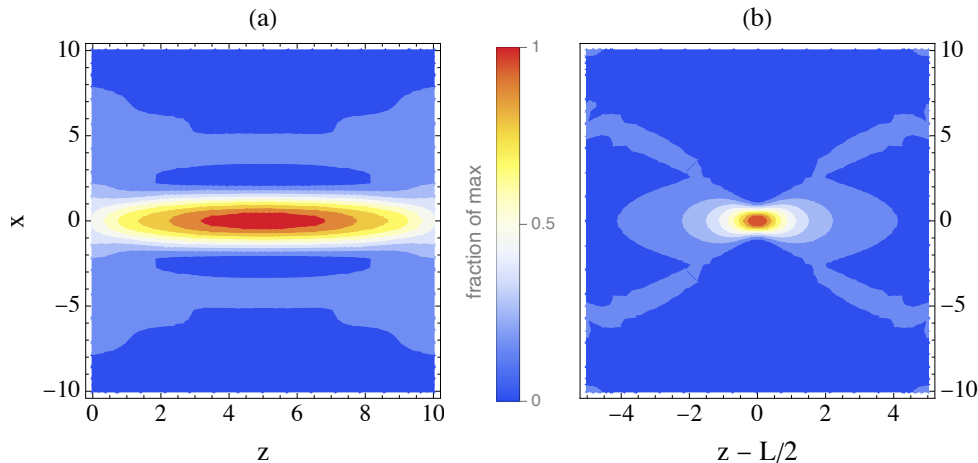


FIG. 1: (a) Contour plot of  $|S_{\text{inst}}(x, z)|^2$  for  $S_{\text{inst}}(x, z)$  in equation (10). (b) Contour plot of the autocorrelation profile (or ‘hot spot profile’)  $|C(x, z - L/2)|^2$  for  $C(x, z - z')$  in equation (9). (Color scale legend : 1 corresponds to the maximum of the plotted function.)

Define  $\hat{S}_{\text{inst}} = S_{\text{inst}} / \|S_{\text{inst}}\|_2$  and  $\hat{S} = S / \|S\|_2$  where  $\|\cdot\|_2$  denotes the  $L^2$ -norm on  $\Lambda \times [0, L]$ . Write  $S_{\parallel} = (\hat{S}_{\text{inst}}, \hat{S}) \hat{S}_{\text{inst}}$  the component of  $\hat{S}$  along  $S_{\text{inst}}$ . We measure the departure of the realizations of  $S$  from the instanton through the  $L^2$ -distance

$$\mathfrak{D} = \|\hat{S} - S_{\parallel}\|_2 = \sqrt{1 - |(\hat{S}, \hat{S}_{\text{inst}})|^2}. \quad (12)$$

Using the Fourier representations (6) and (10) in which we write  $\mathbf{s}$  (with components  $s_n$ ) as  $\mathbf{s} = \|\mathbf{s}\| \hat{\mathbf{s}}$  and  $\mathbf{s} = \|\mathbf{s}\| \hat{\mathbf{s}}$  with  $\hat{\mathbf{s}} = e^{i \arg(c_1)} \mathbf{e}^{(1)}$ , where  $\|\cdot\|$  is the usual Euclidean norm and

where we have used  $\mathbf{s} = \mu_{\max}^{-1/2} c_1 \mathbf{e}^{(1)}$  (see the first equation (52) in [1]), one gets

$$\mathfrak{D} = \sqrt{1 - \frac{|\sum'_n \varsigma_n \hat{s}_n \mathbf{e}_n^{(1)*}|^2}{(\sum'_n \varsigma_n |\hat{s}_n|^2) (\sum'_n \varsigma_n |\mathbf{e}_n^{(1)}|^2)}}, \quad (13)$$

which is the counterpart of the equation (66) in [1] for fixed  $y = 0$ <sup>2</sup>.

From equations (4), (6), and (9) it can be checked that the  $s_n$ s are also the components of  $T_C^{-1/2}S$  in the orthonormal function basis  $(1/\sqrt{\ell L}) \exp i[2\pi n x/\ell + (2\pi n/\ell)^2 z/2]$ , which trivially defines an isomorphism between the  $T_C^{-1/2}S$ -space with  $L^2$ -inner product and the  $\mathbf{s}$ -space with usual dot product. As we will see shortly, for  $S$  given by equation (6) our results come out naturally in terms of  $\mathbf{s}$ , and therefore  $T_C^{-1/2}S$ , which explains why it is  $T_C^{-1/2}S$  that appears in the summary at the end of section I, rather than  $S$ .

We now have everything we need to move on to the presentation of our biased sampling method and its application to check the validity of the results obtained in [1]. This is the subject of the next two sections.

### III. PRELIMINARY RESULTS

The problem is to sample large values of  $\log_{10} U$ . To this end, we need a biased sampling of the realizations of  $S$  or, equivalently, of the random vector  $\mathbf{s}$ . The relationship between  $U$  and  $\mathbf{s}$  is extremely intricate and no simple general expression of the form  $U = U(\mathbf{s})$  can be written explicitly. All we can say at this stage is that a large  $\log_{10} U$  implies a large  $\|\mathbf{s}\|$ , while the converse is not true<sup>3</sup>. So, sampling large  $\|\mathbf{s}\|$ 's alone is not sufficient.

According to instanton analysis, the larger  $\log_{10} U$  the more  $\mathbf{s}$  tends to align with  $\mathbf{e}^{(1)}$ . Testing the instanton analysis in the large  $\log_{10} U$  regime therefore requires as a prerequisite to be able to control the direction of  $\mathbf{s}$ , which can be achieved by a change of variables making the direction and amplitude of  $\mathbf{s}$  explicit.

Write  $N$  the number of terms in the sum on the right-hand side of equation (6) ( $N = 100$ ). A given realization of  $S$  corresponds to a given realization of the (complex)  $N$ -dimensional vector  $\mathbf{s}$  with coordinates  $s_n$ , and conversely. Let  $\hat{\mathbf{u}}$  be a given  $N$ -dimensional (complex) unit

<sup>2</sup> Note the typo on the right-hand side of equation (66) in [1]: in the denominator, it should read  $\hat{s}_n$  and  $\hat{\mathbf{s}}_n = e^{i \arg(c_1)} \mathbf{e}_n^{(1)}$  instead of  $s_n$  and  $\mathbf{s}_n$ .

<sup>3</sup> A realization of  $\mathbf{s}$  with a large  $\|\mathbf{s}\|$  can yield a large  $\log_{10} |\psi(x, L)|^2 \gg 1$  at some  $x$  away from zero and  $\log_{10} U = \log_{10} |\psi(0, L)|^2 = O(1)$  even though  $\|\mathbf{s}\|$  is large.



vector, not necessarily equal to  $\mathbf{e}^{(1)}$ . Define  $\mathbf{s}_{\parallel} = (\mathbf{s} \cdot \hat{\mathbf{u}}^*) \hat{\mathbf{u}} = r_{\parallel} e^{i\varphi} \hat{\mathbf{u}}$  and  $\mathbf{s}_{\perp} = \mathbf{s} - \mathbf{s}_{\parallel} = r_{\perp} \hat{\mathbf{s}}_{\perp}$  with  $r_{\perp} = \|\mathbf{s}_{\perp}\|$ . Switching to polar coordinates  $r_{\parallel} = \sqrt{\eta} \cos \theta$  and  $r_{\perp} = \sqrt{\eta} \sin \theta$ , with  $\eta = \|\mathbf{s}\|^2$  and  $0 \leq \theta \leq \pi/2$ , we characterize the realizations of  $S$  by the new variables  $\theta$ ,  $\eta$ ,  $\varphi$ , and  $\hat{\mathbf{s}}_{\perp}$ . The polar angle  $\theta$  measures how close the direction of  $\mathbf{s}$  is to that of  $\hat{\mathbf{u}}$ :  $\theta = 0$  means that  $\mathbf{s}$  is aligned with  $\hat{\mathbf{u}}$  ( $\mathbf{s}_{\perp} = 0$ ), whereas  $\theta = \pi/2$  means that  $\mathbf{s}$  is orthogonal to  $\hat{\mathbf{u}}$  ( $\mathbf{s}_{\parallel} = 0$ ).

The statistical properties of the new variables are deduced from the ones of the  $s_n$ s. Since the  $s_n$ s are i.i.d. standard complex Gaussian random variables, the projection of  $\mathbf{s}$  onto any given direction is also a standard complex Gaussian random variable independent of the projections onto the orthogonal directions. This applies in particular to  $s_{\parallel} = (\mathbf{s} \cdot \hat{\mathbf{u}}^*)$  and the components of  $\mathbf{s}_{\perp}$ . After some straightforward algebra, one finds that the probability distribution functions (pdf) of  $\theta$  and  $\eta$  are respectively given by

$$f(\theta) = 2(N-1) (\sin \theta)^{2N-3} \cos \theta \mathbf{1}_{0 \leq \theta \leq \pi/2}, \quad (14)$$

and

$$h(\eta) = \Gamma(N)^{-1} \eta^{N-1} e^{-\eta} \mathbf{1}_{0 \leq \eta}. \quad (15)$$

The random phase  $\varphi$  and tip of  $\hat{\mathbf{s}}_{\perp}$  (i.e. the direction of  $\mathbf{s}_{\perp}$ ) are uniformly distributed over  $[0, 2\pi)$  and the sphere  $\|\mathbf{s}_{\perp}\| = 1$ , respectively.

### A. Closeness rule

The class of  $S$  we consider is defined by the condition that, for fixed  $\eta \gg 1$ , there is only one direction of  $\mathbf{s}$  that globally maximizes  $\log_{10} U(\eta, \hat{\mathbf{s}})$ . Call this direction ‘the maximizing direction’ and the class of  $S$  ‘the one-max class’. It is shown in the appendix that  $S$  as given by the equations (6) and (7) and, more generally, any  $S$  admitting a single and non-degenerate instanton, belongs to this class. For a given  $\hat{\mathbf{u}}$ , there are two possibilities: either  $\hat{\mathbf{u}}$  is the maximizing direction or it is not. The way we distinguish between these two possibilities is based on the following reasoning. If  $\hat{\mathbf{u}}$  is the maximizing direction, then all the directions other than  $\hat{\mathbf{u}}$  contribute negligibly to the amplification compared with  $\hat{\mathbf{u}}$  (in the large  $\eta$  limit we consider), and the computed  $\log_{10} U$  depends negligibly on the perpendicular components  $\hat{\mathbf{s}}_{\perp}$ . In addition, at  $\theta = 0$ ,  $\mathbf{s}$  is aligned with  $\hat{\mathbf{u}}$  ( $\mathbf{s}_{\perp} = 0$ ) and  $\log_{10} U$  reaches its maximum, i.e.,  $\log_{10} U_{\theta=0} = \max_{\theta} \log_{10} U_{\theta}$  for all realizations of  $\hat{\mathbf{s}}_{\perp}$ . By contraposition,

if the computed  $\log_{10} U$  is found to be sensitive to  $\hat{\mathbf{s}}_{\perp}$  or if there is a realization of  $\hat{\mathbf{s}}_{\perp}$  for which  $\log_{10} U_{\theta=0} < \max_{\theta} \log_{10} U_{\theta}$ , then  $\hat{\mathbf{u}}$  is not the maximizing direction. Conversely, if  $\hat{\mathbf{u}}$  is not the maximizing direction, then  $\log_{10} U$  does not reach its maximum at  $\theta = 0$  and there are realizations of  $\hat{\mathbf{s}}_{\perp}$  for which  $\log_{10} U_{\theta=0} < \max_{\theta} \log_{10} U_{\theta}$ . In addition, for  $\theta > 0$ , there are realizations of  $\hat{\mathbf{s}}_{\perp}$  with a non negligible component along the maximizing direction, which makes the computed  $\log_{10} U$  sensitive to  $\hat{\mathbf{s}}_{\perp}$ . Observing one of these two characteristics in the behavior of  $\log_{10} U$  is sufficient to conclude that  $\hat{\mathbf{u}}$  is not the maximizing direction. By contraposition, if the computed  $\log_{10} U$  is observed to depend negligibly on  $\hat{\mathbf{s}}_{\perp}$  and if  $\log_{10} U_{\theta=0} = \max_{\theta} \log_{10} U_{\theta}$  for each realization of  $\hat{\mathbf{s}}_{\perp}$ , then  $\hat{\mathbf{u}}$  is the maximizing direction.

For our purposes, we do not need to know whether  $\hat{\mathbf{u}}$  is exactly the maximizing direction, but rather whether it is close enough to the maximizing direction, so that a good sampling of  $\hat{\mathbf{s}}$  around  $\hat{\mathbf{u}}$  ensures a good sampling of  $\hat{\mathbf{s}}$  around the nearby maximizing direction, thereby providing access to large values of  $\log_{10} U$  (for  $\eta \gg 1$ ). This does not change the reasoning above which leads to the following ‘closeness rule’ that we will use for deciding whether  $\hat{\mathbf{u}}$  is close to the maximizing direction.

*Closeness rule:* for  $\eta \gg 1$ ,  $\hat{\mathbf{u}}$  is close to the maximizing direction *iff*  $\log_{10} U$  for  $\hat{\mathbf{s}}$  around  $\hat{\mathbf{u}}$  is observed to depend very little on  $\hat{\mathbf{s}}_{\perp}$  and  $\log_{10} U_{\theta=0} = \max_{\theta} \log_{10} U_{\theta}$  for all the sampled realizations of  $\hat{\mathbf{s}}_{\perp}$  to within a negligible fraction of them.

The point is that by applying this rule to numerical outcomes, we can check whether  $\hat{\mathbf{u}}$  is close to the maximizing direction numerically without having to compare the values of  $\log_{10} U$  for different  $\hat{\mathbf{u}}$ 's. It is worth noticing that the closeness rule applies to any  $S$  in the one-max class, whatever the maximizing direction. It is not necessary to know it beforehand. For this class of  $S$ , the closeness rule does not presuppose any knowledge of the instanton solution to be tested.

## B. Setting a reference direction $\hat{\mathbf{u}}$

It is instructive to compare the values of  $\log_{10} U$  and the compliance with the closeness rule for various reference directions  $\hat{\mathbf{u}}$ . We have made this comparison for  $\hat{\mathbf{u}} = \mathbf{e}^{(1)}$  and the orthogonal directions  $\hat{\mathbf{u}} = \mathbf{e}^{(q \geq 2)}$ , where  $\mathbf{e}^{(q)}$  is the normalized eigenvector of  $M$  associated with the  $q$ th largest eigenvalue  $\mu_q$  ( $1 \leq q \leq N$ ). (For the Gaussian spectrum (7), we have

checked numerically that none of the  $\mu_q$ s is degenerate.) Figure 2 shows scatter plots of  $\theta$  and numerically computed  $\log_{10} U$  for a given realization of  $\varphi$  and  $\hat{\mathbf{s}}_{\perp}$ , fixed  $\eta = 1060$ , and  $10^3$  ascending values of  $\theta$  regularly spaced by  $\Delta\theta = 0.263 \cdot 10^{-3}$  starting from  $\theta = 0$ . Figures 2(a) and (b) correspond to  $\hat{\mathbf{u}} = \mathbf{e}^{(1)}$  and  $\hat{\mathbf{u}} = \mathbf{e}^{(2)}$ , respectively (with  $\mu_1 = \mu_{\max} = 4.34984$  and  $\mu_2 = 2.0126$ ). It is clear that for a given  $\eta \gg 1$ ,  $\log_{10} U$  is significantly smaller in the case  $\hat{\mathbf{u}} = \mathbf{e}^{(2)}$  (figure 2(b)) than for  $\hat{\mathbf{u}} = \mathbf{e}^{(1)}$  (figure 2(a)). To have comparable values of  $\log_{10} U$  we need a larger  $\eta$  for  $\hat{\mathbf{u}} = \mathbf{e}^{(2)}$  than for  $\hat{\mathbf{u}} = \mathbf{e}^{(1)}$ . To be more precise, for  $\hat{\mathbf{u}} = \mathbf{e}^{(2)}$  we have checked that data with  $\log_{10} U \gtrsim 1820$  — the smallest value in figure 2(a) — require  $\eta \gtrsim 1760$ . For instance, it can be seen in figure 3 that it takes  $\eta \simeq 1880$  to bring the data for  $\hat{\mathbf{u}} = \mathbf{e}^{(2)}$  in the same range of  $\log_{10} U$  as for  $\hat{\mathbf{u}} = \mathbf{e}^{(1)}$  in figure 2(a). Given the fast decreasing tail of  $h(\eta)$  in equation (15), an important consequence of this result is that among data falling in the same  $(\log_{10} U, \theta)$  region as in figure 2(a), the proportion of realizations of  $\hat{\mathbf{s}}$  biased towards  $\mathbf{e}^{(2)}$  is completely negligible compared to that of realizations biased towards  $\mathbf{e}^{(1)}$ , typically by a factor less than  $h(1760)/h(1060) = O(10^{-282})$ . We have checked that the proportion of realizations of  $\hat{\mathbf{s}}$  biased towards  $\mathbf{e}^{(q \geq 3)}$  is by far even smaller. From these first results, we can already conclude that the maximizing direction is not orthogonal to  $\mathbf{e}^{(1)}$ . To go further we will use the closeness rule.

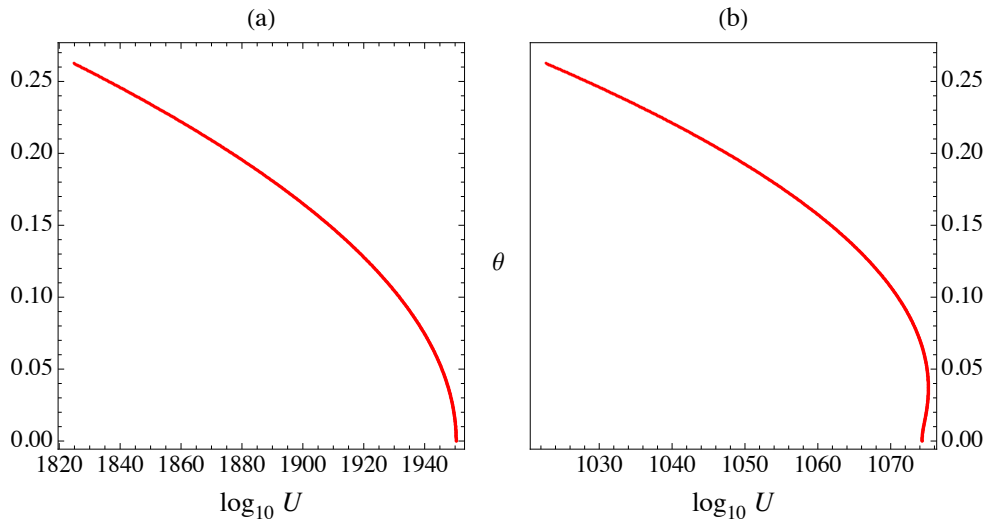


FIG. 2: Scatter plots of  $\theta$  and  $\log_{10} U$  for a given realization of  $\varphi$  and  $\hat{\mathbf{s}}_{\perp}$ , fixed  $\eta = 1060$ , and two different (orthogonal) reference directions: (a)  $\hat{\mathbf{u}} = \mathbf{e}^{(1)}$  and (b)  $\hat{\mathbf{u}} = \mathbf{e}^{(2)}$ .

Figure 3 shows the same kind of scatter plots as in the figure 2 for  $\hat{\mathbf{u}} = \mathbf{e}^{(1)}$  with  $\eta = 1060$  and  $\hat{\mathbf{u}} = \mathbf{e}^{(2)}$  with larger  $\eta = 1880$ , each for two different realizations of  $\varphi$  and  $\hat{\mathbf{s}}_{\perp}$ . For better legibility, only 50 of the  $10^3$  points actually computed are shown. Both figures 2 and 3 show that for a given realization of  $\varphi$  and  $\hat{\mathbf{s}}_{\perp}$ , the data collapse on a single well-defined curve in the  $(\log_{10} U, \theta)$  plane. For  $\hat{\mathbf{u}} = \mathbf{e}^{(2)}$ , it is clear that (i)  $\log_{10} U_{\theta=0} < \max_{\theta} \log_{10} U_{\theta}$  and (ii) a change in the realization of  $\hat{\mathbf{s}}_{\perp}$  visibly affects the curve (see up and down green triangles in figure 3), which means that the contribution of  $\mathbf{s}_{\perp}$  to the amplification is not negligible in this case. We have observed the same behavior for  $\hat{\mathbf{u}} = \mathbf{e}^{(q \geq 3)}$  with an even more pronounced sensitivity to  $\hat{\mathbf{s}}_{\perp}$ . Furthermore, as  $q$  increases, a dispersion of the data around the curve in the  $(\log_{10} U, \theta)$  plane becomes visible and increases with  $q$ . These results do not meet the closeness rule, which implies that none of the orthogonal directions  $\mathbf{e}^{(q \geq 2)}$  is close to the maximizing direction, in agreement with the discussion of figure 2.

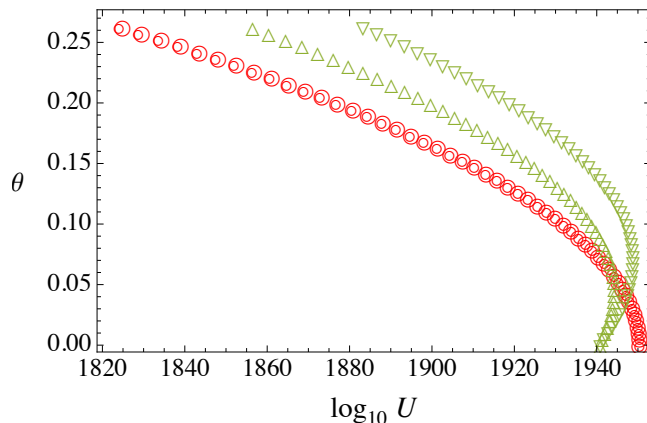


FIG. 3: Scatter plots of  $\theta$  and  $\log_{10} U$  for two different realizations of  $\varphi$  and  $\hat{\mathbf{s}}_{\perp}$ , each represented by a specific marker. Large and small red circles are for  $\hat{\mathbf{u}} = \mathbf{e}^{(1)}$  with  $\eta = 1060$ , up and down green triangles correspond to  $\hat{\mathbf{u}} = \mathbf{e}^{(2)}$  with  $\eta = 1880$ . (For better legibility, only 50 of the  $10^3$  points actually computed are shown.)

In contrast, for  $\hat{\mathbf{u}} = \mathbf{e}^{(1)}$  we observe that (i)  $\log_{10} U_{\theta=0}$  is indistinguishable from  $\max_{\theta} \log_{10} U_{\theta}$  (see also figures 2 and 4) and (ii) the curve depends very little on  $\hat{\mathbf{s}}_{\perp}$  (small and large red circles in figure 3). We have checked that the same was always true for several different subsamples of realizations of  $\hat{\mathbf{s}}_{\perp}$ . It can then be concluded from the closeness rule that  $\mathbf{e}^{(1)}$  is close to the maximizing direction, so that sampling large  $\eta$ 's and  $\hat{\mathbf{s}}$  around  $\mathbf{e}^{(1)}$

simultaneously will yield a good sampling of the large values of  $\log_{10} U$ . Thus, it is legitimate to take  $\hat{\mathbf{u}} = \mathbf{e}^{(1)}$  as a reference direction, which will be the case from now on.

The fact that the numerical results for  $\hat{\mathbf{u}} = \mathbf{e}^{(1)}$  meet the closeness rule and validate  $\mathbf{e}^{(1)}$  as being close to the maximizing direction is a necessary but not sufficient condition for the validity of the instanton analysis. It remains to find out how to use this result to achieve a good sampling of the large values of  $\log_{10} U$  from which numerical results and analytical predictions can be compared. This is the subject of the sections III C and IV.

*Notational remark:* it may be useful to briefly come back to the notations used in sections I and III. As explained at the end of section II,  $\mathbf{s}$  and  $\mathbf{s}$  are respectively isomorphic to  $T_C^{-1/2}S$  and  $T_C^{-1/2}S_{\text{inst}}$ . Consequently,  $r_{\parallel}$  and  $r_{\perp}$  are also the  $L^2$ -norms of the components of  $T_C^{-1/2}S$  parallel and perpendicular to  $T_C^{-1/2}S_{\text{inst}}$  (in the  $L^2$  sense) which is close to the maximizing direction, as written in section I. The quantities characterizing  $T_C^{-1/2}S$  other than  $\theta$  and  $\eta$ , denoted by  $\Sigma_{\text{orv}}$  in section I, are the random variable  $\varphi$  and vector  $\hat{\mathbf{s}}_{\perp}$ .

### C. Highly accurate approximation of $\ln U(\mathbf{s}) \gg 1$ and applications

In figure 4 we show scatter plots of  $\theta$  and  $\log_{10} U$  for a given realization of  $\varphi$  and  $\hat{\mathbf{s}}_{\perp}$ , and four different values of  $\eta$  (see caption for details). Solid lines correspond to the nonlinear fit

$$\theta = \cos^{-1} \left( \frac{\log_{10} U}{a\eta - b} \right)^{\alpha(\mathbf{s})}, \quad (16)$$

where  $a = 1.86428$ ,  $b = 25.7163$ , and  $\alpha(\mathbf{s})$  is a realization dependent exponent. Data points and nonlinear fits match remarkably well. We have checked on  $10^2$  realizations of  $\varphi$  and  $\hat{\mathbf{s}}_{\perp}$ , 10 values of  $\eta$  between 1010 and 1100, and  $10^3$  values of  $\theta$  like in figure 2 (which represents a total of  $10^6$  different realizations of  $S$ ), that for each  $\varphi$ ,  $\hat{\mathbf{s}}_{\perp}$ , and  $\eta$ , the  $10^3$  data points and the nonlinear fit are practically indistinguishable over the whole range  $1820 \leq \log_{10} U \leq 2030$ . Numerical results also show that (i) no systematic (monotonic) variation of  $\alpha(\mathbf{s})$  with  $\eta$  at fixed  $\varphi$  and  $\hat{\mathbf{s}}_{\perp}$  is observed in the range of  $\eta$  considered, and (ii) the relative dispersion of  $\alpha(\mathbf{s})$  for different values of  $\eta$  at fixed  $\varphi$  and  $\hat{\mathbf{s}}_{\perp}$  is one order of magnitude less than for different realizations of  $\varphi$  and  $\hat{\mathbf{s}}_{\perp}$  at fixed  $\eta$ :  $\Delta\alpha(\mathbf{s})/\langle\alpha(\mathbf{s})\rangle = 10^{-3}$  and  $10^{-2}$ , respectively. Thus, with a good accuracy level, it is not unreasonable to ignore the dependence of  $\alpha(\mathbf{s})$  on  $\eta$  and write  $\alpha(\mathbf{s}) = \alpha(\varphi, \hat{\mathbf{s}}_{\perp})$ . We have checked by replacing  $\alpha(\mathbf{s})$  in equation (16) with values obtained for different  $\eta$  and the same realization of  $\varphi$  and  $\hat{\mathbf{s}}_{\perp}$ , all other things being

equal, that the error made in the nonlinear fit is indeed imperceptible.

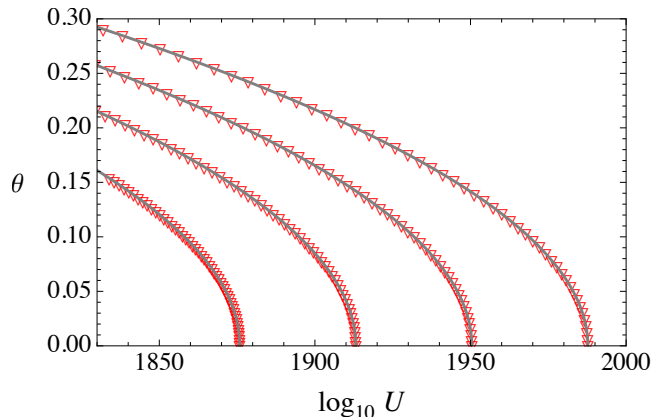


FIG. 4: Scatter plots of  $\theta$  and  $\log_{10} U$  for a given realization of  $\varphi$  and  $\hat{\mathbf{s}}_{\perp}$ , and four different values of  $\eta$  (down red triangles). Solid lines are plots of the nonlinear fit (16) for the corresponding  $\eta$  and numerically computed  $\alpha(\mathbf{s})$ . Parameter values are  $(\eta, \alpha(\mathbf{s})) = (1020, 0.52569)$ ,  $(1040, 0.524924)$ ,  $(1060, 0.524862)$ , and  $(1080, 0.525024)$ , from left to right. (In each plot, only 50 of the  $10^3$  points actually computed are shown.)

We now use these results to write  $p(U)$  and conditional probability  $P(\mathcal{A}|U)$ , where  $\mathcal{A}$  is a set of realizations of  $\mathbf{s}$ , in forms suitable to numerical estimates in the asymptotic regime  $\log_{10} U \gg 1$ . Write  $U(\mathbf{s}) = |\psi_{\mathbf{s}}(0, L)|^2$ ,  $\psi_{\mathbf{s}}(x, z)$  being the solution to equation (1) for a given  $S$  (hence  $\mathbf{s}$ ). Our numerical results show that inverting equation (16) gives a highly accurate approximation of  $U(\mathbf{s})$  when  $\log_{10} U \gg 1$ . Replacing  $U(\mathbf{s})$  with this approximation in the exact expression  $p(U) = \langle \delta(U - U(\mathbf{s})) \rangle_{\mathbf{s}}$ , where  $\langle \cdot \rangle_{\mathbf{s}}$  denotes the average over the realizations of  $\mathbf{s}$ , and integrating out  $\eta$ , one obtains

$$\begin{aligned}
p(U) &= \frac{1}{\ln 10} \frac{1}{U} \left\langle \int_0^{+\infty} \int_0^{\pi/2} \delta[\log_{10} U - (\cos \theta)^{1/\alpha(\varphi, \hat{\mathbf{s}}_{\perp})} (a\eta - b)] f(\theta) h(\eta) d\theta d\eta \right\rangle_{\varphi, \hat{\mathbf{s}}_{\perp}} \\
&= \frac{1}{\ln 10} \frac{1}{U} \left\langle \int_0^{\pi/2} \frac{f(\theta)}{a (\cos \theta)^{1/\alpha(\varphi, \hat{\mathbf{s}}_{\perp})}} h \left( \frac{\log_{10} U}{a (\cos \theta)^{1/\alpha(\varphi, \hat{\mathbf{s}}_{\perp})}} + \frac{b}{a} \right) d\theta \right\rangle_{\varphi, \hat{\mathbf{s}}_{\perp}} \\
&= \left\langle \int_0^{\pi/2} p(U, \theta | \varphi, \hat{\mathbf{s}}_{\perp}) d\theta \right\rangle_{\varphi, \hat{\mathbf{s}}_{\perp}}, \tag{17}
\end{aligned}$$

with

$$p(U, \theta | \varphi, \hat{\mathbf{s}}_{\perp}) = \frac{1}{\ln 10} \frac{f(\theta)}{U a (\cos \theta)^{1/\alpha(\varphi, \hat{\mathbf{s}}_{\perp})}} h \left( \frac{\log_{10} U}{a (\cos \theta)^{1/\alpha(\varphi, \hat{\mathbf{s}}_{\perp})}} + \frac{b}{a} \right). \tag{18}$$

The expression for  $p(\mathcal{A}, U)$  is similar to equation (17) with integrand  $\mathbf{1}_{\mathbf{s} \in \mathcal{A}} p(U, \theta | \varphi, \hat{\mathbf{s}}_{\perp})$ , where  $\mathbf{s} = \sqrt{\eta} \cos \theta e^{i\varphi} \mathbf{e}^{(1)} + \sqrt{\eta} \sin \theta \hat{\mathbf{s}}_{\perp}$  with  $\eta = a^{-1}(b + \log_{10} U / (\cos \theta)^{1/\alpha(\varphi, \hat{\mathbf{s}}_{\perp})})$ . Dividing this expression by  $p(U)$  one obtains

$$P(\mathcal{A}|U) = \left\langle w_{\alpha(\varphi, \hat{\mathbf{s}}_{\perp})}(U) \int_0^{\pi/2} \mathbf{1}_{\mathbf{s} \in \mathcal{A}} p(\theta|U, \varphi, \hat{\mathbf{s}}_{\perp}) d\theta \right\rangle_{\varphi, \hat{\mathbf{s}}_{\perp}}, \quad (19)$$

with

$$p(\theta|U, \varphi, \hat{\mathbf{s}}_{\perp}) = \frac{p(U, \theta | \varphi, \hat{\mathbf{s}}_{\perp})}{p(U | \varphi, \hat{\mathbf{s}}_{\perp})}, \quad (20)$$

and

$$w_{\alpha(\varphi, \hat{\mathbf{s}}_{\perp})}(U) = \frac{p(U | \varphi, \hat{\mathbf{s}}_{\perp})}{p(U)}. \quad (21)$$

In equations (20) and (21),  $p(U | \varphi, \hat{\mathbf{s}}_{\perp})$  is obtained by integrating equation (18) numerically over  $0 \leq \theta \leq \pi/2$  (for fixed  $U$  and a given realization of  $\varphi$  and  $\hat{\mathbf{s}}_{\perp}$ ). Equations (17) to (21) give the expressions of  $p(U)$  and  $P(\mathcal{A}|U)$  valid for  $\log_{10} U \gg 1$ .

#### IV. BIASED SAMPLING AND NUMERICAL VALIDATION OF INSTANTON ANALYSIS

Let  $\{\varphi, \hat{\mathbf{s}}_{\perp}\}$  denote the same sample of  $10^2$  independent realizations of  $\varphi$  and  $\hat{\mathbf{s}}_{\perp}$  as the one we used to check the validity of the nonlinear fit in equation (16). Figure 5 shows plots of  $p(\theta|U, \varphi, \hat{\mathbf{s}}_{\perp})$  for three different realizations in  $\{\varphi, \hat{\mathbf{s}}_{\perp}\}$  and  $\log_{10} U = 1890$ . Curves (a), (b), and (c) correspond to the realizations of  $\{\varphi, \hat{\mathbf{s}}_{\perp}\}$  yielding the smallest, middle, and largest values of  $p(U | \varphi, \hat{\mathbf{s}}_{\perp})$ , respectively. Diamonds correspond to  $p(\theta|U)$  computed as the sample mean of  $p(\theta|U, \varphi, \hat{\mathbf{s}}_{\perp})$  for the realizations in  $\{\varphi, \hat{\mathbf{s}}_{\perp}\}$ .

For a given  $\log_{10} U \gg 1$  and realizations of  $\varphi$  and  $\hat{\mathbf{s}}_{\perp}$  in  $\{\varphi, \hat{\mathbf{s}}_{\perp}\}$ , the statistically significant realizations of  $S$  are those for which  $\theta$  is in the bulk of  $p(\theta|U, \varphi, \hat{\mathbf{s}}_{\perp})$ . Figure 5 and similar results obtained for different values of  $U$  with  $1820 \leq \log_{10} U \leq 1940$  show that a good sampling of the corresponding significant values of  $\theta$ , requires probing the whole range  $0.25 \leq \theta \leq 0.38$ . Since the equation (14) yields  $P(0.25 \leq \theta \leq 0.38) = 5.2 \cdot 10^{-86}$ , it is clear that having  $\theta$  in that range is an extremely rare event, virtually impossible to sample directly by drawing  $\theta$  from  $f(\theta)$ . By contrast, it can easily be achieved by drawing  $\theta$  from  $p(\theta|U, \varphi, \hat{\mathbf{s}}_{\perp})$  rather than from  $f(\theta)$ . This is what defines our biased sampling procedure which consists of the following four steps:

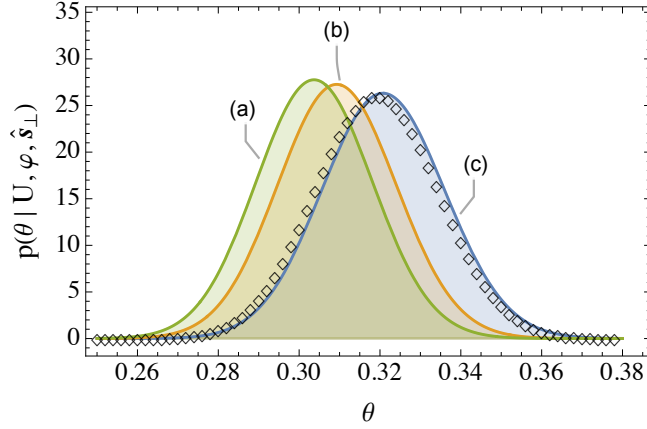


FIG. 5: Plots of  $p(\theta|U, \varphi, \hat{\mathbf{s}}_{\perp})$  for fixed  $\log_{10} U = 1890$  and three particular realizations in  $\{\varphi, \hat{\mathbf{s}}_{\perp}\}$  corresponding to the smallest (a), middle (b), and largest (c) values of  $p(U|\varphi, \hat{\mathbf{s}}_{\perp})$ . Plot of  $p(\theta|U)$  for the same value of  $\log_{10} U = 1890$  (diamonds).

- (A) draw  $\varphi$  and  $\hat{\mathbf{s}}_{\perp}$  from the uniform distributions over  $[0, 2\pi)$  and the sphere  $\|\mathbf{s}_{\perp}\| = 1$ , respectively;
- (B) compute  $\alpha(\varphi, \hat{\mathbf{s}}_{\perp})$  by making the graph of the function in equation (16) fit the data in the  $(\log_{10} U, \theta)$  region of interest, for some fixed  $\eta$ . (We have checked that for  $1820 \leq \log_{10} U \leq 2030$ ,  $\eta$  can be chosen arbitrarily between 1010 and 1100;)
- (C) use the result in equation (18) to get  $p(U, \theta|\varphi, \hat{\mathbf{s}}_{\perp})$ . For fixed  $U$  (with  $\log_{10} U \gg 1$ ), integrate the result numerically over  $0 \leq \theta \leq \pi/2$  to get  $p(U|\varphi, \hat{\mathbf{s}}_{\perp})$ . Then, draw  $\theta$  from  $p(\theta|U, \varphi, \hat{\mathbf{s}}_{\perp})$  in equation (20) and set  $\eta$  by inverting equation (16);
- (D) the outcome defines a realization of  $\mathbf{s} = \sqrt{\eta} \cos \theta e^{i\varphi} \mathbf{e}^{(1)} + \sqrt{\eta} \sin \theta \hat{\mathbf{s}}_{\perp}$  which, once injected onto the right-hand side of equation (6), gives a realization of  $S(x, z)$ .

From the data for  $p(U|\varphi, \hat{\mathbf{s}}_{\perp})$  obtained as explained in step (C) for each realization in  $\{\varphi, \hat{\mathbf{s}}_{\perp}\}$  and fixed  $U$ , we have estimated  $p(U)$  as the sample mean

$$p(U) = \frac{1}{\text{card}\{\varphi, \hat{\mathbf{s}}_{\perp}\}} \sum_{(\varphi, \hat{\mathbf{s}}_{\perp}) \in \{\varphi, \hat{\mathbf{s}}_{\perp}\}} p(U|\varphi, \hat{\mathbf{s}}_{\perp}). \quad (22)$$

Figure 6 shows the results in the  $(\log_{10} U, \log_{10} p(U))$  plane for five different values of  $U$  with  $\log_{10} U$  between 1840 and 1940. Black dots correspond to  $p(U)$ . The dispersion of



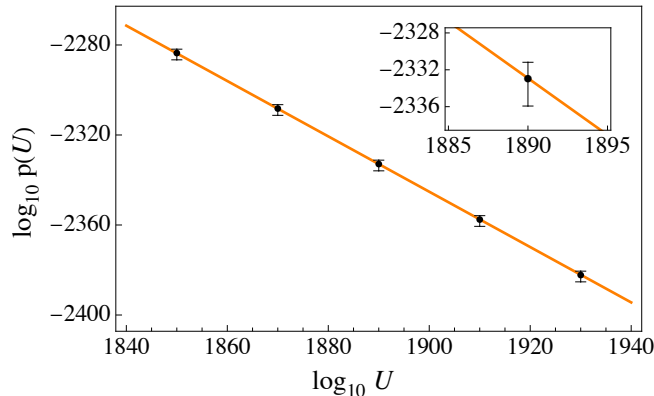


FIG. 6:  $\log_{10} p(U)$  as a function of  $\log_{10} U$  from 1850 to 1930 by steps of 20, with  $p(U)$  given by equation (22) (black dots). Smallest and largest  $p(U|\varphi, \hat{\mathbf{s}}_{\perp})$  in the sample are indicated by the ends of vertical bars. Plot of  $\log_{10} p(U) = -\zeta \log_{10} U + \text{Const}$  with  $\zeta = (1 + 1/2\mu_{\max}g) = 1.22989$  given by instanton analysis and  $\text{Const} = -8.41526$  adjusted to get the best fit to numerical data (solid line). Inset: enlargement of the same plot near  $\log_{10} U = 1890$ .

$p(U|\varphi, \hat{\mathbf{s}}_{\perp})$  around  $p(U)$  is indicated by vertical bars the ends of which correspond to the smallest and largest values of  $p(U|\varphi, \hat{\mathbf{s}}_{\perp})$  in the sample. The error bars corresponding to the standard deviation of the sample mean (22) are found to be eight times shorter, in the case of figure 6 (not shown). Instanton analysis predicts a leading algebraic tail of  $p(U) \propto U^{-\zeta}$  with  $\zeta = (1 + 1/2\mu_{\max}g) = 1.22989$ . The solid line is the plot of  $\log_{10} p(U) = -\zeta \log_{10} U + \text{Const}$ , where  $\text{Const} = -8.41526$  has been adjusted to get the best fit to the data. We observe an almost perfect alignment of numerical data along a straight line with slope  $-\zeta$ , which validates the result of instanton analysis numerically in the considered range of  $\log_{10} U$ . Note also the extremely small value of  $p(U) < 10^{-2270}$  which confirms, if need be, the absolute impossibility of sampling the extreme upper tail of  $p(U)$  directly, without a specific bias procedure.

The question then arises of the realizations of  $S$  behind the results in figure 6. According to the instanton analysis in [1], these realizations should be instanton realizations with the same profile as in figure 1(a). For a given  $U$ , the way  $S$  differs from the instanton can be characterized by the conditional pdf  $p(\mathfrak{D}|U)$ , where  $\mathfrak{D}$  is the  $L^2$ -distance defined in equation (13). For each element of  $\{\varphi, \hat{\mathbf{s}}_{\perp}\}$  and fixed  $U$ , we drew  $10^4$  independent realizations

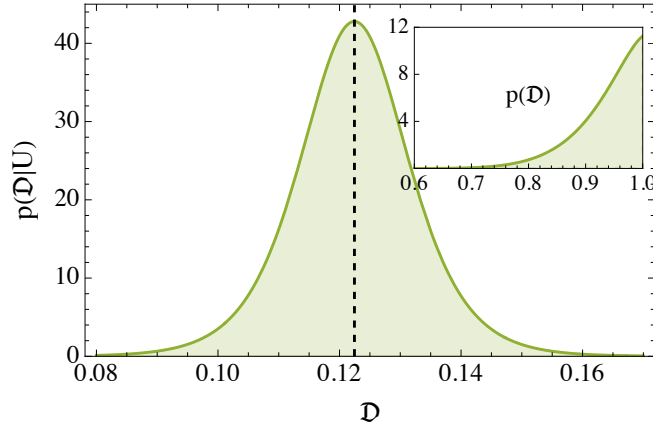


FIG. 7:  $p(\mathfrak{D}|U)$  for fixed  $\log_{10} U = 1890$  estimated from a biased sample of  $10^6$  realizations of  $S$  drawn according to the biased sampling procedure defined in steps (A) to (D) (see the text for details). The median of  $p(\mathfrak{D}|U)$  is at  $\mathfrak{D} = 0.122457$  (dashed line). Inset:  $p(\mathfrak{D})$  estimated from an unbiased sample of  $10^5$  realizations of  $S$ .

of  $\theta$  — denoted in the following by  $\{\theta|\varphi, \hat{\mathbf{s}}_{\perp}\}$  — as explained in step (C). For definiteness, we took  $\log_{10} U = 1890$  at the center of the range considered in figure 6. There are  $10^2$  subsamples  $\{\theta|\varphi, \hat{\mathbf{s}}_{\perp}\}$  of  $10^4$  elements each, representing a total of  $10^6$  different realizations of  $\mathbf{s}$ . For each  $\mathbf{s}$ , we have used the equation (13) to compute the corresponding value of  $\mathfrak{D}$ . We have then estimated  $P(\mathfrak{D} \leq \delta|U)$  in equation (19) with  $\mathcal{A} = \{\mathbf{s}|\mathfrak{D} \leq \delta\}$  as the sample mean

$$P(\mathfrak{D} \leq \delta|U) = \frac{1}{\text{card}\{\varphi, \hat{\mathbf{s}}_{\perp}\}} \sum_{(\varphi, \hat{\mathbf{s}}_{\perp}) \in \{\varphi, \hat{\mathbf{s}}_{\perp}\}} \frac{w_{\alpha(\varphi, \hat{\mathbf{s}}_{\perp})}(U)}{\text{card}\{\theta|\varphi, \hat{\mathbf{s}}_{\perp}\}} \sum_{\theta \in \{\theta|\varphi, \hat{\mathbf{s}}_{\perp}\}} \mathbf{1}_{\mathfrak{D} \leq \delta}, \quad (23)$$

from which  $p(\mathfrak{D}|U)$  is obtained by (numerical) derivation with respect to  $\delta$  at  $\delta = \mathfrak{D}$ . Figure 7 shows the result for  $\log_{10} U = 1890$ . For comparison, we show in inset the pdf of  $\mathfrak{D}$  obtained from an unbiased sample of  $10^5$  realizations of  $S$ . The vertical dashed line indicates the median of  $p(\mathfrak{D}|U)$  at  $\mathfrak{D} = \tilde{\mathfrak{D}} = 0.122457$ . (The median, mean and maximum of  $p(\mathfrak{D}|U)$  are all at  $\tilde{\mathfrak{D}}$ , to within numerical accuracy.) It can be seen that the realizations of  $S$  conditioned to a large  $\log_{10} U \gg 1$  are significantly closer to the instanton than unconditioned realizations:  $0.08 < \mathfrak{D} < 0.16$  and  $0.6 < \mathfrak{D} \leq 1$ , respectively, in the case of figure 7. This is in agreement with the instanton analysis in [1] which predicts  $\mathfrak{D} \rightarrow 0$  in probability, as  $\log_{10} U \rightarrow +\infty$ .

In figure 8, we show  $|\hat{S}(x, z)|^2 = |S(x, z)|^2 / \|S\|_2^2$  for two realizations of  $S$  with  $\mathfrak{D} \simeq 0.08$

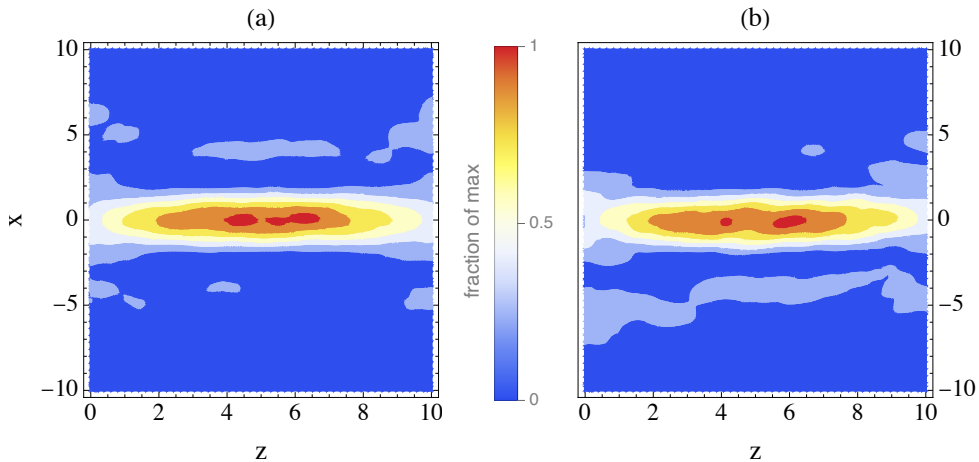


FIG. 8: Contour plots of  $|\hat{S}(x, z)|^2$  for two realizations of  $S$  with  $\mathfrak{D} \simeq 0.08$  (a) and  $\mathfrak{D} \simeq 0.16$  (b).

and  $\mathfrak{D} \simeq 0.16$ , on both sides of the bulk of  $p(\mathfrak{D}|U)$ . Figure 9(a) shows the same quantity for a realization with  $\mathfrak{D} \simeq \tilde{\mathfrak{D}}$ , right at the center of the bulk of  $p(\mathfrak{D}|U)$ . The results in figures 8(a), 8(b), and 9(a) are very similar and typical of the realizations generated by the biased sampling procedure for  $\log_{10} U = 1890$ . These realizations are the superposition of an elongated cigar-shaped component along  $x = 0$  and fluctuations of comparatively small amplitude. Fluctuations can be smoothed out by averaging realizations of  $|\hat{S}|^2$  close together along the  $\mathfrak{D}$  axis, bringing out the underlying cigar-shaped component. To this end, we have constructed five subsamples  $\{S\}_{\mathfrak{D}=\delta}$  of  $10^2$  realizations of  $S$  selected by picking in the total (biased) sample the 50 realizations with largest values of  $\mathfrak{D} \leq \delta$  and the 50 realizations with smallest values of  $\mathfrak{D} > \delta$ , for  $\delta = 0.1, 0.11, \tilde{\mathfrak{D}} (= 0.122457), 0.13$ , and  $0.14$ . Then, we have computed the sample means of  $|\hat{S}(x, z)|^2$  for the realizations in each  $\{S\}_{\mathfrak{D}=\delta}$ . In figure 9(b), we show the result for  $\delta = \tilde{\mathfrak{D}}$ . (The smallest and largest values of  $\mathfrak{D}$  for  $S$  in  $\{S\}_{\mathfrak{D}=\tilde{\mathfrak{D}}}$  are  $0.122403$  and  $0.122504$ , respectively.) We have obtained identical results for the five different values of  $\delta$  we have considered.

The close resemblance between figures 9(b) and 1(a) is obvious and the dominant cigar-shaped component of  $|\hat{S}|^2$  along  $x = 0$  visible in figures 8(a), 8(b), and 9(a) can clearly be identified as the theoretical instanton profile. These results, together with the small values of  $\mathfrak{D}$  observed in figure 7, show that the realizations of  $S$  conditioned on a large but finite  $\log_{10} U$  (here,  $\log_{10} U = 1890$ ) are low-noise instanton realizations. This is in agreement

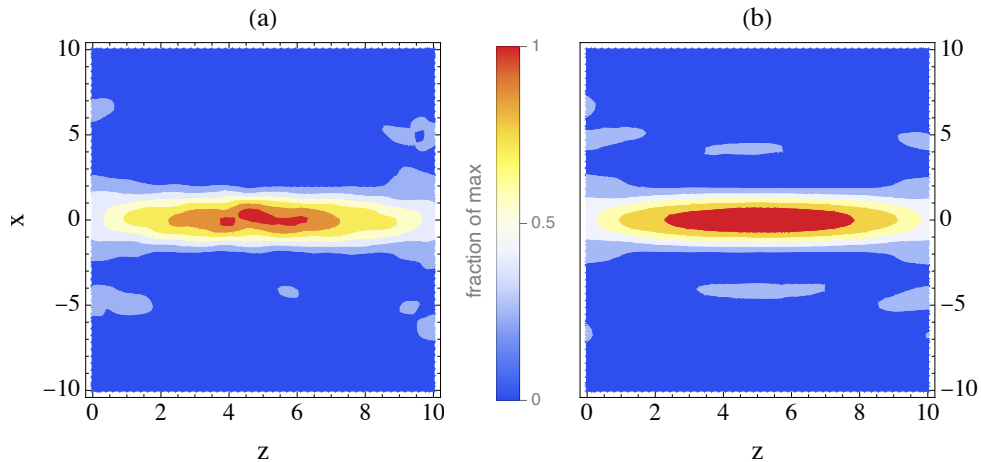


FIG. 9: (a) Contour plot of  $|\hat{S}(x, z)|^2$  for a realization of  $S$  with  $\mathfrak{D} \simeq \tilde{\mathfrak{D}}$ . (b) Contour plot of the sample mean of  $|\hat{S}(x, z)|^2$  for the  $10^2$  realizations in  $\{S\}_{\mathfrak{D}=\tilde{\mathfrak{D}}}$  (with  $0.122403 \leq \mathfrak{D} \leq 0.122504$ ).

with the instanton analysis in [1] which predicts noiseless instanton realizations in the limit  $\log_{10} U \rightarrow +\infty$ .

The numerical results in figures 6 to 9 show remarkable agreement with the analytical predictions of instanton analysis. Data points for  $\log_{10} p(U)$  and  $\log_{10} U$  line up almost perfectly along the predicted algebraic tail  $\log_{10} p(U) \simeq -\zeta \log_{10} U$  with  $\zeta = (1 + 1/2\mu_{\max}g)$  (figure 6), and the corresponding realizations of  $S$  are observed to be heavily biased towards the predicted instanton realizations (figures 7 to 9). In conclusion, we can say that our results provide a convincing numerical validation of the instanton analysis in the large amplification regime  $\log_{10} U \gg 1$ .

## V. SUMMARY AND DISCUSSION

In this paper, we have numerically tested the validity of the instanton analysis approach to study the diffraction-amplification problem (1) in the large amplification regime  $\ln U \gg 1$ , where  $U = |\psi(0, L)|^2$ ,  $\psi(x, z)$  being the solution to equation (1). By analyzing a large number of numerical solutions to equation (1) with  $S$  in the ‘one-max class’ defined at the beginning of Section III A, we have identified a nonlinear fit to numerical data from which a highly accurate approximation of  $U$  as a function of  $S$  can be obtained, when  $\ln U$  is large

(equation (16)). We have then used this result to devise a sampling procedure of  $S$  giving access to large values of  $\ln U$ .

As a first application, we have obtained  $p(U)$  numerically over a large range of  $U$  with  $\ln U \gg 1$ , down to probability density less than  $10^{-2270}$  in the tail. We have found a near-perfect agreement with the algebraic tail of  $p(U)$  theoretically predicted by the instanton analysis in [1] (figure 6). Then, we have determined the conditional probability distribution of  $\mathfrak{D}$  given  $U$  for a large  $\ln U$ , where  $\mathfrak{D}$  is the  $L^2$ -distance measuring the departure of  $S$  from the theoretical instanton normalized such that  $0 \leq \mathfrak{D} \leq 1$ . We have found that the realizations of  $S$  in the far right tail of  $p(U)$  are significantly closer to the instanton than unconditioned realizations:  $0.08 < \mathfrak{D} < 0.16$  and  $0.6 < \mathfrak{D} \leq 1$  respectively, in the case of figure 7. As a confirmation of this result, plots of  $|S|^2/\|S\|_2^2$  clearly show that the realizations of  $S$  in the far right tail of  $p(U)$  are low-noise instanton realizations around the predicted, noiseless, theoretical instanton, with residual noise due to  $\ln U$  being finite (figures 8 and 9). To summarize, our numerical results validate the instanton analysis of the diffraction-amplification problem (1) in the large  $\ln U$  limit for  $S$  in the one-max class.

In conclusion we briefly discuss the possibility of using more standard sampling methods to deal with the same problem. What makes our approach particularly efficient (compared with standard ones) is that we do not have to know the maximizing direction prior to testing whether any given direction is close to it or not. All we need is to look at how  $\ln U$  depends on  $\mathbf{s}$  around the tested direction (here, the instanton direction). This saves a considerable amount of time as we do not have to set off in search of where the maximizing direction is on the  $2N$ -dimensional unit sphere (with  $N = 100$  for  $S$  in equation (6)). Either the observed behavior of  $\ln U$  indicates that the instanton direction is close to the maximizing direction, and we can move on to the biased sampling procedure around the instanton direction, as explained in Section IV, or it does not and the instanton solution must be rejected. By contrast, a standard algorithm like in, e.g., [7–9] involves probing the  $2N$ -dimensional unit sphere until it converges to the maximizing direction, which requires massive numerical simulations. For each  $\hat{\mathbf{s}}$  probed, problem (1) with the corresponding  $S$  must be solved numerically, and a great many  $\hat{\mathbf{s}}$ 's are likely to be probed before converging to the maximizing direction. (Unless prior knowledge of the instanton to be tested is used to pick a starting direction not too far from the maximizing one.)

On the other hand, more standard approaches can in principle give access to large values

of  $\ln U$  in cases where no analytical prediction is available, unlike the method presented here. They may therefore be used when instanton analysis is too difficult (if not impossible) to carry out. In such cases, the landscape of  $\ln U(\eta, \hat{\mathbf{s}})$  to be explored is a priori totally unknown, which can pose tricky problems the solution of which is likely to make the algorithm even more numerically demanding. In this respect, other general sampling algorithms like, e.g., subset simulation [13, 14], may also be of interest as possible alternatives to important sampling algorithms.

To the best of the author's knowledge, general iterative algorithms for sampling rare event sets have never been used so far to deal with stochastic amplifiers in the large amplification limit. Unless there are hidden fundamental reasons preventing such use, it would open up a very interesting new field to explore. This will be the subject of a future work.

### Acknowledgments

The author warmly thanks Denis Pesme for his interest and valuable advice about the manuscript. He also thanks the anonymous referee of reference [1] whose pertinent and constructive remarks have motivated this work.

### Appendix A: A sufficient condition for $S$ to belong to the one-max class

The class of  $S$  considered in the paper is defined by the condition that, in the large  $\eta$  limit, there is only one direction of  $\mathbf{s}$  that globally maximizes  $\ln U(\eta, \hat{\mathbf{s}})$ . In this appendix, we give a sufficient condition ensuring that a given  $S$  do belong to this class.

For any spectral density  $\varsigma_n$  (not necessarily Gaussian), we take

$$S(x, z) = \sum_{n=-N/2}^{N/2} s_n \sqrt{\varsigma_n} e^{2i\pi n x / \ell} \Phi_n(z), \quad (\text{A1})$$

which generalizes the equation (6), where  $N$  is an even integer and the  $\Phi_n$ 's are continuous function of  $0 \leq z \leq L$ . Let  $B(0, L)$  denote the set of all the continuous paths in  $\Lambda$  satisfying  $x(L) = 0$  and define  $M[x(\cdot)]$  the  $N \times N$  positive definite matrix with components

$$M_{nm}[x(\cdot)] = \sqrt{\varsigma_n \varsigma_m} \int_0^L e^{2i\pi(m-n)x(z)/\ell} \Phi_n(z)^* \Phi_m(z) dz, \quad (\text{A2})$$

in which  $x(\cdot) \in B(0, L)$ . Writing the solution to the equation (1) formally as a Feynman-Kac integral over  $B(0, L)$  and using the equations (A1) and (A2), one gets

$$\begin{aligned}\psi(0, L) &= \int_{x(\cdot) \in B(0, L)} e^{\int_0^L \left[ \frac{im}{2} \dot{x}(\tau)^2 + g |S(x(\tau), \tau)|^2 \right] d\tau} \mathcal{D}x \\ &= \int_{x(\cdot) \in B(0, L)} e^{\int_0^L \left[ \frac{im}{2} \dot{x}(\tau)^2 + g\eta \hat{\mathbf{s}}^\dagger M[x(\cdot)] \hat{\mathbf{s}} \right] d\tau} \mathcal{D}x.\end{aligned}\quad (\text{A3})$$

In the large  $\eta$  limit, the amplification of  $\psi$  is dominated by the most amplified paths, and one has

$$|\psi(0, L)|^2 \asymp \exp \left[ 2g\eta \sup_{x(\cdot) \in B(0, L)} \hat{\mathbf{s}}^\dagger M[x(\cdot)] \hat{\mathbf{s}} \right], \quad (\text{A4})$$

where the symbol  $\asymp$  means asymptotic logarithmic equivalence; i.e., the ratio of the logarithms of the two sides tends to 1 as  $\eta \rightarrow +\infty$ . Writing  $\mu_1[x(\cdot)]$  the largest eigenvalue of  $M[x(\cdot)]$ , it follows immediately from the equation (A4) that

$$\begin{aligned}\max_{\hat{\mathbf{s}}} \ln U(\eta, \hat{\mathbf{s}}) &= \max_{\hat{\mathbf{s}}} \ln |\psi(0, L)|^2 \\ &\sim 2g\eta \max_{\hat{\mathbf{s}}} \sup_{x(\cdot) \in B(0, L)} \hat{\mathbf{s}}^\dagger M[x(\cdot)] \hat{\mathbf{s}} \\ &= 2g\eta \sup_{x(\cdot) \in B(0, L)} \mu_1[x(\cdot)] \quad (\eta \rightarrow +\infty),\end{aligned}\quad (\text{A5})$$

the maximum being reached at  $\hat{\mathbf{s}}$  along the fundamental eigenvectors of  $M[x_{\text{inst}}(\cdot)]$ , where  $x_{\text{inst}}$  is a path of  $B(0, L)$  maximizing  $\mu_1[x(\cdot)]$ . We recall that for the  $S$ 's considered in [1], all the paths maximizing  $\mu_1[x(\cdot)]$  are in  $B(0, L)$  and there is a finite number of such paths. The interested reader will find a mathematically rigorous demonstration of (A5) in [12].

If none of the  $\mu_1[x_{\text{inst}}(\cdot)]$ 's is degenerate and all the corresponding normalized eigenvectors are equal, then  $S$  admits a single and non-degenerate instanton (by definition). In this case, there is only one direction  $\hat{\mathbf{s}}$  at which the maximum of  $\ln U(\eta, \hat{\mathbf{s}})$  in (A5) is reached. Namely, the direction of the fundamental eigenvector common to all the  $\mu_1[x_{\text{inst}}(\cdot)]$ 's. Therefore, admitting a single and non-degenerate instanton is a sufficient condition for  $S$  to belong to the one-max class. It is shown in [1] that for  $S$  given by the equations (6) and (7), there is only one path,  $x_{\text{inst}}(\cdot) \equiv 0$ , that maximizes  $\mu_1[x(\cdot)]$ , and the corresponding eigenvalue  $\mu_1[x_{\text{inst}}(\cdot) \equiv 0]$  is non-degenerate. Thus,  $S$  in equations (6) and (7) admits a single and non-degenerate instanton, hence it belongs to the one-max class.

---

[1] Mounaix Ph 2023 *J. Phys. A: Math. Theor.* **56** 305001

- [2] Rajaraman R 1982 *Solitons and Instantons: an Introduction to Solitons and Instantons in Quantum Field Theory* (Amsterdam: North-Holland)
- [3] Schäfer T and Shuryak E V 1998 *Rev. Mod. Phys.* **70** 323
- [4] Falkovich G, Kolokolov I, Lebedev V and Migdal A 1996 *Phys. Rev. E* **54** 4896
- [5] Rose H A and DuBois D F 1994 *Phys. Rev. Lett.* **72** 2883
- [6] Hammersley J M and Morton K W 1956 *Math. Proc. Cambridge Philos. Soc.* **52** 449
- [7] Hartmann A K 2014 *Phys. Rev. E* **89** 052103
- [8] Hartmann A K, Le Doussal P, Majumdar S N, Rosso A and Schehr G 2018 *Europhys. Lett.* **121** 67004
- [9] Hartmann A K, Meerson B and Sasorov P 2019 *Phys. Rev. Research* **1** 032043
- [10] Rose H A and DuBois D F 1993 *Phys. Fluids B* **5** 590
- [11] Strang G 1968 *SIAM J. Numer. Anal.* **5** (3) 506
- [12] Mounaix Ph, Collet P and Lebowitz J L 2006 *Commun. Math. Phys.* **264** 741 and 2008 *Commun. Math. Phys.* **280** 281
- [13] Au S-K and Beck J L 2001 *Probabilistic Engineering Mechanics* **16** 263
- [14] Papaioannou I, Betz W, Zwirgmaier K and Straub D 2015 *Probabilistic Engineering Mechanics* **41** 89