Towards overcoming the Monte Carlo sign problem with tensor networks

Mari Carmen Bañuls, Krzysztof Cichy, J. Ignacio Cirac, Karl Jansen, Stefan Kühn, and Hana Saito

Abstract. The study of lattice gauge theories with Monte Carlo simulations is hindered by the infamous sign problem that appears under certain circumstances, in particular at non-zero chemical potential. So far, there is no universal method to overcome this problem. However, recent years brought a new class of non-perturbative Hamiltonian techniques named tensor networks, where the sign problem is absent. In previous work, we have demonstrated that this approach, in particular matrix product states in 1+1 dimensions, can be used to perform precise calculations in a lattice gauge theory, the massless and massive Schwinger model. We have computed the mass spectrum of this theory, its thermal properties and real-time dynamics. In this work, we review these results and we extend our calculations to the case of two flavours and non-zero chemical potential. We are able to reliably reproduce known analytical results for this model, thus demonstrating that tensor networks can tackle the sign problem of a lattice gauge theory at finite density.

1 Introduction

Lattice methods have proven to be indispensable when addressing questions in the non-perturbative regime of quantum chromodynamics (QCD). Lattice QCD (LQCD), initially introduced in the 1970s by Wilson [1], has succeeded in elucidating many aspects of the strong interaction from first principles, including e.g. the computation of quark and hadron masses, hadron structure related quantities as well as non-zero temperature properties. The standard tool of LQCD are Monte Carlo (MC) simulations, performed on world’s most powerful supercomputers. Despite the great progress achieved in the last decades, there are still areas where LQCD fails to give precise quantitative answers. This concerns, in particular, parameter regimes where MC simulations encounter a sign problem. Arguably the most important such case, at least from the point of view of this paper, is at non-zero chemical potential, i.e. at finite baryon density. When the chemical potential is non-zero, the standard Boltzmann factor of LQCD in the Euclidean formulation becomes complex and can no longer

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).
be interpreted as a probability measure. This undermines the whole principle of MC simulations. Although partial solutions to this problem are known, they only allow to simulate relatively small baryon densities, in particular much smaller than the one corresponding to the conjectured critical endpoint in the density-temperature phase diagram. As a consequence, there is broad search for alternative approaches, including Lefschetz thimbles [2], complex Langevin simulations [3] and density of states methods [4]. Yet another thread of research, one that we concentrate on in this paper, is related to the tensor networks (TN) approach.

The TN approach\(^1\), originally introduced in the context of condensed matter physics and further developed thanks to quantum information theory, has been successfully applied to the description of quantum many-body systems, including, in last years, also lattice field theory. The latter included computations of spectra [7–10], thermal states [11–15], phase diagrams [16–19] and real-time evolution [8, 20–22]. In this paper, we review the approach and our results for the Schwinger model. Moreover, we go one step further by simulating the two-flavour Schwinger model in a setup where MC simulations would suffer from the sign problem, at non-zero chemical potential.

The paper is organized as follows. We formulate the Schwinger model for a general number of flavours in Sec. 2. In Sec. 3, we describe the basics of the tensor network method. Sec. 4 summarizes our results for the chiral condensate in the one-flavour model at zero temperature. In Sec. 5, we review our results at finite temperature, considering again the chiral condensate. In Sec. 6, we discuss our findings for the chiral condensate in the one-flavour model at zero temperature. In Sec. 7 summarizes and a short discussion of the prospects of the TN approach for lattice gauge theories is given.

## 2 Multi-flavour Schwinger model

The Schwinger model [23] is quantum electrodynamics in 1+1 dimensions (QED\(_2\)). It has been widely used as a toy model for testing new lattice methods, before applying them e.g. to QCD. Interestingly, QED\(_2\) shares certain properties with QCD, such as confinement and chiral symmetry breaking.

Our lattice formulation uses Kogut-Susskind staggered fermions [24], leading to the following Hamiltonian for QED\(_2\) used as a toy model for testing new lattice methods, before applying them e.g. to QCD. Interestingly, QED\(_2\) shares certain properties with QCD, such as confinement and chiral symmetry breaking.

\[
\sum_{n=0}^{N-1} \sum_{f=0}^{F-1} \left( \phi_{n,f}^\dagger e^{i\theta_n} \phi_{n+1,f} - \text{H.c.} \right) + \sum_{n=0}^{N-1} \sum_{f=0}^{F-1} (m_f(-1)^n + \kappa_f) \phi_{n,f}^\dagger \phi_{n,f} + \frac{ag^2}{2} \sum_n L_n^2, \tag{1}
\]

where: \(a\) – lattice spacing, \(g\) – coupling, \(m_f\) – \(f\)-th flavour fermion mass, \(\kappa_f\) – its chemical potential, \(\phi_{n,f}\) – fermionic fields describing flavour \(f\) on site \(n\), \(\theta_n\) – operator that gives the electric flux of link \(n\), \(\theta_n \in [0, 2\pi]\) – conjugate operator such that \(e^{i\theta_n}\) acts as the electric flux raising/lowering operator.

The Gauss law is expressed by:

\[
L_n = L_{n-1} + \sum_{f=0}^{F-1} \left( \phi_{n,f}^\dagger \phi_{n,f} - \frac{1}{2} (1 - (-1)^n) \right) \tag{2}
\]

and allows to integrate out the gauge fields, leaving only the electric field on the left boundary as an independent parameter (taken to be zero in this work).

We write the above Hamiltonian in dimensionless form and we use a Jordan-Wigner transformation, \(\phi_n = \sum_{l<0} (i\sigma^z_l) \sigma^-_n, \phi_n^\dagger = \sum_{l<0} (-i\sigma^z_l) \sigma^+_n\), where \(\sigma_{l/z}^\pm\) are Pauli matrices acting on site \(n\) and where

\(^1\)For a pedagogical introduction, see e.g. Refs. [5, 6].
we order the fermions on each site such that \( \phi_{n,f} = \phi_{nF+f} \). The Hamiltonian in this equivalent spin formulation describes a spin chain of length \( NF \) and reads:

\[
W = -x \sum_{p=0}^{NF-1} \left( \sigma_p^+ (i \sigma_{p+1}^z) \ldots (i \sigma_{p+N-1}^z) \sigma_{p+N}^- + \text{h.c.} \right) + \sum_{n=0}^{N-1} \sum_{f=0}^{F-1} (\mu_f \langle -1 \rangle_n^+ + v_f) \frac{1 + \sigma_{nF+f}^z}{2} 
\]

where \( x = 1/(ag)^2 \), \( \mu_f = 2 \sqrt{m_f/g} \), \( v_f = 2 \sqrt{\kappa_f/g} \) are dimensionless parameters. Note that the gauge symmetry (manifested in the Gauss law) has effectively generated long-range interactions, of the form \( \sigma_{nF}^z \sigma_{m}^z \) (with arbitrary site-flavour index \( n, m \)) in the spin language, and also a site-dependent magnetic field. The hopping term always involves fermions of the same flavour, i.e. it acts on pairs of spins at a distance \( F \) in the effective \( NF \)-site chain.

In the following, we consider the 1-flavour and 2-flavour cases. With the former, one can demonstrate, for example, calculations of the spectrum, of ground state expectation values and of thermal properties. However, the chemical potential has no physical effect with only one flavour – therefore, to exhibit that the TN approach can tackle the model at non-zero density, we employ the 2-flavour case. In this way, we show that in the regime where a MC simulation would suffer from a sign problem, the TN methods can still produce reliable results.

### 3 Tensor network approach

Among the tensor network methods, the most successful one for one-dimensional systems (1+1-dimensional quantum field theories) is the matrix product states (MPS) [25–30]. Below, we use MPS for all our investigations at zero temperature. The MPS ansatz for a system of \( N \) sites has the form:

\[
|\Psi\rangle = \sum_{i_0 \ldots i_N=0} |i_0 \ldots i_{N-1}\rangle \text{Tr} (A_{i_0}^{i_1} \ldots A_{i_{N-2}}^{i_{N-1}})|i_0 \ldots i_{N-1}\rangle,
\]

where \( |i_n\rangle \) are basis states for each site \( n \) and \( d \) is the dimension of the one-site Hilbert space. Each matrix \( A_n^{i_k} \) is \( D \)-dimensional and \( D \) is called the bond dimension and determines the number of parameters in the MPS ansatz. It has been shown that MPS describe accurately ground states of local gapped Hamiltonians, but their practical application extends to more general models.

The task of finding the MPS approximation to the ground state (GS) boils down to finding the components of the tensors \( A_n \) that minimize the expectation value of the Hamiltonian, \( E = \langle \Psi | H | \Psi \rangle / \langle \Psi | \Psi \rangle \). This is performed variationally, by minimizing \( E \) with respect to one tensor \( A_n \) at a time and changing \( n \) successively (sweeping from left to right and back) until global convergence is achieved. Such local minimization problem consists in solving the eigenvalue problem of an effective Hamiltonian acting on site \( n \) and its neighbouring virtual bonds. This variational algorithm [29] is related to the original density matrix renormalization group (DMRG) formulation [31, 32], which was better understood in the language of MPS. Note that we use OBC, which is, in general, more efficient and results in more stable numerical behaviour. The commonly used graphical representation of an MPS and its contractions that give the energy \( E \) is shown in Fig. 1. A circle corresponds to one tensor \( A_n \) and the legs represent indices (vertical ones 0, \ldots, \( d-1 \); horizontal ones 0, \ldots, \( D-1 \)). The lines connecting two tensors represent contracted indices, while open legs of an MPS correspond to physical indices of each site. The contraction of an MPS \( |\Psi\rangle \) with itself (contracting the physical
indices) represents the norm \( \langle \Psi | \Psi \rangle \) (lower right of Fig. 1), while an insertion of an operator, such as the Hamiltonian, allows to compute its expectation value in the state represented by an MPS, e.g. \( \langle \Psi | H | \Psi \rangle \) (upper right).

Apart from the GS, the variational algorithm can find excited states. Knowing the ground state MPS, \( | \Psi_0 \rangle \), this proceeds by constructing a projector onto the orthogonal subspace and finding the GS in this subspace. Successively projecting out further excited states, one can find any number of them (however, the accumulated error may become large for higher excited states). In this way, one can access the vector and scalar particles in the Schwinger model spectrum. Note, however, that the usage of OBC violates translational invariance and hence identification of particles, which are zero-momentum excitations, is more difficult. In particular, the scalar meson, naively the second excitation of the one-flavour Schwinger model, is above momentum excitations of the vector particle. Nevertheless, unambiguous identification of particles is still possible.

The application of the MPS approach to non-zero temperatures requires its extension allowing for a description of operators, in this case density matrix operators [33–35]. One introduces matrix product operators (MPO) of the form:

\[
\rho = \sum_{\{i_n, j_n\}} \text{Tr} \left( M_0^{i_0,j_0} \cdots M_{N-1}^{i_{N-1},j_{N-1}} \right) |i_0 \cdots i_{N-1}\rangle \langle j_0 \cdots j_{N-1}| ,
\]

where the tensors \( M_n \) play an analogous role to the tensors \( A_n \) in the MPS ansatz. For density matrices appropriate for the description of a system at finite temperature, an MPO approximation can be obtained using imaginary time evolution of the identity operator [33], \( \rho(\beta) \approx e^{-\beta H} = e^{-\frac{\beta}{2} H} I e^{-\frac{\beta}{2} H} \), with \( \beta \equiv 1/T \) being the inverse temperature. For a detailed description of the technical steps for the one-flavour Schwinger model, we refer to our original papers [11, 12]. The efficient application of the imaginary time evolution to find thermal states requires specific algorithms, for instance applying discrete time steps in a Suzuki-Trotter expansion [36, 37] plus using appropriate approximations for the long-range terms in the Hamiltonian. We start with the identity operator, \( \rho(0) \), which can be represented trivially by an MPO with bond dimension one. Then, we evolve the density matrix and approximate each step by an MPO with the desired maximum bond dimension, using the so-called Choi isomorphism [38], \( |i\rangle \langle j| \rightarrow |i\rangle \otimes |j\rangle \) that vectorizes the density operators, i.e. transforms the MPO into an MPS with physical dimension \( d^2 \). Finally, we minimize the Euclidean distance between the original and final MPS and repeat the procedure until reaching the inverse temperature \( \beta \).
4 Results – 1-flavour Schwinger model at $T = 0$

We start reviewing our results with a zero-temperature calculation of the ground state in the one-flavour Schwinger model. We illustrate the main steps concerning an example of the determination of the ground state expectation value of the chiral condensate, $\Sigma = \langle \bar{\psi} \psi \rangle$, which can be written using spin operators as:

$$\frac{\Sigma}{g} = \frac{\sqrt{\pi}}{N} \sum_n (-1)^n \frac{1 + \sigma_n^z}{2}.$$  \hspace{1cm} (6)

In the massless case, this expectation value can be computed analytically, yielding the result $\Sigma_{T=0,m=0}/g = e^{\gamma_E}/2\pi^{3/2} \approx 0.159929$ [39], where $\gamma_E \approx 0.577216$ is the Euler-Mascheroni constant. The analytical computation at non-zero fermion mass is no longer possible. The non-perturbative MPS determination requires subtraction of a logarithmic divergence, which is present already in the free theory and can be removed with:

$$\Sigma_{\text{subtr}}(m/g, x) = \Sigma(m/g, x) - \Sigma_{\text{free}}(m/g, x),$$ \hspace{1cm} (7)

where $\Sigma_{\text{free}}(m/g, x)$ is the free theory condensate at fermion mass $m/g$ and lattice spacing corresponding to the inverse coupling $x$. 

Figure 2. Successive extrapolations needed to obtain the continuum result for the $T = 0$ chiral condensate at fermion mass $m/g = 0.125$. Upper left: extrapolation in the inverse bond dimension ($x = 200, N = 368$). Upper right: extrapolation to infinite volume ($x = 200$). Lower: continuum extrapolation.
Table 1. Continuum values of the $T = 0$ chiral condensate $\Sigma/g$ for different fermion masses. We compare with results of Ref. [9] and with the analytical result in the massless case or the approximated result from Ref. [40] in the massive case.

<table>
<thead>
<tr>
<th>$m/g$</th>
<th>Subtracted condensate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Our result (MPS)</td>
</tr>
<tr>
<td>0</td>
<td>0.159929(7)</td>
</tr>
<tr>
<td>0.0625</td>
<td>0.1139657(8)</td>
</tr>
<tr>
<td>0.125</td>
<td>0.0920205(5)</td>
</tr>
<tr>
<td>0.25</td>
<td>0.0666457(3)</td>
</tr>
<tr>
<td>0.5</td>
<td>0.0423492(20)</td>
</tr>
<tr>
<td>1.0</td>
<td>0.0238535(28)</td>
</tr>
</tbody>
</table>

An MPS simulation is performed for a finite bond dimension $D$, with a chain of finite length $N$ and at a finite lattice spacing, given by the parameter $x$. In Fig. 2, we illustrate the sequence of extrapolations needed to obtain the continuum result, which is the aim of the computation. First, we extrapolate to infinite bond dimension, thus removing the effect of approximating the ground state with an MPS of finite $D$. The typical behaviour, illustrated in the upper left plot of Fig. 2, is saturation of approximating the ground state, one can perform the infinite volume extrapolation, as illustrated in Fig. 2 (upper right). In this case, and we estimate its uncertainty to be of order of the difference between the results at $D = 160$ and $D = 140$. Note that, given the approximately exponential approach to the value at $1/D = 0$ that we observe in many cases, this error estimate is rather conservative. It should be emphasized that the central value obtained in this way is always compatible with the result of exact diagonalization (ED), in case the latter is feasible.\(^2\) Having the estimates of the $1/D = 0$ values of the condensate for a few values of the volume $N$, one can perform the infinite volume extrapolation, as illustrated in Fig. 2 (upper right). In order that the leading order $1/N$ behaviour is reliable, one needs to ensure that the ratio $N/\sqrt{x}$ is large enough, in this case of order 20-30. Finally, one can subtract the free theory condensate (Eq. (7)) and perform the continuum extrapolation. In the illustrated case (lower plot of Fig. 2), we use Eq. (4.2) of Ref. [12], which is a fitting ansatz quadratic in the lattice spacing $1/\sqrt{x}$ and including logarithmic corrections of $O(\log(x)/\sqrt{x})$. Note that we also test sensitivity to higher order corrections of $O(1/x^{3/2})$ and to the choice of the fitting interval in $x$. In the end, we obtain a continuum result with all sources of possible uncertainties fully estimated and quantified.

Our final results are given in Tab. 1, together with a comparison to another recent MPS computation [9] and the exact result ($m = 0$) or an approximate result from Ref. [40] ($m > 0$). Note that the former comparison is highly non-trivial, since the authors of Ref. [9] used a different approach of infinite MPS, which works directly in the thermodynamic limit, instead of our approach of extrapolating to this limit from selected values of $N$. In general, the obtained precision is, in both cases, very good, at the level of around five-six significant digits.

Apart from the GS expectation value of the condensate, we also computed the spectrum of the one-flavour Schwinger model, reaching similar precision. For details, we refer to Ref. [7].

\(^2\)ED is possible for spin systems including up to around 20-24 spins with a naive approach and up to 40-48 spins when symmetries of the problem are fully exploited. For $N = 368$, the exponential complexity of the problem renders the ED computation practically impossible, as the dimension of the Hamiltonian matrix is of $O(10^{110})$. 
Figure 3. Dependence of the chiral condensate on the inverse temperature $\beta$ in the massless case. The black solid line is the analytical result of Ref. [39]. Left: the red triangles correspond to our continuum results obtained without flux truncation. Right: in addition to the red triangles representing the results without flux truncation, we plot also the result with flux truncation applied (blue circles).

5 Results – 1-flavour Schwinger model at $T > 0$

As described above, the thermal computation is somewhat more complicated. However, the necessary steps in the numerical procedure are very similar, with the addition of another extrapolation that has to be performed. Since the imaginary time interval $[0, \beta]$ is divided into discrete steps of length $\delta$, one needs to take the limit $\delta \to 0$ to eliminate the error necessarily introduced by non-zero $\delta$. This extrapolation is performed as the second one, after the bond dimension extrapolation (see Ref. [12] for explicit plots).

The continuum results for the massless condensate are shown in Fig. 3. The long-range terms in the Hamiltonian require specific approximations to implement the corresponding steps of the evolution. The left plot shows the result when a Taylor expansion is used. We reproduced the analytical curve of Ref. [39] in a wide range of temperatures. Note, however, that the high temperature region with $g\beta \leq 0.5$ does not agree with the analytical result in this plot. This fact, resulting from enhanced cut-off effects in this region, required further investigation. It is possible to increase the precision of the computation by applying instead a truncation to the evolution operator, which turns out to be equivalent to a truncated model, with maximum flux on the link set to a specified value $L_{\text{cut}}$. We found that $L_{\text{cut}} = 10$ effectively corresponds to the untruncated model, but the presence of the flux cut-off leads to a reduced computational effort. Thus, we could simulate at smaller lattice spacings and obtain more reliable continuum results, illustrated in the right plot of Fig. 3. In addition to reducing the uncertainties in the whole range of temperatures, we obtained agreement with the analytical curve even for small $g\beta$.

In Fig. 4, we also show our continuum result for two selected fermion masses, $m/g = 0.125$ (left) and $m/g = 0.5$ (right). We note that, as expected, the thermal curves approach the $T = 0$ result when $g\beta$ is increased. The agreement with the approximate result of Ref. [40] is moderate. However, after our work, a cross-check was provided by the authors of Ref. [15] using the infinite MPS approach and full agreement was observed.
Figure 4. Dependence of the chiral condensate on the inverse temperature $\beta$ for two selected fermion masses, $m/g = 0.125$ (left) and $m/g = 0.5$ (right). Our continuum results (with flux truncation) are shown with red data points. The cyan solid curve is the reference curve from Ref. [40] (generalized Hartree-Fock approximation) and the brown horizontal line is our continuum result for $T = 0$.

Figure 5. Continuum estimate for the isospin number, $\Delta N$, as a function of $\mu_I/2\pi$ ($m = 0$), for volumes $N/\sqrt{x} = 2$ (red solid line), 6 (green dashed line) and 8 (blue dash-dotted line). The dotted vertical lines are theoretical predictions for the phase transitions. Upper inset: GS energy as a function of $\mu_I/2\pi$ for $N/\sqrt{x} = 8$, $x = 16$, $D = 220$ for the two phases with smallest $\Delta N$ (blue X’s indicate $\Delta N = 0$ and red crosses $\Delta N = 2$). The crossing of these energies determines the transition location. Lower inset: Continuum estimate of the location of the first phase transitions vs. inverse volume $\sqrt{x}/N$ for the first (red X’s), second (green crosses), third (blue asterisks) and fourth (magenta dots) transition.

6 Results – 2-flavour Schwinger model at $T = 0$ and non-zero chemical potential

The essential motivation for the TN approach comes from the possibility of overcoming the notorious sign problem that plagues MC simulations of QCD at non-zero chemical potential. The results presented in the previous sections have demonstrated the feasibility of the MPS/MPO approach when applied to lattice gauge theories. However, the analyzed system, the one-flavour Schwinger model both at zero and non-zero temperatures, does not suffer from the sign problem when simulated with MC. Now, we move on to a case where the sign problem indeed appears – the two-flavour model at non-zero chemical potential. For a more detailed description of our results, see Refs. [41, 42].

Our aim was to reproduce the analytical result of Narayanan [43], who computed the isospin number, $\Delta N \equiv N_0 - N_1$, i.e. the difference between the numbers of particles of different flavours, as a function of the isospin chemical potential, $\mu_I$, which corresponds in our notation to $N(v_1 - v_0)/2x$. It was shown that $\Delta N$ undergoes jumps that correspond to first-order phase transitions. In our MPS setup, we could calculate the GS energy of phases with different isospin numbers. The upper inset of
Fig. 5 shows the GS energies of the phases with $\Delta N = 0$ and $\Delta N = 2$ and the crossing of these energies determines the location of the first transition, which corresponds to $\mu_I/2\pi = 1/2$ in the continuum. Our result is, obviously, influenced by effects of finite $D$, $N$ and $x$. We extrapolate away these effects in a similar manner as for the chiral condensate. In the end, we obtain transition locations as illustrated in the main plot of Fig. 5, provided the volume is large enough (see the lower inset for effects of too small $N/\sqrt{x}$). Note that the result of Ref. [43] is volume-independent. In our case, residual finite volume effects can occur, related to the fact that $N$ is finite and thus $\Delta N$ can become of the order of $N$. As soon as we satisfy the condition $\Delta N \ll N$, we no longer observe any dependence on the volume. In this way, our result is in full agreement with the analytical one. We also investigated the massive case, for which no analytical result is available. Contrary to the zero fermion mass case, volume-independence does not hold any more, since the fermion mass provides an additional energy scale in the problem, which affects the locations of the phase transitions.

7 Conclusion and prospects

Tensor networks are a powerful tool for treatment of general physical systems, especially in lower dimensions. The research of the past few years has demonstrated that they are appropriate for the description of lattice gauge theories and they can provide precise and reliable results both at zero and non-zero temperature. In particular, they are free of the sign problem and hence remain as an interesting direction for future simulations of QCD at non-zero chemical potential. For this, extension of their efficiency for higher dimensions is obviously necessary. Theoretical progress in higher dimensional tensor networks is steadily being made. The natural generalization of MPS to higher dimensions is Projected Entangled Pair States (PEPS) [44, 45]. PEPS and other related developments give good prospects of tackling systems in more than one dimension, although they are computationally much more demanding than MPS. Currently, TN methods have already provided one of the most accurate results for certain classes of spin systems in two dimensions, especially for cases where a quantum MC simulation encounters a sign problem. Nevertheless, the application of TN for 2+1-dimensional lattice gauge theories remains difficult. However, it is our aim and the aim of the TN lattice gauge theory community to continue in this direction, which in the future can lead to a full application to 3+1-dimensional QCD.

Acknowledgments. K.C. was supported by the Deutsche Forschungsgemeinschaft (DFG), project nr. Cl 236/1-1.

References

[38] M.D. Choi, Linear Algebra and its Applications 10, 285 (1975)