PHYSICAL REVIEW D **83,** 045021 (2011)

Schwinger-Dyson equations in large-N quantum field theories and nonlinear random processes

P. V. Buividovich[1,2,*]

[1]*ITEP, Bolshaya Cheremushkinskaya 25, 117218 Moscow, Russia*
[2]*JINR, Joliot-Curie 6, 141980 Dubna, Moscow region, Russia*

(Received 27 December 2010; published 23 February 2011)

We propose a stochastic method for solving Schwinger-Dyson equations in large-N quantum field theories. Expectation values of single-trace operators are sampled by stationary probability distributions of the so-called nonlinear random processes. The set of all the histories of such processes corresponds to the set of all planar diagrams in the perturbative expansions of the expectation values of singlet operators. We illustrate the method on examples of the matrix-valued scalar field theory and the Weingarten model of random planar surfaces on the lattice. For theories with compact field variables, such as sigma models or non-Abelian lattice gauge theories, the method does not converge in the physically most interesting weak-coupling limit. In this case one can absorb the divergences into a self-consistent redefinition of expansion parameters. A stochastic solution of the self-consistency conditions can be implemented as a ''memory'' of the random process, so that some parameters of the process are estimated from its previous history. We illustrate this idea on the two-dimensional $O(N)$ sigma model. The extension to non-Abelian lattice gauge theories is discussed.

DOI: 10.1103/PhysRevD.83.045021 PACS numbers: 02.70.−c, 02.50.Ey, 11.15.Pg

I. INTRODUCTION

Modern lattice QCD simulations are mostly based on the direct evaluation of the path integral of the theory. Such an approach, while being very general and efficient for many applications, suffers from a number of problems, most notable of which are the sign problem at finite chemical potential, the critical slowing down at small quark masses, and large finite-volume effects as well as small signal-to-noise ratios in the analysis of excited states. These problems are inherent to standard Monte Carlo simulations and cannot be efficiently solved by simply increasing the computation power, since the required computing time quickly increases (in the worst cases, exponentially) with the required precision. Such a situation makes it tempting to devise alternative simulation algorithms for non-Abelian lattice gauge theories.

One of the efficient alternative numerical methods is the so-called diagrammatic Monte Carlo method, a method based on the stochastic summation of all the terms in the strong- or weak-coupling expansion of the observable of interest [1,2]. Such a method in some cases allows one to reduce or avoid completely the sign problem in the original path integral, and does not suffer from finite-volume effects. Furthermore, one can construct algorithms which yield particular correlation functions in terms of probability distributions of some random variables, which greatly facilitates the analysis of excited states [1,2]. This is the idea of the ''worm'' algorithm by Prokof'ev and Svistunov [1], in which the probability distribution of the positions x, y of the ''head'' and the ''tail'' of the worm yields the two-point Green function $G(x, y)$. The diagrammatic Monte Carlo method and the worm algorithm have been successfully applied to a number of statistical models with discrete symmetry groups, such as the Ising model, the XY model, and unitary Fermi gas, and showed practically no critical slowing down near quantum phase transitions.

However, the application of such methods to lattice field theories with continuous field variables [such as two-dimensional $O(N)$ and $CP(N)$ sigma models, Abelian gauge theories, and the ϕ^4 theory] has resulted so far in quite complicated and model-dependent algorithms [2]. A generalization of such algorithms to $SU(N)$ sigma models or to non-Abelian gauge theories has not yet been found. These algorithms are, in essence, based on the strong-coupling expansion, and while their applicability is not limited by the strong-coupling regime, one can expect that algorithms based on the weak-coupling expansion might show better performance near the continuum limit.

Typically, the weak-coupling expansion in such lattice theories is either quite complicated or nonconvergent. Up to now, the divergent behavior of the weak-coupling perturbative expansions strongly limits the applicability of the diagrammatic Monte Carlo model to field theories with continuous field variables. In a recent paper [3] a method was proposed to construct convergent series which approximate the nonanalytic path integrals with the desired precision. This method, however, is difficult to generalize to physically interesting field theories such as non-Abelian lattice gauge theories.

Another way to obtain convergent series while preserving important physical properties of the theory is to sum over only diagrams with a certain topology. This corresponds to the large-N limit in quantum field theories and matrix models, that is, the limit of infinite dimensionality of an internal symmetry group, such as $O(N)$ or $SU(N)$.

*buividovich@itep.ru

For such theories, each Feynman diagram acquires a factor N^χ, where χ is the Euler character of this diagram [4]. In the limit $N \to \infty$, the contribution of planar diagrams with $\chi = 2$ dominates, and the sum over all planar diagrams typically has a finite convergence radius [5,6].

In this paper we describe a stochastic method for summing over all planar diagrams in large-N quantum field theories. The method is based on the stochastic solution of Schwinger-Dyson equations, so that the correlators of field variables are obtained as stationary probability distributions of certain random variables. In this way, we implement the idea of importance sampling, so that numerically small observables correspond to unlikely events. These probability distributions are sampled by the so-called nonlinear random processes. In contrast to conventional Markov chains, stationary probability distributions of such random processes satisfy nonlinear equations, and hence they can be called ''nonlinear random processes'' or ''nonlinear Markov chains'' in the terminology of [7,8]. The factorization of single-trace operators in the large-N limit of quantum field theories corresponds to the phenomena of ''chaos propagation'' in random processes [9].

While in the diagrammatic Monte Carlo method and in the worm algorithm the diagrams are stored in computer memory as a whole and are updated in such a way that the detailed balance condition is satisfied at each step, the method described in this paper works only with external lines. In contrast to the standard Metropolis algorithm, one should not know explicitly the weight of each diagram, and the transition probabilities do not satisfy any detailed balance condition. Unlike the quite popular ''numerical functional methods'' in continuum gauge theories (see [10] for a review), the proposed method does not require any truncation of the hierarchy of Schwinger-Dyson equations, and works only with singlet operators with respect to the internal symmetry group. Another distinct feature is that the computational complexity of the method does not depend on N, while the standard Monte Carlo method, the functional methods, and the worm algorithm all require infinite computational resources in the limit $N \to \infty$. This feature might be advantageous for numerical checks of the predictions of the holographic models which are dual to large-N quantum field theories [11].

In Sec. II we analyze the general structure of Schwinger-Dyson equations in large-N quantum field theories using the example of a scalar matrix-valued field theory. When large-N factorization is taken into account, Schwinger-Dyson equations become nonlinear equations with infinitely many unknowns. In Sec. III we describe nonlinear random processes of recursive type [7], which can be used to stochastically solve such equations. In Sec. IV we apply such random processes to solve Schwinger-Dyson equations in several large-N theories. In Sec. IV A we consider the scalar matrix-valued field theory, for which the perturbative expansion yields the conventional Feynman diagrams in momentum space. In Sec. IV B this solution is compared with the exact solution of the simplest quantum field theory in zero dimensions, that is, the Hermitian matrix model [6]. The convergence of such a solution and the strength of the sign problem are discussed. In Sec. IV C we consider the Weingarten model [12,13] and demonstrate how the proposed method can be used to simulate random surfaces on the hypercubic lattice. In this case, our method reproduces an ensemble of open, rather than closed, random surfaces, with critical behavior which is quite different from that of the closed planar random surfaces. Since the structure of Schwinger-Dyson equations in the Weingarten model is similar to the loop equations in large-N non-Abelian lattice gauge theories [14], studying this model might be helpful for further extensions of the present approach to non-Abelian gauge theories.

While the method described in Sec. III works well for noncompact field variables, for field theories with compact field variables, such as nonlinear sigma models or non-Abelian lattice gauge theories, a straightforward stochastic interpretation of Schwinger-Dyson equations is only possible in the strong-coupling limit. In the weak-coupling limit one expects the field correlators to contain both the perturbative part in the coupling constant g and the nonperturbative corrections of the form $\exp(-c/g^2)$ with some constant c. Moreover, a perturbative expansion in powers of g typically results in asymptotic series, and nonperturbative corrections appear as a result of the resummation of such series [15].

In Sec. V we show how such nonperturbative corrections can be taken into account by a further relaxation of the Markov property of the random process. The basic idea is to absorb the divergent part of the series into a self-consistent redefinition of the expansion parameter. These redefined parameters play the role of nonperturbative ''condensates'' [15,16]. It turns out that the redefined expansion parameters can be estimated with increasing precision from the previous history of the random process which solves the Schwinger-Dyson equations, thus leading to the emergence of the memory of the random process. The approach of the redefined parameters to their self-consistent values is reminiscent somehow of the renormalization-group flow [10]. Such a dependence on the previous history makes the random process essentially non-Markovian, so that the stationary probability distribution also satisfies some nonlinear equation.

We illustrate this idea using the example of an $O(N)$ sigma model in two dimensions, which is equivalent to a bosonic random walk with a self-consistent mass. The random process which simulates this model has memory but no ''recursive'' structure. Presumably, in order to sum up both perturbative and nonperturbative corrections which arise in non-Abelian lattice gauge theories or $U(N)$ sigma models, one should devise the recursive nonlinear random process (which would sum up perturbative

corrections) with memory (which would generate nonperturbative quantities in a renormalization-group-like way).

Finally, in the concluding section we summarize the present work and discuss its extension to non-Abelian lattice gauge theories in the limit of large N.

II. GENERAL STRUCTURE OF SCHWINGER-DYSON EQUATIONS FOR LARGE-N QUANTUM FIELD THEORIES

In order to analyze the general structure of Schwinger-Dyson equations for large-N quantum field theories, let us first consider the theory of a Hermitian $N \times N$ matrix-valued field $\phi(x)$ with the following Lagrangian:

$$\mathcal{L}[\phi(x)] = N \operatorname{Tr}\phi(x)(m^2 - \Delta)\phi(x) + \frac{N\lambda}{4} \operatorname{Tr}\phi^4(x). \quad (1)$$

This theory is most convenient to illustrate the method described in this paper, since its perturbative expansion leads to conventional Feynman diagrams in momentum space. Since this theory should be somehow regularized, let us assume from the very beginning that the action (1) is defined on the Euclidean hypercubic D-dimensional lattice with total volume V in lattice units. Thus, the coordinates x take integer values, and Δ is the lattice Laplacian (for definiteness, with periodic boundary conditions).

Schwinger-Dyson equations for a theory with the action (1) read [17]

$$(m^2 - \Delta_1)G(x_1, x_2) = \delta(x_1, x_2) + \lambda G(x_1, x_1, x_1, x_2), \quad (2)$$

$$\begin{aligned}(m^2 - \Delta_1)&G(x_1, \ldots, x_n) \\ &= \delta(x_1, x_2)G(x_3, \ldots, x_n) + \delta(x_1, x_n)G(x_2, \ldots, x_{n-1}) \\ &\quad + \sum_{m=3}^{n-1} \delta(x_1, x_m)G(x_2, \ldots, x_{m-1})G(x_{m+1}, \ldots, x_n) \\ &\quad + \lambda G(x_1, x_1, x_1, x_2, \ldots, x_n), \qquad n > 2, \end{aligned} \quad (3)$$

where the single-trace correlators are $G(x_1, \ldots, x_n) = \langle \frac{1}{N} \operatorname{Tr}(\phi(x_1) \ldots \phi(x_n)) \rangle$, Δ_1 is the Laplacian acting on x_1, and we have already taken into account the factorization property in the limit $N \to \infty$ [4]:

$$\begin{aligned}\left\langle \frac{1}{N} \operatorname{Tr}(\phi(x_1) \ldots \phi(x_n)) \frac{1}{N} \operatorname{Tr}(\phi(y_1) \ldots \phi(y_m)) \right\rangle &= \left\langle \frac{1}{N} \operatorname{Tr}(\phi(x_1) \ldots \phi(x_n)) \right\rangle \left\langle \frac{1}{N} \operatorname{Tr}(\phi(y_1) \ldots \phi(y_m)) \right\rangle \\ &\quad + O\left(\frac{1}{N^2}\right). \end{aligned} \quad (4)$$

These equations hold for any argument of the correlators, but the resulting system is redundant, and it is sufficient to consider only those Schwinger-Dyson equations which were obtained by the variation of the fields at x_1.

It is convenient now to go to the momentum representation, introducing the Green functions in momentum space, $G(k_1, \ldots, k_n) = \sum_{x_1} \ldots \sum_{x_n} \exp(i \sum_m k_m \cdot x_m) G(x_1, \ldots, x_n)$. In order to keep all expressions as symmetric as possible, we do not separate the factor $\delta(\sum_m k_m)$ in $G(k_1, \ldots, k_n)$ explicitly. This condition will be automatically satisfied by the nonlinear random process which we describe in Sec. IV A. Equations (2) and (3) in the momentum representation are

$$\begin{aligned}G(k_1, k_2) &= G_0(k_1)V\delta(k_1 + k_2) + G_0(k_1)\frac{\lambda}{V^2} \\ &\quad \times \sum_{q_1, q_2, q_3} \delta(k_1 - q_1 - q_2 - q_3)G(q_1, q_2, q_3, k_2), \end{aligned} \quad (5)$$

$$\begin{aligned}G(k_1, \ldots, k_n) &= G_0(k_1) \sum_{m=3}^{n-1} \delta(k_1 + k_m)VG(k_2, \ldots, k_{m-1}) \\ &\quad \times G(k_{m+1}, \ldots, k_n) \\ &\quad + G_0(k_1)\delta(k_1 + k_2)VG(k_3, \ldots, k_n) \\ &\quad + G_0(k_1)\delta(k_1 + k_n)VG(k_2, \ldots, k_{n-1}) \\ &\quad + G_0(k_1)\frac{\lambda}{V^2} \sum_{q_1, q_2, q_3} \delta(k_1 - q_1 - q_2 - q_3) \\ &\quad \times G(q_1, q_2, q_3, k_2, \ldots, k_n), \end{aligned} \quad (6)$$

where $G_0(k) = (m^2 + \sum_\mu 4\sin^2(k_\mu/2))^{-1}$ is the free scalar propagator on the hypercubic lattice. All momenta are assumed to lie in the first Brillouin zone $-\pi \le k_\mu \le \pi$ and are added modulo 2π. The structure of these equations is schematically illustrated in Fig. 1, where dashed blobs denote the Green functions $G(k_1, \ldots, k_n)$ and empty blobs denote $G_0(k)$.

Thus we have obtained an infinite system of quadratic functional equations for the set of functions $G(k_1, \ldots, k_n)$ with $n = 2, 4, \ldots$. Such a structure is common for large-N quantum field theories; Schwinger-Dyson equations are quadratic equations for an infinite set of unknown variables. In the case of scalar matrix field theory considered

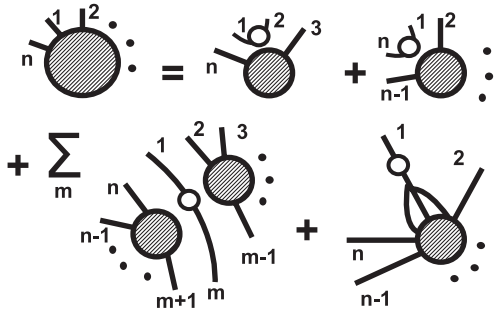

FIG. 1. Schematic illustration of the structure of the Schwinger-Dyson equations (6). Dashed blobs denote the Green functions $G(k_1, \ldots, k_n)$ and empty blobs denote the free propagator $G_0(k)$.

here, the unknown variables are the functions of the sequences of momenta $\{k_1, \ldots, k_n\}$ for any even $n \geq 2$. In the case of lattice gauge theories or string theories, Schwinger-Dyson equations are most naturally formulated in terms of the Wilson loops, which are the functions defined on the discrete space of closed loops on the lattice [12–14]. In this case, the equations are also quadratic with respect to the Wilson loops. For an $O(N)$ sigma model, Schwinger-Dyson equations are also quadratic equations which involve only the two-point function (see Sec. V).

Typically, systems of equations with infinitely many unknowns can be efficiently solved by stochastic methods. It is advantageous to estimate the value of each unknown variable as a probability of observing some state of a random process. In this case the unknowns with numerically small values correspond to unlikely events, and the set of infinitely many unknown variables is automatically truncated to a set of unknowns with sufficiently large values. Such methods are well known mainly in the context of kinetic equations [9]. Recently, they were also discussed in the context of probabilistic programming [7]. In the next section we describe a discrete-time, discrete-space method of such type, which is in our opinion most suitable for solving the Schwinger-Dyson equations in the large-N limit.

III. STOCHASTIC SOLUTION OF NONLINEAR EQUATIONS BY RANDOM PROCESSES OF RECURSIVE TYPE

We consider nonlinear equations of the following form:

$$w(x) = p_c(x) + \sum_y p_e(x|y)w(y) + \sum_{y_1,y_2} p_j(x|y_1, y_2)w(y_1)w(y_2), \tag{7}$$

where x, y, y_1, y_2 are the elements of some space X and $\sum_x$, and $x \in X$ denotes summation or integration over all the elements of this space. We also assume that the functions $p_c(x)$, $p_e(x|y)$, and $p_j(x|y_1, y_2)$ satisfy the inequalities

$$\sum_x |p_c(x)| + |p_e(x|y_1)| + |p_j(x|y_1, y_2)| < 1 \tag{8}$$

for any y_1, y_2.

We would like to find a stochastic process for which $w(x)$ is proportional to the probability of the occurrence of the element x in some configuration space. Obviously, an ordinary Markov process with configuration space X cannot solve such a problem, since stationary distributions of Markov processes obey linear equations. In order to solve the nonlinear equation (7), one can, for example, somehow extend the configuration space. Extensions of Markov processes with stationary probability distributions which obey nonlinear equations have been considered recently in [7,8]. In this section we concentrate on random processes similar to the recursive Markov chains of [7]. The basic idea is that, at any time, one can leave the current chain and start a new one, then return back to the old chain at some time. The initial state of a newly created chain depends on the states of older chains. Thus one does not have a single Markov chain, but rather an infinite stack of chains. The random process which we describe below will be similar to these recursive Markov chains, but instead of referring to "recursion," we will explicitly introduce the underlying stack structure. Here we first consider Eqs. (7) with the coefficients $p_c(x)$, $p_e(x|y)$, and $p_j(x|y_1, y_2)$ all being positive, and in the Appendix we generalize to coefficients with arbitrary signs or complex phases.

Consider an extended configuration space which consists of ordered sequences $\{x_1, \ldots, x_n\}$ for arbitrary $n \geq 1$, with $x_1, \ldots, x_n \in X$. It is illustrative to interpret such configuration space as a stack of elements of the space X, so that x_n is at the top of the stack. The desired random process can be specified by the following prescriptions. At each discrete time step, do one of the following:

Create.—With the probability $p_c(x)$ create a new element $x \in X$ and push it to the stack.

Evolve.—With the probability $p_e(x|y)$ pop the element y from the stack and push the element x to the stack.

Join.—With the probability $p_j(x|y_1, y_2)$ consecutively pop two elements y_1, y_2 from the stack and push a single element x to the stack.

Restart.—With the probability $1 - \sum_x (p_c(x) + p_e(x|y_1) + p_j(x|y_1, y_2))$, where y_1, y_2 are the two topmost elements in the stack, empty the stack and push a single element $x \in X$ into it, with a probability distribution proportional to $p_c(x)$.

The last action is also the procedure used to initialize the random process. The "evolve" action is just the evolution of a single Markov chain at the top of the stack, with transition probabilities proportional to $p_e(x|y)$. The condition (8) and the positivity requirement ensure that $p_c(x)$, $p_e(x|y)$, and $p_j(x|y_1, y_2)$ can be interpreted as probabilities.

Consider now an equation for the stationary probability distribution of such a Markov chain. It has a general form $p(A) = \sum_B P(B \to A)p(B)$, where $p(A)$ is a stationary probability of the occurrence of a state A and $P(B \to A)$ is the transition probability between the states B and A. Let $W(x_1, \ldots, x_n)$ be the stationary probability to find the elements $x_1, \ldots, x_n$ in the stack. This probability distribution function is obviously normalized to unity:

$$\sum_{n=1}^{\infty} \sum_{x_1} \cdots \sum_{x_n} W(x_1, \ldots x_n) = 1.$$

The equation for the stationary probability distribution in our case reads

$$W(x_1) = \mathcal{N}_c^{-1} p_c(x_1)\xi_R + \sum_y p_e(x_1|y)W(y) + \sum_{y_1,y_2} p_j(x_1|y_1,y_2)W(y_1,y_2), \quad (9)$$

$$W(x_1,\ldots,x_n) = p_c(x_n)W(x_1,\ldots,x_{n-1}) + \sum_y p_e(x_n|y)W(x_1,\ldots,x_{n-1},y) + \sum_{y_1,y_2} p_j(x_n|y_1,y_2) \times W(x_1,\ldots,x_{n-1},y_1,y_2), \quad n>1, \quad (10)$$

where

$$\xi_R = \sum_n \sum_{x_1} \ldots \sum_{x_n} (1 - p_c(x_n)W(x_1,\ldots,x_{n-1}) - \sum_y p_e(x_n|y)W(x_1,\ldots,x_{n-1},y) - \sum_{y_1,y_2} p_j(x_n|y_1,y_2)W(x_1,\ldots,x_{n-1},y_1,y_2)) \quad (11)$$

and $\mathcal{N}_c = \sum_x p_c(x)$. By a direct substitution one can check that there is a factorized solution for $W(x_1,\ldots,x_n)$:

$$W(x_1,\ldots,x_n) = w_0(x_1)w(x_2)\ldots w(x_n), \quad (12)$$

where $w(x)$ obeys exactly Eq. (7) and $w_0(x)$ obeys the following inhomogeneous linear equation:

$$w_0(x) = \mathcal{N}_c^{-1} p_c(x)\xi_R + \sum_y p_e(x|y)w_0(y) + \sum_{y_1,y_2} p_j(x|y_1,y_2)w_0(y_1)w(y_2). \quad (13)$$

Thus, for any equation of the form (7) with positive coefficients which satisfy (8), there is a random process whose stationary distribution encodes the solution of this equation as in (12). The factorization of the stationary probability distribution of random processes with such an infinite configuration space is known as the "propagation of chaos" in random processes and was discovered for classical kinetic equations by McKean, Vlasov, and Kac [9]. Comparing Eq. (7) with the Schwinger-Dyson equations (2), (3), and (5), we conclude that this property corresponds to the factorization of single-trace operators in large-N quantum field theories. It is interesting that time reversal of the random process described above leads to the so-called branching random process [9], which has quite different properties. This is due to the fact that for such random processes there is no detailed balance condition, and hence no time reversal symmetry. We do not consider here a subtle mathematical question of the existence of solutions to Eq. (7), since in our case it is ensured by the physical applications of this equation.

Finally, let us describe a practical procedure for finding $w(x)$ by simulating the random process described above. By standard statistical methods, one should sample the probability distribution $p(x_n)$ of the topmost element in the stack [provided there is more than one element in it; otherwise we estimate $w_0(x)$ rather than $w(x)$; see (12)]. From (12), we get $p(x_n) = s^{-1}w(x_n)$, with $s = \sum_x w(x)$. It should be stressed that $w(x)$ is not normalized to unity, but rather satisfies the inequality $\sum_x w(x) = s < 1$. The value of the normalization constant s can also be easily found numerically, since the probability to find n elements in the stack decreases as s^n for $n > 1$.

IV. STOCHASTIC SOLUTION OF SCHWINGER-DYSON EQUATIONS BY RECURSIVE RANDOM PROCESSES

A. Scalar matrix field theory

After presenting the general method in Sec. III, we are ready to describe a stochastic numerical solution of the Schwinger-Dyson equations (5) and (6). For simplicity, let us assume that the coupling constant λ in (1) is negative. This allows us to apply directly the results of Sec. III, where all the coefficients in (7) are assumed to be positive. In the case of positive λ, additional sign variables for each sequence of momenta can be easily introduced following the Appendix. This will be done in the next subsection for the Hermitian matrix model. Note that while at finite N the theory with a negative coupling constant is not defined and the correlators are nonanalytic in λ [3], in the leading order in N perturbative series converge even when the coupling is negative, but not exceeding some critical value [6]. Correspondingly, in the planar approximation the correlators are analytic in λ.

The space X in (7) should be the space of ordered sequences (of any size) of momenta $\{k_1,\ldots,k_n\}$; correspondingly, the extended configuration space is a stack which contains such sequences. It is convenient also to introduce two normalization constants $\mathcal{N}$ and c, so that the functions $w(k_1,\ldots k_n)$ which will be estimated stochastically are defined as

$$G(k_1,\ldots,k_n) = \mathcal{N}V^n c^{n-2} w(k_1,\ldots,k_n), \quad (14)$$

where V is again the total volume of space. The constant c can be thought of as the renormalization constant for the one-particle wave functions, and $\mathcal{N}$ as the overall wave-function normalization.

In terms of the functions $w(k_1,\ldots,k_n)$ the Schwinger-Dyson equations (5) and (6) read

$$w(k_1,k_2) = G_0(k_1)\mathcal{N}^{-1}\frac{\delta(k_1+k_2)}{V} + G_0(k_1)\lambda c^2 \times \sum_{q_1,q_2,q_3} \delta(k_1 - q_1 - q_2 - q_3)w(q_1,q_2,q_3,k_2), \quad (15)$$

$$
\begin{aligned}
w(k_1,\ldots,k_n) = {} & G_0(k_1)c^{-2}\frac{\delta(k_1+k_2)}{V}w(k_3,\ldots,k_n) \\
& + G_0(k_1)c^{-2}\frac{\delta(k_1+k_n)}{V}w(k_2,\ldots,k_{n-1}) \\
& + G_0(k_1)\mathcal{N}c^{-4}\sum_{m=3}^{n-1}\frac{\delta(k_1+k_m)}{V} \\
& \times w(k_2,\ldots,k_{m-1})w(k_{m+1},\ldots,k_n) \\
& - G_0(k_1)\lambda c^2\sum_{q_1,q_2,q_3}\delta(k_1-q_1-q_2-q_3) \\
& \times w(q_1,q_2,q_3,k_2,\ldots,k_n). \qquad (16)
\end{aligned}
$$

Comparing the Schwinger-Dyson equations (15) and (16) with the general equation (7), we arrive at the random process which stochastically solves these equations. This random process is specified by the following probabilistic choice of actions at each discrete time step:

Create.—With the probability $G_0(k)(\mathcal{N}V)^{-1}$ push a new sequence of momenta $\{k,-k\}$ to the stack.

Add.—With the probability $G_0(k)c^{-2}/V$ modify the topmost sequence of momenta $\{k_1,\ldots,k_n\}$ in the stack by adding a pair of momenta $\{k,-k\}$ either as $\{k,k_1,\ldots,k_n,-k\}$ or $\{k,-k,k_1,\ldots,k_n\}$.

Create vertex.—With the probability $|\lambda|G_0(q_1+q_2+q_3)c^2$ replace the topmost sequence $\{q_1,q_2,q_3,k_2,\ldots,k_n\}$ in the stack by $\{q_1+q_2+q_3,k_2,\ldots,k_n\}$. This action can only be performed if the topmost sequence contains more than two elements.

Join.—With the probability $G_0(k)\mathcal{N}c^{-4}/V$ pop the two sequences $\{k_1,\ldots,k_n\}$, $\{q_1,\ldots,q_m\}$ from the stack (provided there are more than two elements in it) and join them into a single sequence as $\{k_1,\ldots,k_n,k,q_1,\ldots,q_n,-k\}$. Push the result to the stack.

Restart.—Otherwise, restart with a stack containing a sequence $\{k,-k\}$, k being distributed with a probability proportional to $G_0(k)$.

Since the momenta are always added to the stack in pairs which sum up to zero, for all sequences in the stack the total sum of all momenta in the sequence is always zero. The V^{-1} factors in (15) and (16) ensure that the probability distributions of the newly created momenta can be normalized to unity.

Let us check whether the inequalities (8) are satisfied for such a process, that is, whether the total probability of all possible actions does not exceed unity. For the free propagator $G_0(k)$, one has the inequalities $G_0(k)<1/m^2$ and $\sum_k G_0(k)<V/m^2$. The total probability of all possible actions can then be estimated as $(\mathcal{N}^{-1}+c^{-2}+|\lambda|c^2+\mathcal{N}c^{-4})/m^2$. Clearly, for sufficiently small $|\lambda|$ this estimate can always be made smaller than unity by increasing c and $\mathcal{N}$. In Secs. IV A and IV B we will analyze such bounds on coupling constants in more detail for the Hermitian matrix model and for the Weingarten model.

Since the constructed process involves no permutations, one can trace the history of each momentum in the stack—from creation to joining into a vertex or a "restart operation." By drawing all the momenta in the stack as points on the vertical lines of some two-dimensional grid and connecting the corresponding points along the horizontal lines, all planar diagrams of the theory (with an arbitrary number of external lines) can be obtained. Note also that the number of vertices in the planar diagrams drawn by this random process cannot exceed the number of time steps from the previous "restart" action. Thus, in order to maximize the mean order of diagrams which are summed up in some fixed number of time steps, it is advantageous to maximally reduce the rate of restart events, that is, to saturate the inequalities (8).

One could also try to devise a random process which would solve the Schwinger-Dyson equations (2) and (3) directly in physical space-time, rather than in momentum space. The configuration space of such a process would be the stack of sequences of points $\{x_1,\ldots,x_n\}$. As compared to the algorithm in momentum space, there would be an additional choice of moving the last point x_n in the topmost sequence to adjacent lattice sites, with the probability proportional to the hopping parameter $\kappa=(2D+m^2)^{-1}$. This would correspond to drawing the worldlines of virtual and real particles by bosonic random walks. Interestingly, such worldlines can be mapped onto the string world sheets in simplicial string theory [18]. In addition, the creation of a new interaction vertex would only be possible if three such random walks intersect in one point. However, this is an unlikely event, with the probability going to zero in the continuum limit. Thus, solving the Schwinger-Dyson equations directly in the coordinate representation would lead to a less efficient numerical algorithm.

Note that for the theory (1) at finite N the Schwinger-Dyson equations are linear equations, which are, however, defined on much larger functional space: the set of unknown functions also includes expectation values of multitrace operators, such as $\langle \mathrm{Tr}(\phi(x_1)\ldots\phi(x_n))\,\mathrm{Tr}(\phi(y_1)\ldots\phi(y_m))\rangle$. One can try to solve these linear equations by interpreting them as the equations for the stationary probability distribution of a Markov process. The configuration space of such a process should be a space of sequences of the form $\{\{x_1,\ldots,x_n\},\ldots,\{y_1,\ldots,y_m\}\}$, thus encoding the expectation values of all multitrace operators. However, such a straightforward procedure leads to non-normalizable transition probabilities, indicating that the series which one tries to sum up are divergent. Only when the terms subleading in $1/N$ are omitted from the Schwinger-Dyson equations can they be interpreted as stochastic equations. At the same time, we obtain the Markov process on the extended configuration space described in Sec. III, which we interpret as the stack of sequences. The property of the propagation of chaos [9] ensures large-N factorization of

single-trace operators [see Eq. (12)]. We are thus led to the random process of recursive type [7].

B. Hermitian matrix model

To check the considerations of the previous subsection, let us consider the theory (1) in zero dimensions, that is, the Hermitian matrix model with the following partition function:

$$\mathcal{Z}(\lambda) = \int \prod_{i,j} d\phi_{ij} \exp\left(-N/2\,\mathrm{Tr}\phi^2 + \frac{\lambda N}{4}\,\mathrm{Tr}\phi^4\right). \quad (17)$$

The Green functions now depend only on one integer n: $G_n \equiv G(n) = \langle \frac{1}{N}\,\mathrm{Tr}\phi^{2n}\rangle$. The Schwinger-Dyson equations (2) and (3) also take a very simple form:

$$\begin{aligned} &G_1 = 1 + \lambda G_2, \\ &G_n = 2G_{n-1} + \sum_{m=1}^{n-2} G_m G_{n-m-1} + \lambda G_{n+1}, \qquad n > 1. \end{aligned} \quad (18)$$

Here we will assume that the coupling constant λ can be both positive and negative, in order to illustrate the method described in the Appendix. Let us again define the "renormalized" Green functions w_n as $G_n = \mathcal{N} c^{n-1} w_n$. In the case of an arbitrary sign of λ, the configuration space of the random process should be the stack which contains integer positive numbers and additional sign variables. Following the Appendix, we introduce the variables $w_n^{(+)}$ and $w_n^{(-)}$, which are proportional to the probabilities to find the elements $\{n, +\}$ or $\{n, -\}$ at the top of the stack (provided the stack contains more than one element). Then $w_n = w_n^{(+)} - w_n^{(-)}$. We thus arrive at the following random process for the stochastic evaluation of $w_n^{(\pm)}$. At each discrete time step, one performs, at random, one of the following actions:

(i) With the probability $\mathcal{N}^{-1}$ add a new element $\{1, +\}$ to the stack.
(ii) With the probability $2c^{-1}$ increase the topmost element in the stack by 1 and do not change its sign.
(iii) With the probability $|\lambda| c$ decrease the topmost element in the stack by 1 (if it is greater than 1) and multiply its sign by the sign of λ.
(iv) With the probability $\mathcal{N} c^{-2}$ pop the two elements $\{n, s_1\}$ and $\{m, s_2\}$ from the stack (provided there are more than two elements) and push the element $\{n + m + 1, s_1 s_2\}$ to the stack.
(v) Otherwise, empty the stack and push a single element $\{1, +\}$ into it.

Note that for positive λ elements with a minus sign are not generated, so that $w_n^{(-)} \equiv 0$ and the random process automatically reduces to the one described in Sec. III.

The inequalities (8) now read

$$\mathcal{N}^{-1} + 2c^{-1} + |\lambda| c + \mathcal{N} c^{-2} \le 1, \qquad \mathcal{N} > 0, \qquad c > 0. \quad (19)$$

As discussed in Sec. IVA, in order to increase the efficiency of the algorithm it is advantageous to saturate this inequality. It is easy to see that, at the same time, we saturate the upper bound on the absolute value of the coupling constant λ. Maximizing this upper bound with respect to $\mathcal{N}$ and c, we see that $|\lambda|$ cannot exceed the value $\bar{\lambda} = 1/16 = 3/4\lambda_c$, where $\lambda_c = 1/12$ is the convergence radius of the planar perturbative expansion which can be found from the exact solution of the matrix model (17) [6]. Thus, the random process described above covers only some finite subrange of coupling constants for which the model (17) is defined. It is easy to understand the origin of this limitation; in fact, this random process simulates an ensemble of diagrams with an arbitrary number of external legs, with the weight of each diagram being proportional to λ^{N_v}, where N_v is the number of vertices. The number of open diagrams with a given number of vertices is obviously larger than the number of closed diagrams; hence the sums over open diagrams have a smaller convergence radius.

For $|\lambda| < \bar{\lambda}$, there is a continuous set of $\mathcal{N}$, c which saturate the inequality (19). One can, for example, fix c and express $\mathcal{N}$ as a function of λ:

$$\mathcal{N} = \frac{c^2 - 2c - |\lambda| c^3 \pm \sqrt{c^3(c|\lambda| - 1)(4 - c + c^2|\lambda|)}}{2}. \quad (20)$$

We call the solution with the minus sign in front of the square root "branch 1" and the other solution "branch 2."

In Fig. 2 we plot the Green functions $G_n(\lambda)$, evaluated using the random process described above, as functions of λ up to $n = 5$. These results were obtained after $N = 10^6$ discrete time steps at fixed $c = 8$ and with $\mathcal{N}$ given by branch 1 of (20). The error bars are smaller than the symbols on the plot. Solid lines are the exact results for $G_n(\lambda)$ in the planar approximation, obtained using the saddle point method [6].

The autocorrelation time and mean stack size for the random process described above are plotted in Figs. 3 and 4, respectively, as functions of the coupling constant λ. The observable used to define the autocorrelation time was the sum of all numbers in the stack. First, we note that branch 1 is more advantageous for simulations, since with larger mean stack size one can gain more statistics. However, in this case the autocorrelation time is also larger. Interestingly, for this branch both the autocorrelation time and the mean stack size have a maximum near $\lambda = 0$ rather than near the "critical point" of the random process $\bar{\lambda}$. For branch 2, these quantities increase slowly towards $\bar{\lambda}$.

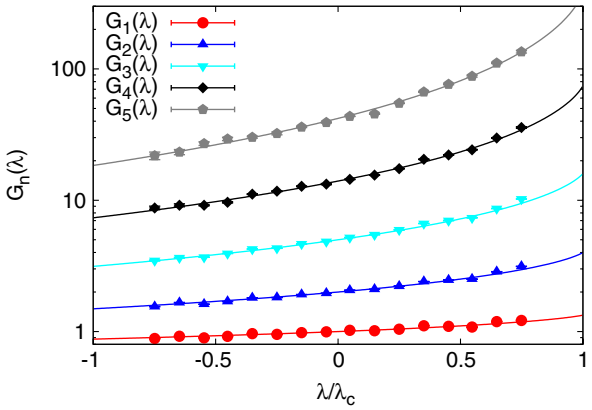

FIG. 2 (color online). Green functions $G_n(\lambda)$ in (18) versus the coupling constant λ for $n = 1, \ldots, 5$, obtained after $N = 10^6$ discrete time steps of the algorithm described above at fixed $c = 8$ and with $\mathcal{N}$ given by branch 1 of (20). The error bars are smaller than the symbols on the plot. Exact results of [6] are plotted with solid lines.

In order to characterize the strength of the sign problem, we consider the quantity

$$\xi_n^s = (w_n^{(+)} - w_n^{(-)})/(w_n^{(+)} + w_n^{(-)}). \tag{21}$$

$\xi_n^s = 1$ if the random process generates only elements with a plus sign, and $\xi_n^s = 0$ if the numbers of pluses and minuses cancel exactly. In practice, it is advantageous to have as large a ξ_n^s as possible, so that the difference $w_n^{(+)} - w_n^{(-)}$ can be estimated with maximal precision. ξ_n^s are plotted in Fig. 5 as functions of λ for $n = 1, 3, 5$. For $\lambda < 0$, $\xi_s(\lambda)$ decreases with λ and n. The sign cancellation is thus moderate for $n = 1$ [$\xi_1^s(-\bar{\lambda}) \approx 0.75$] and becomes more and more important for higher-order correlators—ξ_5^s at $\lambda = -\bar{\lambda}$ is close to zero. It is interesting that ξ_n^s are almost equal for two different choices of $\mathcal{N}$ in (20).

Thus there are no indications of severe critical slowing down in the whole range of possible coupling constants $-\bar{\lambda} < \lambda < \bar{\lambda}$. The sign problem is also moderate for low-order correlators, but becomes more severe for higher-order correlators. It could be extremely interesting to extend the applicability of the random process described above up to $\lambda = \lambda_c$ while preserving these attractive features of the algorithm.

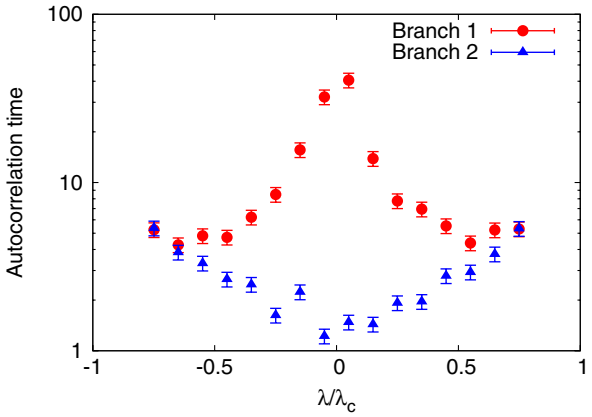

FIG. 3 (color online). Autocorrelation time of the random process described in Sec. IV B as a function of the coupling constant λ at fixed $c = 8$ and for different choices of $\mathcal{N}$ in (20).

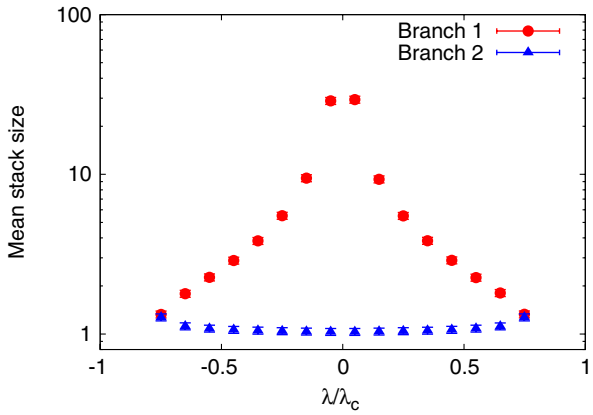

FIG. 4 (color online). Mean stack size of the random process described in Sec. IV B as a function of the coupling constant λ at fixed $c = 8$ and for different choices of $\mathcal{N}$ in (20).

C. Random planar surfaces: The Weingarten model

The Weingarten model [12,13] is a lattice field theory which in the large-N limit reproduces the sum over all closed surfaces with genus one on the hypercubic lattice. The action for each surface is proportional to its area; thus the model can be considered as a lattice regularization of bosonic strings with a Nambu-Goto action. Although this model does not have a nontrivial continuum limit for any space dimensionality [19], the structure of the functional integral and of the Schwinger-Dyson equations in this model are similar to those in large-N non-Abelian lattice gauge theory, and the analysis of this model might be helpful for the extension of the approach described here

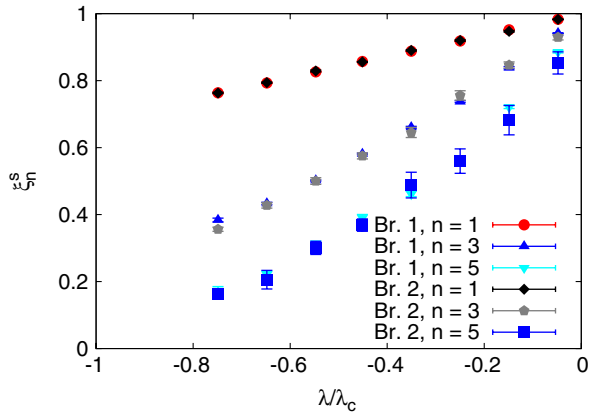

FIG. 5 (color online). The quantity ξ_n^s in (21) for $n = 1, 3, 5$, as a function of the coupling constant λ at fixed $c = 8$ and for different choices of $\mathcal{N}$ in (20). "Br. 1,2" is for "Branch 1,2."

to non-Abelian gauge theories. In order to derive the Schwinger-Dyson equations, it is convenient to consider the reduced Weingarten model [13], which in the large-N limit is equivalent to the original model, similar to the Eguchi-Kawai model for non-Abelian lattice gauge theory. It can be shown that, in contrast to reduced lattice gauge theories, for the reduced Weingarten model additional twisting is not necessary [20].

The reduced model is defined by an integral over complex $N \times N$ matrices $U_\mu \equiv U^\dagger_{-\mu}$ with $\mu = 1, \ldots, D$:

$$\mathcal{Z}(\beta) = \int \mathcal{D}U_\mu \exp\Big(-N \sum_{\mu=1}^{D} \mathrm{Tr}(U_\mu U_\mu^\dagger) + N\beta \sum_{\mu\neq\nu=1}^{D} \mathrm{Tr}(U_\mu U_\nu U_\mu^\dagger U_\nu^\dagger)\Big). \quad (22)$$

If one treats the second term in the exponent in (22) as a perturbation and expands $Z(\beta)$ in powers of β, the resulting sum over planar diagrams is equivalent to the sum over all possible closed surfaces of genus one on the lattice with weight $\beta^{|S|}$, where $|S|$ is the area of each surface.

A basic observable in this model is the sum over all planar surfaces which are bounded by some closed loop C. The loop C can be uniquely specified by a sequence $\{\mu_1, \ldots, \mu_n\}$, where the μ's take the values $\pm 1, \ldots, \pm D$. In order to reconstruct the loop C from the sequence, one should start from an arbitrary point on a hypercubical lattice and move along one link in the direction μ_1, forward if μ_1 is positive and backward if μ_1 is negative. From this new position one should similarly move in the direction μ_2, and so on. From the diagrammatic expansion one can see that such a sum over surfaces is given by the following correlator:

$$W(C) = W(\mu_1, \ldots, \mu_n) = \left\langle \frac{1}{N} \mathrm{Tr}(U_{\mu_1} \ldots U_{\mu_n}) \right\rangle, \quad (23)$$

where one takes the conjugate variable $U^\dagger_{|\mu_A|}$ if μ_A is negative. This observable is similar to the Wilson loop in lattice gauge theory, but, unlike the Wilson loop, it does not have a "zigzag symmetry" [11]: passing a link forward and immediately backward changes the value of the Wilson loop. In the large-N limit the single-loop observables factorize, which allows us to obtain a closed set of Schwinger-Dyson equations for $W(\mu_1, \ldots, \mu_n)$ [12,13]:

$$\begin{aligned} W(\mu_1, \mu_2) &= \delta(\mu_1, -\mu_2) + \beta \sum_{|\mu|\neq|\mu_1|} W(\mu, \mu_1, -\mu, \mu_2) W(\mu_1, \ldots, \mu_n) \\ &= \delta(\mu_1, -\mu_2) W(\mu_3, \ldots, \mu_n) + \delta(\mu_1, -\mu_n) W(\mu_2, \ldots, \mu_{n-1}) \\ &\quad + \sum_{A=3}^{n-1} W(\mu_2, \ldots, \mu_{A-1}) W(\mu_{A+1}, \ldots, \mu_n) \delta(\mu_1, -\mu_A) + \beta \sum_{|\mu|\neq|\mu_1|} W(\mu, \mu_1, -\mu, \mu_2, \ldots, \mu_n), \qquad n > 2. \end{aligned} \quad (24)$$

These equations should hold for any lattice link μ_k belonging to the loop C, but the resulting system of equations is redundant, and it is sufficient to consider only one link μ_1 on the loop. Equations (24) are schematically illustrated in Fig. 6, where the link μ_1 is marked by a thick line.

We see that Eqs. (24) again take a form similar to (7). Let us now define the renormalized observable $w(\mu_1, \ldots, \mu_n)$ by rescaling $W(\mu_1, \ldots, \mu_n)$ by the factors $\mathcal{N}$ and q as $W(\mu_1, \ldots, \mu_n) = \mathcal{N} q^n w(\mu_1, \ldots, \mu_n)$. One can interpret the factor q^n as the mass attached to the boundaries of random surfaces, somewhat like the bare quark mass in QCD. Equations (24) then take the following form:

$$\begin{aligned} w(\mu_1, \mu_2) &= (\mathcal{N} q^2)^{-1} \delta(\mu_1, -\mu_2) + \beta q^2 \sum_{|\mu|\neq|\mu_1|} w(\mu, \mu_1, -\mu, \mu_2) w(\mu_1, \ldots, \mu_n) \\ &= q^{-2} \delta(\mu_1, -\mu_2) w(\mu_3, \ldots, \mu_n) + q^{-2} \delta(\mu_1, -\mu_n) w(\mu_2, \ldots, \mu_{n-1}) \\ &\quad + \mathcal{N} q^{-2} \sum_{A=3}^{n-1} w(\mu_2, \ldots, \mu_{A-1}) w(\mu_{A+1}, \ldots, \mu_n) \delta(\mu_1, -\mu_A) + \beta q^2 \sum_{|\mu|\neq|\mu_1|} w(\mu, \mu_1, -\mu, \mu_2, \ldots, \mu_n), \qquad n > 2. \end{aligned} \quad (25)$$

Let us now devise a random process of the type described in Sec. III, which stochastically solves these equations. The configuration space is now a stack which contains closed loops, that is, sequences of indices $\mu = \pm 1, \ldots, \pm D$. The desired random process is defined by the following possible actions at each discrete time step:

Create a new loop.—With the probability $2D\mathcal{N}^{-1}q^{-2}$ create a new elementary loop $C = \{\mu, -\mu\}$, where $\mu = \pm 1, \ldots, \pm D$ is random (either positive or negative), and push it to the stack.

Join loops.—With the probability $2D\mathcal{N}q^{-2}$ pop the two loops $C_1 = \{\mu_1, \ldots, \mu_n\}$, $C_2 = \{\nu_1, \ldots, \nu_m\}$ from

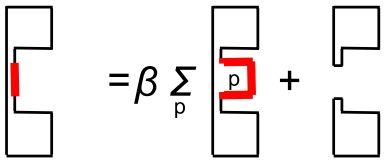

FIG. 6 (color online). Schematic illustration of the structure of the loop equations (24) in the Weingarten model (22). These equations should hold for any link (marked by a thick line) which belongs to the loop.

the stack and form a new loop C by joining the loops C_1, C_2 with a link in the random direction μ (either positive or negative): $C = \{\mu_1, \ldots, \mu_n, \mu, \nu_1, \ldots, \nu_m, -\mu\}$. This action can only be performed if there are more than two loops in the stack.

Flatten loop.—If the three links in the end of the sequence on the top of the stack form a boundary of the plaquette, that is, if the topmost loop has the form $C = \{\mu_1, \ldots, \mu_n, \mu, \nu, -\mu\}$ for some μ and ν, replace these three links by a single link in the direction ν with the probability βq^2: $C' = \{\mu_1, \ldots, \mu_n, \nu\}$.

Append to the loop.—With the probability $4Dq^{-2}$ append a pair $\{\mu, -\mu\}$, where μ is random (either positive or negative), to the topmost sequence $\{\mu_1, \ldots, \mu_n\}$ in the stack as $\{\mu, -\mu, \mu_1, \ldots, \mu_n\}$ or $\{\mu, \mu_1, \ldots, \mu_n, -\mu\}$. The probabilities of these two choices are equal.

Restart.—Otherwise, start with a stack containing an elementary random loop $C = \{\mu, -\mu\}$, where $\mu = \pm 1, \ldots, \pm D$ is chosen randomly.

Again assuming that the sum of the probabilities of all possible actions is equal to 1 and the probability of restart events is minimized, we obtain an equation relating β, $\mathcal{N}$, and q:

$$\beta q^2 + 2Dq^{-2}(\mathcal{N} + \mathcal{N}^{-1} + 2) = 1. \tag{26}$$

Maximization with respect to $\mathcal{N}$ yields the relation between q and β:

$$q = \sqrt{\frac{1 \pm \sqrt{1 - 32D\beta}}{2\beta}}. \tag{27}$$

We call the solution with the minus sign in front of the square root "branch 1" and the other solution "branch 2."

The value of β in (27) cannot exceed the critical value $\bar{\beta}(D) = 1/(32D)$. As we have already seen on the example of the matrix model, this critical value does not necessarily coincide with the true critical point $\beta_c(D)$ at which the sum over planar surfaces diverges. Indeed, $\bar{\beta}(D)$ does not exceed the lower bound $\beta_c(D) > (24(D-1))^{-1}$ obtained in [13], and is significantly lower than the critical values obtained numerically in [20]. In fact, for $\beta = \bar{\beta}_D$ all the observables are still dominated by the lowest-order perturbative contributions. Expectation values of the observables $W_{1\times 0} \equiv W(\mu, -\mu)$ (1×0 loop) and $W_{1\times 1} \equiv W(\mu, \nu, -\mu, -\nu)$ (1×1 loop), which were obtained after 10^7 iterations of the random process described above [with q given by branch 1 of (27)], are plotted in Fig. 7 as functions of the coupling constant β. The solid line corresponds to the first two terms in the perturbative expansion: $W_{1\times 0} = 1 + 2(D-1)\beta^2 + O(\beta^4)$, $W_{1\times 1} = \beta + 8(D-1)\beta^3 + O(\beta^5)$. Within statistical errors, one sees only the lowest-order perturbative contributions. It should be stressed that the proposed random process implements stochastic summation of diagrams of *all* orders, but due to the smallness of $\beta < \bar{\beta}(D)$, a very large computational time is required to see the contributions of higher-order terms.

Note that when the coupling constant β tends to zero (that is, the "bare string tension" of the random surfaces tends to infinity) and q lies between the two solutions of (27), the above random process describes just the growth of "branched polymers," whose branches are bosonic random walks and hence correspond to particles rather than "strings." These branches consist of loops in which every lattice link is passed twice and which hence sweep out zero area.

Taking the limit $\beta \to 0$ in (27), we find that the minimal value of q is $\bar{q} = \sqrt{8D}$. In order to understand this critical value, we first note that in the limit $\beta \to 0$ the observables $W(\mu_1, \ldots, \mu_n)$ are all equal to 1 if the links $\mu_1, \ldots, \mu_n$ form a loop which sweeps zero area, and zero otherwise. The probability to encounter such loops in the random process described above is hence proportional to $w(\mu_1, \ldots, \mu_n) \sim q^{-n}$. Simple examples of such loops, which can be also thought of as the random treelike graphs on the lattice, are shown in Fig. 8. Now imagine adding to some loop k links stemming from some lattice site. Since the loop includes each link twice, the

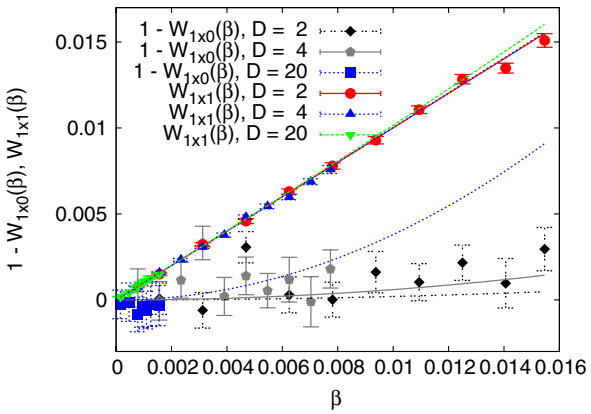

FIG. 7 (color online). The observables $W_{1\times 0} \equiv W(\mu, -\mu)$ (1×0 loop) and $W_{1\times 1} \equiv W(\mu, \nu, -\mu, -\nu)$ (1×1 loop) in the Weingarten model as functions of the coupling constant β for different dimensions D. The solid line corresponds to the first two terms in the perturbative expansion: $W(\mu, -\mu) = 1 + 2(D-1)\beta^2 + O(\beta^4)$, $W(\mu, \nu, -\mu, -\nu) = \beta + 8(D-1)\beta^3 + O(\beta^5)$. The data were obtained after 10^7 iterations of the algorithm described above.

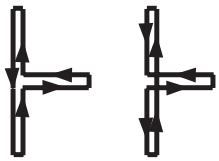

FIG. 8. Two simple configurations of branched polymers on the lattice. The configurations on the left and on the right count as different configurations.

probability of such a configuration decreases by q^{-2k}. The number of possible configurations of k links is $(2 \times 2D)^k$ since each of the k links can point along any of the D directions, both forward and backward. An additional factor of 2 appears since the zero-area loops pass through each point twice, and the new links can be inserted between the links pointing either forward or backward (for example, compare the configurations on the left and on the right of Fig. 8). Finally, one can add any number $k = 1, 2, \ldots, \infty$ of branches to any point belonging to the branched polymer. At the criticality, adding any number of random links to some configuration should not change its overall weight. Therefore, the change of the weight due to the added links times the number of ways to add them should be equal to unity. We are thus led to the following equation for $\bar{q}$:

$$\sum_{k=1}^{+\infty} (4Dq^{-2})^k = \frac{4Dq^{-2}}{1 - 4Dq^{-2}} = 1, \tag{28}$$

or $8D\bar{q}^{-2} = 1$. Thus, in the limit $\beta \to 0$ we indeed reproduce branched polymers with the correct critical behavior.

At nonzero β, deviation from trivial branched polymer configurations can be characterized by the rate of the "flatten loop" events. Indeed, since the probability of n such events is proportional to β^n, such a sequence of events corresponds to a random surface (which is, in general, open) consisting of n lattice plaquettes, plus some number of random trees. One can therefore think of the random process described above as the process of drawing random loops which sweep out random planar surfaces. The average rate of "flattening" events is plotted in Fig. 9 on the right as a function of the coupling constant β for different dimensions D and for different choices of q in (27). One can see that in the whole range of coupling constants, $0 < \beta < \bar{\beta}(D)$, the rate of flattening events is numerically very small. On the other hand, the number of links in the loops, as well as the number of loops stored in the stack, is quite large. The mean stack size and mean length of the topmost loop in the stack are plotted in Fig. 10 as a function of the coupling constant β for different dimensions D and for different choices of q in (27). One can conclude, therefore, that the branched polymers actually dominate in the properties of the random process described above. The critical behavior of these random trees is universal for any dimension D; that is why observables such as the mean stack size or the mean loop length, which are mainly

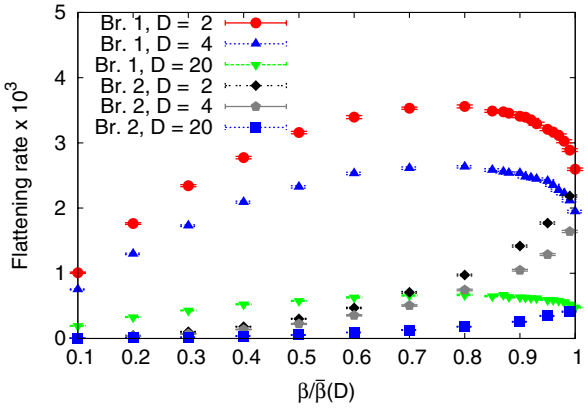

FIG. 9 (color online). Mean rate of flattening events for the random process solving the loop equations in the Weingarten model as a function of the coupling constant β at different dimensions D and for different choices of q in (27).

sensitive to the length of loops rather than to the area of random surfaces, practically do not depend on space dimensionality.

While the closed planar surfaces in the vicinity of the true critical point $\beta_c(D)$ of the Weingarten model are also dominated by branched polymers [19], in our ensemble of

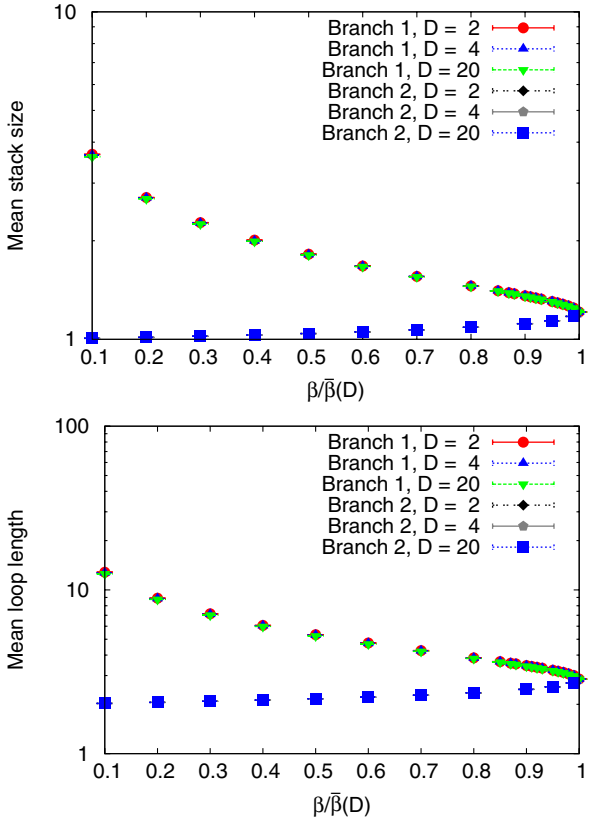

FIG. 10 (color online). Mean stack size (upper panel) and mean length of the topmost loop in the stack (lower panel) for the random process solving the loop equations in the Weingarten model as a function of the coupling constant β at different dimensions D and for different choices of q in (27).

open random surfaces this dominance can be thought of as the manifestation of the tachyonic instability of open, rather than closed, strings. The fact that the critical coupling $\bar{\beta}(D)$ in our case is smaller than the true critical point $\beta_c(D)$ can be explained by the fact that the number of open surfaces with a given area is obviously larger than the number of closed surfaces with the same area. The true critical coupling constant $\beta_c(D)$ can be quite easily found by a very simple reweighting procedure, which will be described in detail in a separate publication.

V. RESUMMATION OF DIVERGENT SERIES AND RANDOM PROCESSES WITH MEMORY

In Sec. IV we have described a stochastic method for the solution of the Schwinger-Dyson equations for theories with noncompact variables. This method works only at small coupling constants and implements stochastic summation of perturbative series. For theories with compact variables, such as nonlinear σ models or lattice gauge theories, the structure of the Schwinger-Dyson equations is such that the method described can be straightforwardly applied only in the strong-coupling regime, where one can stochastically sum all terms in the strong-coupling expansion. However, the continuum limit of such theories typically corresponds to the weak-coupling limit. An additional complication is that for physically interesting theories the observables cannot be expressed as convergent power series in the small coupling constant g, but rather contain a nonanalytic part which is typically of the form $\exp(-c/g^2)$ with some constant c [15].

In this section we point out one possible way to deal with this problem. The basic idea is to absorb nonperturbative corrections into some self-consistent redefinition of the expansion parameter [15,16]. Recently, a similar resummation method was also considered in [21]. Solving the self-consistency condition leads to the concept of a nonlinear random process with memory [8], in which all the previous history of the process is used to estimate the value of the self-consistent expansion parameter.

Let us illustrate this idea on the simplest example of an $O(N)$ sigma model in the limit of large N. The model is defined by the following path integral over unit N-component vectors $n(x)$ living on the sites of the D-dimensional hypercubic lattice:

$$\mathcal{Z} = \int_{|n(x)|=1} \mathcal{D}n(x) \exp\left(\frac{N}{\lambda}\sum_{\langle xy\rangle} n(x)\cdot n(y)\right), \tag{29}$$

where the summation goes over all neighboring lattice sites. Despite its simplicity, this model in $D = 2$ dimensions is asymptotically free and has a mass gap which depends nonperturbatively on the coupling constant λ. Schwinger-Dyson equations in this theory can be written in terms of the two-point function $\xi(x, y) = \langle n(x)\cdot n(y)\rangle$, $\xi(x) \equiv \xi(x, 0)$ as

$$\xi(x) = \frac{1}{\lambda}\sum_{\mu}(\xi(x \pm e_\mu) - \xi(x)\xi(\pm e_\mu)) + \delta(x, 0). \tag{30}$$

Clearly, these equations have a structure similar to (7), but the inequalities (8) are satisfied only for sufficiently large λ, that is, in the strong-coupling regime. Therefore, the continuum limit at $\lambda \to 0$ cannot be reached by the method described in Sec. IV.

Let us, however, rewrite Eq. (30) as

$$\xi(x) = \frac{1}{\lambda + \sum_{\mu}\xi(\pm e_\mu)}\left(\sum_{\mu}\xi(x \pm e_\mu) + \lambda\delta(x)\right), \tag{31}$$

and introduce the "hopping parameter"

$$\kappa = \frac{1}{\lambda + \sum_{\mu}\xi(\pm e_\mu)}. \tag{32}$$

Now Eq. (30) in the form (31) looks like the equation for the free massive scalar propagator on the lattice with the mass $m^2 = \kappa^{-1} - 2D$ in lattice units. Note that $\xi(\pm e_\mu)$ tends to unity as λ tends to zero; hence m^2 also tends to zero, and the continuum limit is approached.

Let us now solve Eq. (31) stochastically, assuming that $\xi(x)$ is proportional to the stationary probability distribution $w(x)$ of some random process: $\xi(x) = cw(x)$, $\sum_x w(x) = 1$. From (31) we get $c = \frac{\lambda\kappa}{1-2D\kappa}$. Equation (31) now looks like

$$w(x) = \kappa\sum_{\mu} w(x \pm e_\mu) + (1 - 2D\kappa)\delta(x). \tag{33}$$

Combining this equation with the definition (32), it is easy to show that κ obeys the following self-consistency condition:

$$\kappa = \frac{1}{2D + \lambda w(0)}. \tag{34}$$

Equation (33) has the form (7) without the nonlinear term and thus can be interpreted as the equation for the stationary probability distribution of the position of an ordinary bosonic random walk, defined by the following possible actions at each discrete time step:

Move.—With the probability $2D\kappa$ move along the random unit lattice vector $\pm e_\mu$.

Restart.—With the probability $(1 - 2D\kappa)$ start again at the origin $x = 0$.

This ensures that $w(0) > 0$ and hence κ never exceeds its critical value $\kappa_c = (2D)^{-1}$. Therefore $\xi(x)$ and $w(x)$ can be expanded in powers of κ.

Thus we have defined a new expansion parameter κ, which should obey the self-consistency equation (34), and we have obtained a well-defined convergent expansion, namely, the sum over all paths on the lattice with the weight κ^L, where L is the length of the path. Note that the quantity $\xi(\pm e_\mu)$ in fact plays a role similar to the gluon

condensate (which is expressed in terms of the mean plaquette in lattice theory) in non-Abelian gauge theory: one can absorb all the divergences into the self-consistent definition of condensates [15,16].

The final step in the construction of the nonlinear random process which solves Eq. (30) is the solution of the self-consistency equation (34). One possible solution is to use the iterations

$$\kappa_{i+1} = \frac{1}{2D + \lambda w(0;\kappa_i)}. \tag{35}$$

Here $w(0;\kappa)$ is the return probability of a bosonic random walk with the hopping parameter κ. In practice, one should simulate the bosonic random walk at fixed $\kappa = \kappa_i$ for some number T of discrete time steps, and then estimate $w(0;\kappa_i)$ as $w(0;\kappa_i) \approx t(0)/T$, where $t(0)$ is the number of discrete time steps spent at $x = 0$. From (35) one then gets κ_{i+1}, and the process is repeated until the value of κ stabilizes with sufficient numerical precision. We call this "algorithm A." One can also consider an ultimate case, for which the return probability is updated and estimated as $t(0)/t$ every time the point $x = 0$ is reached. Now t is the time from the start of the random process and $t(0)$ is the number of time steps spent at $x = 0$. This case will be called "algorithm B."

Mathematically, such random processes are not Markov processes, since the transition probabilities at each next step depend (via κ_i) on the behavior of the process at all previous time steps. Stationary probability distributions of such processes obey nonlinear equations [such as (30)] [8], and hence they are also called nonlinear random processes.

As an interesting side remark, let us discuss such a theory at finite temperature, which is described by a bosonic random walk on the cylinder. Clearly, an ordinary bosonic random walk does not feel this compactification of space, and its stationary probability distribution is just a periodic linear combination of the corresponding distribution in infinite space. Such behavior cannot lead to any nonlinear finite-temperature effects such as phase transitions. On the other hand, if the parameters of the random walk depend on the return probability, as in (34), there is a nonlinear feedback mechanism since in the compactified space the returns are more likely. Thus finite temperature indeed affects the local behavior of the random walker with memory and might lead to interesting critical phenomena.

In order to illustrate such a stochastic solution of Eq. (30), we consider the case $D = 2$. In two dimensions the model (29) is asymptotically free, and one can introduce the lattice spacing by fixing the value of mass in physical units (we set $m_{\text{phys}} = 1$): $m_{\text{phys}} a(\lambda) = m_{\text{latt}}(\lambda) = \sqrt{\kappa^{-1}(\lambda) - 2D}$. The process of convergence of the lattice spacing to its exact value is illustrated in Fig. 11 for both algorithms A and B. For algorithm A we have used $T = 5 \times 10^5$. Algorithm A converges much faster than algorithm B. The values of lattice spacing obtained using both algorithms are compared with the exact solution in Fig. 12. In agreement with asymptotic freedom, lattice spacing quickly decreases with λ. Again, algorithm A yields more precise results in the same number of time steps.

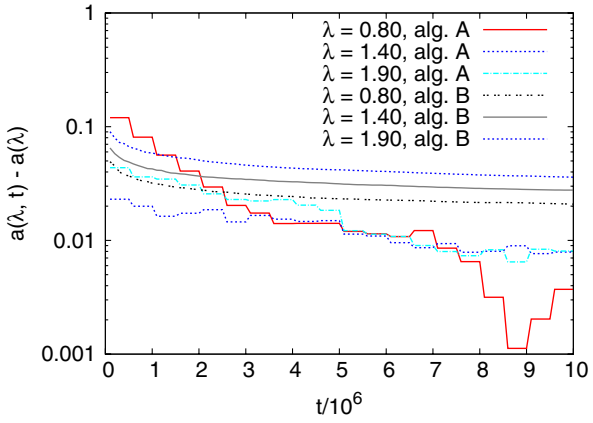

FIG. 11 (color online). The process of convergence of the random process with memory which solves the Schwinger-Dyson equations (30) for $D = 2$. The quantity plotted is the estimate of the lattice spacing $a(\lambda, t)$ after t discrete time steps, with the exact result $a(\lambda, t \to \infty)$ subtracted.

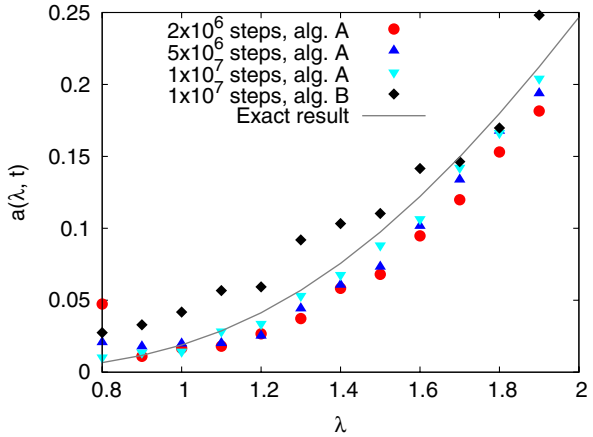

FIG. 12 (color online). Numerical estimates of the lattice spacing as a function of the coupling constant for the two-dimensional large-N $O(N)$ sigma model, compared with the exact result. The estimates were obtained using both algorithms A and B with different numbers of time steps.

VI. DISCUSSION AND CONCLUSIONS

In this paper we have presented numerical strategies for the stochastic summation and resummation of perturbative expansions in large-N quantum field theories. Our basic approach was to interpret the Schwinger-Dyson equations as the equations for the stationary probability distribution of some random process. Since Schwinger-Dyson equations in such theories are nonlinear equations, we had to use so-called nonlinear random processes, rather than

ordinary Markov processes whose stationary probability distributions always obey linear equations.

It is interesting to note that since the configuration spaces of random processes described in this paper are discrete, their numerical implementation require floating-point operations only for the random choice of actions. Thus such algorithms can be potentially much faster than the standard Monte Carlo simulations based on floating-point arithmetic, and can be advantageous for machines based on GPUs.

Our final goal is to extend the presented approach to non-Abelian lattice gauge theories. However, in this case a direct stochastic interpretation of Schwinger-Dyson equations is only possible at strong coupling, while the continuum limit of such theories corresponds to the weak-coupling limit. In Sec. V, we have discussed a way to access the weak-coupling limit, which, however, was implemented numerically only for an $O(N)$ sigma model at large N. The basic idea is to absorb the divergences into a self-consistent redefinition of the expansion parameter and solve the self-consistency conditions using random processes with memory. In some sense, an $O(N)$ sigma model can be thought of as the bosonic random walk in its own condensate, and the approach to the self-consistent value of the mass gap (see Fig. 11)—as a renormalization-group flow.

For non-Abelian gauge theories the redefined expansion parameters can emerge as the Lagrange multipliers for the ''zigzag symmetry'' of the QCD string and should also satisfy some self-consistency conditions [16]. Zigzag symmetry means that when one adds a line which is passed forward and backward to the boundary of the fluctuating string, the amplitudes should not change. In lattice gauge theory, this condition is equivalent to the unitarity of the link variables $U(x, \mu)$, which is similar to the condition $|n(x)| = 1$ in an $O(N)$ sigma model. These redefined parameters can also be related to the gluon condensate [16]. By analogy with the sigma model, one can think that non-Abelian gauge theories are similar to strings moving in some self-consistent condensates. Such a picture is also close to the idea of holographic AdS/CFT duality for non-Abelian gauge theories, where the dual string lives in some self-consistent gravitational background, and the parameters of this background can be related to gluon condensates in gauge theory [11]. In fact, the requirement that the metric of the holographic background approaches that of the AdS space-time ensures the zigzag symmetry of the strings which end on the AdS boundary [11].

In view of these qualitative considerations, our hope is that the loop equations in non-Abelian gauge theories can be solved stochastically by a random process similar to the one which was devised for the Weingarten model of random surfaces (see Sec. IV C), but with some self-consistent choice of parameters, which might be implemented as the memory in the random process.

Among other possible applications of the presented method, one can think of the solution of Schwinger-Dyson equations in continuum gauge theories, combined with the renormalization-group methods [10], numerical analysis of quantum gravity models described by various matrix models, and numerical solutions of hydrodynamical equations [22].

It should be noted here that several attempts at the stochastic solution of the loop equations in large-N gauge theories were already described in the literature quite a long time ago [23]. These algorithms were, in essence, based on the so-called branching random processes, so that the Wilson loop $W(C)$ is proportional to the probability of the transition from the initial loop configuration C to the empty configuration with no loops. In particular, in contrast to the algorithm described in Sec. IV C, where one of the basic steps is to join loops, in the algorithms described in [23] the basic step was to split a self-intersecting loop into two loops. As a result, these algorithms did not implement the importance sampling and were not able to produce any sensible results for the four-dimensional gauge theory. Generally, branching random processes similar to those considered in [23] can be obtained from the recursive nonlinear random process described in this paper by time reversal. However, since such processes do not satisfy any detailed balance condition, they are not invariant under this operation, and they lead to very different numerical algorithms.

ACKNOWLEDGMENTS

I am grateful to Dr. M. I. Polikarpov, Dr. Yu. M. Makeenko, Dr. A. S. Gorsky, Dr. N. V. Prokof'ev, and Dr. I. Ya. Aref'eva for interesting and stimulating discussions. I would also like to thank Dr. F. Bruckmann and Dr. A. Schaefer for their kind hospitality at the University of Regensburg, where part of this work was written. This work was partly supported by Grants No. RFBR 09-02-00338-a and No. RFBR 08-02-00661-a, a grant for the leading scientific schools (No. NSh-6260.2010.2), the Federal Special-Purpose Programme ''Personnel'' of the Russian Ministry of Science and Education, and by personal grants from the ''Dynasty'' Foundation and from the FAIR-Russia Research Center (FRRC).

APPENDIX: STOCHASTIC SOLUTION OF NONLINEAR EQUATIONS WITH COEFFICIENTS OF ARBITRARY SIGN

The random process described in Sec. III was devised under the assumption that the coefficients $p_c(x)$, $p_e(x|y_1)$, and $p_j(x|y_1, y_2)$ are real and positive for any x, y_1, and y_2. In this appendix we show how the solution of Eq. (7), with arbitrary signs or complex phases on the right-hand side, can be reduced to the solution of another equation of the form (7) with all positive coefficients, provided the

inequalities (8) are satisfied. We begin by discussing the case of real but nonpositive coefficients in detail, and finally sketch the extension to complex-valued coefficients.

To this end, let us represent the sign-alternating coefficients in (7) as

$$\begin{aligned} p_c(x) &= p_c^{(+)}(x) - p_c^{(-)}(x), \\ p_e(x|y) &= p_e^{(+)}(x|y) - p_e^{(-)}(x|y), \\ p_j(x|y) &= p_j^{(+)}(x|y_1, y_2) - p_j^{(-)}(x|y_1, y_2), \end{aligned} \tag{A1}$$

where $p_c^{(\pm)}(x)$, $p_e^{(\pm)}(x|y)$, and $p_j^{(\pm)}(x|y_1, y_2)$ are all positive and also obey the following inequality:

$$\sum_{s=\pm} \sum_x p_c^{(s)}(x) + p_e^{(s)}(x|y_1) + p_j^{(s)}(x|y_1, y_2) < 1 \tag{A2}$$

for any y_1, y_2. Obviously, these inequalities can be satisfied if the inequalities (8) are satisfied. Let us now introduce two functions $w^{(+)}(x)$ and $w^{(-)}(x)$, which satisfy the following equations:

$$\begin{aligned} w^{(+)}(x) = {} & p_c^{(+)}(x) + \sum_y p_e^{(+)}(x|y) w^{(+)}(y) + \sum_y p_e^{(-)}(x|y) w^{(-)}(y) + \sum_{y_1,y_2} p_j^{(+)}(x|y_1, y_2) w^{(+)}(y_1) w^{(+)}(y_2) \\ & + \sum_{y_1,y_2} p_j^{(+)}(x|y_1, y_2) w^{(-)}(y_1) w^{(-)}(y_2) + \sum_{y_1,y_2} p_j^{(-)}(x|y_1, y_2) w^{(+)}(y_1) w^{(-)}(y_2) + \sum_{y_1,y_2} p_j^{(-)}(x|y_1, y_2) w^{(-)}(y_1) w^{(+)}(y_2), \\ w^{(-)}(x) = {} & p_c^{(-)}(x) + \sum_y p_e^{(+)}(x|y) w^{(-)}(y) + \sum_y p_e^{(-)}(x|y) w^{(+)}(y) + \sum_{y_1,y_2} p_j^{(+)}(x|y_1, y_2) w^{(+)}(y_1) w^{(-)}(y_2) \\ & + \sum_{y_1,y_2} p_j^{(+)}(x|y_1, y_2) w^{(-)}(y_1) w^{(+)}(y_2) + \sum_{y_1,y_2} p_j^{(-)}(x|y_1, y_2) w^{(+)}(y_1) w^{(+)}(y_2) + \sum_{y_1,y_2} p_j^{(-)}(x|y_1, y_2) w^{(-)}(y_1) w^{(-)}(y_2). \end{aligned} \tag{A3}$$

It is now easy to check that the difference

$$w(x) = w^{(+)}(x) - w^{(-)}(x) \tag{A4}$$

satisfies Eq. (7). On the other hand, Eqs. (A3) again have the form of (7) with all positive coefficients, but with the configuration space X' being the direct product $X \otimes \mathbb{Z}_2$. In other words, each variable x now, in addition, carries the "sign" + or −, which can be written as $\{x, +\}$ or $\{x, -\}$. Based on the results presented in Sec. III, one can devise the random process which solves these equations:

Create.—With the probability $p_c^{(\pm)}(x)$ create a new element $\{x, \pm\} \in X'$ and push it to the stack.

Evolve.—With the probability $p_e^{(+)}(x|y)$ pop the element $\{y, \pm\}$ from the stack and push the element $\{x, \pm\}$ to the stack; with the probability $p_e^{(-)}(x|y)$ do the same but flip the sign of y.

Join.—With the probability $p_j^{(+)}(x|y_1, y_2)$ consecutively pop two elements $\{y_1, s_1\}, \{y_2, s_2\}$ from the stack and push a single element $\{x, s_1 s_2\}$ to the stack. That is, two pluses or two minuses associated with the y's give $\{x, +\}$, but one plus and one minus give $\{x, -\}$. With the probability $p_j^{(-)}(x|y_1, y_2)$ do the same, but flip the resulting sign; that is, push the element $\{x, -s_1 s_2\}$ to the stack.

Restart.—Otherwise, empty the stack and push a single element $\{x, \pm\} \in X'$ into it, with the probability proportional to $p_c^{(\pm)}(x)$.

As in Sec. III, $w^{(+)}(x)$ and $w^{(-)}(x)$ are proportional to the probability of finding the elements $\{x, +\}$ or $\{x, -\}$ on the top of the stack, provided there is more than one element in it.

The extension of this construction to complex-valued coefficients is quite straightforward. We represent the coefficients in (7) as $p_c(x) = \int_0^{2\pi} d\theta\, p_c(x, \theta) \exp(i\theta)$, where $p_c(x, \theta)$ is real and positive, and similarly for the other coefficients. The configuration space X' becomes the direct product $X \otimes S^1$, where S^1 is the unit circle in the complex plane. The stack now contains the pairs $\{x, \theta\}$, with $\theta \in [0, 2\pi]$ being the complex phase. The function $w(x)$ is estimated as $w(x) = \int_0^{2\pi} d\theta w(x, \theta) \exp(i\theta)$, where $w(x, \theta)$ is the probability to find the element $\{x, \theta\}$ at the top of the stack. In the random process, new elements are created with a probability distribution proportional to $p_c(x, \theta)$, and in the "evolve" and the "join" actions the phases of the elements and the coefficients in (7) are added modulo 2π, similarly to signs.

Note that this solution does not, in general, have the property of importance sampling. Indeed, one can have some x for which $w(x)$ is numerically very small, but the random process can spend an almost equally large amount of time in the states $\{x, +\}$ and $\{x, -\}$, so that numerically large $w^{(+)}(x)$ and $w^{(-)}(x)$ nearly cancel. Whether this occurs or not depends on the particular system of equations and on the particular unknown variables, but potentially this feature can make numerical simulations less efficient.

[1] N. V. Prokof'ev and B. V. Svistunov, Phys. Rev. Lett. **81**, 2514 (1998); Phys. Rev. Lett. **87**, 160601 (2001); E. Burovski, N. Prokof'ev, B. Svistunov, and M. Troyer, New J. Phys. **8**, 153 (2006); K. Van Houcke, E. Kozik, N. Prokof'ev, and B. Svistunov, in *Computer Simulation Studies in Condensed Matter Physics XXI*, edited by D. P. Landau, S. P. Lewis, and H. B. Schuttler (Springer Verlag, Heidelberg, Berlin 2008).

[2] U. Wolff, Nucl. Phys. **B824**, 254 (2010); **B832**, 520 (2010); Phys. Rev. D **79**, 105002 (2009); T. Korzec and U. Wolff, Proc. Sci., LATTICE2010 (2010) 029 [arXiv:1011.1359].

[3] L. Pollet, N. V. Prokof'ev, and B. V. Svistunov, Phys. Rev. Lett. **105**, 210601 (2010).

[4] G. t'Hooft, Nucl. Phys. **B72**, 461 (1974).

[5] J. Koplik, A. Neveu, and S. Nussinov, Nucl. Phys. **B123**, 109 (1977).

[6] E. Brezin, C. Itzykson, G. Parisi, and J. B. Zuber, Commun. Math. Phys. **59**, 35 (1978).

[7] K. Etessami and M. Yannakakis, Lect. Notes Comput. Sci. **3404**, 340 (2005).

[8] T. D. Frank, Physica A (Amsterdam) **331**, 391 (2004); T. D. Frank, Phys. Lett. A **372**, 4553 (2008).

[9] M. Kac, *Probability and Related Topics in Physical Sciences* (American Mathematical Society, Providence, Rhode Island, 1976); J. H. P. McKean, Proc. Natl. Acad. Sci. U.S.A. **56**, 1907 (1966).

[10] J. M. Pawlowski, Ann. Phys. (N.Y.) **322**, 2831 (2007); R. Alkofer and L. von Smekal, Phys. Rep. **353**, 281 (2001).

[11] A. M. Polyakov, Int. J. Mod. Phys. A **14**, 645 (1999); S. S. Gubser, I. R. Klebanov, and A. M. Polyakov, Phys. Lett. B **428**, 105 (1998).

[12] D. Weingarten, Phys. Lett. **90B**, 280 (1980).

[13] T. Eguchi and H. Kawai, Phys. Lett. **114B**, 247 (1982).

[14] Y. Makeenko and A. A. Migdal, Nucl. Phys. **B188**, 269 (1981).

[15] G. Parisi, Phys. Lett. **76B**, 65 (1978); A. I. Vainshtein and V. I. Zakharov, Phys. Rev. Lett. **73**, 1207 (1994).

[16] V. Kazakov, Nucl. Phys. **B413**, 433 (1994); V. A. Kazakov, JETP **58**, 1096 (1983); I. K. Kostov, Nucl. Phys. **B265**, 223 (1986).

[17] I. Y. Aref'eva, Phys. Lett. **104B**, 453 (1981).

[18] E. T. Akhmedov, JETP Lett. **80**, 218 (2004).

[19] B. Durhuus, J. Frölich, and T. Jonsson, Nucl. Phys. **B240**, 453 (1984).

[20] H. Kawai and Y. Okamoto, Phys. Lett. **130B**, 415 (1983).

[21] L. Pollet, N. V. Prokof'ev, and B. V. Svistunov, arXiv:1012.5808.

[22] A. A. Migdal, Int. J. Mod. Phys. A **9**, 1197 (1994).

[23] A. A. Migdal, Nucl. Phys. **B265**, 594 (1986); G. Marchesini, Nucl. Phys. **B239**, 135 (1984).