In this paper we describe a new integration method for the groups $U(N)$ and $SU(N)$, for which we verified numerically that it is polynomially exact for $N \leq 3$. The method is applied to the example of one-dimensional QCD with a chemical potential. We explore, in particular, regions of the parameter space in which the sign problem appears due to the presence of the chemical potential. While Markov chain Monte Carlo fails in this region, our new integration method still provides results for the chiral condensate on arbitrary precision, demonstrating clearly that it overcomes the sign problem. Furthermore, we demonstrate that also in other regions of parameter space our new method leads to errors which are reduced by orders of magnitude.

DOI: 10.1103/PhysRevD.94.114508

I. INTRODUCTION

The sign problem in models of statistical and high-energy physics constitutes one of the greatest challenges for computational sciences, because of the difficulty of evaluating such systems [1]. Many attempts using various techniques have been developed but no general solution to overcome the sign problem has been found so far [2]. On the other hand, the sign problem appears in important problems in physics. For example, in high-energy physics, the sign problem prevents the full understanding of the physics of the early Universe and the explaining and interpreting of heavy ion collisions. In order to progress with these questions, simulations within the framework of lattice QCD with a nonzero chemical potential would be required. However, these are impossible with present techniques; see Refs. [3,4] for recent reviews. The reason is that standard computations in lattice QCD employ Markov chain Monte Carlo (MC-MC) methods which need a positive integrand in order to be applicable. However, in the problem just mentioned a chemical potential is required, leading to a complex integrand and therefore to an oscillating function. In particular, if the sign cancellation errors due to the plural oscillations are of significantly higher magnitude than the real integral value, it becomes unfeasible to evaluate such systems.

Therefore, alternative approaches to MC-MC methods need to be developed and in [5,6] we have proposed and tested quasi–Monte Carlo and iterated numerical integration techniques. These methods can improve the convergence of the involved integrations and also have the potential to deal with the sign problem. However, in this paper we discuss yet another method of numerical integration for generic systems with a sign problem. This new method leads to an arbitrarily precise evaluation of the involved integrals and is based on a complete symmetrization of the integrals considered.

This can be achieved through new integration rules on compact groups, as developed in this article, which lead to polynomial exactness. We test the method on the example of one-dimensional QCD with a chemical potential, see, e.g., [7], for which other approaches have already been used to solve the sign problem [8]. Although one-dimensional QCD is a model with an interest in its own as the strong coupling limit of QCD [9], we consider it here only as a benchmark model for testing our approach, especially since it is possible to compute observables analytically and, thus, check the numerical results directly. In particular, we will compute the chiral condensate for a broad range of action parameters, including values of the chemical potential that are impossible (for all practical purposes) to address with standard Monte Carlo techniques.

The idea to symmetrize the involved integrals in a MC-MC simulation to achieve positivity and stable results has also been proposed in Refs. [10,11]. However, in these works only an incomplete symmetrization has been used and still a large number of Monte Carlo samples were necessary to obtain accurate results. In our approach, we perform a polynomially exact integration avoiding the MC-MC step. This way, we only need a very small number of integration points. In fact, we can reach arbitrary (up to machine) precision for the targeted physical observables and avoid the MC error completely.
For our computations, we employ the compact groups $U(N)$ and $SU(N)$ and give a description for a complete symmetrization for $N \leq 3$. As we will demonstrate, for these cases with our new approach the sign problem is completely avoided.

This paper is composed in the following way: In Sec. II, we introduce the model of one-dimensional QCD, show analytic results of the partition function $Z$, and demonstrate the difficulty of computing $Z$ for specific parameters numerically. In Sec. III, we describe the polynomially exact method based on completely symmetrized spherical quadrature rules [12]. In Sec. IV, we explain our numerical computations in more detail, show results for the partition function and the chiral condensate, and explain their behavior for different parameter values. In Sec. V, we finally conclude the paper.

II. ONE-DIMENSIONAL LATTICE QCD

Let us consider the following Dirac operator (cf., e.g., [7]) for a lattice with $n$ points

$$
\mathfrak{D}(U) = \begin{pmatrix}
\frac{m}{2} U_1 & \frac{e^\mu}{2} U_1 \\
-\frac{e^\mu}{2} U^*_1 & \frac{m}{2} U_2 \\
-\frac{e^\mu}{2} U^*_2 & m \\
\vdots & \ddots & \ddots \\
-\frac{e^\mu}{2} U^*_{n-2} & m & \frac{e^\mu}{2} U_{n-1} \\
-\frac{e^\mu}{2} U_{n-1} & m & \frac{e^\mu}{2} U_n
\end{pmatrix}
$$

(1)

where all empty entries are zero and the corresponding one-flavor partition function

$$
Z(m, \mu, G, n) = \int_{G^n} \det \mathfrak{D}(U) dh_G^n(U)
$$

where $G = U(N)$ or $G = SU(N)$, $N \in \mathbb{N}$, and $h_G$ is the corresponding (normalized) Haar measure on $G$.

In order to reduce the numerical effort in calculating $\det \mathfrak{D}$, we will first reduce the dimension using the following theorem.

**Theorem 1:** Let $U_0 := U_n$, $\tilde{m}_1 := m$,

$$
\forall j \in [2, n-1] \cap \mathbb{N}: \tilde{m}_j := m + \frac{1}{4m_{j-1}}
$$

and

$$
\tilde{m}_n := m + \frac{1}{4m_{n-1}} + \sum_{j=1}^{n-1} (-1)^{j+1} 2^{-2j} \tilde{m}_j \tilde{m}_j^{-1} \tilde{m}_j^2.
$$

Then,

$$
\det \mathfrak{D} = \det \left( \prod_{j=1}^n \tilde{m}_j + 2^{-n} e^{-\mu} \left( \prod_{j=0}^{n-1} U_j \right)^* \right) + (-1)^n 2^{-n} e^{\mu} \prod_{j=0}^{n-1} U_j.
$$

(5)

The proof of this theorem can be found in Appendix A.

In particular, in the gauge satisfying $U_j = 1$ except for $U_n = U$, Theorem 1 yields

$$
\det = \det \left( \prod_{j=1}^n \tilde{m}_j + 2^{-n} e^{-\mu} \left( \prod_{j=0}^{n-1} U_j \right)^* \right) + (-1)^n 2^{-n} e^{\mu} \prod_{j=0}^{n-1} U_j.
$$

(6)

with $c_1 = \prod_{j=1}^n \tilde{m}_j$, $c_2 = 2^{-n} e^{-\mu}$, and $c_3 = (-1)^n 2^{-n} e^{\mu}$.

Mathematically speaking, (6) is an application of “Fubini” and translation invariance of the Haar measure since $\det \mathfrak{D}$ only depends on $\prod_{j=1}^n U_j$. We will frequently assume this form of $\det \mathfrak{D}$ in analytic computations and we have implemented this form of $\det \mathfrak{D}$ in order to reduce computational overhead. Similarly, $c_1$, $c_2$, and $c_3$ are standard notations in this paper. Since $U \in U(N)$ or $U \in SU(N)$ $\det \mathfrak{D}$ is a polynomial of degree $N$.

As an observable of the model, we investigate the chiral condensate

$$
\chi(m, \mu, G, n) = \partial_m \ln Z(m, \mu, G, n)
$$

and

$$
\frac{\partial_m Z(m, \mu, G, n)}{Z(m, \mu, G, n)} = \frac{\int_G \partial_m \det \mathfrak{D} dh_G}{\int_G \det \mathfrak{D} dh_G}.
$$

(7)

Since $\det \mathfrak{D}$ is a polynomial of degree $N$ and the derivative $\partial_m$ only acts on the term $\prod_{j=1}^n \tilde{m}_j$ in Theorem 1, $\partial_m \det \mathfrak{D}$ is still a polynomial of degree $N$ and $\partial_m \prod_{j=1}^n \tilde{m}_j$ can be computed using symbolic differentiation.\footnote{Since all our groups are compact, they are unimodular and the Haar measures satisfy $h_{G \times H} = h_G \times h_H$ and $h_{G \ast H} = h_G \times h_H$ (cf., e.g., exercise 2.1.7 in [13]).}
Theorem 2 not only allows us to reduce numerical overhead but we can furthermore calculate the partition function (2) (and therefore also the chiral condensate) analytically.

\textbf{Theorem 2:} Let \( c_1 := \prod_{j=1}^{n} \tilde{m}_j \), \( c_2 = 2^{-n} e^{-\mu \mu} \), and \( c_3 = (-1)^{n} 2^{-n} e^{\mu \mu} \) with \( \tilde{m}_j \) as in Theorem 1. Then,

\[
Z(m, \mu, U(1), n) = \int_{U(1)} \det \mathcal{D}(U) dh_{U(1)}(U) = c_1, \quad (8)
\]

\[
Z(m, \mu, U(2), n) = \int_{U(2)} \det \mathcal{D}(U) dh_{U(2)}(U) = c_1^2 - c_2 c_3, \quad (9)
\]

\[
Z(m, \mu, SU(2), n) = \int_{SU(2)} \det \mathcal{D}(U) dh_{SU(2)}(U) = c_1^2 + c_2^2 - c_2 c_3 + c_3^2, \quad (10)
\]

\[
Z(m, \mu, U(3), n) = \int_{U(3)} \det \mathcal{D}(U) dh_{U(3)}(U) = c_1^3 - 2 c_1 c_2 c_3, \quad (11)
\]

and

\[
Z(m, \mu, SU(3), n) = \int_{SU(3)} \det \mathcal{D}(U) dh_{SU(3)}(U) = c_1^3 - 2 c_1 c_2 c_3 + c_2^2 + c_3^2. \quad (12)
\]

For the proof of this theorem, see Appendix B.

In addition, we can deduce the behavior of \( Z \) for \( m \leq 0 \).

\textbf{Corollary 1:} Let \( \tilde{m}_1 := m, \ \tilde{m}_j := m + \frac{1}{4 m_{j-1}}, \ \tilde{m}_n := m + \frac{1}{4 m_{n-1}} + \sum_{j=1}^{n-1} (-1)^{j+1} \frac{1}{m \Pi_{m=1}^{j} m_{m+1}}, \) and \( c_1 := \prod_{j=1}^{n} \tilde{m}_j \). Then,

\[
\lim_{m \to 0} c_1 = \begin{cases} 2^{1-n}, & n \in 2\mathbb{N} \\ 0, & n \in 2\mathbb{N} - 1 \end{cases}. \quad (13)
\]

In particular,

\[
\lim_{m \to 0} Z(m, \mu, U(1), n) = \begin{cases} 2^{1-n}, & n \in 2\mathbb{N} \\ 0, & n \in 2\mathbb{N} - 1 \end{cases}, \quad (14)
\]

\[
\lim_{m \to 0} Z(m, \mu, U(2), n) = \begin{cases} 3 \cdot 2^{-2n}, & n \in 2\mathbb{N} \\ -2^{-2n}, & n \in 2\mathbb{N} - 1 \end{cases}, \quad (15)
\]

\[
\lim_{m \to 0} Z(m, \mu, SU(2), n) = \begin{cases} 3 \cdot 2^{-2n} + 2^{1-2n} \cosh(2 \mu \mu), & n \in 2\mathbb{N} \\ 2^{1-2n} \sinh(2 \mu \mu) - 2^{-2n}, & n \in 2\mathbb{N} - 1 \end{cases}, \quad (16)
\]

\[
\lim_{m \to 0} Z(m, \mu, U(3), n) = \begin{cases} 4 \cdot 2^{-3n}, & n \in 2\mathbb{N} \\ 0, & n \in 2\mathbb{N} - 1 \end{cases}, \quad (17)
\]

\[
\lim_{m \to 0} Z(m, \mu, SU(3), n) = \begin{cases} 4 \cdot 2^{-3n} + 2^{1-3n} \cosh(3 \mu \mu), & n \in 2\mathbb{N} \\ 2^{1-3n} \sinh(3 \mu \mu), & n \in 2\mathbb{N} - 1 \end{cases}. \quad (18)
\]

For the proof of this corollary, see Appendix C.

If \( n \mu \) is large and \( m \) small, we can see clearly why the integrals in Theorem 2 are difficult to treat numerically, especially the \( U(N) \) cases. If we assume a stochastic approach, e.g., a Monte Carlo method, then each evaluation of \( \det \mathcal{D} \) in the form (6) is a value in the vicinity of \( |c_2|^N + |c_3|^N \approx |c_3|^N = 2^{-N} e^{N \mu \mu} \). However, performing the integration (or taking the limit of infinitely many samples), there is a very high degree of cancellations to be observed. Since discrete Markov chain Monte Carlo methods perform poorly with respect to such cancellations, they have to overcome an initial error in the vicinity of \( e^{N \mu \mu} \). In other words, as \( n \mu \) grows larger, we need very good algorithms to suppress the initial error and the convergence of Monte Carlo methods is simply not viable anymore. For example, in Fig. 1, we compare a Monte Carlo method (using reweighting) to our new, polynomially exact method proposed in Sec. III (details of the numerical tests can be found in Sec. IV). The error bars, the known rate of convergence \( \frac{1}{\sqrt{\text{sample size}}} \), and the relative error of order 1

\[
\frac{2 |c_2|^N + |c_3|^N}{2n \mu \mu} = |c_3|^N, \quad \text{due to the fact that} \quad e^x > e^{-x} \quad \text{for} \quad x \in \mathbb{R}_\geq 0 \quad \text{and the (anti)} \quad \text{symmetric shape of} \quad e^x \pm e^{-x}.
\]
seen here show that the Monte Carlo method cannot reach the same level of precision with a reasonable number of samples (note the different scales for the Monte Carlo and polynomially exact results).

### III. EFFICIENT QUADRATURE RULES OVER THE COMPACT GROUPS

Consider $Z(m, \mu, U(1), n)$ for the moment. As we have mentioned before, the problem is that the integral $\int_{U(1)} (-1)^n 2^{2n} e^{\mu U} dh_{U(1)}(U)$ in (8) vanishes but the modulus of each evaluation $|(-1)^n 2^{2n} e^{\mu U}|$ is large. However, if we were also to evaluate at $-U$ (or, more generally, at $t$ equally spaced points along the unit circle), the two terms would cancel. However, the (geometric) idea of taking opposite points or equally spaced points on circles holds, where $t$ suffices to consider $U$-designs for $N > 1$ and, especially, why we base the $U(N)$ and $SU(N)$ quadrature on the quadrature rules for the spheres $S^3$.

Since

$$U(N) \cong SU(N) \ltimes U(1),$$

(20)

holds, where $\ltimes$ denotes the (inner) semidirect product, we may construct a (weighted) $t$-design $Q_{U(N)}$ over $U(N)$ by considering two different (weighted) $t$-design rules $Q_{SU(N)}$ and $Q_{U(1)}$ over $SU(N)$ and $U(1)$ correspondingly, and then define the product rule $Q_{U(N)} = Q_{SU(N)} \times Q_{U(1)}$. It is clear that by defining $Q_{SU(N)}$ as a product rule in this way, we obtain a (weighted) $t$-design over $U(N)$. Since $t$-designs over $U(1)$ are easy to construct [see (19)], the entire problem of constructing (weighted) $t$-designs for the compact groups considered here reduces to the one of constructing (weighted) $t$-designs over $SU(N)$.

Starting with $SU(2)$, we have a measure preserving diffeomorphism $SU(2) \cong S^3$. An explicit mapping can be given by

$$\Phi: \mathbb{C}^2 \to \mathbb{C}^2; \quad (\alpha, \beta) \mapsto \left( \begin{array}{c} \alpha - \beta^* \\ \beta \alpha^* \end{array} \right)$$

(21)

whose restriction $\Phi_{SU(2)}: S^3 \to SU(2)$ is the mentioned measure preserving diffeomorphism. Thus, for this case we can resort to already well-known (weighted) $t$-designs over the 3-sphere (see [12,15]) for obtaining (weighted) $t$-designs over $SU(2)$ through the mapping $\Phi$.

Moving on to $SU(3)$, we note that there is a correspondence between $SU(3)$ and $S^6$. More specifically, we consider first the covering $\Phi_1: [0, 2\pi)^3 \times [0, \frac{\pi}{2})^2 \to S^5$ defined by

$$\begin{align*}
x_1 &= \cos(\alpha_1) \sin(\varphi_1) \\
x_2 &= \sin(\alpha_1) \sin(\varphi_1) \\
x_3 &= \sin(\alpha_2) \cos(\varphi_1) \sin(\varphi_2) \\
x_4 &= \cos(\alpha_2) \cos(\varphi_1) \sin(\varphi_2) \\
x_5 &= \sin(\alpha_3) \cos(\varphi_1) \cos(\varphi_2) \\
x_6 &= \cos(\alpha_3) \cos(\varphi_1) \cos(\varphi_2)
\end{align*}$$

and note that the restriction $\Phi_1: [0, 2\pi)^3 \times (0, \frac{\pi}{2})^2 \to S^5_1$, $S^5_1 := \Phi_1([0, 2\pi)^3 \times (0, \frac{\pi}{2})^2]$, is a diffeomorphism. Furthermore, the set $S^5_3 \setminus S^5_1$ is a null set. On the other hand, we have the mapping $\Phi_2: ([0, 2\pi)^3 \times [0, \frac{\pi}{2})^2) \times SU(2) \to SU(3)$ defined by

$$\begin{pmatrix} 0 \\ e^{i\alpha_1} \cos(\varphi_1) \\ -e^{i\alpha_2} \sin(\varphi_1) \sin(\varphi_2) \\ -e^{i\alpha_3} \sin(\varphi_1) \cos(\varphi_2) \\ e^{-i\alpha_1} \cos(\varphi_1) \sin(\varphi_2) \\ e^{-i\alpha_2} \cos(\varphi_1) \cos(\varphi_2) \end{pmatrix} \begin{pmatrix} U & 0 \\ 0 & 1 \end{pmatrix},$$

(22)

is a bijection and the set $SU(3)_0 := SU(3) \setminus SU(3)_1$ is a Haar null set. Thus, starting with a (weighted) $t$-design rule $Q_{S^5_1}$ over $S^3$ and a (weighted) $t$-design $Q_{S^5_3}$ over $S^5$, such that each point of $Q_{S^5_3}$ lies in $S^5_1$, and considering the mapping

$$\Phi_3: S^5_3 \times S^3 \to SU(3); \quad (x, y) \mapsto \Phi_2(\Phi^{-1}_1(x), \Phi(y)),$$

(24)

More precisely, $SU(N)$ is a principal $SU(N-1)$ bundle over $S^{2N-1}$; cf., e.g., Eq. (22.18) of [16].
we obtain a quadrature rule $Q_{SU(3)}$ over $SU(3)$ by setting $Q_{SU(3)} = \Phi_3[Q_{SU(3)} \times Q_{SU(3)}]$. In fact, by considering (randomized) fully symmetric interpolatory rules $Q^{1,3}$ and $Q^{1,5}$ from [12] as weighted $t$-designs $Q_{SU(3)}$ and $Q_{SU(3)}$, we checked numerically that the resulting quadrature rule $Q_{SU(3)}$ is again a weighted $t$-design over $SU(3)$, for $t \leq 3$. The latter observation drove us to investigate a procedure in more detail for constructing weighted $t$-design rules over $SU(N)$, for arbitrary positive integers $N$ and $t$. This procedure is based on a generalization of the mapping $\Phi_3$ as stated above and relies on the correspondence between $SU(N)$ and $\chi_{j=1}^{N-1}S^{2j+1}$. This new construction of quadrature rules over $SU(N)$ is subject of current research by the authors, but the potential applications of this new method exceed the scope of this article and will be not reported at this point.

Note that the procedure above is not the only way of writing integrals over $SU(N)$ as an integral over spheres. For instance, using the eigenvalue decomposition of $SU(N)$ we obtain integrals over products of $U(1)$. However, it should be noted that the eigenvalue decomposition is not an isomorphism and, thus, a functional determinant will appear (which in itself is a polynomial). In other words, the polynomial degree of the integrand will increase. Nonetheless this is a viable approach to obtain polynomial exactness, and using the points in Eq. (19) yields a method equivalent to [17].

**IV. NUMERICAL RESULTS**

In this section we will provide a comparison of the evaluation of the partition function $Z$ and the chiral condensate $\chi$ using MC-MC and our new polynomially exact integration rules. First we will concentrate on the partition function $Z$. We will have a short look at the behavior of the analytic values of $Z$ before comparing them to the quadrature results of $Z$ using the Monte Carlo and polynomially exact method in terms of a relative error. To present the real power of the polynomially exact method, we will show computational results for two different floating point number precisions. Then we will investigate the relative error behavior of the chiral condensate. Since we compute the relative error as the deviation of the quadrature result from the computation using analytic formulas, we explicitly differentiate these ways of computation in the following using the terms $Z_{\text{quadrature}}$ and $Z_{\text{analytic}}$.

As stated above, for the here-considered model both $Z$ and $\chi$ can be computed analytically for the groups $U(N)$ and $SU(N)$. In particular, the expression of the partition functions in Theorem 2 for $SU(N)$ can be related to the one for $U(N)$ through

$$Z_{\text{analytic}}(m, \mu, SU(N), n) = Z_{\text{analytic}}(m, \mu, U(N), n) + c_2^N + c_3^N$$

$$= Z_{\text{analytic}}(m, \mu, U(N), n)$$

$$+ \left\{ \begin{array}{ll}
2^{1-Nn} \cosh(Nn\mu), & n \in 2\mathbb{N} \\
-2^{1-Nn} \sinh(Nn\mu), & n \in 2\mathbb{N} - 1
\end{array} \right. \quad (25)$$

We note that for $U(N)$ the partition function smoothly approaches a much smaller value than $c_2^N + c_3^N$ when decreasing the mass parameter $m$, while for $SU(N)$ it approaches a constant near $c_2^N + c_3^N$ as given in Theorem 2; see also Corollary 1. The behavior of $Z_{\text{analytic}}(m, \mu, G, n)$ as a function of the mass parameter $m$ for $G \in \{U(3), SU(3)\}, n = 6$, $\mu = 1$ is shown in Fig. 2 and there we can clearly see the different behaviors of $Z_{\text{analytic}}$ for $U(3)$ and $SU(3)$ for $m > 0$.

For the groups $U(N)$ and $SU(N)$ each point evaluation of the quadrature rule is of order $O(2^{-Nn}e^{Nn\mu})$; that is, a double precision computation cannot resolve values below $10^{-16}2^{-Nn}e^{Nn\mu}$. Since the behavior of the partition function in comparison to the constant $c_2^N + c_3^N$ will be important in order to understand the relative error $|Z_{\text{quadrature}} - Z_{\text{analytic}}|/Z_{\text{analytic}}$, we also show the value of $|c_2^N + c_3^N| \approx 2^{-Nn}e^{Nn\mu}$ in Fig. 2 (see discussion at the end of Sec. II above, as well) for the examples of $U(3)$ and $SU(3)$. In Fig. 2, we furthermore distinguished three regions with different behavior, indicated in the following as region I, II, and III.
Let us first discuss the group $U(3)$. For large values of $m$ (region III) $2^{-3n}e^{3np}$ is negligible compared to $Z_{\text{analytic}}$. We therefore expect a small deviation of $Z_{\text{quadrature}}$ from $Z_{\text{analytic}}$ and hence a small relative error. On the other hand, for small values of $m$ (region I) $Z_{\text{analytic}}$ becomes much smaller than $2^{-3n}e^{3np}$ and we expect a significant relative error due to rounding errors. There is also a transition regime in $m$ (region II) in which the values of $Z_{\text{analytic}}$ and $2^{-3n}e^{3np}$ have the same order of magnitude. Hence, we expect a significant increase in the relative error while decreasing $m$, but the smooth behavior of $Z_{\text{analytic}}$ for $U(3)$ suggests that there will be a similarly smooth increase of the relative error as a function of $m$. As we will discuss below, this expectation is indeed verified in our numerical tests.

In the $SU(3)$ case, we have the additional constant $c_3^2 + c_3^3$ which, for small $m$, is significantly larger than $Z_{\text{analytic}}(U(3))$, see (25). Looking at Fig. 2, we expect a relative error similar to the $U(3)$ case in region III. In region I, though, the relative error should be much less now due to the fact that the analytic value and order of magnitude of each point evaluation are closer together than in the $U(3)$ case. In the transition region II, the behavior may be different from $U(3)$ as well, although this is not deduced from the figure per se but from the differences in the formulas of $Z_{\text{analytic}}$ (25). There, the $m$-dependent term of $Z_{\text{analytic}}$, the constant $c_3^2 + c_3^3$, and the point evaluation in the quadrature rules are of the same order of magnitude $O(2^{-3n}e^{3np})$. Thus, this additional term $c_3^2 + c_3^3$, not present at $U(N)$, could lead to competing effects for the relative error and, hence, an irregular behavior of the relative error (at least in the MC-MC case).

Let us now move on to our numerical experiments. In Fig. 3, we compare the quadrature rule

$$Z_{\text{quadrature}}^{\text{MCMC}}(m, \mu, G, n) = \int_G \det \mathcal{D} d h_G \approx \frac{1}{\#Q_G} \sum_{k=1}^{\#Q_G} \det \mathcal{D}(U_k)$$

(26)

where each $U_k$ is chosen randomly in $G$ (uniformly with respect to the Haar measure) and the polynomially exact version

$$Z_{\text{quadrature}}^{\text{poly. exact}}(m, \mu, G, n) = \int_G \det \mathcal{D} d h_G \approx \frac{1}{\#Q_G} \sum_{V \in Q_G} \det \mathcal{D}(VU_1)$$

(27)

where $U_1$ is the $U_1$ sampled in the nonexact version in (26).\(^5\) Here, we chose

\(^5\)Any $U_1 \in G$ would be perfectly fine; in fact, choosing the identity for $U_1$ would be a good canonical choice. However, we chose $U_1$ randomly (uniformly with respect to the Haar measure) in order to approximate the error.
we observe large relative errors in region I of Fig. 3. Note that the polynomially exact integration (error bars) have been computed from 50 independent computations for \( U(1), U(2), \) and \( SU(2) \), and from 10 independent computations for \( SU(3) \) and \( U(3) \). Again \( n = 20, \mu = 1 \), and \( m \in [0.001, 100] \) are used.

For very small values of \( m \) (region I) \( Z_{\text{analytic}}(m, \mu, U(N), n) \) is significantly smaller than \( 2^{-Nn \epsilon^{Nn \mu}} \); hence, \( Z_{\text{analytic}}(m, \mu, U(N), n) \) is negligible compared to the machine error and we observe large relative errors in region I of Fig. 3. Note that the polynomially exact computation still sums values of magnitude \( 2^{-Nn \epsilon^{Nn \mu}} \); i.e., the relative error of the exact method cannot be below \( 10^{-16} \) times the error of the nonexact method which is, indeed, what we see in Fig. 3. Returning to Fig. 2 and the \( U(N) \) discussion above, the observed smooth increase of the relative error in region II matches our expectations.

In the \( SU(N) \) case, the relative error is comparable to the \( U(N) \) case in regions I and III; we simply obtain smaller errors in region I since \( 2^{-Nn \epsilon^{Nn \mu}} \) does not dominate \( Z_{\text{analytic}} \) as is the case for \( U(N) \). However, in the transition region II of Fig. 3, we can see a rather irregular behavior, the possibility for which to occur we already mentioned in the discussion of \( Z_{\text{analytic}} \) above. This can be attributed to the fact that the mass-dependent term of \( Z_{\text{analytic}}(m, \mu, U(N), n) \) and the constant \( c_2^2 + c_3^2 \), see (25), as well as the point evaluations in \( Z_{\text{quadrature}} \), are of the same order of magnitude. Hence neither term can suppress the error of the other, which we interpret as the cause of the peak in the relative error.

Figure 4 shows the same comparison as Fig. 3 but computations were performed with 1024-bit floating point arithmetic,\(^6\) i.e., approximately 307-digit precision. Again, we observe that the polynomially exact method operates on machine precision (as to be expected). The averages and standard deviations of the relative error were computed from 50 independent computations for \( G \in \{ U(1), SU(2), U(2) \} \) and from 10 independent computations for \( G \in \{ SU(3), U(3) \} \). All computations were performed on an IBM laptop in less than an hour.\(^7\) The behavior of the relative error, for both, Monte Carlo and polynomially exact method, is very similar to the double precision case in Fig. 3. Note that the polynomially exact integration always leads to machine precision results even in this extreme case of 1024-bit precision, whereas the relative error of the MC-MC results does not notably decrease in regions I and II when replacing double precision floats in Fig. 3 with 1024-bit extended floats in Fig. 4.

\(^6\) These are 1024 mantissa bits; double precision (about 15-digit precision) corresponds to 53 bits.

\(^7\) For \( SU(3) \) and \( U(3) \) only, runtime was considerably longer than a few minutes.
In general, we observe in Figs. 3 and 4 that the polynomially exact quadrature rule always provides machine error results.

In order to test our new polynomially exact method against an actual physical observable, Fig. 5 shows the comparison of the relative error of the chiral condensate (again using 1024-bit extended floats). The analytic values of the chiral condensate have been obtained through symbolic differentiation of the formulas in Theorem 2; the numerical values by symbolic differentiation of (6). We observe that the relative error follows the trend we have already seen for the partition function in the three different regions.\(^8\)

Let us discuss the relative error in Fig. 5 in a bit more detail. A first observation is that the polynomially exact method operates on the level of machine precision and, as such, reduces the relative error by (many) orders of magnitude for all values of \(m\). Even more interesting and striking is the size of the relative error of the chiral condensate in the small-\(m\) region. As pointed out in [7], in this region of parameter space there is a severe sign problem. Indeed, for the MC-MC method the relative error becomes \(O(1)\) for sufficiently small \(m\); i.e., no statistically significant result for the chiral condensate can be obtained with standard MC-MC calculations. [In Fig. 5, this behavior can only be seen for \(U(N)\) but it is also present and was observed by us for \(SU(N)\) for \(m\)-values smaller than the ones shown here.\(^9\)] This is a clear manifestation of the infamous sign problem.

In contrast, our polynomially exact method again provides results on machine precision. Thus, the polynomially exact method completely overcomes the sign problem and can lead to very accurate results even in regions where MC-MC computations are unfeasible.

V. CONCLUSION

In this work, we have developed and tested a new integration method for the groups \(U(N)\) and \(SU(N)\). As a major outcome of our work, we could in fact provide a numerical verification that the method developed here leads to polynomial exactness of the integration for \(N \leq 3\). We have applied the method to the one-dimensional QCD with a chemical potential where for certain values of the action parameters a sign problem appears with MC-MC methods. Using the groups \(U(1), U(2), U(3)\), and \(SU(2), SU(3)\) we have demonstrated that even for cases when the sign problem is most severe, the chiral condensate of this model can be computed to arbitrary precision with the new method. In contrast, standard MC-MC methods show large \(O(1)\) relative errors and do not give any statistically significant result. For this comparison, we even went to 1024-bit extended precision and were able to show that our new method still achieves results on the level of machine precision. We therefore conclude that our polynomially exact integration method can completely avoid the sign problem. Furthermore, it is important to point out that it also leads to errors reduced by orders of magnitude compared to MC-MC even in regions of parameter space where no sign problem occurs.

The fact that our new integration method overcomes the sign problem and leads in general to errors reduced by orders of magnitude in the one-dimensional QCD considered here is certainly a very promising finding and stands as a result by itself. However, this benchmark model can only be regarded as a toy example. It will be necessary to demonstrate that the method can also be applied in higher dimensions. To this end, we are presently considering the Schwinger model as an example of a quantum field theory in two dimensions.

In addition, so far we do not have proof yet of the polynomial exactness for the groups \(U(N)\) and \(SU(N)\) with general \(N\). Although we are very confident that our integration method leads to polynomial exactness for general \(N\), we are working on a proof to substantiate this statement.

ACKNOWLEDGMENTS

The authors wish to express their gratitude to Prof. Andreas Griewank for inspiring comments and conversations, which helped to develop the work in this article. H. L. and J. V. acknowledge financial support by the Projects No. JA 674/6-1 and No. GR 705/13, funded by the Deutsche Forschungsgemeinschaft.

APPENDIX A: PROOF OF THEOREM 1

Let

\[
Y = \begin{pmatrix} A & B \\ C & D \end{pmatrix}
\]

be a block decomposition where \(A\) and \(D\) are square matrices and \(A\) is invertible. Then,

\[
\det Y = \det A \det (D - CA^{-1}B).
\]

Here, we are considering matrices of the form...
\[
X = \begin{pmatrix}
  m_1 & \varepsilon U_1 & & \\
  -\varepsilon^2 U_1^* & m_2 & \varepsilon U_2 & \\
  & -\varepsilon^2 U_2^* & m_3 & \varepsilon U_3 & \\
  & & \ddots & \ddots & \ddots \\
  & & & -\varepsilon^2 U_{n-2}^* & m_{n-1} & \varepsilon U_{n-1} \\
  & & & & -\varepsilon^2 U_{n-1}^* & m_n
\end{pmatrix}
\]

where all \( m_i \) are positive. Choosing \( A \) to be the \( m_1 \) block in \( X \), we obtain

\[
D = \begin{pmatrix}
  m_2 & \varepsilon U_2 & & \\
  -\varepsilon^2 U_2^* & m_3 & \varepsilon U_3 & \\
  & -\varepsilon^2 U_3^* & m_4 & \varepsilon U_4 & \\
  & & \ddots & \ddots & \ddots \\
  & & & -\varepsilon^2 U_{n-2}^* & m_{n-1} & \varepsilon U_{n-1} \\
  & & & & -\varepsilon^2 U_{n-1}^* & m_n
\end{pmatrix}
\]

and

\[
-CA^{-1}B = -\frac{1}{m_1} \begin{pmatrix}
  -\frac{1}{4} & 0 & -\varepsilon^2 U_1^* U_n^* \\
  0 & 0 & 0 \\
  -\varepsilon^2 U_n U_1 & 0 & -\frac{1}{4}
\end{pmatrix}
\]

In other words, \( D - CA^{-1}B \) is of the initial form again and

\[
\det X = m_1^n \det(D - CA^{-1}B)
\]

Let \( U_0 := U_n, \tilde{m}_1 := m_1 \),

\[
\forall j \in [2, n-1] \cap \mathbb{N}: \tilde{m}_j := m_j + \frac{1}{4\tilde{m}_{j-1}}
\]

and

\[
\tilde{m}_n := m_n + \frac{1}{4\tilde{m}_{n-1}} + \sum_{j=1}^{n-1} \frac{(-1)^j j! 2^{-2j}}{\tilde{m}_j \prod_{k=1}^{j-1} \tilde{m}_k^2}
\]
The following identities are well known (cf., e.g., [18]):

\[
\det X = \prod_{j=1}^{n-3} \tilde{m}_j^n \det \begin{pmatrix}
\tilde{m}_{n-2} & \frac{\epsilon^{\nu}}{2} U_{n-2} \\
\frac{(1)_{n-2} e^{-\nu}}{\prod_{j=1}^{n-3} \tilde{m}_j} \prod_{j=1}^{n-2} U_{j-1} & m_{n-1} & \frac{\epsilon^{\nu}}{2} U_{n-1} \\
\frac{\epsilon^{\nu}}{2} U_{n-1} & m_n & \frac{\epsilon^{\nu}}{2} U_{n-1} \\
\end{pmatrix}
\]

(A10)

\[
= \prod_{j=1}^{n-1} \tilde{m}_j^n \det \left( \tilde{m}_n + \frac{(-1)^n 2^n e^{\nu \mu}}{\prod_{j=1}^{n-1} \tilde{m}_j} \prod_{j=1}^n U_{j-1} + \frac{2^n e^{-\nu \mu}}{\prod_{j=1}^{n-1} \tilde{m}_j} \left( \prod_{j=1}^n U_{j-1} \right)^n \right)
\]

(A11)

which finally yields

\[
\det X = \det \left( \prod_{j=1}^n \tilde{m}_j + \frac{(-1)^n 2^n e^{\nu \mu}}{\prod_{j=1}^n \tilde{m}_j} \prod_{j=1}^n U_{j-1} + \frac{2^n e^{-\nu \mu}}{\prod_{j=1}^n \tilde{m}_j} \left( \prod_{j=1}^n U_{j-1} \right)^n \right).
\]

(A12)

**APPENDIX B: PROOF OF THEOREM 2**

Note that the $U(1)$ case is trivial. Hence, we will start considering $U(N)$ with $N \geq 2$ and use the notations

\[
U_{ij}^* := (U^*)_{ij}
\]

and

\[
\forall p \in \mathbb{N}_0 \quad \forall I, \quad J \in \mathbb{N}_0^p : U_{IJ} := \prod_{k=0}^{p-1} U_{I_k J_k} \wedge U_{IJ}^* := \prod_{k=0}^{p-1} (U^*)_{I_k J_k}.
\]

(B2)

Furthermore, we set $\forall p, q \in \mathbb{N}_0 \quad \forall I, J \in \mathbb{N}_0^p \quad \forall K, L \in \mathbb{N}_0^q$,

\[
\langle I, J | K, L \rangle := \int_{U(N)} U_{IJ} U_{KL} \, dh_{U(N)}(U)
\]

(B3)

and use abbreviations for empty sets or singletons similar to

\[
\langle 0, 1 \rangle := (0), (1)|(, )
\]

(B4)

The following identities are well known (cf., e.g., [18]):

(i) $p \neq q \Rightarrow \langle I, J | K, L \rangle = 0$

(ii) $\langle | \rangle = 1$

(iii) $\langle i, j | k, l \rangle = \frac{\delta_{ij} \delta_{kl}}{N}$.

For $N = 2$, we may expand the determinant in

\[
\int_{U(2)} \det \mathfrak{D} \, dh_{U(2)} = \int_{U(2)} \det (c_1 + c_2 U^* + c_3 U) \, dh_{U(2)}(U)
\]

(B5)

\[
= \int_{U(2)} \det \begin{pmatrix}
c_1 + c_2 U_{00}^* + c_3 U_{00} & c_2 U_{01}^* + c_3 U_{01} \\
c_2 U_{10}^* + c_3 U_{10} & c_1 + c_2 U_{11}^* + c_3 U_{11}
\end{pmatrix} \, dh_{U(2)}(U)
\]

(B6)

directly and, using the identities above, we obtain

\[
\int_{U(2)} \det \mathfrak{D} \, dh_{U(2)} = c_1^2 - c_2 c_3.
\]

(B7)

Similarly, we can expand the determinant in
OVERCOMING THE SIGN PROBLEM IN ONE-...

\[ \int_{U(3)} \det \left( \begin{array}{ccc}
  c_1 + c_2 U_{00} + c_3 U_{00} & c_2 U_{01} + c_3 U_{01} & c_2 U_{02} + c_3 U_{02} \\
  c_2 U_{10} + c_3 U_{10} & c_1 + c_2 U_{11} + c_3 U_{11} & c_2 U_{12} + c_3 U_{12} \\
  c_2 U_{20} + c_3 U_{20} & c_2 U_{21} + c_3 U_{21} & c_1 + c_2 U_{22} + c_3 U_{22}
\end{array} \right) dh_{U(3)}(U) \]  

(B8)

using Sarrus’s rule, which yields

\[ \int_{U(3)} \det \mathfrak{D} dh_{U(3)} = c_1^3 - 2c_1c_2c_3 \]  

(B9)

using the identities above.

For SU(N), we have

i) \( p \neq q \Rightarrow \langle I, J | K, L \rangle = 0 \)

(ii) \( \langle I \rangle = 1 \)

(iii) \( \langle i, j | k, l \rangle = \delta_{ij} \delta_{kl} \)

(iv) \( \text{in SU}(2) \)

\( \langle (i, j), (k, l) \rangle = \langle (i, j), (k, l) \rangle = \langle (i, j), (k, l) \rangle \)

\( \langle (i, j, k), (l, m, n) \rangle = \langle (i, j, k), (l, m, n) \rangle = \frac{\epsilon_{ijklmn}}{6} \)

in \( \text{SU}(3) \).

Hence [analogous to the \( \text{U}(N) \) computations],

\[ \int_{SU(2)} \det \mathfrak{D} dh_{SU(2)} = c_1^2 + c_2^2 - c_2c_3 + c_3^2 \]  

(B10)

and

\[ \int_{SU(3)} \det \mathfrak{D} dh_{SU(3)} = c_1^3 - 2c_1c_2c_3 + c_2^3 + c_3^3 \]  

(B11)

**APPENDIX C: PROOF OF COROLLARY 1**

By induction, we note for \( 2j < n \)

\[ \lim_{m \to 0} \frac{\tilde{m}_2}{jm} = 1 \quad \text{and} \quad \lim_{m \to 0} \frac{\tilde{m}_2}{jm} = 1. \]  

(C1)

This is trivially true for \( \tilde{m}_1 = m \) and \( \tilde{m}_2 = m + \frac{1}{4m} \). Then, we observe for \( j > 1 \)

\[ \lim_{m \to 0} \frac{\tilde{m}_2}{jm} = \lim_{m \to 0} \frac{m + \frac{1}{4m}}{jm} = \lim_{m \to 0} \frac{1}{jm} + \frac{\frac{1}{4m}}{jm} \]

\[ = \lim_{m \to 0} \frac{1}{jm} + \frac{\frac{1}{4m}}{jm} \]

\[ = \frac{1}{j} + \frac{1}{j} = 1 \]  

(C2)

and

\[ \lim_{m \to 0} \frac{\tilde{m}_2}{jm} = \lim_{m \to 0} \frac{m + \frac{1}{4m}}{jm} = \lim_{m \to 0} \frac{4jm^2 +jm}{jm} = 1. \]  

(C3)

\[ \text{Similarly, for } n \in 2\mathbb{N} - 1, \]
Finally, the asserted identities for $Z(m, \mu, G, n)$ with $G \in \{U(1), SU(2), U(2), SU(3), U(3)\}$ are a trivial corollary substituting $\lim_{m \to 0} c_1$ into the formulas given in Theorem 2.


