PHYSICAL REVIEW D **92,** 085030 (2015)

Thimble regularization at work: From toy models to chiral random matrix theories

F. Di Renzo and G. Eruzzi

Dipartimento di Fisica e Scienze della Terra, Università di Parma and INFN, Gruppo Collegato di Parma I-43100 Parma, Italy

(Received 22 July 2015; published 20 October 2015)

We apply the Lefschetz thimble formulation of field theories to a couple of different problems. We first address the solution of a complex zero-dimensional ϕ^4 theory. Although very simple, this toy model makes us appreciate a few key issues of the method. In particular, we will solve the model by a correct accounting of all the thimbles giving a contribution to the partition function and we will discuss a number of algorithmic solutions to simulate this (simple) model. We will then move to a chiral random matrix (CRM) theory. This is a somehow more realistic setting, giving us once again the chance to tackle the same couple of fundamental questions: How many thimbles contribute to the solution? How can we make sure that we correctly sample configurations on the thimble? Since the exact result is known for the observable we study (a condensate), we can verify that, in the region of parameters we studied, only one thimble contributes and that the algorithmic solution that we set up works well, despite its very crude nature. The deviation of results from phase quenched ones highlights that in a certain region of parameter space there is a quite important sign problem. In view of this, the success of our thimble approach is quite a significant one.

DOI: 10.1103/PhysRevD.92.085030 PACS numbers: 11.15.Ha

I. INTRODUCTION

The so-called sign problem is one of the current big challenges for lattice field theories. It is in fact the major obstacle to tackling a nonperturbative study of the QCD phase diagram. Following pioneering work by Witten [1], Lefschetz thimble regularization has been proposed as a possible solution [2,3] (for more recent contributions, see also [4–7]): the functional integral is defined in terms of fields taking values on nontrivial manifolds on which the imaginary part of the action stays piecewise (i.e., on the distinct thimbles attached to different critical points) constant. It is an elegant, although with many respects nontrivial alternative to the standard formulation of field theories. It has intriguing connections with resurgence theory, a few results of which motivate the conjecture that *the semiclassical expansion of the path integral can be geometrized as a sum over Lefschetz thimbles*[1] [8]. Morse theory [9] is the natural framework for discussing the thimble regularization, even though it could be that it does not necessarily have the last word on the subject. In this work we will discuss the thimble solution of two different models, having in mind two big issues. First of all, since there is a thimble attached to every critical point of the (complexified) theory one is considering, we need to understand how many thimbles do give contribution to the solution of a given theory. A second relevant problem is that of devising Monte Carlo algorithms to correctly sample thimbles, which are manifolds for which we lack a local description.

Toy models can be a precious tool to approach hard problems in theoretical physics, with the hope that a simplified model can nevertheless capture the relevant issues. Given that the sign problem is both relevant and hard, it does not come as a surprise that toy models have been around for a while. Quite interestingly, some of these have been resisting the efforts to solve them in much the same way as the real problem is still far from being fully solved. The first application we discuss is the solution of a toy model that dates back to almost 30 years ago [10]. It can be regarded as a zero-dimensional version of a ϕ^4 field theory. This simple model became a benchmark for the complex Langevin treatment of the sign problem, and quite interestingly only partial success has been claimed over the years. We will show that the thimble regularization completely solves the model. In this case we will show how in different regimes one or more thimbles give contribution to the solution. Moreover, it is possible to perform numerical simulations on the thimble(s): in this simple case we will have a number of algorithmic solutions at work. A partial account of these results has already been given in [11].

We will then address the solution of a chiral random matrix theory. This is a somehow more realistic problem, for which once again the application of the complex Langevin method has been shown to be nontrivial: in a given parametrization it fails [12], while in a different one it works [13]. Actually one can show that the sign problem can be quite severe for this model. The theory has an adjustable parameter (the dimension N of the matrices) which controls the dimensionality of the problem one has to

[1]This is a literal citation from [8]; we will have more to quote later on the subject of resurgence.

solve. Since the analytical solution for the observable we will study (a mass condensate) is known, it appeared to us a perfect setting for testing a conjecture: it could well be that more than one thimble contributes (just like in the zero-dimensional ϕ^4 theory) in low dimensions, but as N grows it could be that a single thimble dominates in the thermodynamic limit. While we were ready for a richer scenario, in the region of parameters we studied we actually did not find any other thimble but the one attached to the global minimum. We did not find any problem related to the parametrization of the theory; in particular, the parametrization that was failing in the case of the complex Langevin was absolutely fine for the thimble treatment of the theory. In this case algorithmic problems are nontrivial. We will show how a natural parametrization of the thimble can be the starting point for an algorithmic solution that in this case works in its simplest version (admittedly a very crude one).

The paper is organized as follows. In Sec. II we give a brief account of the Lefschetz thimble approach to field theories: this is a short review of results that are collected to facilitate the reader. Section III is dedicated to the zero-dimensional ϕ^4 toy model, showing that thimbles provide a complete solution; in particular we show that we can effectively numerically simulate the model, making use of different algorithms. In Sec. IV we address the chiral random matrix theory, showing that we can numerically solve it by thimble regularization: all the analytical results are correctly reconstructed. In the final section we draw a few conclusions and mention natural steps forward.

II. THIMBLE REGULARIZATION IN A NUTSHELL

In the following we collect the basics of thimble regularization: the interested reader is referred to [1–3] for further details and references.

A conceptual starting point to approach the thimble regularization is that of generalizing saddle point integration. The latter displays a couple of features which appear as good candidates to tackle the sign problem: stationary phase and localization. The full generalization of saddle point techniques is formulated in the framework of Morse theory [9].

(i) One starts with an integral on a real domain of the form

$$\int_{\mathcal{C}} d^n x g(x_1, \ldots, x_n) \mathrm{e}^{-S(x_1,\ldots,x_n)}, \tag{1}$$

in which $\mathcal{C}$ is a real domain of real dimension n[2] and both $S(x_1, \ldots, x_n) = S_R(x_1, \ldots, x_n) + iS_I(x_1, \ldots, x_n)$ and $g(x_1, \ldots, x_n)$ are holomorphic functions. The notation for the exponential makes it clear that we have already in mind a functional integral (even if the normalizing factor Z^{-1} is missing). For such an integral the following decomposition holds:

$$\int_{\mathcal{C}} d^n x g(x_1, \ldots, x_n) \mathrm{e}^{-S(x_1,\ldots,x_n)} = \sum_\sigma n_\sigma \int_{\mathcal{J}_\sigma} d^n z g(z_1, \ldots, z_n) \mathrm{e}^{-S(z_1,\ldots,z_n)}, \tag{2}$$

in which an extension from a real domain to a complex one has been performed (see the complex variables $z_i = x_i + iy_i$ as opposed to the real ones x_i). Equation (2) holds in the homological sense, i.e., $\mathcal{C} = \sum_\sigma n_\sigma \mathcal{J}_\sigma$.

(ii) The index σ labels the critical points of the complex (ified) $S(z_1, \ldots, z_n)$ and to each critical point p_σ a stable thimble $\mathcal{J}_\sigma$ is attached. Each $\mathcal{J}_\sigma$ is defined as the union of all the steepest ascent (SA; we will also write SD for steepest descent) paths falling into p_σ at (minus) infinite time, i.e., the union of the solutions of

$$\frac{dx_i}{d\tau} = \frac{\partial S_R(x,y)}{\partial x_i}$$
$$\frac{dy_i}{d\tau} = \frac{\partial S_R(x,y)}{\partial y_i}, \tag{3}$$

satisfying $z(\tau=-\infty)=x(\tau=-\infty)+iy(\tau=-\infty)=p_\sigma$. The real dimension of each thimble $\mathcal{J}_\sigma$ is n. It is quite natural to regard thimbles as manifolds embedded in $\mathbb{C}^n$ (which is instead of real dimension $2n$).

(iii) For each critical point one also defines unstable thimbles $\mathcal{K}_\sigma$ as the union of all flows satisfying Eq. (3) and going to the critical point in the opposite time limit, i.e., such that $z(\tau = \infty) = p_\sigma$. The coefficients n_σ count the intersections of the $\mathcal{K}_\sigma$ with the original domain of integration $n_\sigma = \langle \mathcal{C}, \mathcal{K}_\sigma \rangle$.

(iv) The imaginary part S_I stays constant on thimbles; i.e., there is a phase associated to each thimble.

Note that in the framework of field theories a natural picture of universality emerges. A single thimble can give us a formulation of a field theory with the same degrees of freedom, the same symmetries[3] and symmetry

[2]In what follows we will be a little bit sloppy in our notation: whenever there is not subscript attached to a symbol (e.g., x), that will denote an n-dimensional coordinate.

[3]Since we have discussed the case of a nondegenerate Hessian, one could wonder how the method can be applied in the case where spontaneous symmetry breaking (SSB) is in place. In [14] a solution has been described and shown to be effective: one introduces an explicit symmetry breaking term h and studies the limit $h \to 0$. This is not the only way to proceed: for a different thimble approach to SSB the reader is referred to [3]. Symmetries are dealt in yet another way in the case of gauge theories: the construction of thimbles was discussed in [15] and reviewed in [1]; see also the discussion in [2,16].

representations, and the same perturbation theory and naive continuum limit of the original formulation (see [2] for details). In force of universality we expect that these properties essentially determine the behavior of physical quantities in the continuum limit. Moreover a simple argument suggests that in the thermodynamic limit only thimbles attached to global minima can survive, as it is easily seen (we now call ϕ_σ the critical points, having in mind field configurations, and consider the partition function of the field theory)

$$Z = \sum_\sigma n_\sigma e^{-iS_I(\phi_\sigma)} \int_{\mathcal{J}_\sigma} d^n z e^{-S_R(z)} = \sum_\sigma n_\sigma e^{-S(\phi_\sigma)} \int_{\mathcal{J}_\sigma} d^n z e^{-(S_R(z)-S_R(\phi_\sigma))}.$$

In the end, it could well be that a full resolution in terms of all the thimbles could turn out to be in many respects overkilling the original problem. Having said this, we nevertheless stress that both the universality argument and the thermodynamic argument cannot be regarded as conclusive. It is worth noting that resurgence theory [17] in many respects even *asks for* more than one thimble in view of the interpretation of the semiclassical approximations as trans-series.[4] All in all, the fact that in certain cases a single thimble dominance can take place has to be regarded as a conjecture: as we will see, this was in a sense one of the motivations of this work on a random matrix model. The subject deserves deeper investigation and we think it will certainly receive it.

It is good to have a somehow more constructive approach to the thimble formulation. We therefore now sketch a few more technical details that the reader will see at work in the following sections. The integral we have in mind will be the functional integral of a field theory. Let us first of all parametrize the field in the vicinity of a critical point as $\Phi_i = \phi_i - \phi_{\sigma,i}$. Here and in the following i is a multi-index; in particular it can refer to a real or imaginary part. The real part of the action can be expressed as

$$S_R(\phi) = S_R(\phi_\sigma) + \frac{1}{2}\Phi^T H \Phi + \mathcal{O}(\phi^3), \tag{4}$$

where the $2n \times 2n$ matrix H is the Hessian evaluated at the critical point

$$H_{ij} = \left.\frac{\partial^2 S_R}{\partial\phi_i \partial\phi_j}\right|_{\phi=\phi_\sigma}.$$

H can be put in diagonal form,

$$\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n, -\lambda_1, \ldots, -\lambda_n),$$

by a transformation $H = W\Lambda W^T$ defined by the orthogonal matrix W whose columns are given by the normalized eigenvectors of H, that is, $\{v^{(i)}\}_{i=1\cdots 2n}$. Half of the eigenvalues of H are positive; the corresponding eigenvectors span the tangent space to the thimble at the critical point. Any combination of these vectors is a direction along which the real part of the action grows. If we leave the critical point along these directions integrating the SA equations we span the thimble. On the other side, the other directions (which are attached to the negative eigenvalues) would take us along the unstable thimble.

At a generic point $Z \in \mathcal{J}_\sigma$ we miss *a priori* the knowledge of the tangent space $T_Z\mathcal{J}_\sigma$; in general we expect that the latter is not parallel to the canonical basis of $\mathbb{C}^n$ whose duals appear in $\mathrm{d}^n z = \mathrm{d}z_1 \wedge \cdots \wedge \mathrm{d}z_n$. We thus want to perform the relevant change of coordinates from the canonical ones (of $\mathbb{C}^n$) to the basis of $T_Z\mathcal{J}_\sigma$, given by the (complex) vectors $\{U^{(i)}\}_{i=1\cdots n}$ (these are orthonormal with respect to the standard Hermitian metric of $\mathbb{C}^n$). Let $\varphi: N \subset \mathcal{J}_\sigma \to \mathbb{R}^n$ be a local chart in a neighborhood $N \subset \mathcal{J}_\sigma$ of Z:

$$\varphi\left(Z + \sum_{i=1}^n U^{(i)} y_i\right) = Y + \mathcal{O}(y^2) \in \mathbb{R}^n.$$

If we denote U the $n \times n$ complex unitary matrix whose columns are the vectors $\{U^{(i)}\}$, we can express the integral of a generic function $f(Z)$ on the thimble as

$$\int_N \mathrm{d}^n z f(Z) = \int_{\varphi(N)} \prod_{i=1}^n \mathrm{d}y_i f(\varphi^{-1}(Y)) \det U(\varphi^{-1}(Y)). \tag{5}$$

In this expression the quantity $\det U = e^{i\omega}$ (U is unitary) has appeared; this is what has been termed the residual phase [2] (see [19] for further details). This could in principle reintroduce a sign problem in the thimble formulation, but it is expected that this is not the case. Not every phase gives rise to a serious sign problem, and in particular one expects that a phase changing rather smoothly can be safely taken into account by reweighting. This expectation could appear optimistic, but it has been till now confirmed (see [3]) and will be confirmed also in this work.

We end this brief introduction to the thimble formulation by going back to the constructive point of view: we can span the thimble by integrating the SA equations for the field ϕ[5]:

[4] A nice account of many issues connected to resurgence has been recently provided in [18].

[5] It is worth recalling here that the subscript i is a multi-index, in which real and imaginary parts are on the same footing.

$$\frac{d\phi_i}{dt} = \frac{\partial S_R}{\partial \phi_i} \quad i = 1 \cdots 2n.$$

This has a counterpart in parallel-transport equations for the n basis vectors which defines the tangent space to the thimble (see [2,3]):

$$\frac{dV_j^{(i)}}{dt} = \sum_{k=1}^{2n} \frac{\partial^2 S_R}{\partial \phi_k \partial \phi_j} \quad V_k^{(i)} i = 1 \cdots n \quad j = 1 \cdots 2n. \tag{6}$$

We can set up a similar equation for any other vector with an initial condition on the tangent space at the critical point; (6) expresses the parallel transport of a vector along the gradient flow. In the vicinity of the critical point one knows the asymptotic ($t \to -\infty$) solutions

$$t \ll 1 \begin{cases} \phi_j(t) \approx \phi_{\sigma,j} + \sum_{i=1}^{n} v_j^{(i)} e^{\lambda_i t} n_i & j = 1 \cdots 2n \quad |\vec{n}|^2 = 1 \\ V_j^{(i)}(t) \approx v_j^{(i)} e^{\lambda_i t} & j = 1 \cdots 2n \quad i = 1 \cdots n \end{cases}. \tag{7}$$

Note that from a practical point of view the former parametrization is viable only provided one introduces a reference time $t_0 \ll 1$ at which the former asymptotic solution holds. We will make extensive use of the former equations, which in particular can be regarded as initial conditions for a given flow on the thimble, e.g., for Eq. (6). A natural picture thus emerges in which a generic point $\Phi \in \mathcal{J}_\sigma$ is unambiguously defined by a choice of $\hat{n}$[6] and the time t: $\Phi = \Phi(\hat{n}, t)$ (this has been very effectively discussed in [3]). Note also that one could insist on regarding (7) as valid all over; this would in turn mean one is considering a purely quadratic action (i.e., the free field approximation).

III. THE ZERO-DIMENSIONAL ϕ^4 TOY MODEL

We will now put to work what we have just seen, applying the thimble regularization to the study of the action

$$S(\phi) = \frac{1}{2}\sigma\phi^2 + \frac{1}{4}\lambda\phi^4,$$

with $\phi \in \mathbb{R}$, $\lambda \in \mathbb{R}^+$, and $\sigma = \sigma_R + i\sigma_I \in \mathbb{C}$. This is obviously a toy model, and one regards as correlators plain one-dimensional integrals such as

$$\langle \phi^n \rangle = \frac{1}{Z} \int_{\mathbb{R}} d\phi \phi^n e^{-S(\phi)}, \tag{8}$$

[6]The hat notation reminds us of the normalization condition; i.e., $\hat{n}$ singles out a direction in the tangent space at the critical point.

with the partition function given by

$$Z = \int_{\mathbb{R}} d\phi e^{-S(\phi)}.$$

The solution is given in terms of a modified Bessel function, i.e., $Z = \sqrt{\frac{\sigma}{2\lambda}} e^{\frac{\sigma^2}{8\lambda}} K_{-\frac{1}{4}}(\frac{\sigma^2}{8\lambda})$, differentiating appropriately which one can get any of Eq. (8). The choice of a complex σ is a prototypal case of the sign problem: with a complex action, we miss a positive semidefinite measure and hence a probability distribution to start with; in particular, a direct access to Monte Carlo methods is ruled out.

It was realized a long time ago that a solution to the sign problem could be sought in the context of stochastic quantization: the Langevin equation admits a formal solution also for complex actions, in particular via the Fokker-Planck formulation [20,21]. Turning the formal arguments into a rigorous proof eventually turned out to be hard and numerical instabilities (suggesting problems) were in particular discussed in the context of the theory at hand [10]. Much experience has been gained over the years and much progress has been done [22]. The question of convergence of the complex Langevin equation remains a subtle one, and quite interestingly even the simple model at hand displays delicate issues. For a recent and thorough study of the complex Langevin dynamics of this model, the reader can refer to [23]. One peculiar feature of this model is that complex Langevin simulations display divergences for $\langle \phi^n \rangle$ with $n > 4$ in a certain region of parameters. The relation between complex Langevin and thimbles has been investigated in [24,25].

We can complexify the field by setting $\phi = x + iy$. As a result, the real and imaginary parts of the action read

$$S^R = \frac{1}{2}[\sigma_R(x^2 - y^2) - 2\sigma_I xy] + \frac{1}{4}\lambda(x^4 + y^4 - 6x^2y^2)$$
$$S^I = \frac{1}{2}[\sigma_I(x^2 - y^2) + 2\sigma_R xy] + \lambda(x^3 y - xy^3).$$

The Hessian is built from the second derivatives of S_R and takes the form

$$H(x,y) = \begin{pmatrix} \sigma_R + 3\lambda x^2 - 3\lambda y^2 & -\sigma_I - 6\lambda xy \\ -\sigma_I - 6\lambda xy & -\sigma_R - 3\lambda x^2 + 3\lambda y^2 \end{pmatrix}. \tag{9}$$

There are three critical points: $\phi_0 = 0$ and $\phi_\pm = \pm\sqrt{-\frac{\sigma}{\lambda}}$ (which are the two, complex valued "Higgs vacua"). The question is now which thimbles do give a contribution to the integrals we want to compute, and the answer is quite different in the three cases $\sigma_R > 0$, $\sigma_R < 0$, and $\sigma_R = 0$: in each case we computed the stable and unstable thimbles associated to each critical point. This can be done by putting to work the constructive definition of thimbles we discussed in the previous section.

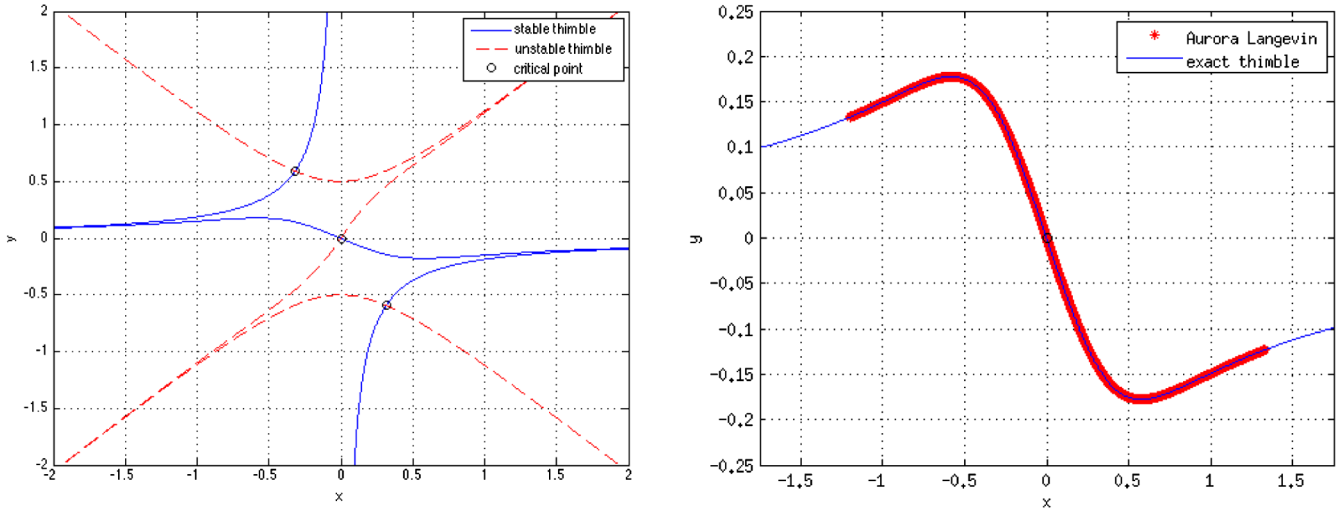

FIG. 1 (color online). Thimbles structure for $\sigma = 0.5 + i0.75$, $\lambda = 2$ (left panel). In this case only the unstable thimble attached to $z = 0$ intersects the real axis and thus only one critical point contributes. On the right we can see how the Langevin simulation correctly covers the relevant thimble.

In practice, we want to integrate the equations of SA starting in the vicinity of the critical point ϕ_σ for an arbitrarily long flow time t. We can do this provided that the initial condition is chosen correctly: for the stable thimble this means we leave the critical point along the direction (in the xy plane) which is given by the eigenvector of positive eigenvalue of the Hessian (9) computed at the critical point. Once we have singled out the relevant direction, we can ascend in two ways (namely, increasing or decreasing x), both of which we have to take to cover the entire thimble. By holomorphicity the Hessian has two eigenvalues opposite in sign. Since S_R always increases along the flow, $\exp(-S_R)$ goes to 0 as $t \to +\infty$, thus ensuring convergence of the integrals along the thimble. To obtain the unstable thimble $\mathcal{K}_\sigma$, we can repeat the same procedure described above, but picking up the eigenvector of the Hessian of S_R with negative eigenvalue. Note that the unstable thimble is needed because the coefficient n_σ in our master equation (2) counts the intersection of such thimbles with the original domain of integration, which in our case is the real axis (the sign ambiguity is not resolved just by this definition, but it can be deduced by means of other considerations). Figures 1 (left panel) and 2 show the results for the three cases $\sigma_R > 0$, $\sigma_R < 0$, and $\sigma_R = 0$ (see also [26]).

From Fig. 1, we see that when $\sigma_R > 0$ the unstable thimbles related to the Higgs vacua do not intersect the real axis. Therefore these points do not contribute to the integrals, that is, $n_\pm = 0$ and $n_0 = 1$. By integrating along

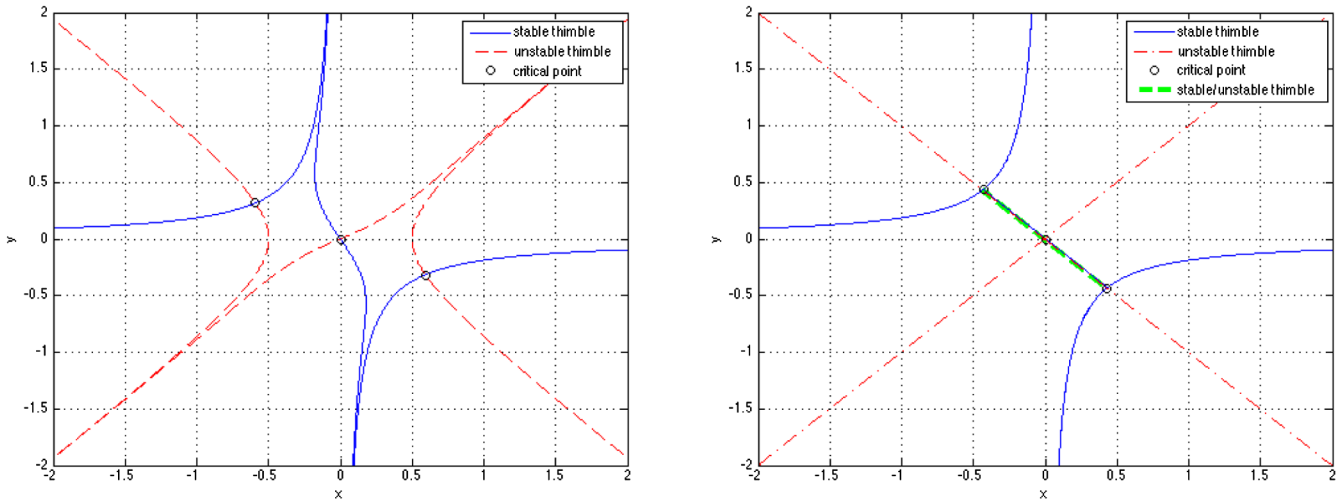

FIG. 2 (color online). Thimbles structure for $\sigma = -0.5 + i0.75$ (left panel) and $\sigma = i0.75$ (right); in both cases $\lambda = 2$. For $\sigma_R < 0$ (left) all three critical points contribute. $\sigma_R = 0$ (right) is an example of a Stokes phenomenon.

the stable thimble attached to ϕ_0, we recover the correct results for, say, $Z = \int e^{-S}$ (the integration can be easily carried on along the real axis both analytically and numerically). The case $\sigma_R < 0$ depicted in the left panel of Fig. 2 is a totally different matter, as we cross the Stokes ray $\sigma_R = 0$ while changing sign to σ_R. Now we see that the unstable thimbles connected to the Higgs vacua do intersect the real axis and therefore $n_\pm \neq 0$, as well as $n_0 \neq 0$. The correct combination which recovers the expected results for the integrals turns out to be $n_0 = -1$ and $n_\pm = +1$. What is the origin of this discontinuity? Above all, if we had not known the correct result from the beginning, how would we have calculated the n_σ? The answer lies in considering the case $\sigma_R = 0$, shown in the right panel of Fig. 2. The stable thimble connected to 0 exhibits the Stokes phenomenon: in fact it "collapses" into the Higgs vacua, from which it does not "move" any more; the unstable thimble continues to say that $n_0 \neq 0$. The stable thimbles connected to the Higgs vacua display the same shape, but their unstable counterparts collapse into 0 (by overlapping its stable thimble) and therefore there is intersection with the real axis; so, $n_\pm \neq 0$. However, there is no integer-valued combination of n_σ that recovers the correct results for $\sigma_R = 0$. This is quite expected, as the Morse decomposition along thimbles is not legitimate when we are on a Stokes ray, on which we clearly are (the imaginary axis in the complex σ plane is a "Stokes ray").[7] Now, the original integral is continuous (in fact, it is holomorphic) in σ and therefore there cannot be any discontinuity in the computation of the partition function Z in $\sigma_R = 0$. Thus, we must have $Z[\sigma_R \to 0^+] = Z[\sigma_R \to 0^-] = Z[\sigma_R = 0]$. By examining the integration along the thimble connected to 0, we find that it is discontinuous in $\sigma_R = 0$, and again, this is not surprising as the thimble shape undergoes a radical change between the two cases. The change in sign of the n_σ is precisely the only one which keeps the original integral continuous while crossing the Stokes ray.

A. A variety of algorithmic solutions

Within the thimble regularization we were able to perform numerical simulations of the quartic toy model, making use of different algorithms. In particular, we were able to numerically compute all the possible moments (8).

It was observed in [2] that the Langevin algorithm is the obvious candidate for sampling configurations on the thimble. In

$$\frac{\mathrm{d}\phi_i}{\mathrm{d}t} = -\frac{\partial S_R}{\partial \phi_i} + \eta_i, \quad i = 1 \cdots 2n, \tag{10}$$

the drift term constrains the field on the thimble by definition, so that the problem boils down to extracting a convenient noise, i.e., a noise tangent to the thimble. We do not discuss here the original solution which was put forward in [2] (the Aurora algorithm); there will be a convenient time for such a discussion when we later approach the CRM model. Here it suffices to say that, the thimble being one dimensional, at every point the tangent space reduces to the direction singled out by the drift term itself. As a matter of fact, Langevin works pretty well; in the right panel of Fig. 1 one can see how the simulation correctly samples configurations on the thimble. Here parameters are the same as in the left panel, so one thimble is relevant, i.e., the one attached to the origin [which in the notation of (2) we denote p_0]. The (Aurora) Langevin algorithm samples points according to the measure normalized by[8]

$$\mathcal{Z}^{(0)} \equiv \int_{\mathcal{J}_0} \mathrm{d}\tau e^{-S_R}. \tag{11}$$

We now denote

$$\langle O \rangle_0 \equiv \frac{\int_{\mathcal{J}_0} \mathrm{d}\tau O e^{-S_R}}{\mathcal{Z}^{(0)}} \tag{12}$$

and stress that this is not what we have to compute. Properly including the residual phase, the correct result was computed as

$$\langle O \rangle = \frac{\langle e^{i\omega} O \rangle_0}{\langle e^{i\omega} \rangle_0}. \tag{13}$$

When $\sigma_R < 0$ the thimbles associated to all three critical points[9] contribute and we have to compute

$$\langle O \rangle = \frac{\sum_{i=0}^{2} n_i e^{-iS_I(p_i)} \int_{\mathcal{J}_i} \mathrm{d}\tau e^{-S_R} O e^{i\omega}}{\sum_{i=0}^{2} n_i e^{-iS_I(p_i)} \int_{\mathcal{J}_i} \mathrm{d}\tau e^{-S_R} e^{i\omega}}, \tag{14}$$

which can be written

$$\langle O \rangle = \frac{\langle e^{i\omega} O \rangle_0 + \alpha_1 \langle e^{i\omega} O \rangle_1 + \alpha_2 \langle e^{i\omega} O \rangle_2}{\langle e^{i\omega} \rangle_0 + \alpha_1 \langle e^{i\omega} \rangle_1 + \alpha_2 \langle e^{i\omega} \rangle_2}, \tag{15}$$

with

$$\alpha_i = \frac{n_i e^{-iS_I(p_i)}\ \mathcal{Z}^{(i)}}{n_0 e^{-iS_I(p_0)}\ \mathcal{Z}^{(0)}}, \quad i = 1, 2. \tag{16}$$

On each thimble $\mathcal{J}_i$ ($i = 0, 1, 2$) the quantities $\langle e^{i\omega} O \rangle_i$ and $\langle e^{i\omega} \rangle_i$ can be computed via (Aurora) Langevin simulations. The (complex) unknown coefficients α_i can then be fixed by relations which can be regarded as *renormalization conditions in a physical scheme*, i.e.,

[7] See [1] for a detailed explanation of the Stokes phenomenon with respect to the Airy integral.

[8] τ is the real coordinate on the (one-dimensional) thimble.

[9] In the notation of (2) we now denote $\phi_+ = p_1$ and $\phi_- = p_2$.

$$\frac{\langle e^{i\omega} O_i\rangle_0 + \alpha_1\langle e^{i\omega} O_i\rangle_1 + \alpha_2\langle e^{i\omega} O_i\rangle_2}{\langle e^{i\omega}\rangle_0 + \alpha_1\langle e^{i\omega}\rangle_1 + \alpha_2\langle e^{i\omega}\rangle_2} = X_i, \quad i = 1, 2, \tag{17}$$

where the X_i are known values of given observables O_i [e.g., in the case of moments (8), two of them]. As always in such an approach, one gives up predicting everything, but after normalizing results to a (minimum) number of external inputs, one has full predictive power for (all the) other quantities. Of course computing the moments (8) for the toy model at hand is not such a big numerical success; nevertheless the outline of the method is quite general. In particular, we will refer to it in Sec. IV C.

Another algorithmic solution for this simple setting is provided by the Metropolis algorithm which is described in [27]. The method relies on a correspondence between the full model one has to simulate and a Gaussian approximation associated to it. The latter is obtained by diagonalizing the Hessian at a critical point and truncating the expansion of the action around it, i.e.,

$$S_R(\eta) = S_R(\phi_\sigma) + \frac{1}{2}\sum_{k=1}^{2n} \lambda_k \eta_k^2 \equiv S_R(\phi_\sigma) + S_G(\eta), \tag{18}$$

where the η_k are the $\Phi_k = \phi_k - \phi_{\sigma,k}$ of Eq. (4) expressed in the basis provided by the eigenvectors of the Hessian, with a convenient ordering in which $\lambda_k > 0$ for $k = 1 \ldots n$ and $\lambda_k < 0$ for $k = n+1 \ldots 2n$. For the Gaussian action (18) it is very simple to construct the associated stable thimble. It is a *flat* thimble in which the tangent space is known once and for all, i.e., the span of the eigenvectors associated to $\{\lambda_k | k = 1 \ldots n\}$: we term it a *Gaussian thimble*. For the Gaussian thimble the solution in the right-hand side of (7) is valid all over the manifold.

The simulation is run as a quite standard Metropolis algorithm controlled by an accept/reject test, with a mechanism for proposing configurations which is dictated by the correspondence between the thimble one has to sample and its Gaussian approximation. We sketch the method in the case of more than one thimble contributing to the final result, to stress how also in this case we were able to run numerical simulations on thimbles, for both $\sigma_R > 0$ and $\sigma_R < 0$.

The method always handles a couple of configurations, i.e., one ϕ field on the thimble we have to sample and one auxiliary η field on the associated Gaussian thimble. In order to extract a new ϕ' field one proceeds as follows:

(i) One proposes a thimble σ' (i.e., a critical point) with a probability

$$\frac{|n_{\sigma'}|}{\sum_\sigma |n_\sigma|}.$$

(ii) One extracts a configuration η' on the Gaussian thimble associated to that critical point according to the weight e^{-S_G}. This is trivial, given the Gaussian form.

(iii) One starts a SD on the Gaussian thimble with η' as initial condition. The integration is carried on over a time extent $\bar{\tau}$ such that one ends up close enough to the critical point, namely, at a point where the Gaussian thimble and the thimble one has to sample effectively sit on top of each other [this means that the solution (7) holds for both thimbles]. We call $\bar{\eta}$ the configuration that has been obtained in this way.

(iv) Taking $\bar{\eta}$ as the initial condition, one integrates the SA equations for the complete theory over the same time extent $\bar{\tau}$. This generates the new configuration ϕ'.

(v) ϕ' is accepted with probability

$$P_{acc} = \min\left\{1, e^{-[S_R(\phi') - S_R(\phi)] + [S_G(\eta') - S_G(\eta)]}\right\}.$$

The result for a given observable O is obtained as

$$\langle O\rangle = \frac{\frac{1}{T}\sum_{t=1}^{T} e^{i\omega(\phi_t)} O(\phi_t)}{\frac{1}{T}\sum_{t=1}^{T} e^{i\omega(\phi_t)}},$$

where the index t runs over all the configurations sampled by Metropolis.

Notice that the previous accounting of the Metropolis algorithm is technically different from the proposal of [27].[10] The latter relies on an exponential mapping for the integration time (i.e., $r = e^{-t}$) and an adjustable parameter is introduced to control convergence properties (the interested reader can refer to Fig. 4 of [11]). For a given, effective choice of this parameter Fig. 3 displays how the three thimbles giving contribution in the region $\sigma_R < 0$ are sampled in a Metropolis simulation.

We stress that also in this case one could think of situations in which the weights n_σ are unknown. However, here the situation is different from that of Langevin, since we only need to know a few integers' values. In other terms, given the knowledge of the set of relevant integers $\{n_\sigma\}$, the problem is solved on an entire region of parameter space: in the case at hand, for each $\sigma_R < 0$ (technically, over the entire region which ends up in a point where a Stokes phenomenon shows up). Notice that in principle there can be different ways of finding the relevant set of integers (e.g., known asymptotic solutions in a convenient region).[11]

Both (Aurora) Langevin and Metropolis could correctly compute the moments (8), in both regions $\sigma_R > 0$ and $\sigma_R < 0$. For example, Fig. 3 (right panel) displays the

[10] We decided to enlighten the rationale of the algorithm, leaving out the technicalities.

[11] The α_i introduced for the (Aurora) Langevin algorithm entail instead the values of partition functions and are given at a given point of parameter space; in the case at hand, for a given value of $\sigma_R < 0$.

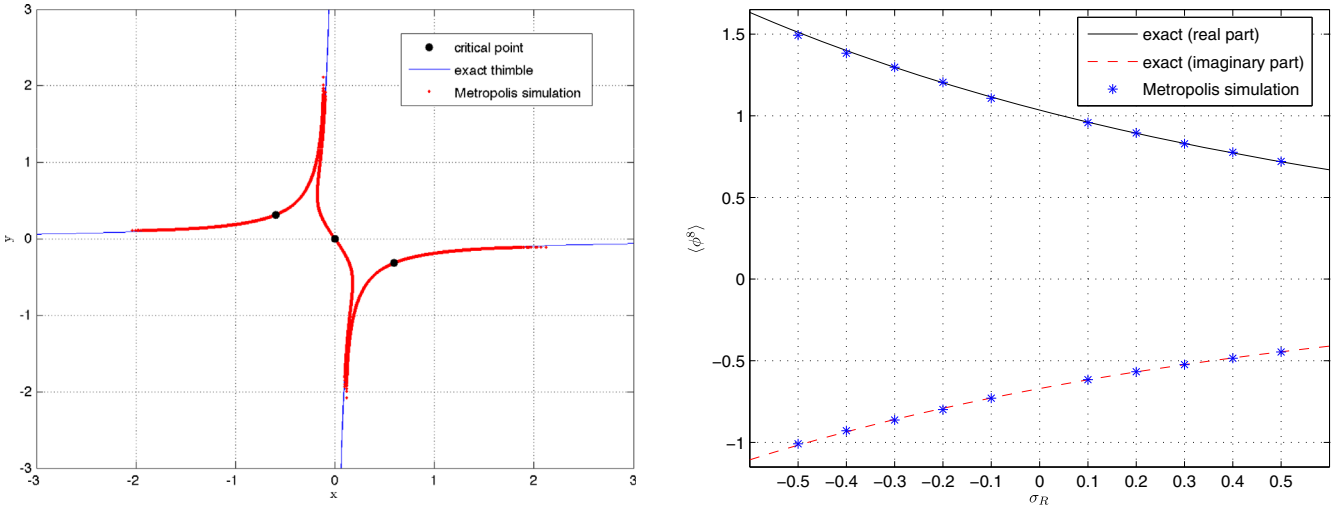

FIG. 3 (color online). Left panel: The three thimbles associated to $\sigma = -0.5 + i0.75$ correctly sampled by a Metropolis simulation. Right panel: For the same choice of parameters, the computed values of $\langle\phi^8\rangle$ over a range of both $\sigma_R > 0$ and $\sigma_R < 0$.

computed values of $\langle\phi^8\rangle$ over a range of both $\sigma_R > 0$ (one thimble being relevant) and $\sigma_R < 0$ (three thimbles to be taken into account) values.

Note that there is another natural way of computing on a thimble, and this takes advantage of the fact that on a thimble there is a one-to-one correspondence in between configurations and values of S_R. To make things simple, let us consider the case in which only one thimble is relevant and let us write the partition function $Z = Z_{\rm up} + Z_{\rm down}$: these are the two contributions resulting from the two pieces of the thimble we have already referred to. Namely, they are associated to leaving the critical point along the direction dictated by the eigenvector of the Hessian in one of the two possible ways (i.e., increasing or decreasing x values). Each $Z_{\rm up/down}$ has the global phase $e^{-iS_I(p_\sigma)}$ as a factor and features an integrand which is the residual phase times a monotonic function of S_R. It thus can be written taking the action as the integration variable, e.g.,

$$Z_{\rm up} = e^{-iS_I(p_\sigma)} \int_{S_{p_\sigma}}^{\infty} dS e^{-S_R} |\nabla S_R|^{-1} e^{i\tan^{-1}(\partial_y S_R/\partial_x S_R)}. \quad (19)$$

We could have written the integral by taking the flow time as the integration variable (also in this case there is a one-to-one correspondence with the configurations along the thimble, each reached at a given flow time). Note that in computing (19) one proceeds by integrating the SA. We illustrated the issue by taking into account the Z, but we showed that all the moments (8) can be successfully computed in this way. In a sense, (19) is the prototype of a parametrization we will see at work for the CRM model.

All in all, we think that the simple toy model we discussed is a perfect playground to see thimble regularization at work: it is instructive both from the point of view of inspecting the structure of relevant thimbles and from the algorithmic point of view (we can compute on thimbles).

IV. CHIRAL RANDOM MATRIX MODEL

We now address the chiral random matrix model defined by the partition function

$$Z_N^{N_f}(m) = \int d\Phi d\Psi \det{}^{N_f}(D(\mu) + m) \times \exp\left(-N \cdot \mathrm{Tr}[\Psi^\dagger\Psi + \Phi^\dagger\Phi]\right), \quad (20)$$

where

$$D(\mu) + m = \begin{pmatrix} m & i\cosh(\mu)\Phi + \sinh(\mu)\Psi \\ i\cosh(\mu)\Phi^\dagger + \sinh(\mu)\Psi^\dagger & m \end{pmatrix}. \quad (21)$$

The degrees of freedom of the model are $N \times N$ general complex matrices Ψ and Φ. Since its introduction it has attracted attention due to the many features which it shares with QCD [28–30]: they both have in their functional integral the determinant of a Dirac operator and the flavor symmetries and explicit breaking hereof are identical. Chiral perturbation theory at leading order in the ϵ-domain is the relevant low energy theory in the microscopic limit for both theories, which resulted in a lot of interesting insights into QCD coming from the (much simpler) random matrix theory. The microscopic limit in which contact is made with the ϵ-regime of chiral perturbation theory is that of $N \to \infty$ with $\tilde{m} \equiv Nm$ and $\tilde{\mu} \equiv \sqrt{N}\mu$ kept constant.

A sign problem is there for this theory as it is for QCD. This sign problem can be a severe one, as it is made

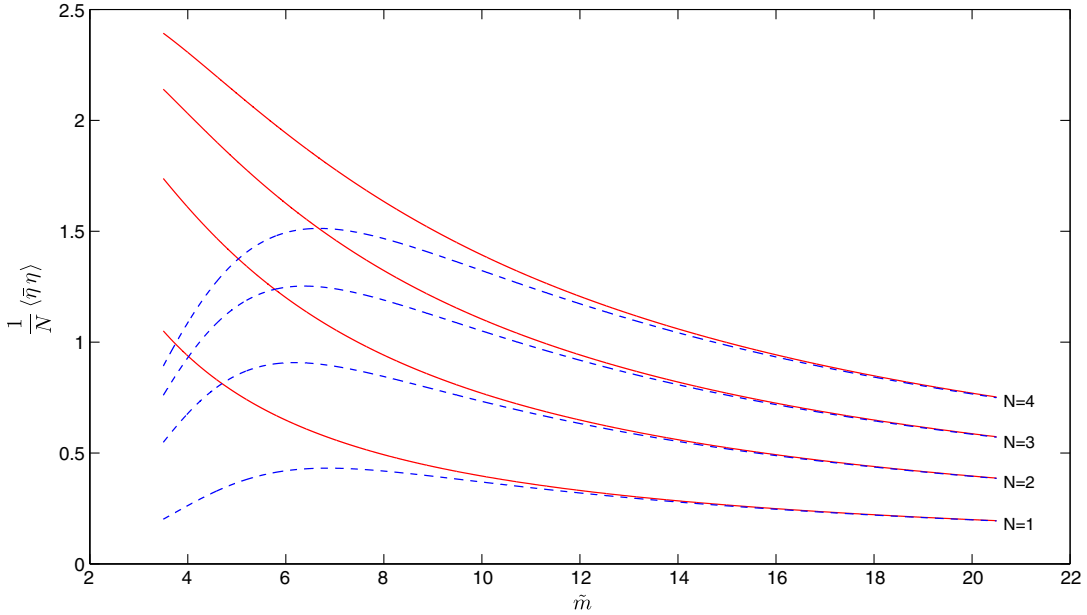

FIG. 4 (color online). Exact (solid red line) and phase quenched (dashed blue line) results for the condensate, at fixed $N_f = 2$, $\tilde{\mu} = 2$.

manifest by considering the observable we will be concerned with, i.e., the mass dependent chiral condensate

$$\frac{1}{N}\langle\bar{\eta}\eta\rangle = \frac{1}{N}\partial_m \log(Z). \tag{22}$$

Figure 4 displays both the exact (solid red line) and the phase quenched (dashed blue line) results for $\frac{1}{N}\langle\bar{\eta}\eta\rangle$ as a function of $\tilde{m}$ for $N = 1, 2, 3, 4$ and fixed $N_f = 2$, $\tilde{\mu} = 2$. As it can be seen, the sign problem is indeed severe in certain regimes of small (rescaled) masses.[12]

Our interest in the model was triggered by [12,13]: the nature of the sign problem (which is due to the determinant) has a counterpart in a nontrivial success of the complex Langevin method (which needs to take the logarithm of the determinant to define the standard effective action dictating the drift term of the Langevin equation). While the application of the complex Langevin in the most direct parametrization of the theory fails [12], a different parametrization (resulting in a different complexification) reproduces the right results [13].

A. How many thimbles should we take into account?

We take the most direct path to complexification; i.e., for each field (we directly deal with the matrix elements) $\Phi_{ij} = a_{ij} + ib_{ij}$ and $\Psi_{ij} = \alpha_{ij} + i\beta_{ij}$, each real component gets complexified [e.g., $\beta_{ij} = \beta_{ij}^{(R)} + i\beta_{ij}^{(I)}$]. We adhere to the notation of [12] and denote the action as

$$S(a,b,\alpha,\beta) = N\sum_{i,j}(a_{ij}^2 + b_{ij}^2 + \alpha_{ij}^2 + \beta_{ij}^2) - N_f \mathrm{Tr}\,\log(m^2\mathbb{1}_{N\times N} - XY),$$

with

$$X_{ij} = i\cosh\mu(a_{ij} + ib_{ij}) + \sinh\mu(\alpha_{ij} + i\beta_{ij})$$
$$Y_{ij} = i\cosh\mu(a_{ji} - ib_{ji}) + \sinh\mu(\alpha_{ji} - i\beta_{ji}).$$

Once we have complexified the degrees of freedom, the first step for the thimble approach is the identification of critical points of the resulting action. The first candidate is the absolute minimum which is already there for the real formulation, i.e., $\Psi = \Phi = 0$. All the relevant formulas for the spectral analysis of the Hessian of S_R are collected in Appendix A. Here we simply state that the Hessian in 0 has the expected number of positive eigenvalues, i.e., the real dimension of the thimble attached to 0 is $4N^2$. Note that there is a huge degeneracy: we have only two different eigenvalues, with the two eigenspaces having the same dimension. As the (rescaled) mass $\tilde{m}$ gets smaller, the gap between the two eigenvalues gets larger. Some insight can now be gained from Eq. (7): in first approximation, the closer the eigenvalues, the more isotropic we expect the thimble to be. This expectation turned out to be correct in view of the results of our simulations.

We tried to identify other critical points. In our study we explored different values of $\tilde{m}$ (at different values of N) while keeping fixed $N_f = 2$ and $\tilde{\mu} = 2$. One approach was solving $\nabla S = 0$ via the Newton-Raphson method. We cross-checked results by applying the Nelder-Mead simplex method to minimize $\|\nabla S\|^2$. We found two classes of

[12]The value of the condensate is real. This has to be understood later when we will compare to our results; we will always plot only the real part, the imaginary one having been correctly verified to be zero within errors.

extrema, both outside the original domain and featuring an action smaller than $S_R(\Psi = \Phi = 0)$, which turns out to be the absolute minimum in the original domain. Under such conditions, since the unstable thimbles attached to the extrema we found cannot intersect the original domain of integration, we expect no contribution from their stable thimbles ($n_\sigma = 0$; see Sec. II.B.3 of [2] for a more extensive discussion).

B. Algorithmic issues for the CRM model

While for the zero-dimensional toy model the original algorithmic solution proposed in [2] is trivial, this is not the case for the CRM model. However, previous experience with the Bose gas [14] taught us that there can be lucky cases. Let us remind the reader of the Aurora algorithm and of its Gaussian approximation (which successfully deals with the lucky cases we were referring to).

We want to extract a proper noise vector for the Langevin dynamics

$$\frac{\mathrm{d}\phi_i}{\mathrm{d}t} = -\frac{\partial S_R}{\partial \phi_i} + \eta_i, \quad i = 1 \cdots 2n.$$

We can proceed as follows [2]:

(i) We extract a Gaussian noise $\eta_i^{(0)}(0)$, where the superscript qualifies this quantity as an initial proposal and the argument has to be thought as a flow time in a sense that will be clear soon.

(ii) We evolve it following the flow (6) downwards [i.e., with a change of sign with respect to (6)], aiming at getting close enough to the critical point in order to make contact with the regime of (7). This will hold at a given descent time τ^*.

(iii) We then project with

$$\eta_i^{\parallel} = P_{ij}\eta_j^{(0)}(\tau^*), \qquad P \equiv \frac{1}{2}\left(\frac{H}{\sqrt{H^2}} + 1\right), \tag{23}$$

and normalize the result:

$$\eta(\tau^*) = r\frac{\eta^{\parallel}}{\|\eta^{\parallel}\|}, \tag{24}$$

r being extracted according to the n-dimensional χ distribution.

(iv) We then ascend along the flow, covering again a time interval of length τ^*. The result is the noise η_i we will put in our Langevin equation.

Extracting the noise vector is not yet the end of the story, since any finite order approximation to Langevin equation, e.g., the Euler scheme

$$\phi_i' = \phi_i - \delta t\frac{\partial S_R}{\partial \phi_i} + \sqrt{\delta t}\eta_i,$$

will introduce systematic effects; since the manifold is not flat, the final point ϕ_i' will be moved away from the thimble. The obvious remedy for this effect is to repeat the same procedure that we carried out for the noise vector (move the configuration along the flow downward, close to the critical point, project it onto the tangent space, and move it upward along the flow for the same time length). Note that it is expected that, in all the descent/ascent mechanisms we have just described, the downward flow, i.e., the SD, will be numerically delicate. It is thus much better to formulate the descent as a boundary value problem (BVP) rather than as an initial value problem, as it was observed in [14].

A much more appealing observation was also made in [14]: there are lucky cases in which a quite rough approximation holds; with a slight linguistic abuse we call it a *Gaussian approximation*.[13] Roughly speaking, this means taking the minimum value for the τ^* technical parameter, i.e., $\tau^* = 0$. This formally relies on the assumption that integrating the system on the vector space defined by the tangent space at the critical point actually takes into account the relevant configurations giving the most important contribution to the functional integral. This was actually holding in the case of [14].

Does the Gaussian approximation hold true also for the CRM model? Figure 5 reveals that there is actually a regime in which it can do pretty well. Not surprisingly, it is a regime in which results are not that far away from the phase quenched approximation; we know that this is a regime in which the two different eigenvalues of the Hessian at the critical point are quite close to each other and the problem appears all in all quite symmetric and not that far away from the regime of (7). Note that the value of the (rescaled) mass at which the solution provided by the Gaussian approximation departs from the correct one varies with N.

The next step was to leave the Gaussian approximation aside and try to implement the full Aurora algorithm. There are a couple of issues one should be aware of: we need a solid estimate for τ^*; also, within a time length of order τ^* we have to make sure we have under good numerical control both the SA and the SD. The latter is the critical one, for which we have already made clear that a BVP formulation is the choice to go for. Our implementation was along the same lines of the code available at [31]. All in all, our experience with the complete Aurora algorithm for the CRM model was at first somehow inconclusive: a clear-cut indication of values of τ^* at the same time safe and manageable was missing. We will come back to this

[13] One should not confuse this with the Gaussian approximation described in Sec. III A. In that case the action is approximated with its leading (Gaussian) term and the entire thimble analysis is performed consequently. In this (algorithmic) Gaussian approximation one pretends that the thimble manifold we are interested in and the thimble associated to the Gaussian approximation of the action sit on top of each other also away from the critical point.

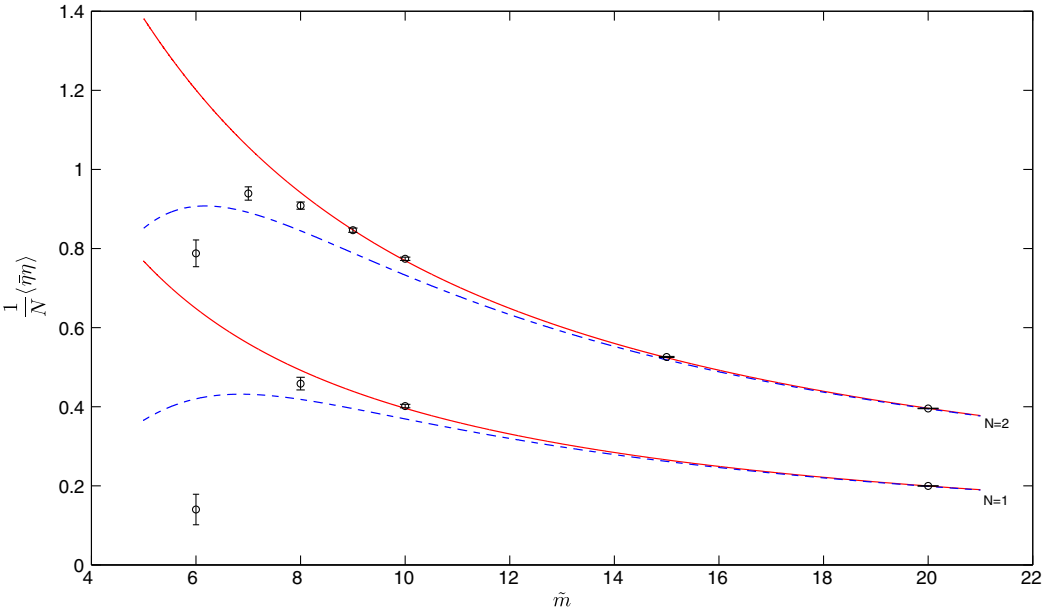

FIG. 5 (color online). Exact (solid red line), phase quenched (dashed blue line) and Gaussian approximation results for the condensate, at fixed $N_f = 2$, $\tilde{\mu} = 2$.

observation later, in the framework of the other numerical approach that we chose to implement.

C. A different numerical approach

We now want to take advantage of the parametrization

$$\Phi \in \mathcal{J}_\sigma \leftrightarrow (\hat{n}, t)$$

(a few of the formulas we will need to implement our strategy were clearly stated in [3], while the strategy itself we will see at work in the following was first described in [32]). Basically one describes a generic point by locating it on the SA curve it lies on. This means providing $\hat{n}$ (the direction one is taking while leaving the critical point) and the time t at which one reaches Φ while integrating the SA equations. The first goal is now to rewrite the contribution to the partition function which is attached to one thimble. In full detail this reads

$$\begin{aligned} Z^{(\sigma)} &= \int_{\mathcal{J}_\sigma} dz_1 \wedge \ldots \wedge dz_n \mathrm{e}^{-S} \\ &= \sum_{\text{charts } c} \int_{\Gamma_c} \prod_i^n dy_i^c \det(U) \mathrm{e}^{-S} \\ &= \mathrm{e}^{-iS_I} \sum_{\text{charts } c} \int_{\Gamma_c} \prod_i^n dy_i^c \mathrm{e}^{i\omega} \mathrm{e}^{-S_R}. \end{aligned} \tag{25}$$

In (25) we have taken into account that S_I is constant on $\mathcal{J}_\sigma$. Moreover, there could be more than one relevant chart and on a given chart we have to take into account the residual phase. For the sake of simplicity in notation, we now take a few shortcuts. First of all, we discard the overall phase e^{-iS_I}; it will be easy to account for it when we come back to the actual computation of an observable. We also discard the fine detail of more than one chart, since in practice this is not a issue.[14] Finally, we leave the residual phase aside, having in mind that we can take it into account *a posteriori* by reweighting. We thus write a new quantity, which is the one we will further manipulate in order to single out the contributions from the single ascents. We define[15]

$$\mathcal{Z}^{(\sigma)} \equiv \int \prod_{i=1}^n \mathrm{d}y_i e^{-S_R}. \tag{26}$$

Roughly speaking, this is the quantity that can have a probabilistic interpretation. The key point is now to write an expression for 1:

$$1 = \Delta_{\hat{n}}(t) \int \prod_{k=1}^n \mathrm{d}n_k \delta(|\vec{n}|^2 - 1) \int \mathrm{d}t \prod_{i=1}^n \delta(y_i - y_i(\hat{n}, t)), \tag{27}$$

where $\{y_i(\hat{n}, t)\}$ are the coordinates of the field as expressed in the local (orthonormal) basis $\{U^{(i)}_{\hat{n}}(t)\}$

[14]We will be concerned with single ascents, for which we will revert to a different, smooth parametrization.

[15]For the sake of notational simplicity we also omit the explicit indication that the integration is on the thimble, as it is easy to recognize we are assuming. Notice that (26) is the generalization of (11) of Sec. III A.

parallel-transported along the SA defined by $\hat{n}$ until time t. The solution for $\Delta_{\hat{n}}(t)$ is in terms of (the module of) a determinant

$$\Delta_{\hat{n}}(t) = \left| \det \begin{pmatrix} \frac{\delta(|\vec{n}'|^2-1)}{\delta t} & \frac{\delta(|\vec{n}'|^2-1)}{\delta n_1} & \cdots & \frac{\delta(|\vec{n}'|^2-1)}{\delta n_n} \\ \frac{\delta(y_1-y_1(\hat{n},t))}{\delta t} & \frac{\delta(y_1-y_1(\hat{n},t))}{\delta n_1} & \cdots & \frac{\delta(y_1-y_1(\hat{n},t))}{\delta n_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\delta(y_n-y_n(\hat{n},t))}{\delta t} & \frac{\delta(y_n-y_n(\hat{n},t))}{\delta n_1} & \cdots & \frac{\delta(y_n-y_n(\hat{n},t))}{\delta n_n} \end{pmatrix} \right|$$

or

$$\Delta_{\hat{n}}(t) = \left| \det \begin{pmatrix} 0 & 2n_1 & \cdots & 2n_n \\ \frac{\delta y_1(\hat{n},t)}{\delta t} & \frac{\delta y_1(\hat{n},t)}{\delta n_1} & \cdots & \frac{\delta y_1(\hat{n},t)}{\delta n_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\delta y_n(\hat{n},t)}{\delta t} & \frac{\delta y_n(\hat{n},t)}{\delta n_1} & \cdots & \frac{\delta y_n(\hat{n},t)}{\delta n_n} \end{pmatrix} \right|. \quad (28)$$

The first column of this determinant can be easily related to the gradient of the action. It turns out that to compute the generic matrix element we need to do the following:

(i) We need to evolve not only the field, but the entire basis by integrating (6).
(ii) We construct the $2n \times n$ matrix V whose columns are the $\{V^{(i)}(t)\}$.
(ii) We construct the $2n \times n$ matrix u whose columns are the vectors $\{u^{(i)}\}_{i=1\cdots n}$ which are obtained from the $\{V^{(i)}(t)\}$ by means of the Gram-Schmidt orthonormalization procedure.
(iii) The relation $V = uE$ holds, with

$$E_{ij} = \begin{cases} V^{(j)} \cdot u^{(i)} & j \geq i \\ 0 & j < i \end{cases}.$$

(iv) The entries of the determinant we are looking for are now given by

$$\begin{cases} \frac{\delta y_i}{\delta t} = \sum_{k=1}^{n} \lambda_k n_k E_{ik} \\ \frac{\delta y_i}{\delta n_j} = E_{ij} \end{cases}. \quad (29)$$

Not surprisingly, there is a lot of information in the (tremendous amount of) computations we have just sketched. In particular, if we now introduce the $n \times 2n$ complex space projector P,

$$P = (\, \mathbb{1}_{n\times n} \quad i\mathbb{1}_{n\times n}), \quad (30)$$

then the $n \times n$ complex matrix $U = Pu$ is unitary; this is precisely the matrix of Eq. (5), whose determinant is the residual phase $e^{i\omega}$.

The details of the previous computation of $\Delta_{\hat{n}}(t)$ are given in Appendix B. We now proceed to make use of the expression for the identity encoded in (27). Inserting it in (26) we get

$$\begin{aligned} \mathcal{Z}^{(\sigma)} &= \int \prod_{i=1}^{n} \mathrm{d}y_i e^{-S_R} \\ &= \int \prod_{i=1}^{n} \mathrm{d}y_i e^{-S_R} \Delta_{\hat{n}}(t) \int \prod_{k=1}^{n} \mathrm{d}n_k \delta(|\vec{n}|^2 - 1) \\ &\quad \times \int \mathrm{d}t \prod_{i=1}^{n} \delta(y_i - y_i(\hat{n}, t)) \\ &= \int \prod_{k=1}^{n} \mathrm{d}n_k \delta(|\vec{n}|^2 - 1) \int \mathrm{d}t \\ &\quad \times \int \prod_{i=1}^{n} \mathrm{d}y_i \delta(y_i - y_i(\hat{n}, t)) \Delta_{\hat{n}}(t) e^{-S_R} \\ &= \int \prod_{k=1}^{n} \mathrm{d}n_k \delta(|\vec{n}|^2 - 1) \int \mathrm{d}t \Delta_{\hat{n}}(t) e^{-S_R(\hat{n},t)}, \end{aligned}$$

which has a possible interpretation in terms of

$$\mathcal{Z}^{(\sigma)} = \int \prod_{k=1}^{n} \mathrm{d}n_k \delta(|\vec{n}|^2 - 1)\ \mathcal{Z}^{(\sigma)}_{\hat{n}}; \quad (31)$$

i.e., there is a contribution to the partition function for each SA path

$$\mathcal{Z}^{(\sigma)}_{\hat{n}} = \int_{-\infty}^{+\infty} \mathrm{d}t \Delta_{\hat{n}}(t) e^{-S_R(\hat{n},t)}. \quad (32)$$

Note that the procedure naturally defines a probability, i.e., that for a point reached at time t on the SA defined by $\hat{n}$:

$$P_{\hat{n}}(t) = \frac{\Delta_{\hat{n}}(t) e^{-S_R(\hat{n},t)}}{\mathcal{Z}^{(\sigma)}_{\hat{n}}}. \quad (33)$$

One can also naturally define the cumulative distribution function (it is manifestly nondecreasing, positive definite, and has the correct normalization)

$$F_{\hat{n}}(t) = \frac{1}{\mathcal{Z}^{(\sigma)}_{\hat{n}}} \int_{-\infty}^{t} \mathrm{d}t' \Delta_{\hat{n}}(t') e^{-S_R(\hat{n},t')}. \quad (34)$$

Since we can easily invert this function numerically, we have a tool to ideally sample configurations on a single SA. Namely, we extract a random number $\xi \in [0, 1]$ and then get the point on the SA (rather, the time at which the point is reached) by $t = F^{-1}_{\hat{n}}(\xi)$. Actually this is not that useful. The fact that we ascend all the way along a given SA in order to compute $\mathcal{Z}^{(\sigma)}_{\hat{n}}$ suggests that it is rather convenient to compute the entire contribution which is attached to that given ascent. On the other hand, the relative weight of a given SA [within the complete partition function $\mathcal{Z}^{(\sigma)}$] is given by $\mathcal{Z}^{(\sigma)}_{\hat{n}} / \mathcal{Z}^{(\sigma)}$.

We now want to take advantage of the parametrization $\Phi \in \mathcal{J}_\sigma \leftrightarrow (\hat{n}, t)$ in the computation of an observable. In the following, we will assume we are in a case in which only one single thimble is relevant. This is not the general case, but for what we want to obtain it is not a limitation. In the cases in which more than one thimble contributes, we can address the problem using the same strategy described in Sec. III A in the context of the quartic toy model: it will be easy for the reader to generalize Eq. (14) and the discussion following it. With this caveat in mind, we first of all write

$$\langle O \rangle = \frac{\int_{\mathcal{J}_\sigma} dz_1 \wedge \ldots \wedge dz_n O \mathrm{e}^{-S}}{Z^{(\sigma)}} = \frac{\int \prod_{i=1}^n \mathrm{d}y_i O e^{i\omega} e^{-S_R}}{\int \prod_{i=1}^n \mathrm{d}y_i e^{i\omega} e^{-S_R}} = \frac{\langle e^{i\omega} O \rangle_\sigma}{\langle e^{i\omega} \rangle_\sigma},$$

where $\langle \ldots \rangle_\sigma \equiv \mathcal{Z}^{(\sigma)-1} \int \prod_{i=1}^n \mathrm{d}y_i \ldots e^{-S_R}$. We have till now simply generalized Eq. (13). We can go further by making use of the new parametrization we introduced[16]:

$$\begin{aligned}\langle O \rangle &= \frac{\int \mathcal{D}\hat{n} \int dt \Delta_{\hat{n}}(t) e^{-S_R(\hat{n},t)} e^{i\omega(\hat{n},t)} O(\hat{n},t)}{\int \mathcal{D}\hat{n} \int dt \Delta_{\hat{n}}(t) e^{-S_R(\hat{n},t)} e^{i\omega(\hat{n},t)}} \\ &= \frac{\int \mathcal{D}\hat{n} \mathcal{Z}_{\hat{n}}^{(\sigma)} (\mathcal{Z}_{\hat{n}}^{(\sigma)-1} \int dt \Delta_{\hat{n}}(t) \mathrm{e}^{-S_R(\hat{n},t)} e^{i\omega(\hat{n},t)} O(\hat{n},t))}{\int \mathcal{D}\hat{n} \mathcal{Z}_{\hat{n}}^{(\sigma)} (\mathcal{Z}_{\hat{n}}^{(\sigma)-1} \int dt \Delta_{\hat{n}}(t) \mathrm{e}^{-S_R(\hat{n},t)} e^{i\omega(\hat{n},t)})} \\ &\equiv \frac{\int \mathcal{D}\hat{n} \mathcal{Z}_{\hat{n}}^{(\sigma)} \langle e^{i\omega} O \rangle_{\hat{n}}}{\int \mathcal{D}\hat{n} \mathcal{Z}_{\hat{n}}^{(\sigma)} \langle e^{i\omega} \rangle_{\hat{n}}}. \end{aligned} \tag{35}$$

Equation (35) is in a sense a new average. Namely, the different directions $\hat{n}$ can now be regarded as the new degrees of freedom of the overall integral, the quantities to be measured are the $\langle O e^{i\omega} \rangle_{\hat{n}}$ and $\langle e^{i\omega} \rangle_{\hat{n}}$ (i.e., partial averages attached to single SA), and the weights are given by the $\mathcal{Z}_{\hat{n}}^{(\sigma)}$. It is rather obvious that

(i) The basic building blocks are complete ascents. This is good, since we can have their computation under good numerical control. In other words, sampling *on* the thimble is not a problem: we stay on the thimble by definition.

(ii) The way to importance sampling now appears tricky. This is easy to understand, since picking up a contribution means picking up a $\hat{n}$, whose weight $\mathcal{Z}_{\hat{n}}^{(\sigma)}$ is not known *a priori*, but only after the SA path associated to $\hat{n}$ has been obtained.

(iii) The crudest approach one can think of is of course a uniform sampling of the $\hat{n}$-space; this is a static, crude Monte Carlo, which can easily become inefficient (in particular for large systems).

[16]We denote $\mathcal{D}\hat{n} \equiv \prod_{k=1}^n \mathrm{d}n_k \delta(|\vec{n}|^2 - 1)$.

In the following we will just be satisfied with the last approach of static, crude Monte Carlo; this will be enough to show that we can reproduce the correct results for the model at hand (even in regions where the sign problem is quite severe) and this holds true taking into account the contribution of the single critical point we found. From this very basic approach there will be something to learn also with respect to the Aurora algorithm (and on the computation of density of states as well). We will finally report on a few speculations on smarter algorithms which we are trying to devise.

D. Results for the CRM model

Figure 6 displays the results we obtained from simulations performed in the static Monte Carlo approach we have just discussed. All these results come from the contribution of one single thimble. As we have already pointed out, one original motivation of ours turned out to be not relevant: we were ready for looking for dominance of one thimble in some asymptotic regime (in the thermodynamic limit) and, in the region of parameters we studied, we actually found no other thimble but the trivial one. Also, complexifying the theory in the parametrization that was shown to be problematic for the complex Langevin [12] did not result in any problem. Results are shown for $N = 1, 2, 3, 4$ and fixed $N_f = 2$, $\tilde{\mu} = 2$.

It is interesting to regard the parametrization we have employed from another point of view. In the upper panel of Fig. 7 we plot the quantity $\Delta_{\hat{n}}(t) \mathrm{e}^{-S_R(\hat{n},t)} / \mathcal{Z}_{\hat{n}}^{(\sigma)}$ as a function of S_R along a given ascent (remember that on each ascent one single value of S_R is only read once). This is the real weight of the functional integral for the configurations which lie on that given ascent: it can be thought of as a different way of looking at the density of states (namely, this is the contribution attached to a given ascent).

In the lower panel of Fig. 7 we plot the same quantity as a function of the flow time. A first remark is due for the long initial flat region. Consider Eq. (7). When we want to compute the contribution from a single ascent (that is, from a single $\hat{n}$) we need an initial condition, i.e., an initial value t_0 at which the asymptotic regime holds. In principle, the more back in time we take t_0, the better initial condition we prepare. There are of course accuracy issues one has to live with. The flat initial region reflects the fact that we do our best to ensure we stay on the thimble. On the other side, one would like to know till what value of the flow time the asymptotic regime holds to a reasonable confidence. We can think of more than one indicator for the latter condition, e.g., the Gaussian approximation of the action is very close to the actual value of S, or the factor $\Delta_{\hat{n}}(t)$ is very close to its Gaussian approximation (see Appendix B.1). We mark with a (red) star a value of flow time which can be assumed to be the boundary of the region we have just described. Now, an efficient dynamic Monte Carlo is supposed to sample configurations in regions where the weight is

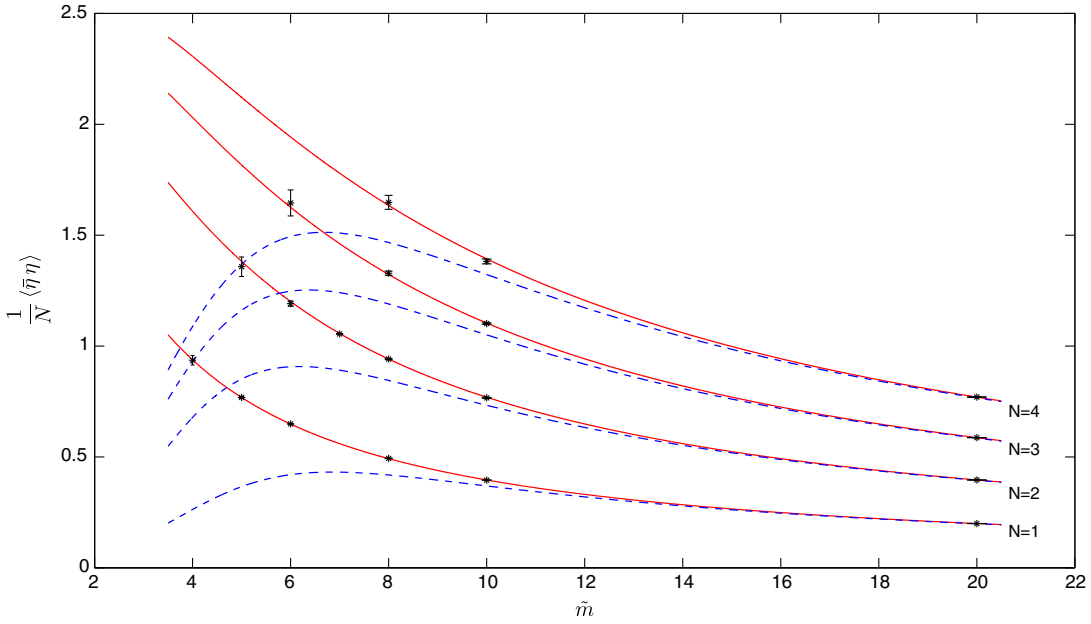

FIG. 6 (color online). Exact (solid red line), phase quenched (dashed blue line), and thimble simulation results for the condensate, at fixed $N_f = 2$, $\tilde{\mu} = 2$.

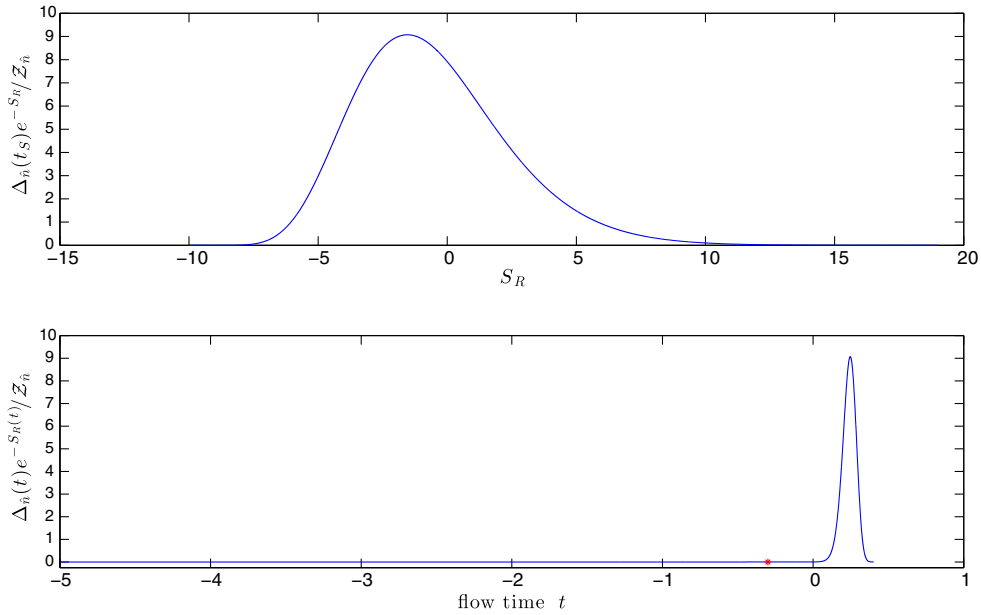

FIG. 7 (color online). Real weight of the functional integral for the configurations which lie on the SA defined by a particular $\hat{n}$ as a function of S_R (upper panel) and t (lower panel) for $N = 2$, $\tilde{m} = 7$, $N_f = 2$, $\tilde{\mu} = 2$.

concentrated. In this sense we can say that the distance (in flow time) between the region around the maximum of $\Delta_{\hat{n}}(t)\mathrm{e}^{-S_R(\hat{n},t)}/\mathcal{Z}_{\hat{n}}^{(\sigma)}$ and the flow time marked with the star is a reasonable indicator of the τ^* parameter of the Aurora algorithm.

A comment is due concerning the residual phase: we encountered no problem in taking it into account by reweighting. As it was expected, it is a smooth function on the ascents, so that $\langle Oe^{i\omega}\rangle_{\hat{n}}$ and $\langle e^{i\omega}\rangle_{\hat{n}}$ can be safely computed and wild cancellations are never there at any stage of our computations. The fact that the residual phase is smooth does not of course mean it has no net effect, as can be seen in Fig. 8. In the upper panel we show the effect of neglecting the contribution of the denominator of Eq. (35):

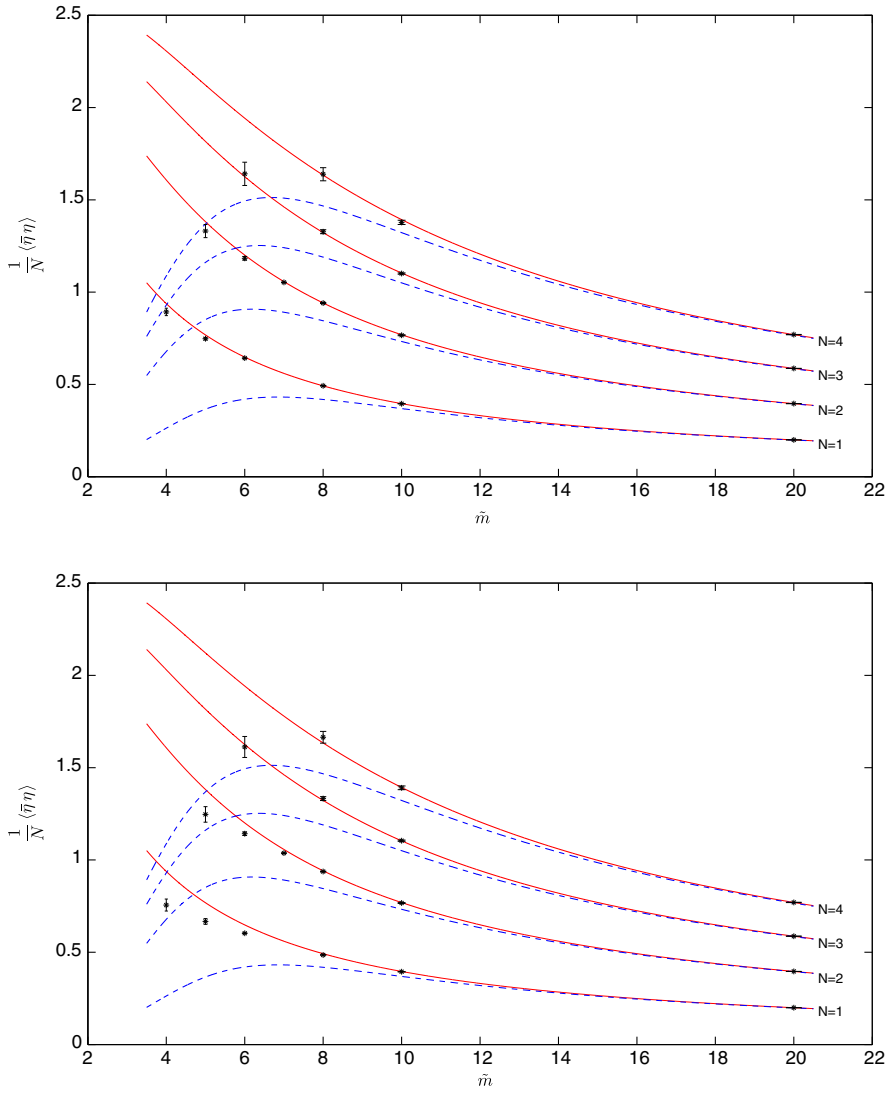

FIG. 8 (color online). The effect of not accounting for the residual phase. In the upper panel its contribution is not accounted for in the denominator of Eq. (35); i.e., we compute $\mathcal{Z}^{(\sigma)-1}\int \mathcal{D}\hat{n}\mathcal{Z}^{(\sigma)}_{\hat{n}}\langle e^{i\omega}O\rangle_{\hat{n}}$. In the lower panel we omit the residual phase completely and compute $\mathcal{Z}^{(\sigma)-1}\int \mathcal{D}\hat{n}\mathcal{Z}^{(\sigma)}_{\hat{n}}\langle O\rangle_{\hat{n}}$.

this amounts to computing $\mathcal{Z}^{(\sigma)-1}\int \mathcal{D}\hat{n}\mathcal{Z}^{(\sigma)}_{\hat{n}}\langle e^{i\omega}O\rangle_{\hat{n}}$.[17] In the lower panel we show yet another phase quenched computation, namely, what we could term a *residual phase quenched* result. In this case we simply omit the contribution of the residual phase and compute $\mathcal{Z}^{(\sigma)-1}\int \mathcal{D}\hat{n}\mathcal{Z}^{(\sigma)}_{\hat{n}}\langle O\rangle_{\hat{n}}$. Both plots show that reweighting for the residual phase is essential to get the correct results; this happens to be the case in particular for low dimensions.

We admittedly made use of the crudest possible application of the parametrization contained in Eq. (35), i.e., a static, crude Monte Carlo sampling of the integral. It is static; i.e., it is not based on a stochastic process. It is crude, because it does not implement importance sampling.[18] In general, importance sampling computes an integral $I = \int dx f(x)$ as an average $I = \int dx \rho(x)\frac{f(x)}{\rho(x)}$ and the entire point is being able to extract configurations distributed according to $\rho(x)$. The natural importance sampling for (35) would try to sample the space of ascents (i.e., the different $\hat{n}$) according to the weights $\mathcal{Z}^{(\sigma)}_{\hat{n}}/\mathcal{Z}^{(\sigma)}$ and we have already made the point that this is tricky. Crude Monte Carlo simply extracts $\hat{n}$ with constant probability. As it is well known, the more nontrivial the profile of $\mathcal{Z}^{(\sigma)}_{\hat{n}}/\mathcal{Z}^{(\sigma)}$ is, the more inefficient one expects

[17]Note that in this case we take the residual phase into account in the numerator.

[18]We recall that the Monte Carlo methods we are mostly familiar with are dynamic Monte Carlo in which importance sampling is obtained via convergence to the equilibrium distribution of a stochastic process, which in virtually all the cases is a Markov chain.

the crude Monte Carlo to be. This can be clearly seen at low masses, where the problem is less symmetric with respect to different choices of the $\hat{n}$: we have already made this point while commenting on the failure of the Gaussian approximation for the Aurora algorithm. The interested reader is referred to Appendix B.1 to get more insight on this in the case of the Gaussian action. Here we want to stress that the problem with crude Monte Carlo does not necessarily relate to the dimension N (of the matrices): going low enough in mass at any fixed N can already result in a difficult computation.[19] As the mass gets lower, one indeed clearly sees larger error bars in Fig. 6. Needless to say, this is a region in which we had to collect many ascents: hundreds of thousands, actually, by definition all statistically independent. This is a huge numerical effort. On the other side, even accepting that at some point our crude Monte Carlo has to give up in front of a nontrivial profile of the weights $\mathcal{Z}_{\hat{n}}^{(\sigma)}/\mathcal{Z}^{(\sigma)}$, it is reassuring to see that we nevertheless solved a nontrivial problem: Fig. 6 clearly shows that we were able to solve the problem also in regions in which the sign problem shows up as quite severe.

All in all, we saw that implementing importance sampling is hard, since the relative weights $\mathcal{Z}_{\hat{n}}^{(\sigma)}/\mathcal{Z}^{(\sigma)}$ are only known after the complete SA associated to $\hat{n}$ has been computed. This could sound like a very pessimistic conclusion. On the other hand, the Gaussian approximation of the $\mathcal{Z}_{\hat{n}}^{(\sigma)}$ can be easily computed (see Appendix B.1), which suggests the idea of making use of them to formulate proposals for the $\hat{n}$. This is something we are currently investigating.

V. CONCLUSIONS AND PROSPECTS

We discussed the solution of a simple toy model via thimble regularization. Quite interestingly, this model, which dates back to some 30 years ago and was proposed as a sort of benchmark for the complex Langevin, was still missing a full solution in the context of the latter. In thimble regularization the solution is clear and can be implemented numerically by a number of simulation algorithms.

We then investigated the chiral random matrix model and showed that thimble regularization can successfully deal with the sign problem that the system displays. In the region of parameters we studied, a single thimble accounts for the results. We made use of a parametrization in terms of contributions attached to SA (which are the basic building blocks to define the thimble). This was done by crude Monte Carlo, leaving open the problem of devising a smarter algorithm (importance sampling) to take advantage of the parametrization we made use of. This is the subject we are currently investigating in view of other applications.

[19]Of course going to higher values of N would result in extra computational effort, but what is really crucial is to see at each value of N the threshold in mass below which we have a *too-nontrivial* profile of the $\mathcal{Z}_{\hat{n}}^{(\sigma)}/\mathcal{Z}^{(\sigma)}$.

ACKNOWLEDGMENTS

We warmly thank Luigi Scorzato for many valuable discussions and for all the common work on the subject in recent years. We are indebted to Kim Splittorff for very fruitful conversations and for having introduced us to the subject of the CRM model. We also acknowledge useful discussions with Marco Cristoforetti and Michele Brambilla and we are grateful to Christian Torrero who has collaborated with us at an early stage of this work. We finally thank A. Alexandru for a useful conversation which made us consider a more extended comment on the effects of the residual phase. This research was initially supported by the Research Executive Agency (REA) of the European Union under Grant Agreement No. PITN-GA-2009-238353 (ITN STRONGnet) and by Italian MIUR under Contract No. PRIN2009 (20093BMNPR 004). We acknowledge partial support from I.N.F.N. under the research project *i.s. QCDLAT*.

APPENDIX A: THE HESSIAN FOR THE CRM MODEL

We want to compute the Hessian at the critical point $a_{ij} = b_{ij} = \alpha_{ij} = \beta_{ij} = 0$. We need the following second derivatives (the fields are complexified by setting, e.g., $a_{ij} = a_{ij}^R + i a_{ij}^I$, etc.):

$$\left.\frac{\partial^2 S_R}{\partial a_{mn}^R \partial a_{ij}^R}\right|_0 = -\left.\frac{\partial^2 S_R}{\partial a_{mn}^I \partial a_{ij}^I}\right|_0 = A_- \delta_{mi}\delta_{nj}$$

$$\left.\frac{\partial^2 S_R}{\partial a_{mn}^R \partial a_{ij}^I}\right|_0 = \left.\frac{\partial^2 S_R}{\partial a_{mn}^I \partial a_{ij}^R}\right|_0 = 0$$

$$\left.\frac{\partial^2 S_R}{\partial b_{mn}^R \partial b_{ij}^R}\right|_0 = -\left.\frac{\partial^2 S_R}{\partial b_{mn}^I \partial b_{ij}^I}\right|_0 = A_- \delta_{mi}\delta_{nj}$$

$$\left.\frac{\partial^2 S_R}{\partial b_{mn}^R \partial b_{ij}^I}\right|_0 = \left.\frac{\partial^2 S_R}{\partial b_{mn}^I \partial b_{ij}^R}\right|_0 = 0$$

$$\left.\frac{\partial^2 S_R}{\partial \alpha_{mn}^R \partial \alpha_{ij}^R}\right|_0 = -\left.\frac{\partial^2 S_R}{\partial \alpha_{mn}^I \partial \alpha_{ij}^I}\right|_0 = A_+ \delta_{mi}\delta_{nj}$$

$$\left.\frac{\partial^2 S_R}{\partial \alpha_{mn}^R \partial \alpha_{ij}^I}\right|_0 = \left.\frac{\partial^2 S_R}{\partial \alpha_{mn}^I \partial \alpha_{ij}^R}\right|_0 = 0$$

$$\left.\frac{\partial^2 S_R}{\partial \beta_{mn}^R \partial \beta_{ij}^R}\right|_0 = -\left.\frac{\partial^2 S_R}{\partial \beta_{mn}^I \partial \beta_{ij}^I}\right|_0 = A_+ \delta_{mi}\delta_{nj}$$

$$\left.\frac{\partial^2 S_R}{\partial \beta_{mn}^R \partial \beta_{ij}^I}\right|_0 = \left.\frac{\partial^2 S_R}{\partial \beta_{mn}^I \partial \beta_{ij}^R}\right|_0 = 0$$

$$\left.\frac{\partial^2 S_R}{\partial a_{mn}^R \partial b_{ij}^R}\right|_0 = \left.\frac{\partial^2 S_R}{\partial a_{mn}^I \partial b_{ij}^I}\right|_0 = \left.\frac{\partial^2 S_R}{\partial a_{mn}^R \partial b_{ij}^I}\right|_0 = \left.\frac{\partial^2 S_R}{\partial a_{mn}^I \partial b_{ij}^R}\right|_0 = 0$$

$$\left.\frac{\partial^2 S_R}{\partial a^R_{mn}\partial\alpha^R_{ij}}\right|_0 = \left.\frac{\partial^2 S_R}{\partial a^I_{mn}\partial\alpha^I_{ij}}\right|_0 = 0$$
$$\left.\frac{\partial^2 S_R}{\partial a^R_{mn}\partial\alpha^I_{ij}}\right|_0 = \left.\frac{\partial^2 S_R}{\partial a^I_{mn}\partial\alpha^R_{ij}}\right|_0 = B\delta_{mi}\delta_{nj}$$
$$\left.\frac{\partial^2 S_R}{\partial a^R_{mn}\partial\beta^R_{ij}}\right|_0 = \left.\frac{\partial^2 S_R}{\partial a^I_{mn}\partial\beta^I_{ij}}\right|_0 = \left.\frac{\partial^2 S_R}{\partial a^R_{mn}\partial\beta^I_{ij}}\right|_0 = \left.\frac{\partial^2 S_R}{\partial a^I_{mn}\partial\beta^R_{ij}}\right|_0 = 0$$
$$\left.\frac{\partial^2 S_R}{\partial b^R_{mn}\partial\alpha^R_{ij}}\right|_0 = \left.\frac{\partial^2 S_R}{\partial b^I_{mn}\partial\alpha^I_{ij}}\right|_0 = \left.\frac{\partial^2 S_R}{\partial b^R_{mn}\partial\alpha^I_{ij}}\right|_0 = \left.\frac{\partial^2 S_R}{\partial b^I_{mn}\partial\alpha^R_{ij}}\right|_0 = 0$$
$$\left.\frac{\partial^2 S_R}{\partial b^R_{mn}\partial\beta^R_{ij}}\right|_0 = \left.\frac{\partial^2 S_R}{\partial b^I_{mn}\partial\beta^I_{ij}}\right|_0 = 0$$
$$\left.\frac{\partial^2 S_R}{\partial b^R_{mn}\partial\beta^I_{ij}}\right|_0 = \left.\frac{\partial^2 S_R}{\partial b^I_{mn}\partial\beta^R_{ij}}\right|_0 = B\delta_{mi}\delta_{nj}$$
$$\left.\frac{\partial^2 S_R}{\partial\alpha^R_{mn}\partial\beta^R_{ij}}\right|_0 = \left.\frac{\partial^2 S_R}{\partial\alpha^I_{mn}\partial\beta^I_{ij}}\right|_0 = \left.\frac{\partial^2 S_R}{\partial\alpha^R_{mn}\partial\beta^I_{ij}}\right|_0 = \left.\frac{\partial^2 S_R}{\partial\alpha^I_{mn}\partial\beta^R_{ij}}\right|_0 = 0$$

where use of the Cauchy-Riemann equations has been made (the other derivatives are trivially related to these by the Schwarz theorem, e.g., $\frac{\partial^2 S_R}{\partial a^R_{mn}\partial a^R_{ij}} = \frac{\partial^2 S_R}{\partial a^R_{ij}\partial a^R_{mn}}$) and the coefficients are given by

$$A_- = 2\left(N - N_f\frac{\cosh^2\mu}{m^2}\right) \qquad A_+ = 2\left(N + N_f\frac{\sinh^2\mu}{m^2}\right)$$
$$B = -2N_f\frac{\cosh\mu\sinh\mu}{m^2}.$$

The Hessian for $N = 1$ is (with the conventional choice of ordering: a^R, b^R, α^R, β^R, a^I, b^I, α^I, β^I)

$$H^{(1)} = \begin{pmatrix} A_- & 0 & 0 & 0 & 0 & 0 & B & 0 \\ 0 & A_- & 0 & 0 & 0 & 0 & 0 & B \\ 0 & 0 & A_+ & 0 & B & 0 & 0 & 0 \\ 0 & 0 & 0 & A_+ & 0 & B & 0 & 0 \\ 0 & 0 & B & 0 & -A_- & 0 & 0 & 0 \\ 0 & 0 & 0 & B & 0 & -A_- & 0 & 0 \\ B & 0 & 0 & 0 & 0 & 0 & -A_+ & 0 \\ 0 & B & 0 & 0 & 0 & 0 & 0 & -A_+ \end{pmatrix}.$$

As the second derivatives are manifestly diagonal with respect to the indices i, j, m, n, the Hessian for a generic N is block-diagonal:

$$H^{(N)} = \bigoplus^N H^{(1)}.$$

The model thus features a huge degeneracy of eigenvalues, as the spectrum of $N = 1$ is repeated N times. The spectrum for $N = 1$ features four positive eigenvalues and four eigenvalues opposite in sign (as expected by holomorphicity). We are interested in the positive part of the spectrum. An explicit computation shows that the distinct positive eigenvalues of $H^{(1)}$ are actually two and they are

$$\lambda_\pm = \frac{1}{2m^2}\Big|2N_f\cosh(2\mu) \pm \sqrt{2}\sqrt{8m^4 - 8N_f m^2 + N_f^2 + N_f^2\cosh(4\mu)}\Big|.$$

We note in passing that (here and in many other places) we could have written formulas in the complex notation of the Takagi factorization theorem (see [3]), which we decided not to employ here or in the entire paper to stick to a completely real notation.

APPENDIX B: COMPUTING $\Delta_{\hat{n}}(t)$

We want to compute $\Delta_{\hat{n}}(t)$, which is defined in (27). The main point is that we are ascending along a given flow, and while doing that we are also transporting the basis vectors along the flow; i.e., we are integrating (6) as well. Near the critical point Eq. (7) holds, in which the parametrization $\Phi \in \mathcal{J}_\sigma \leftrightarrow (\hat{n}, t)$ is manifest. Given a reference point $t_0 \ll 1$, Eq. (7) can be regarded as the initial conditions for the flow associated to $\hat{n}$. Near a generic point, under infinitesimal variations of t and $\hat{n}$, the variation of the point $\delta\Phi$ is given by

$$\delta\Phi = \delta\Phi(\hat{n}, t) = \sum_{i=1}^n V^{(i)}(t)\delta c^{(i)}.$$

This is so because $\delta\Phi$ is itself a vector belonging to the tangent space $T_\Phi\mathcal{J}_\sigma$. The (constant) coefficients $\delta c^{(i)}$ can be worked out from the asymptotic form of $\Phi(t)$ near the critical point

$$\begin{aligned}\delta\Phi &\approx \delta\left(\phi_\sigma + \sum_{i=1}^n v^{(i)}e^{\lambda_i t}n_i\right) \\ &= \sum_{i=1}^n v^{(i)}\left(\sum_{j=1}^n \delta n_j\frac{\partial}{\partial n_j} + \delta t\frac{\partial}{\partial t}\right)e^{\lambda_i t}n_i \\ &= \sum_{i=1}^n v^{(i)}e^{\lambda_i t}(\delta n_i + \lambda_i n_i\delta t) \\ &\approx \sum_{i=1}^n V^{(i)}(t)(\delta n_i + \lambda_i n_i\delta t),\end{aligned}$$

from which it follows

$$\delta c^{(i)} = \delta n_i + \lambda_i n_i\delta t.$$

Given that $\delta\Phi$ is a vector of $T_\Phi\mathcal{J}_\sigma$, we can write it as a decomposition on the (orthonormal) u-basis:

$$\delta\Phi = \sum_{i=1}^{n} u^{(i)} \delta y_i,$$

and from this we have

$$\delta y_i = \sum_{j=1}^{2n} u_j^{(i)} \delta\phi_j.$$

Let us now consider the terms $\frac{\delta y_i}{\delta\star}$ appearing in $\Delta_{\hat{n}}(t)$, where $\star$ is either t or n_j. For these we have

$$\begin{aligned}\frac{\delta y_i}{\delta\star} &= \sum_{j=1}^{2n} u_j^{(i)} \frac{\delta\phi_j}{\delta\star} = \sum_{j=1}^{2n} u_j^{(i)} \sum_{k=1}^{n} V_j^{(k)} \frac{\delta c^{(k)}}{\delta\star} \\ &= \sum_{j=1}^{2n} u_j^{(i)} \sum_{k=1}^{n}\sum_{l=1}^{n} u_j^{(l)} E_{lk} \frac{\delta c^{(k)}}{\delta\star} \\ &= \sum_{k=1}^{n}\sum_{l=1}^{n} E_{lk} \frac{\delta c^{(k)}}{\delta\star} \sum_{j=1}^{2n} u_j^{(i)} u_j^{(l)} = \sum_{k=1}^{n}\sum_{l=1}^{n} E_{lk} \frac{\delta c^{(k)}}{\delta\star} \delta_{il} \\ &= \sum_{k=1}^{n} E_{ik} \frac{\delta c^{(k)}}{\delta\star}.\end{aligned}$$

Now we make use of the explicit form of $\delta c^{(k)}$, which gives $\frac{\delta c^{(k)}}{\delta t} = \lambda_k n_k$ and $\frac{\delta c^{(k)}}{\delta n_j} = \delta_{kj}$, from which one can easily derive Eq. (29).

1. $\Delta_{\hat{n}}(t)$ in the Gaussian approximation

We can compute $\Delta_{\hat{n}}(t)$ for the Gaussian case (purely quadratic action), where the asymptotic form for the SA and parallel-transport equations is correct arbitrarily far away from the critical point. In that case the entries of $\Delta_{\hat{n}}(t)$ are ($E = \mathbb{1}_{n\times n}$)

$$\frac{\delta y_i}{\delta t} = \lambda_i n_i e^{\lambda_i t}$$
$$\frac{\delta y_i}{\delta n_j} = e^{\lambda_i t}\delta_{ij}$$

and the determinant is

$$\Delta_{\hat{n}}(t) = 2\left(\sum_{i=1}^{n} \lambda_i n_i^2\right) e^{\left(\sum_{i=1}^{n}\lambda_i\right)t}.$$

For the Gaussian action $S_R(\hat{n},t) = S_R(\phi_\sigma) + \frac{1}{2}\sum_{k=1}^{n}\lambda_k n_k^2 e^{2\lambda_k t}$, so that collecting everything we can write an expression for the weight itself:

$$\mathcal{Z}_{\hat{n}}^{(\sigma)} = 2\left(\sum_{i=1}^{n}\lambda_i n_i^2\right) e^{-S_R(\phi_\sigma)} \int_{-\infty}^{\infty} dt e^{\sum_{i=1}^{n}\lambda_i t} e^{-\frac{1}{2}\sum_{i=1}^{n}\lambda_i n_i^2 e^{2\lambda_i t}}.$$

From this expression it is easy to understand that the more the eigenvalues differ from each other (which in our case happens for low values of the mass parameter), the more various $\mathcal{Z}_{\hat{n}}^{(\sigma)}$ can differ.

2. A useful consistency relation

Since in the end we are performing quite a lot of computations, it is useful to have a consistency relation to be checked while ascending along the flow. The gradient of the action is yet another vector belonging to the tangent space $T_\Phi\mathcal{J}_\sigma$. Let us write the decomposition

$$\nabla_\Phi S^R = \sum_{i=1}^{n} V^{(i)}(t) g^{(i)}.$$

The coefficients $g^{(i)}$ can be found with the aid of the asymptotic form of the action

$$\begin{aligned}\nabla_\Phi S^R \approx \nabla_\Phi\left(S^R(\phi_\sigma) + \frac{1}{2}\Phi^T H\Phi\right) = H\Phi \approx H\sum_{i=1}^{n} v^{(i)} n_i e^{\lambda_i t} \\ = \sum_{i=1}^{n} v^{(i)} e^{\lambda_i t} n_i\lambda_i \approx \sum_{i=1}^{n} V^{(i)}(t) n_i\lambda_i,\end{aligned}$$

where we have used the symmetry of the Hessian and the fact that $Hv^{(i)} = \lambda_i v^{(i)}$. We have found that $g^{(i)} = n_i\lambda_i$, so while integrating the flow equations, we can keep checked the norm

$$\left|\nabla_\Phi S^R - \sum_{i=1}^{n} V^{(i)}(t) n_i\lambda_i\right|$$

and make sure that it is small with respect to the size of the system.

[1] E. Witten, Analytic continuation of Chern-Simons theory, AMS/IP Stud. Adv. Math. **50**, 347 (2011).
[2] M. Cristoforetti, F. Di Renzo, and L. Scorzato (AuroraScience Collaboration), New approach to the sign problem in quantum field theories: High density QCD on a Lefschetz thimble, Phys. Rev. D **86**, 074506 (2012).
[3] H. Fujii, D. Honda, M. Kato, Y. Kikukawa, S. Komatsu, and T. Sano, Hybrid Monte Carlo on Lefschetz thimbles: A

study of the residual sign problem, J. High Energy Phys. 10 (2013) 147.
[4] Y. Tanizaki and T. Koike, Real-time Feynman path integral with Picard-Lefschetz theory and its applications to quantum tunneling, Ann. Phys. (Amsterdam) **351**, 250 (2014).
[5] Y. Tanizaki, Lefschetz-thimble techniques for path integral of zero-dimensional $O(n)$ sigma models, Phys. Rev. D **91**, 036002 (2015).
[6] T. Kanazawa and Y. Tanizaki, Structure of Lefschetz thimbles in simple fermionic systems, J. High Energy Phys. 03 (2015) 044.
[7] Y. Tanizaki, H. Nishimura, and K. Kashiwa, Evading the sign problem in the mean-field approximation through Lefschetz-thimble path integral, Phys. Rev. D **91**, 101701 (2015).
[8] A. Cherman, D. Dorigoni, and M. Unsal, Decoding perturbation theory using resurgence: Stokes phenomena, new saddle points and Lefschetz thimbles, arXiv:1403.1277.
[9] F. Pham, Vanishing homologies and the n variable saddle point method, Proceedings of Symposia in Pure Mathematics **40**, 321 (1983).
[10] J. Ambjorn and S. K. Yang, Numerical problems in applying the Langevin equation to complex effective actions, Phys. Lett. **165B**, 140 (1985).
[11] G. Eruzzi and F. Di Renzo, Solution of simple toy models via thimble regularization of lattice field theory, *Proc. Sci.*, LATTICE (2014) 202.
[12] A. Mollgaard and K. Splittorff, Complex Langevin dynamics for chiral random matrix theory, Phys. Rev. D **88**, 116007 (2013).
[13] A. Mollgaard and K. Splittorff, Full simulation of chiral random matrix theory at nonzero chemical potential by complex Langevin, Phys. Rev. D **91**, 036007 (2015).
[14] M. Cristoforetti, F. Di Renzo, A. Mukherjee, and L. Scorzato, Monte Carlo simulations on the Lefschetz thimble: Taming the sign problem, Phys. Rev. D **88**, 051501 (2013).
[15] M. F. Atiyah and R. Bott, The Yang-Mills equations over Riemann surfaces, Phil. Trans. R. Soc. A **308**, 523 (1983).
[16] G. Eruzzi and F. Di Renzo, Thimble regularization at work for gauge theories: From toy models onwards, *Proc. Sci.*, LATTICE (2015) 189.
[17] G. Basar, G. V. Dunne, and M. Unsal, Resurgence theory, ghost-instantons, and analytic continuation of path integrals, J. High Energy Phys. 10 (2013) 041.
[18] M. Unsal, What is QFT? Resurgent trans-series, Lefschetz thimbles, and new exact saddles (to be published).
[19] M. Cristoforetti, F. Di Renzo, G. Eruzzi, A. Mukherjee, C. Schmidt, L. Scorzato, and C. Torrero, An efficient method to compute the residual phase on a Lefschetz thimble, Phys. Rev. D **89**, 114505 (2014).
[20] G. Parisi, On complex probabilities, Phys. Lett. **131B**, 393 (1983).
[21] J. R. Klauder, Coherent state Langevin equations for canonical quantum systems with applications to the quantized Hall effect, Phys. Rev. A **29**, 2036 (1984).
[22] G. Aarts, L. Bongiovanni, E. Seiler, D. Sexty, and I. O. Stamatescu, Controlling complex Langevin dynamics at finite density, Eur. Phys. J. A **49**, 89 (2013).
[23] G. Aarts, P. Giudice, and E. Seiler, Localised distributions and criteria for correctness in complex Langevin dynamics, Ann. Phys. (Amsterdam) **337**, 238 (2013).
[24] G. Aarts, Lefschetz thimbles and stochastic quantization: Complex actions in the complex plane, Phys. Rev. D **88**, 094501 (2013).
[25] G. Aarts, L. Bongiovanni, E. Seiler, and D. Sexty, Some remarks on Lefschetz thimbles and complex Langevin dynamics, J. High Energy Phys. 10 (2014) 159.
[26] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevD.92.085030 for an (avi) movie. We show how the thimbles structure evolves while we go from positive values of σ_R to negative ones (one passes the situation in which a Stokes phenomenon occurs, i.e., $\sigma_R = 0$).
[27] A. Mukherjee, M. Cristoforetti, and L. Scorzato, Metropolis Monte Carlo on the Lefschetz thimble: Application to a one-plaquette model, Phys. Rev. D **88**, 051502 (2013).
[28] E. V. Shuryak and J. J. M. Verbaarschot, Random matrix theory and spectral sum rules for the Dirac operator in QCD, Nucl. Phys. **A560**, 306 (1993).
[29] J. J. M. Verbaarschot and I. Zahed, Spectral Density of the QCD Dirac Operator near Zero Virtuality, Phys. Rev. Lett. **70**, 3852 (1993).
[30] J. J. M. Verbaarschot, The Spectrum of the QCD Dirac Operator and Chiral Random Matrix Theory: The Threefold Way, Phys. Rev. Lett. **72**, 2531 (1994).
[31] https://bitbucket.org/ggscorzato/thimble-monte-carlo.
[32] F. Di Renzo, G. Eruzzi, and M. Brambilla, An algorithm for thimble regularization of lattice field theories (and possibly not only for that), *Proc. Sci.*, LATTICE (2014) 046.