

Evading the sign problem in random matrix simulations

Jacques Bloch

*Institute for Theoretical Physics, University of Regensburg, 93040 Regensburg, Germany**

(Dated: March 17, 2011)

We show how the sign problem occurring in dynamical simulations of random matrices at nonzero chemical potential can be avoided by judiciously combining matrices into subsets. For each subset the sum of fermionic determinants is real and positive such that importance sampling can be used in Monte Carlo simulations. The number of matrices per subset is proportional to the matrix dimension. We measure the chiral condensate and observe that the statistical error is independent of the chemical potential and grows linearly with the matrix dimension, which contrasts strongly with its exponential growth in reweighting methods.

PACS numbers: 02.10.Yn, 02.70.Tt, 12.38.Gc

Dynamical simulations of QCD at nonzero chemical potential are hampered by the sign problem (see ref. [1] for a review). The problem occurs when the probabilistic weight used in Monte Carlo methods becomes complex, thus precluding the use of standard importance sampling techniques. Methods to study QCD at small chemical potential, where the sign problem is mild, include reweighting methods, Taylor expansions and analytic continuation from imaginary chemical potential [1]. QCD at zero and nonzero density can also be investigated using random matrix theory because of the universality properties of the Dirac operator to leading order in the epsilon regime. Based on this equivalence, the severity of the sign problem has been studied using the average phase of the fermion determinant [2–6]. As dynamical simulations of random matrices at nonzero chemical potential also suffer from the sign problem they can be used as a playground for its study. In this letter we present a subset method which avoids the sign problem in this case.

In chiral random matrix theory at nonzero chemical potential μ the Dirac operator for a fermion of mass m can be represented by [7]

$$D_{\mu,m}(\phi) = \begin{pmatrix} m & i\phi_1 + \mu\phi_2 \\ i\phi_1^\dagger + \mu\phi_2^\dagger & m \end{pmatrix}, \quad (1)$$

where the configurations $\phi = (\phi_1, \phi_2)$ are pairs of complex random matrices ϕ_1 and ϕ_2 of dimension $(N+\nu) \times N$ with ν the number of zero modes of D (for $m = 0$). The partition function of the chiral random matrix model with N_f dynamical quarks with equal mass m is

$$Z = \int d\phi_1 d\phi_2 w(\phi_1) w(\phi_2) \det^{N_f} D_{\mu,m}(\phi), \quad (2)$$

with Gaussian weights

$$w(\phi_i) = (N/\pi)^{N(N+\nu)} \exp(-N \operatorname{tr} \phi_i^\dagger \phi_i) \quad (3)$$

and integration over the real and imaginary parts of all matrix components. In the presence of a chemical potential the fermion determinant becomes complex, the

sign problem arises and the work needed to make reliable measurements on the statistical ensemble grows exponentially with the volume. In reweighting methods this exponential increase comes from the need to compute exponentially small reweighting factors from a statistical sampling of largely canceling contributions [1].

Herein we propose a method to avoid the sign problem in dynamical simulations of random matrices. The idea is to rewrite the partition function (2) as an integral over subsets $\Omega = \{\phi^i : i = 1, \dots, N_s\}$, each containing N_s configurations $\phi^i = (\phi_1^i, \phi_2^i)$, such that

$$Z = \int d\Omega W(\Omega) \sigma_{\mu,m}(\Omega), \quad (4)$$

where by construction the Gaussian weights $w(\phi_1^i)w(\phi_2^i)$ will be invariant for all ϕ^i in the subset and can thus be denoted by $W(\Omega)$, and we define the fermionic subset weight

$$\sigma_{\mu,m}(\Omega) = \sum_{i=1}^{N_s} \det^{N_f} D_{\mu,m}(\phi^i), \quad (5)$$

which depends on m and μ . The construction of the subsets will be described below, but its crucial property is that the sum $\sigma_{\mu,m}(\Omega)$ over all the determinants in a subset is *real* and *positive*, for every Ω of the ensemble, and can thus be used as a probabilistic weight in a Monte Carlo method. The sample average of an observable O over a sample of N_{MC} subsets Ω_k is computed as

$$\bar{O}_{\mu,m} = \frac{1}{N_{\text{MC}}} \sum_{k=1}^{N_{\text{MC}}} \sum_{i=1}^{N_s} \frac{\det^{N_f} D_{\mu,m}(\phi^{ki})}{\sigma_{\mu,m}(\Omega_k)} O_{\mu,m}(\phi^{ki}), \quad (6)$$

where $\phi^{ki} = (\phi_1^{ki}, \phi_2^{ki}) \in \Omega_k$.

We now sketch how the subsets are constructed. As the partition function (2) is real and positive (for $\mu < 1$) the basic idea is to create subsets of matrices for which $\sigma_{\mu,m}$ has the same property. For $\mu = 0$ the Dirac matrices are anti-Hermitian and their determinants are real and positive. When μ is increased the matrices become non-Hermitian, the determinants complex and the sign

problem sets in. For $\mu = 1$ and $m = 0$ the ensemble becomes maximally non-Hermitian, the average phase factor is zero and the sign problem maximal. For this limiting case we ideally want to find a partner configuration $\psi = (\psi_1, \psi_2)$ for any given ϕ of the ensemble, such that their fermionic determinants explicitly cancel. This is easily achieved if, given some angle θ , we can find a ψ for which $D_{1,0}(\psi) = e^{i\theta} D_{1,0}(\phi)$. Using (1) the latter can be written as

$$e^{i\theta} D_{1,0}(\phi) = \begin{pmatrix} 0 & i\psi_1 + \psi_2 \\ i\psi_1^\dagger + \psi_2^\dagger & 0 \end{pmatrix} = D_{1,0}(\psi), \quad (7)$$

where

$$\begin{cases} \psi_1(\phi; \theta) = \cos \theta \phi_1 + \sin \theta \phi_2 \\ \psi_2(\phi; \theta) = \cos \theta \phi_2 - \sin \theta \phi_1 \end{cases} \quad (8)$$

are orthogonal rotations of (ϕ_1, ϕ_2) , which emerge when $e^{i\theta} D_{1,0}(\phi)$ is explicitly rewritten in an anti-Hermitian and a Hermitian part. From (3) it follows that $w(\psi_1)w(\psi_2) = w(\phi_1)w(\phi_2)$ such that ψ and ϕ have the same Gaussian weights in the partition function, for any rotation θ . An exact pairwise cancellation of the fermionic determinants will happen for any θ obeying $e^{i2NN_f\theta} = -1$. The idea of canceling determinants can be extended from pairs (ϕ, ψ) to subsets $\{\psi(\phi; \theta_n)\}$, where the θ_n are such that the total sum of fermionic determinants cancels for $\mu = 1$ and $m = 0$, i.e., $\sum_n e^{i2NN_f\theta_n} = 0$.

This idea is now ported to arbitrary chemical potential μ and mass m . In the following we will consider subsets

$$\Omega(\phi) = \{\psi(\phi; \theta_n) : n = 0, \dots, N_s - 1\}, \quad (9)$$

where $\psi = (\psi_1, \psi_2)$ is defined in (8) with $\theta_n = \pi n/N_s$ and N_s a positive integer. To each configuration $\phi = (\phi_1, \phi_2)$ of the random matrix ensemble corresponds a subset $\Omega(\phi)$, and the set of all subsets forms an N_s -fold covering of the random matrix ensemble. The configurations $\psi(\phi; \theta)$ have Gaussian weights independent of θ , but different fermionic weights $\det^{N_f} D_{\mu,m}(\psi(\phi; \theta))$.

The method presented in this paper is based on the following *theorem*: For any Ω constructed according to (9), and for arbitrary $\mu < 1$ and m , the fermionic subset weight $\sigma_{\mu,m}(\Omega)$ defined in (5), i.e., the sum of the fermionic determinants of the N_s configurations $\psi(\phi; \theta_n)$, is *real* and *positive* if $N_s > NN_f$.

More specifically, for $m = 0$ it can be shown that

$$\sigma_{\mu,0}(\Omega) = (1 - \mu^2)^{NN_f} \sigma_{0,0}(\Omega), \quad (10)$$

for any μ . As $\sigma_{0,0}$ is real and positive, $\sigma_{\mu,0}$ is also real and positive for $\mu < 1$. For $\mu = 1$ the sum of determinants is exactly zero. For nonzero mass we can show that $\sigma_{\mu,m}$ is real and $\sigma_{\mu,m}(\Omega) > (1 - \mu^2)^{NN_f} \sigma_{0,m}(\Omega)$ for $\mu < 1$, such that the weight $\sigma_{\mu,m}$ is positive. The weights $\sigma_{\mu,m}$ can thus be used to generate subsets of random matrices using a Metropolis algorithm and compute observables

using (6). In practice, N_s will be set to its minimum value, i.e., $N_s = NN_f + 1$.

This theorem can be proven analytically and was thoroughly tested numerically. The proof will be given in a forthcoming paper.

We applied the subset method to compute the chiral condensate of the chiral random matrix ensemble, which is defined as $\Sigma = \frac{1}{2N} \text{tr} D^{-1}$ [8]. We use the Metropolis algorithm to generate subsets Ω according to their weights $W(\Omega)\sigma_{\mu,m}(\Omega)$, where the determinants are computed numerically. Successive subsets in the Markov chain are generated as follows: randomly choose a configuration in the current subset, generate a new configuration by making a random step, construct the subset corresponding to this new configuration, apply an accept-reject step to the newly proposed subset. In each Markov chain we generate 100,000 subsets. The chiral condensate is computed for each configuration and its average is computed using (6). As usual, successive configurations in the Markov chains are correlated and the number of independent subsets is smaller by a factor 2τ , where τ is the integrated autocorrelation time. The statistical errors on the measurements are determined using the standard error formula corrected for the autocorrelations.

To compare the subset method with standard reweighting methods, all the simulations were repeated using quenched, phase quenched and sign quenched reweighting [6]. In those cases, we used $(NN_f + 1) \times 100,000$ random matrices in the Markov chains, such that the total number of matrices generated in the reweighting methods is the same as in the subset method. For the sake of clarity, we only show the results of phase quenched reweighting in the figures below, as its results are representative for the various reweighting methods.

We performed simulations for $N = 2, \dots, 34$, choosing $N_f = 1$ and $m = 0.1/2N$, so that the mass is small w.r.t. the magnitude of the smallest eigenvalue. In fig. 1 the chiral condensate is shown as a function of the chemical potential for matrices with $N = 2, 4, 8$. We compare the results obtained using the subset method with those from phase quenched reweighting. We also show the exact results computed in ref. [8]. The data are shown in the top row and the corresponding relative statistical errors in the bottom row. As the matrix size increases the reweighting method fails for smaller and smaller μ^2 due to the sign and the overlap problem. As expected, the error of the reweighting method grows exponentially, until the method fails when the set of sampled matrices no longer overlaps with the relevant configurations for the given value of μ^2 and N . This strongly contrasts with the results of the subset method which are reliable up to much larger values of μ^2 and agree with the analytical predictions. Moreover, the error is independent of the chemical potential. For larger N we observed that the subset method breaks down at an N -dependent value $\mu_c(N)$, when the positivity of $\sigma_{\mu,m}$ is violated due to the

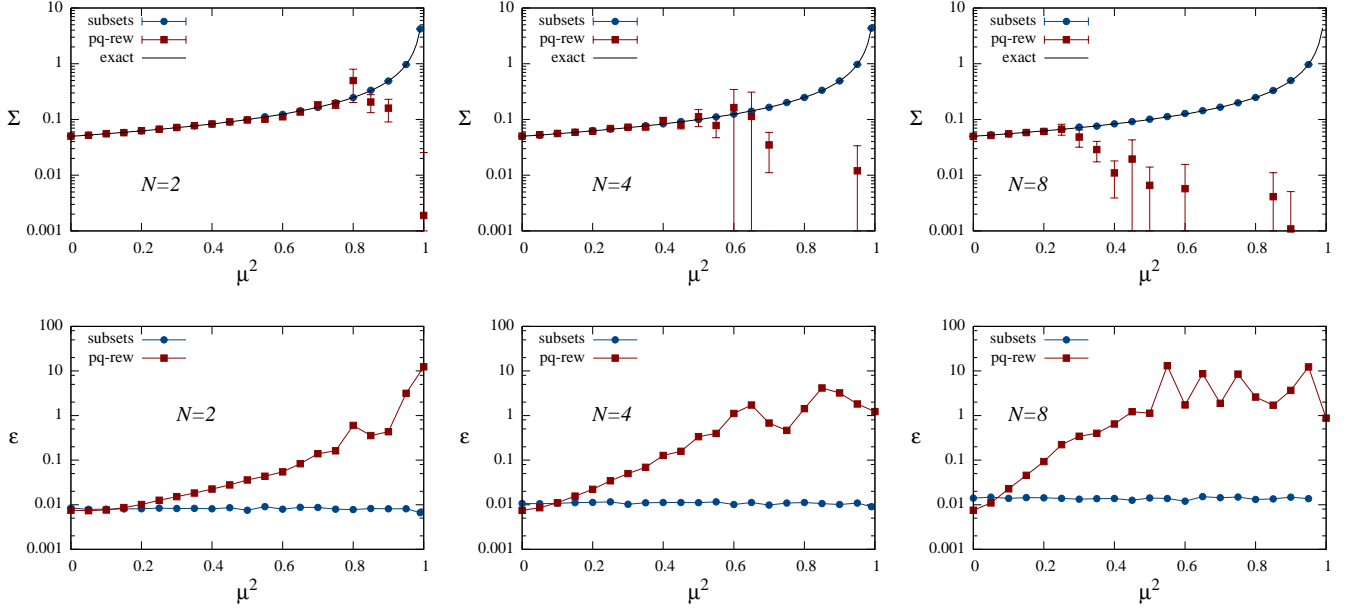


FIG. 1. Chiral condensate Σ (top row) as a function of the chemical potential μ^2 for the subset method and the phase-quenched reweighting method for $N = 2, 4, 8$. The full line shows the exact analytical result of ref. [8]. The reweighting method fails for ever smaller μ^2 when N grows (some data points are negative and are left out of the semi-log plots). The corresponding relative statistical error ε is shown in the bottom row. For the subset method, the error is independent of the chemical potential. The error for the reweighting method grows very rapidly and should only be trusted as long as the method works (see top row).

finite accuracy of the numerical calculations. (This will be discussed later in fig. 3.) Below μ_c the accuracy of the subset method is independent of the chemical potential and the chiral condensate can be computed to good accuracy.

We also measured the chiral condensate as a function of N for fixed values of μ^2 . The N -dependence of the relative statistical error on the chiral condensate is shown in fig. 2 (left: subset method, right: phase quenched reweighting) for different values of μ^2 . For a fixed number of subsets the error in the subset method (left panel) increases approximately as \sqrt{N} and is independent of μ (the latter also follows from fig. 1). If we fix the number of matrices, rather than the number of subsets, the error will increase with an additional factor \sqrt{N} (as the subset size itself grows with $N + 1$), such that the overall relative error will grow linearly with N . Turning the argument around, to achieve a constant error the number of subsets would have to grow proportionally to N , i.e., the total number of matrices should approximately grow as N^2 . The right plot shows the same quantity for phase quenched reweighting (on semi-log scale). We observe that the error grows exponentially with N until the method completely fails when the error stagnates around one and is no longer reliable. Note that for both methods the additional cost for the numerical computation of the determinants is proportional to N^3 .

We return to the observation that the subset method breaks down at large values of N or μ^2 because the finite

precision arithmetic limits the accuracy of the computation of $\sigma_{\mu,m}$. The breakdown occurs when the sum of the determinants in a subset becomes of the order of the accuracy of the individual determinants (which for small m happens when $\sigma_{\mu,m}/\sigma_{0,m} \approx (1 - \mu^2)^{N N_f} \approx 10^{-16}$). We measured the approximate value μ_c^2 , where the breakdown occurs, as a function of N and show the results in fig. 3. The data are fitted with $\mu_c^2 \approx 1 - \exp(-38/N^{1.2})$, which compares well with the rough analytic approximation $\mu_c^2 \approx 1 - \exp(-16 \log(10)/N)$. For large N the fitted curve goes like $1/N^{1.2}$. Note that the breakdown can easily be recognized during simulations by monitoring $\sigma_{\mu,m}$ and the magnitude of the individual determinants. When the ratio of both quantities is of the order of the machine precision the method no longer gives sensible results. Usually this is accompanied by a numerical violation of the positivity of $\sigma_{\mu,m}$. For simulations at $m = 0$ this breakdown can be avoided altogether by using the analytic formula (10) to compute the sampling weight $\sigma_{\mu,0}$ from $\sigma_{0,0}$.

Note that the breakdown of the subset method at $\mu_c(N)$, which is caused by the finite numerical accuracy, will equally well show up in standard reweighting methods, even if an exponentially large effort is invested to keep the statistical error under control. This accuracy problem is therefore different from the sign problem, as the latter is of a pure statistical nature and does not depend on the use of finite-precision or exact arithmetic. The former is an additional problem of numerical nature

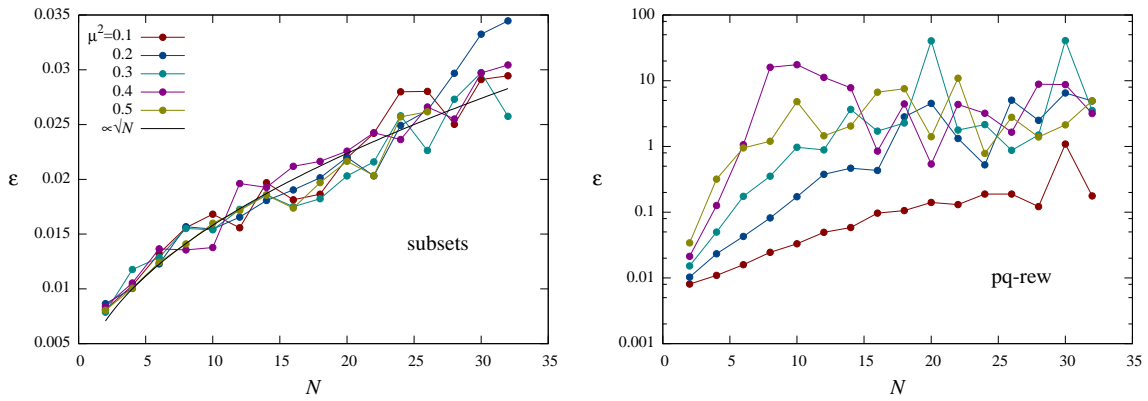


FIG. 2. Relative error ε on the chiral condensate versus matrix size N for various values of chemical potential $\mu^2 = 0.1, 0.2, 0.3, 0.4, 0.5$. The left plot shows the results of the subset method. The full curve $\varepsilon(N) \propto \sqrt{N}$ serves to guide the eye. As a comparison, the right plot shows the relative error for phase quenched reweighting on a semi-log plot (the color coding for μ^2 is the same as in the left plot). The error increases exponentially with N , until the reweighting method fails.

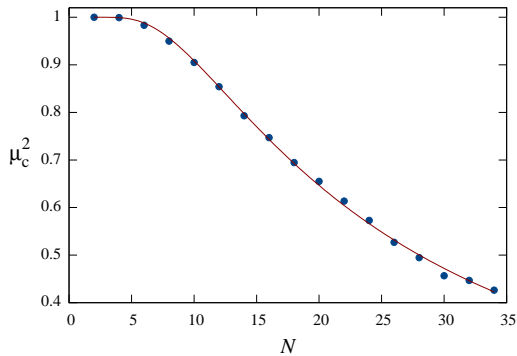


FIG. 3. Chemical potential $\mu_c^2(N)$ where the subset method breaks down due to the finite precision in the computation of $\sigma_{\mu,m}$. The full line shows the fit $\mu_c^2 = 1 - \exp(-38/N^{1.2})$.

which can be improved upon if necessary by using more sophisticated numerical methods.

What happened to the sign problem in the subset method? In standard reweighting methods the large cancellations, inherent to simulations at real chemical potential, happen through statistical sampling of the partition function. In the subset method these cancellations are removed from the statistical sampling procedure and become confined inside the subsets, which are constructed in a deterministic way and whose partial sums $\sigma_{\mu,m}$ yield net real and positive weights. Thus, a fundamental difference between both methods is that the number of configurations in reweighting grows exponentially with the volume to maintain the necessary statistical accuracy on the average weight factor [1], whereas only $NN_f + 1$ matrices per subset are needed in the subset method to compute the positive subset weights $\sigma_{\mu,m}$, independently of its magnitude. This is how the subset method avoids the large statistical cancellations characterizing the sign

problem.

From a practical point of view, even though finite-precision arithmetic in the numerical simulations eventually leads to a breakdown of the subset method, our numerical tests have shown that the method yields a vast improvement over the standard reweighting methods for the dynamical simulation of random matrices. With the new method higher values of μ^2 and N can be accessed, without any loss of accuracy when increasing the chemical potential and with a measurement error growing proportionally to N when increasing the matrix size.

The crucial question whether the subset method is also applicable to physically relevant systems will be investigated in future research.

I would like to thank Philippe de Forcrand and Tilo Wettig for useful discussions. This work was supported by the DFG collaborative research center SFB/TR-55.

* jacques.bloch@physik.uni-regensburg.de

- [1] P. de Forcrand, PoS, **LAT2009**, 010 (2009), arXiv:1005.0539 [hep-lat].
- [2] K. Splittorff and J. J. M. Verbaarschot, Phys. Rev. Lett., **98**, 031601 (2007), arXiv:hep-lat/0609076.
- [3] K. Splittorff and J. J. M. Verbaarschot, Phys. Rev., **D75**, 116003 (2007), hep-lat/0702011.
- [4] J. Bloch and T. Wettig, JHEP, **03**, 100 (2009), arXiv:0812.0324 [hep-lat].
- [5] M. P. Lombardo, K. Splittorff, and J. J. M. Verbaarschot, Phys. Rev., **D80**, 054509 (2009), arXiv:0904.2122 [hep-lat].
- [6] J. Bloch and T. Wettig, (2011), arXiv:1102.3715 [hep-lat].
- [7] J. C. Osborn, Phys. Rev. Lett., **93**, 222001 (2004), arXiv:hep-th/0403131.
- [8] J. C. Osborn, K. Splittorff, and J. J. M. Verbaarschot, Phys. Rev., **D78**, 065029 (2008), arXiv:0805.1303 [hep-th].