

Progress on Complex Langevin simulations of a finite density matrix model for QCD

Jacques Bloch¹, Jonas Glesaaen², Jacobus Verbaarschot³, and Savvas Zafeiropoulos^{4,5,6,*}

¹*Institute for Theoretical Physics, University of Regensburg, 93040 Regensburg, Germany*

²*Department of Physics, College of Science, Swansea University, Swansea, SA2 8PP, United Kingdom*

³*Department of Physics and Astronomy, Stony Brook University, Stony Brook, NY, 11794*

⁴*Thomas Jefferson National Accelerator Facility, Newport News, VA 23606, USA*

⁵*Department of Physics, College of William and Mary, Williamsburg, VA 23187-8795, USA*

⁶*Institute for Theoretical Physics, Universität Heidelberg, Philosophenweg 12, D-69120 Germany*

Abstract. We study the Stephanov model, which is an RMT model for QCD at finite density, using the Complex Langevin algorithm. Naive implementation of the algorithm shows convergence towards the phase quenched or quenched theory rather than to intended theory with dynamical quarks. A detailed analysis of this issue and a potential resolution of the failure of this algorithm are discussed. We study the effect of gauge cooling on the Dirac eigenvalue distribution and time evolution of the norm for various cooling norms, which were specifically designed to remove the pathologies of the complex Langevin evolution. The cooling is further supplemented with a shifted representation for the random matrices. Unfortunately, none of these modifications generate a substantial improvement on the complex Langevin evolution and the final results still do not agree with the analytical predictions.

1 Introduction

Many interesting physical systems have a complex action which impedes Monte Carlo simulations due to the notorious sign problem. In this category belongs QCD at finite baryon density, Yang-Mills theories in the presence of a θ -term, systems of strongly correlated electrons and many other interesting physical systems. The Complex Langevin (CL) algorithm has been advocated as one of the most promising routes for a solution of the sign problem [1, 2]. Despite some successes in toy models, the algorithm seems to have pathological issues such as converging to the wrong theory. Detailed studies have also revealed that the criteria which were put forward in order to guarantee a correct result are not fulfilled in practice in many cases of interest such as cold and dense QCD close to the chiral limit. An update of the current status can be found in [3]. In this talk we discuss a Random Matrix Theory (RMT) model of QCD at nonzero baryon density, which has an exponentially hard sign problem, but serves as an excellent test bed for new algorithms since it can be solved analytically. This model is based on a model for QCD at nonzero nonzero temperature or imaginary chemical potential [4] and was extended to QCD at nonzero chemical potential by Stephanov in [5].

*Speaker, e-mail: savvas@jlab.org

It has a rich phenomenological structure, in particular, it shows a first order transition to a phase of nonzero baryon density. RMT has provided a great deal of analytical results for non-perturbative aspects of QCD, including finite density, lattice cutoff effects and topology [6–13]. First results of our studies were presented in [14].

2 Random Matrix Model

We study an RMT model for QCD at finite density, proposed by Stephanov [5]. This model is defined by its partition function

$$\mathcal{Z}_N^{N_f} = e^{N\mu^2} \int dW dW^\dagger \det^{N_f}(D + m) e^{-N \text{Tr} W W^\dagger}. \quad (1)$$

Originally this model at imaginary chemical potential was introduced as a model for QCD at nonzero temperature [4] in which case it has a second order chiral symmetry restoration transition. The block structure of the Dirac operator D is

$$D = \begin{pmatrix} m & iW + \mu \\ iW^\dagger + \mu & m \end{pmatrix}, \quad (2)$$

where a term proportional to the baryon chemical potential $\mu\gamma_0$ has been added to the chRMT Dirac operator first proposed in [15]. The matrix W is a random complex matrix of size $N \times (N + \nu)$, consisting of Gaussian distributed random numbers, N is the size of the block matrix W , and ν is the analogue of the topological charge. Here we focus on the $\nu = 0$ case, because the topological charge has a negligible effect on the observables that we study. In direct analogy to QCD, RMT simulations are similarly hampered by the sign problem, which becomes exponentially hard when the quark mass is inside the support of the Dirac spectrum.

We mainly focus on two observables, the mass dependent chiral condensate $\Sigma(m) = \langle \bar{\eta}\eta \rangle$ and the baryon number density $n_B = \langle \eta^\dagger \eta \rangle$ which are defined as

$$\langle \bar{\eta}\eta \rangle = \frac{1}{2N} \frac{\partial \log \mathcal{Z}_N^{N_f}}{\partial m}, \quad \text{and} \quad \langle \eta^\dagger \eta \rangle = \frac{1}{2N} \frac{\partial \log \mathcal{Z}_N^{N_f}}{\partial \mu}. \quad (3)$$

In order to understand better the success and failure of the complex Langevin algorithm we will, in parallel to the Stephanov model, also study a similar matrix model for QCD at finite density which turns out not to have a transition to a phase with nonzero baryon density. This is the Osborn model [16], which is a model for QCD at low baryon density. It is a two-matrix model which is structurally quite similar to the Stephanov model and its partition function reads

$$\mathcal{Z}_N^{N_f} = e^{N\mu^2} \int dW dW' dW^\dagger dW'^\dagger \det^{N_f}(D + m) e^{-N \text{Tr}(W W^\dagger + W' W'^\dagger)}. \quad (4)$$

The Dirac operator D has the form

$$D = \begin{pmatrix} m & iW + \mu W' \\ iW^\dagger + \mu W'^\dagger & m \end{pmatrix}. \quad (5)$$

However this model has a quite trivial dependence on the baryon chemical potential since its partition function at $\mu \neq 0$ is related to the one at $\mu = 0$ by a multiplicative factor, as follows [16, 17]

$$\mathcal{Z}_N^{N_f}(m, \mu) = (1 - \mu^2)^{N_f N} \mathcal{Z}_N^{N_f} \left(\frac{m}{\sqrt{1 - \mu^2}}, 0 \right). \quad (6)$$

Due to the above factorization property of the partition function the Osborn model does not possess a phase transition to a phase with non-zero baryon density and thus can only describe properties of QCD for low values of the baryon chemical potential and is expected to have a simpler sign problem compared to the model of Stephanov. In this work, we will check in detail some of the proposed techniques which could in principle fix the problematic issues of the CL algorithm.

The partition function of the Stephanov model can be brought to a form that allows for an easy numerical evaluation at finite N , or can be treated completely analytically by a saddle point approximation when the matrix size $N \rightarrow \infty$ [18]. For the case of one dynamical flavor, $N_f = 1$, the partition function, in units where the chiral condensate $\Sigma = 1$, can be reduced to a σ -model via bosonization

$$\mathcal{Z}_N^{N_f=1}(m, \mu) = e^{N\mu^2} \int d\sigma d\sigma^* e^{-N\sigma^2} (\sigma\sigma^* + m(\sigma + \sigma^*) + m^2 - \mu^2)^N, \quad (7)$$

where σ is the bosonized version of $\bar{\psi}_L\psi_R$. If one changes variables to polar coordinates, the angular integral can be computed analytically and it evaluates to a modified Bessel function. In the end, the partition function becomes a one-fold integral

$$\mathcal{Z}_N^{N_f=1}(m, \mu) = \pi e^{-Nm^2 + N\mu^2} \int_0^\infty du (u - \mu^2)^N I_0(2mN\sqrt{u}) e^{-Nu}. \quad (8)$$

We perform numerical simulations of the Stephanov model employing the Complex Langevin algorithm and we test the algorithm by comparing the numerical data from our simulations to analytical results for the chiral condensate and the baryon density that we can derive by differentiating the partition function (8) with respect to the corresponding sources, m and μ . We also perform simulations of the Osborn model in order to better sketch how the CL algorithm can handle models (or theories in general) that have a strong or a mild sign problem.

3 Implementation of the algorithm

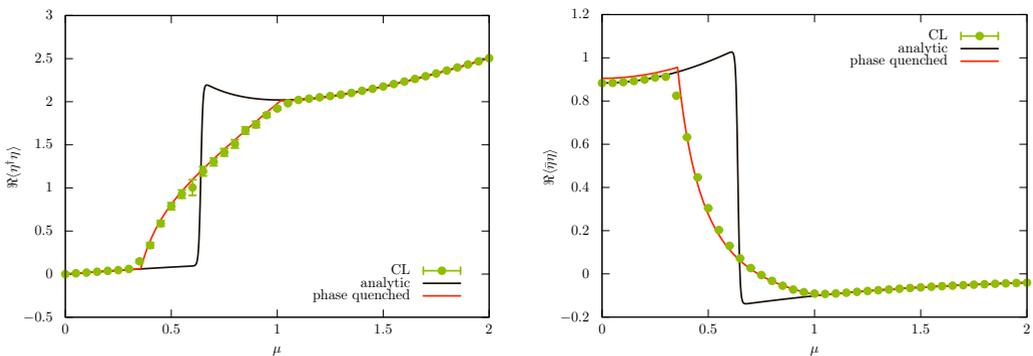


Figure 1. μ -scan for $m = 0.2$ and $N = 48$. We show the baryon number density on the left and the chiral condensate on the right.

When employing the CL algorithm, the real degrees of freedom of the original model will take on complex values; this is called *complexification*. To clearly distinguish between pre- and post-complexification variables, we introduce the following notation,

$$W = X, \quad \text{and} \quad W^\dagger = Y, \quad (9)$$

where W describes the RMT system before complexification, while X and Y are general complex matrices describing the complexified system. At zero Langevin time, $\tau = 0$, we have $Y^\dagger = X$, but this will no longer be the case for $\tau > 0$. In the Cartesian representation the complex matrices X and Y are given by

$$X = A + iB, \quad \text{and} \quad Y = A^T - iB^T, \quad (10)$$

where $\{A, B\} \in \mathbb{R}$ for $\tau = 0$, while $\{A, B\} \in \mathbb{C}$ for $\tau > 0$. As we already showed in [14] a naive implementation of the algorithm leads to convergence to the phase quenched theory, as is clearly demonstrated in figure 1.

4 Gauge cooling

The RMT action is invariant under a $U(N) \times U(N+\nu)$ gauge transformation, whereas after the variables have been complexified, this invariance is enhanced to a $Gl(N) \times Gl(N+\nu)$ symmetry. One can utilize this additional gauge symmetry in order to modify the Langevin evolution in the hope to retrieve the correct results for this model. This method of treating the pathologies of the algorithm has been coined as *gauge cooling*, and has been employed with success on certain models [19–21]. Closely related to our study is its successful application to the Osborn model [20].

4.1 Cooling norms

The matrices $h \in Gl(N)$, so that $\{X, Y\} \rightarrow h\{X, Y\}h^{-1}$, are constructed in a way that reduces a *cooling norm*. Various norms are built in ways that parametrize pathological features of the Complex Langevin evolution. The natural norm is the Hermiticity norm which quantifies the difference between X^\dagger and Y ,

$$\mathcal{N}_H = \frac{1}{N} \text{tr} \left[(X - Y^\dagger)^\dagger (X - Y^\dagger) \right]. \quad (11)$$

One can also introduce an eigenvalue norm [20]

$$\mathcal{N}_{\text{ev}} = \sum_{i=1}^{n_{\text{ev}}} e^{-\xi \gamma_i}, \quad (12)$$

where γ_i are the n_{ev} lowest eigenvalues of the positive definite matrix $D^\dagger D$, and ξ is a real positive parameter. Finally we can define an anti-Hermiticity norm which reads

$$\mathcal{N}_{AH} = \frac{1}{N} \text{tr} \left[\left((\varphi + \psi^\dagger)^\dagger (\varphi + \psi^\dagger) \right)^p \right], \quad (13)$$

The matrices ψ and φ are the off-diagonal elements of D : $\psi = iY + \mu$, $\varphi = iX + \mu$. The first two norms as well as the anti-Hermiticity norm for $p = 1$ were first proposed in [20]. For the Stephanov model, we are using the anti-Hermiticity norm with $p = 2$ because for $p = 1$ the derivatives with respect to the similarity transformation are independent of the chemical potential. For the Osborn model, where this is not a problem, we use $p = 1$. Usually different norms attempt to fix different problems of the Langevin evolution so one can also combine them to make aggregate norms. For example one could combine the Hermiticity norm, which gives the magnitude of drift into the imaginary plane, with either the anti-Hermiticity or the eigenvalue norm, which both deal with the singularities of the drift,

$$\mathcal{N}_{\text{agg}} = (1 - s)\mathcal{N}_{AH/\text{ev}} + s\mathcal{N}_H, \quad \text{where } s \in [0, 1]. \quad (14)$$

We refer the reader to [20, 22] for the details on how to construct the matrix h .

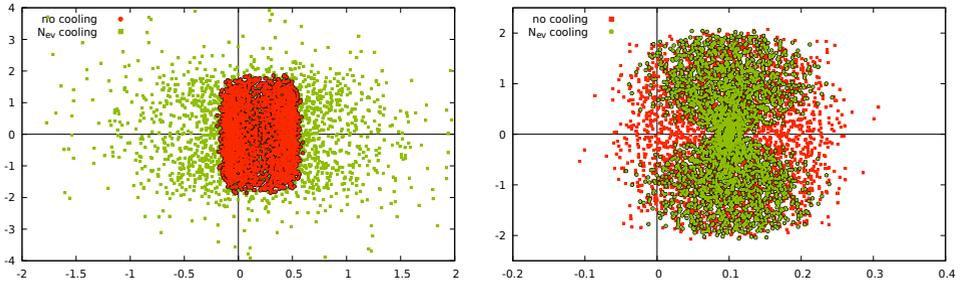


Figure 2. Scatter plots of the eigenvalues of the fermion matrix for a standard CL run (red points) together with the ones from a gauge cooled run (green points). We chose the parameters $\{\xi = 100, n_{ev} = 2\}$ for N_{ev} . The plots show the eigenvalues from the last 60 trajectories, separated by 100 updates. The left hand plot shows the Stephanov model, while the Osborn model is shown to the right.

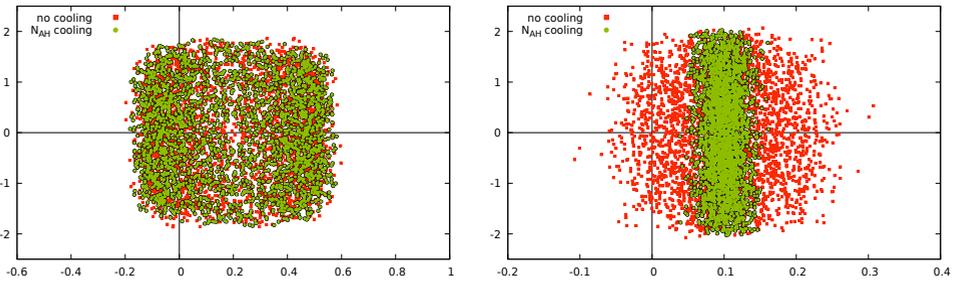


Figure 3. Scatter plots of the Dirac eigenvalues for a standard CL run (red points) together with the ones from a run cooled with the N_{AH} cooling norm (green points). The plots show the eigenvalues from the last 60 trajectories, separated by 100 updates. The LHS plot shows the Stephanov model, while the RHS shows the results of the Osborn model.

4.2 Gauge cooling results

Next, we present our results when applying the method of gauge cooling to the two aforementioned matrix models. In these runs we have used a block matrix of size $N = 24$, Langevin stepsize $dt = 10^{-4}$, and runtime $t_{end} = 1$. Between each Langevin update we perform ten cooling transformations. In what follows the Stephanov model is simulated at parameters $\{m = 0.2, \mu = 0.5\}$, while we chose $\{m = 0.1, \mu = 0.25\}$ for the Osborn model. Parameters are chosen in a way that the two models have a seemingly equally severe sign problem to overcome.

We start by looking at scatter plots of the eigenvalues of the Dirac matrix, shown in figure 2. In direct analogy with the results obtained by [21] the RHS plot shows the effects of cooling on the Osborn model. For this model, the effect of cooling is that the eigenvalue distribution deforms and avoids the origin. On the other hand, this is not the case for the Stephanov model shown on the LHS. Figure 3 shows the same models cooled via N_{AH} . Here again, we observe that cooling cures the pathologies of the Osborn model but not of the Stephanov model.

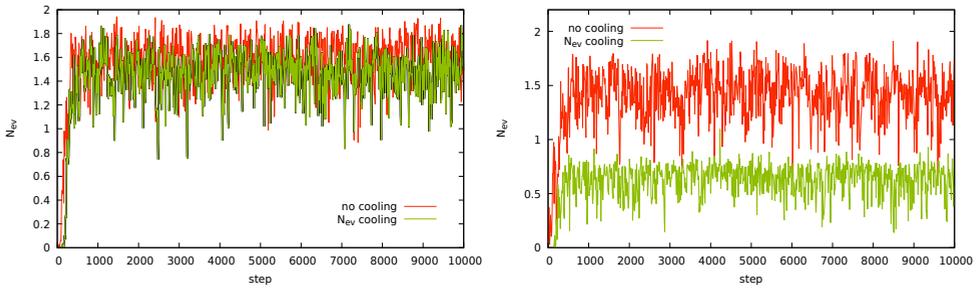


Figure 4. Value of the eigenvalue norm, N_{ev} , with (green curves) and without cooling (red curves) as a function of Langevin time. Stephanov model (LHS), Osborn model (RHS).

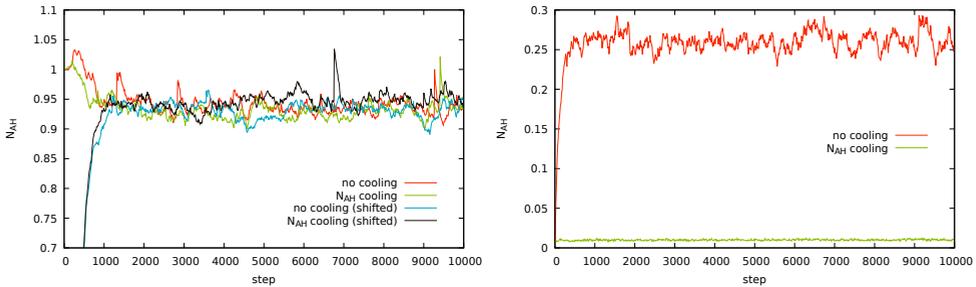


Figure 5. Value of the anti-hermiticity norm, N_{AH} , with (green curves) and without cooling (red curves) as a function of Langevin time. Stephanov model (LHS), Osborn model (RHS). The Stephanov plot also includes the values from the shifted representation with (purple curve) and without cooling (blue curve). These start at 0 for $t = 0$, but very quickly join the unshifted curves.

We now focus on N_{AH} (figure 4) and N_{ev} (figure 5) as a function of Langevin time. We observe that in the case of the Osborn model the norms are heavily reduced by the application of cooling but the Stephanov model is unaffected. Even in the extreme case of the shifted representation (see the next section), which through its advantageous initial conditions could have been a promising fix, we see that we obtain results similar to those without cooling and/or shifting. In figure 6 we show our results for $N = 8$ matrices where we plot the norm and zoom in, in order to see better the effects of cooling and Langevin updates. Once thermalization has been reached we see that the cooling does not have a great enough impact on the simulation as to overcome a single Langevin update; this does not change when decreasing the CL algorithm's stepsize, indicating that the solutions we are interested in are out of reach of the $GI(N)$ transformations.

5 Shifted representation

Another possible resolution of the problematic issues of the algorithm is to shift the chemical potential out of the fermion determinant to the "gauge" degrees of freedom. This can be achieved by a simple

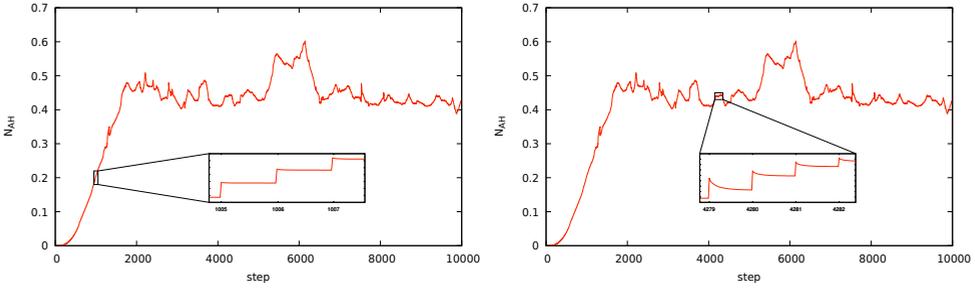


Figure 6. The anti-Hermiticity norm, \mathcal{N}_{AH} , for the shifted representation of the Stephanov model as a function of Langevin time for 8×8 block matrices. The magnification in the inset shows the effect of cooling on \mathcal{N}_{AH} . In this specific analysis, we use 50 gauge cooling transformations between each Langevin step.

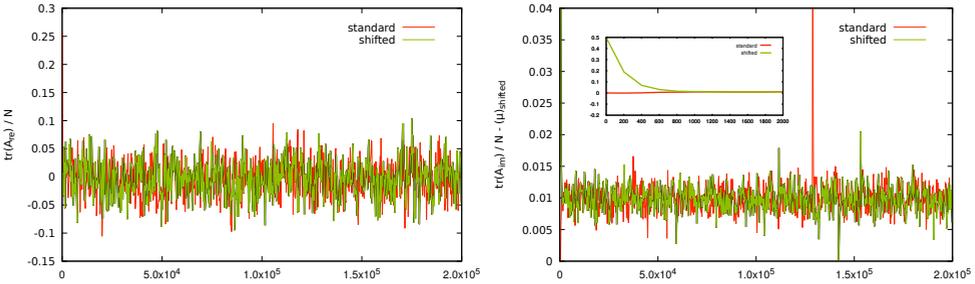


Figure 7. Average diagonal entry of A (red curves) and $A' + i\mu$ (green curves) as a function of the algorithmic steps. The real component of the trace is shown on the left and the imaginary component on the right. The inset in the right figure shows the first 2000 Langevin steps.

change of variables. The Dirac operator initially reads

$$D = \begin{pmatrix} m & iA - B + \mu \\ iA^T + B^T + \mu & m \end{pmatrix}. \quad (15)$$

One can absorb μ into A as follows, $A' = A - i\mu$. In terms of the matrices A' and B the action reads

$$S = N \text{tr} (A'^T A' + 2i\mu A' - \mu^2 + B^T B) - N_f \text{tr} \log (m^2 + X' Y'), \quad (16)$$

where $X' = A' + iB$ and $Y' = A'^T - iB^T$. This result in different Langevin forces [22].

To analyze the dynamics of the shifted representation we analyze the matrix elements of A and A' during a typical CL simulation; these are shown in figure 7. Although the two matrices start out very different, they do coincide after thermalization. This means that

$$\langle A' \rangle_{\text{CL,shifted}} = \langle A \rangle_{\text{CL,standard}} - i\mu, \quad (17)$$

and thus they converge to the same solution.

6 Conclusions

In this contribution we have performed a detailed comparison of the effects of gauge cooling on two different matrix models for QCD at finite density. We have shown that gauge cooling can fix the pathologies of the Osborn model, where the sign problem is artificial and can even be removed by careful redefinition of the parameters, but fails to overcome the problems of the Stephanov model which is a true matrix model for QCD at finite baryon density. The Stephanov model, possessing a phase transition to a phase with non-zero baryon density, is an excellent candidate for this type of investigations since dedicated studies have shown that CL fails when one approaches the deconfinement transition region of the QCD phase diagram, cf. [23] and [24] for an attempt to fix this problem.

7 Acknowledgements

SZ acknowledges support by the National Science Foundation (USA) under grant PHY-1516509 and by the Jefferson Science Associates, LLC under U.S. DOE Contract # DE-AC05-06OR23177. JV acknowledges partial support from U.S. DOE Grant No. DEFAG-88FR40388. JB was supported by the Deutsche Forschungsgemeinschaft (SFB/TRR-55). The authors would like to thank G. Aarts, I. O. Stamatescu, K. Nagata and J. Nishimura for fruitful discussions.

References

- [1] G. Parisi, Phys. Lett. **131B**, 393 (1983)
- [2] G. Aarts, Phys. Rev. Lett. **102**, 131601 (2009), [0810.2089](#)
- [3] E. Seiler, [1708.08254](#)
- [4] A.D. Jackson, J.J.M. Verbaarschot, Phys. Rev. **D53**, 7223 (1996)
- [5] M.A. Stephanov, Phys. Rev. Lett. **76**, 4472 (1996)
- [6] P.H. Damgaard, K. Splittorff, J.J.M. Verbaarschot, Phys. Rev. Lett. **105**, 162002 (2010)
- [7] G. Akemann, P.H. Damgaard, K. Splittorff, J.J.M. Verbaarschot, Phys. Rev. **D83**, 085014 (2011)
- [8] M. Kieburg, J.J.M. Verbaarschot, S. Zafeiropoulos, Phys. Rev. Lett. **108**, 022001 (2012)
- [9] M. Kieburg, J.J.M. Verbaarschot, S. Zafeiropoulos, Phys. Rev. **D88**, 094502 (2013)
- [10] M. Kieburg, J.J.M. Verbaarschot, S. Zafeiropoulos, Phys. Rev. **D92**, 045026 (2015)
- [11] K. Cichy, K. Splittorff, S. Zafeiropoulos (2016), [1612.01289](#)
- [12] A.M. Halasz, A.D. Jackson, R.E. Shrock, M.A. Stephanov, J.J.M. Verbaarschot, Phys. Rev. **D58**, 096007 (1998)
- [13] J.B. Kogut, M.A. Stephanov, D. Toublan, J.J.M. Verbaarschot, A. Zhitnitsky, Nucl. Phys. **B582**, 477 (2000)
- [14] J. Bloch, J. Glesaaen, O. Philipsen, J. Verbaarschot, S. Zafeiropoulos, EPJ Web Conf. **137**, 07030 (2017)
- [15] E.V. Shuryak, J.J.M. Verbaarschot, Nucl. Phys. **A560**, 306 (1993)
- [16] J.C. Osborn, Phys. Rev. Lett. **93**, 222001 (2004)
- [17] J. Bloch, Phys. Rev. **D86**, 074505 (2012)
- [18] A.M. Halasz, A.D. Jackson, J.J.M. Verbaarschot, Phys. Rev. **D56**, 5140 (1997)
- [19] E. Seiler, D. Sexty, I.O. Stamatescu, Phys. Lett. **B723**, 213 (2013)
- [20] K. Nagata, J. Nishimura, S. Shimasaki, JHEP **07**, 073 (2016)
- [21] K. Nagata, H. Matsufuru, J. Nishimura, S. Shimasaki, PoS **LATTICE2016**, 067 (2016)
- [22] J. Bloch, J. Glesaaen, J. Verbaarschot, S. Zafeiropoulos, [17XX.XXXXX](#)
- [23] Z. Fodor, S.D. Katz, D. Sexty, C. Török, Phys. Rev. **D92**, 094516 (2015)
- [24] J. Bloch, O. Schenk, [1707.08874](#)