

# A Method to Calculate Conserved Currents and Fermionic Force for the Lanczos Approximation to the Overlap Dirac Operator

---

**Matthias Puhr\***

*Institut für Theoretische Physik, Universität Regensburg*

*E-mail:* [matthias.puhr@physik.uni-regensburg.de](mailto:matthias.puhr@physik.uni-regensburg.de)

**Pavel Buividovich**

*Institut für Theoretische Physik, Universität Regensburg*

*E-mail:* [pavel.buividovich@physik.uni-regensburg.de](mailto:pavel.buividovich@physik.uni-regensburg.de)

The overlap Dirac operator obeys the Ginsparg-Wilson equation and offers a possibility to introduce chiral symmetry on the lattice. Evaluating the overlap operator is numerically very expensive and one has to rely on approximation methods. At finite chemical potential the overlap operator can be efficiently computed with the two-sided Lanczos algorithm. To calculate conserved currents on the lattice, or to evaluate the fermionic force in HMC calculations, one needs to compute derivatives of the Dirac operator with respect to gauge fields. In this paper we present a method to simultaneously compute the action of the overlap operator and its derivative on a source vector.

*The 32nd International Symposium on Lattice Field Theory,  
23-28 June, 2014  
Columbia University New York, NY*

---

\*Speaker.

## 1. Motivation

Quantum chromodynamics (QCD) is a theory that is chirally symmetric in the limit of massless quarks. Chiral symmetry and its spontaneous breaking play an important role in QCD phenomenology. Many observables depend strongly on the chiral properties of QCD. When one studies QCD on a finite space-time lattice it is therefore desirable to use a discretisation of the Dirac operator that respects chiral symmetry and is free of doublers.

A chiral symmetry preserving and doubler free discretisation of the Dirac operator has to obey the Ginsparg-Wilson equation [1]. Finding such a lattice Dirac operator is a non trivial task, but today several solutions are known. One of them is the overlap Dirac operator, which is an exact solution to the Ginsparg-Wilson equation. Unfortunately the definition of the overlap operator includes the matrix sign function, which is numerically very expensive. Evaluating the overlap operator exactly is therefore not feasible for reasonably large lattice sizes and one has to rely on approximation methods.

A very efficient method is the two-sided Lanczos (TSL) algorithm. It computes an approximation to the action of the overlap operator on a source vector. A major advantage of the TSL algorithm is that it works for general complex matrices. This is important when one considers finite chemical potential, where the sign function of a non-Hermitian matrix has to be evaluated[2]. While the TSL method is well suited to compute the overlap operator, the evaluation of the fermionic force in HMC calculations and the calculation of conserved currents make it necessary to additionally compute the derivative of the Dirac operator. The purpose of this paper is to introduce a numerical method to simultaneously compute the action of the overlap Dirac operator and its derivative on a source vector.

## 2. The overlap operator and the matrix sign function

At finite quark chemical potential  $\mu$  the massless overlap Dirac operator is defined as[2]

$$D_{\text{ov}} := \frac{1}{a} (\mathbb{1} + \gamma_5 \text{sgn}[\gamma_5 D_w(\mu)]), \quad (2.1)$$

where  $D_w(\mu)$  is the Wilson Dirac operator at non-zero chemical potential,  $\text{sgn}$  stands for the matrix sign function and  $a$  is the lattice spacing. The function  $f$  of a matrix  $A \in \mathbb{C}^{n \times n}$  can be defined in several equivalent ways [3]. For a diagonalisable<sup>1</sup>  $A$ , that is  $A = U\Lambda U^{-1}$  with the diagonal eigenvalue matrix  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ , one can employ the particularly simple and convenient spectral decomposition:

$$f(A) := U f(\Lambda) U^{-1}, \quad f(\Lambda) := \text{diag}(f(\lambda_1), \dots, f(\lambda_n)) \quad (2.2)$$

In general the argument of the sign function  $\gamma_5 D_w(\mu)$  is a non-Hermitian matrix with complex eigenvalues  $\lambda_i$  and we need a generalisation of the sign function for complex arguments. It is important that the sign function satisfies  $\text{sgn}(z)^2 = 1$  for any complex number  $z$ , since this ensures that the overlap operator respects the Ginsparg-Wilson relation. Moreover for  $x \in \mathbb{R} \setminus \{0\}$  the

<sup>1</sup>This can be generalised to non-diagonalisable matrices using the Jordan canonical form[3].

complex sign function should reduce to the standard definition  $\text{sgn}(x) = \pm 1$ . A choice that has the requested properties is

$$\text{sgn}(z) := \frac{z}{\sqrt{z^2}} = \text{sgn}(\Re(z)), \quad (2.3)$$

where the cut of the square root is chosen along the negative real axis, so that the cut of the sign function is along the imaginary axis.

For numerical matrix computations it is often more convenient to use an iterative method to evaluate the sign function, the so called Roberts iteration:

$$X_{k+1} := \frac{1}{2} (X_k + X_k^{-1}), \quad X_0 := A \quad (2.4)$$

This is Newton's method applied to the matrix equation  $X^2 = 1$ . If  $A$  has no purely imaginary eigenvalues the  $X_k$  converge quadratically to  $\text{sgn}(A)$ [3].

The run-time complexity of both the spectral decomposition (2.2) and the Roberts iteration (2.4) is approximately  $\mathcal{O}(n^3)$ . Therefore the numerical cost of an evaluation of the overlap operator becomes prohibitively large very quickly. Even for relatively small lattice sizes it is not feasible to compute the matrix sign function and one has to resort to approximation methods.

### 3. The two-sided Lanczos algorithm

In many applications it is not necessary to compute a matrix function  $f(A)$  explicitly since it is sufficient to evaluate the action of  $f(A)$  on a vector  $|\psi\rangle$ . The TSL algorithm is a Krylov subspace method and computes an approximation to  $|y\rangle = f(A)|\psi\rangle$ . The Krylov subspace of order  $k$  for a matrix  $A$  and a vector  $|\psi\rangle$  is defined as

$$\mathcal{K}_k(A, |\psi\rangle) := \text{span}(|\psi\rangle, A|\psi\rangle, \dots, A^{k-1}|\psi\rangle). \quad (3.1)$$

The TSL constructs biorthonormal bases  $V_k = (|v_1\rangle, \dots, |v_k\rangle)$  and  $W_k = (|w_1\rangle, \dots, |w_k\rangle)$  of the right ( $\mathcal{K}_k(A, |\psi\rangle)$ ) and left ( $\mathcal{K}_k(A^\dagger, |\psi\rangle)$ ) Krylov subspace, such that

$$T_k := W_k^\dagger A V_k \quad (3.2)$$

is a tridiagonal  $k \times k$  matrix:

$$T_k = \begin{pmatrix} \alpha_1 & \gamma_1 & 0 & \dots & 0 \\ \beta_1 & \alpha_2 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \gamma_{k-1} \\ 0 & \dots & 0 & \beta_{k-1} & \alpha_k \end{pmatrix} \quad (3.3)$$

Using equations (3.2) and (3.3) one can show that  $V_k$  and  $W_k$  can be built with the following recurrence relations [4]:

$$\begin{aligned} \beta_i |v_{i+1}\rangle &= (A - \alpha_i) |v_i\rangle - \gamma_{i-1} |v_{i-1}\rangle \\ \gamma_i^* |w_{i+1}\rangle &= (A^\dagger - \alpha_i^*) |w_i\rangle - \beta_{i-1}^* |w_{i-1}\rangle \end{aligned} \quad (3.4)$$

The seeds of the recurrence relations (3.4) are chosen to be  $|v_1\rangle = |w_1\rangle = |\psi\rangle / \|\psi\|$ , with the norm  $\|\psi\| := \sqrt{\langle \psi | \psi \rangle}$ . The diagonal of  $T_k$  is fixed by  $\alpha_i = \langle w_i | A | v_i \rangle$ , whereas  $\beta_i$  and  $\gamma_i$  are not uniquely determined and can be obtained from the normalisation condition  $\langle w_{i+1} | v_{i+1} \rangle = 1$ .

The matrix  $V_k W_k^\dagger$  is an oblique projector on the space  $\mathcal{K}_k(A, |\psi\rangle)$  and in a first step we approximate  $|y\rangle$  by the oblique projection of  $f(A) |\psi\rangle$  on  $\mathcal{K}_k(A, |\psi\rangle)$ :

$$|y\rangle \approx |y\rangle_{\text{obl}} = V_k W_k^\dagger f(A) V_k W_k^\dagger |\psi\rangle. \quad (3.5)$$

Combined with the approximation  $W_k^\dagger f(A) V_k \approx f(T_k)$  this yields the final result<sup>2</sup>

$$|y\rangle \approx \|\psi\| V_k f(T_k) |e_1\rangle. \quad (3.6)$$

The problem of evaluating  $f(A)$  is now reduced to the calculation of the bases  $V_k$  and  $W_k$  and the computation of  $f(T_k)$  for the tridiagonal matrix  $T_k$ . The complexity of the function evaluation is reduced from  $\mathcal{O}(n^3)$  to  $\mathcal{O}(nk) + \mathcal{O}(k^3)$ . In practice one can obtain very good approximations of the overlap Dirac operator already for  $k \ll n$ . The efficiency of the TSL can be further increased by using deflation methods [5, 6] and a nested version of the algorithm [7].

## 4. Derivatives of the Lanczos Algorithm

### 4.1 The method

It is relatively simple to compute the derivatives  $\partial D_w / \partial \theta_v(x)$  of the Wilson Dirac operator over the (lattice) gauge field  $\theta_v(x)$  by hand. For the overlap operator things are more complicated and if approximation methods like TSL are used the derivatives can only be evaluated numerically. Apart from divided difference methods the most straightforward way to calculate derivatives of the TSL is algorithmic differentiation[8]. However, experiments with random matrices showed that the algorithmic differentiation approach for TSL is numerically unstable, see figure 1(a). The reason for this is most probably the loss of biorthogonality of  $V_k$  and  $W_k$  because of round-off errors. While this is not a big problem for the TSL, it seems to strongly influence the numerical stability of the algorithmic differentiation.

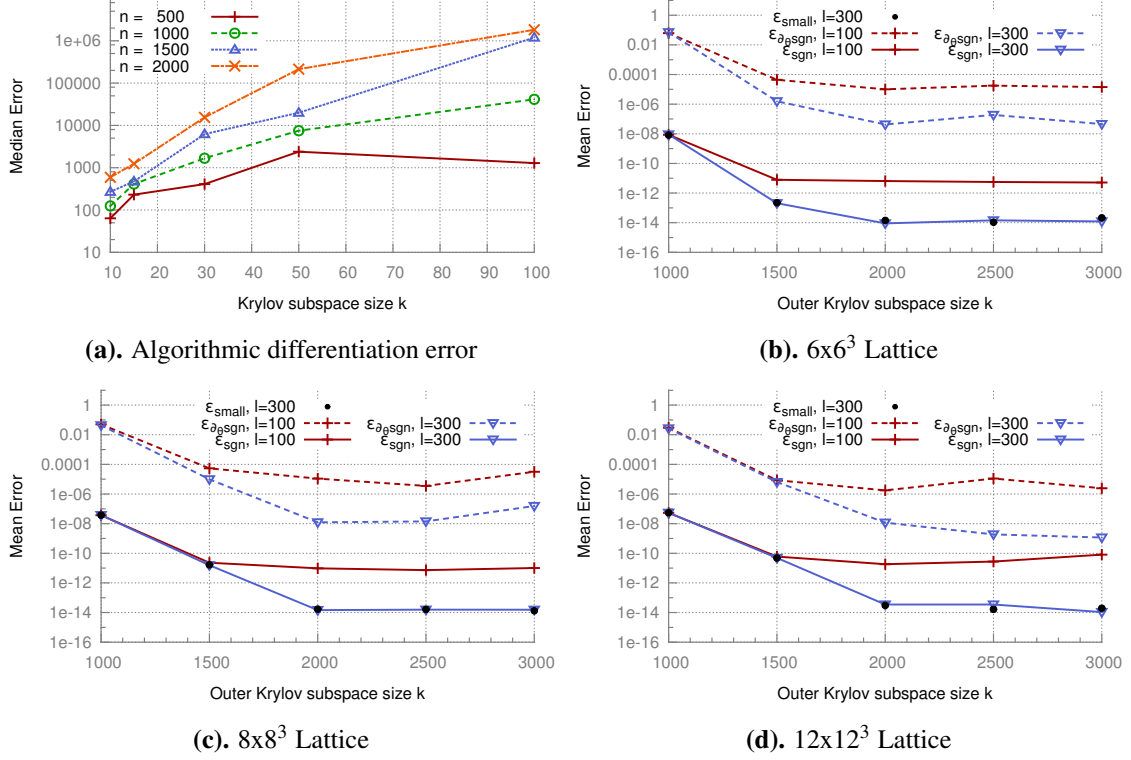
To avoid the numerical stability problems we propose an algorithm based on the following matrix function theorem[9]:

**Theorem 1.** *Let  $A(t) \in \mathbb{C}^{n \times n}$  be differentiable at  $t = 0$  and assume that the spectrum of  $A(t)$  is contained in an open subset  $\mathcal{D} \subset \mathbb{C}$  for all  $t$  in some neighbourhood of 0. Let  $f$  be  $2n - 1$  times continuously differentiable on  $\mathcal{D}$ . Then:*

$$f(\bar{A}) \equiv \begin{bmatrix} f(A(0)) & \frac{d}{dt} \Big|_{t=0} f(A(t)) \\ 0 & f(A(0)) \end{bmatrix}, \quad \bar{A} := \begin{bmatrix} A(0) & \dot{A}(0) \\ 0 & A(0) \end{bmatrix}$$

Theorem 1 relates the derivative of a matrix function to the function of a block matrix. It is remarkable that this enables us to compute the derivative of  $f$  without the knowledge of  $f'$ . This comes at the cost of evaluating the function for a matrix that has twice the dimension of the original matrix  $A$ . Fortunately the block matrix  $\bar{A}$  is sparse and one only needs to store  $A$  and  $\dot{A}$ , as in the

<sup>2</sup>Note that by construction  $\langle w_i | \psi \rangle = \|\psi\| |e_1\rangle \delta_{i1}$



**Figure 1** (a) Algorithmic differentiation of the TSL method for  $n$ -dimensional test matrices. The algorithmic differentiation approach is numerically unstable and the errors increase with the Krylov subspace size. (b) – (d) Results for  $\text{sgn}(\gamma_5 D_w)$  and its derivative at finite chemical potential  $\mu = 0.20$  (in lattice units). We plot the errors (see section 4.2) as a function of the outer Krylov subspace size for two different inner Krylov subspace sizes  $l$ . The points are connected to guide the eye.

algorithmic differentiation approach. Another concern is the convergence of the Lanczos algorithm for the block matrix  $\bar{A}$ . Convergence critically depends on the spectrum of the matrix. It is easy to prove that  $\bar{A}$  has the same spectrum as the matrix  $A$  and every eigenvalue of  $A$  is a (degenerate) eigenvalue of  $\bar{A}$ . In general matrices of the special form  $\bar{A}$  are not diagonalisable. Numerical experiments show that the Jordan normal form of  $\bar{A}$  has non-trivial Jordan blocks of size exactly two, i.e. the Jordan matrix is a block diagonal matrix with 2-dimensional matrices on the main diagonal and zeros everywhere else. Note that the formula for the eigenvectors of  $\bar{A}$  presented in the talk was derived under the assumption of diagonalisability and does not hold in the general case.

The complex sign function is differentiable on  $\mathbb{C} \setminus i\mathbb{R}$  and in practice  $\gamma_5 D_w$  does not have purely imaginary eigenvalues. Therefore theorem 1 holds and we have

$$\text{sgn}(\gamma_5 \bar{D}) \begin{pmatrix} 0 \\ |\psi\rangle \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial \theta_v(x)} \text{sgn}(\gamma_5 D_w) |\psi\rangle \\ \text{sgn}(\gamma_5 D_w) |\psi\rangle \end{pmatrix}, \quad \bar{D} := \begin{bmatrix} D_w & \frac{\partial D_w}{\partial \theta_v(x)} \\ 0 & D_w \end{bmatrix}. \quad (4.1)$$

We can now apply the TSL method to compute the result of the action of  $\text{sgn}(\gamma_5 \bar{D})$  on the vector  $(0, |\psi\rangle)^T$  and to simultaneously obtain an approximation for  $\text{sgn}(\gamma_5 D_w) |\psi\rangle$  and  $\frac{\partial}{\partial \theta_v(x)} \text{sgn}(\gamma_5 D_w) |\psi\rangle$ .

In this way the derivative of the matrix sign function can be computed without the need to modify the TSL algorithm.

## 4.2 Numerical Results

To test the method proposed in the last subsection we compute  $\text{sgn}(\gamma_5 \bar{D})(0, |\psi\rangle)^T$  for random source vectors  $|\psi\rangle$  to get an approximation for  $(\text{sgn}(\gamma_5 D_w) |\psi\rangle, \partial_{\theta_v} \text{sgn}(\gamma_5 D_w) |\psi\rangle)^T$ . The lattice site  $x$  and the direction  $v$  for the derivative  $\partial/\partial\theta_v(x)$  are chosen at random. For the tests we use  $SU(3)$  configurations generated with an improved action[10] on lattices of small to medium size. The inverse coupling is set to  $\beta = 5.95$  in all cases.

To estimate the numerical error  $\varepsilon_{\text{sgn}}$  of the sign function approximation we use the identity  $\text{sgn}(A)^2 = \mathbb{1}$  and define  $\varepsilon_{\text{sgn}} := \|\text{sgn}(A)^2 |\psi\rangle - |\psi\rangle\| / (2\|\psi\|)$ . The factor two enters the definition because we have to apply our approximation twice to compute the square of the sign function. The commutator  $\{\partial_{\theta_v} \text{sgn}(A), \text{sgn}(A)\}$  vanishes, as can be seen most easily by taking the derivative of the squared sign function. We use this fact to define the numerical error  $\varepsilon_{\partial_{\theta} \text{sgn}}$  of the derivative as  $\varepsilon_{\partial_{\theta} \text{sgn}} := \|\{\partial_{\theta_v} \text{sgn}(A), \text{sgn}(A)\} |\psi\rangle\| / (2\|\psi\|)$ . To compare the convergence properties, we also computed the numerical error  $\varepsilon_{\text{small}}$  of a TSL approximation to  $\text{sgn}(\gamma_5 D_w) |\psi\rangle$ .

In our calculations we use a nested version of the TSL method[7] with a single nesting step. The outer Krylov subspace size  $k$  varies between  $k = 1000$  and  $k = 3000$ . For the inner subspace size  $l$  values between  $l = 100$  and  $l = 500$  are used. We find that for our test cases an inner size  $l = 300$  is sufficient and further increasing the size of the inner space does not significantly improve the approximation. The results for different lattice sizes are plotted in figures 1(b) – 1(d). We observe that  $\varepsilon_{\text{small}}$  and  $\varepsilon_{\text{sgn}}$  lie almost on top of each other. Although the matrix dimension increases by a factor of two, it is still sufficient to use the same Krylov subspace size to reach a given precision. The algorithm scales very well with matrix size. To reach a precision of  $10^{-12}$  in the sign function approximation we need a Krylov subspace size of  $k \approx 1400$ ,  $k \approx 1600$  and  $k \approx 1800$  for  $6 \times 6^3$ ,  $8 \times 8^3$  and  $12 \times 12^3$  lattices, respectively.

In all test cases  $\varepsilon_{\text{sgn}}$  is much smaller than  $\varepsilon_{\partial_{\theta} \text{sgn}}$ , but they stem from different error definitions and are not directly comparable. Qualitatively  $\varepsilon_{\text{sgn}}$  and  $\varepsilon_{\partial_{\theta} \text{sgn}}$  show the same behaviour: The error decreases with increasing Krylov subspace size until the optimal subspace size is reached. Then the approximation converges and an additional increase of the subspace size does not improve the results.

## 5. Conclusion and Outlook

We introduce a new numerical method to simultaneously compute the action of a matrix function and its derivative on a source vector and present test results for  $\text{sgn}(\gamma_5 D_w)$  and  $\partial_{\theta_v} \text{sgn}(\gamma_5 D_w)$  on small and medium sized lattices. The method is based on the matrix function identity (4.1) and uses the well established two-sided Lanczos algorithm to compute the action of a matrix function on a vector. Our tests show that the proposed method is reliable and scales very well.

The convergence properties of the TSL for a given matrix  $A$  are closely related to the spectrum of  $A$ . The efficiency of the algorithm can be greatly enhanced by using deflation methods[5, 6]. We are currently working on the implementation of suitable deflation techniques for the proposed method. One problem is that the block matrix constructed as part of the algorithm is in general

not diagonalisable, which makes the adaption of standard deflation methods difficult. A version of theorem 1 for higher order derivatives exists[9] and we are investigating ways to generalise our method to higher order derivatives.

## Acknowledgements

We acknowledge helpful discussions with Andreas Frommer, who pointed us to theorem 1, Jacques Bloch and Simon Heybrock. We thank Oleg Kochetkov for providing the gauge configurations.

## References

- [1] P. H. Ginsparg, K. G. Wilson, *A remnant of chiral symmetry on the lattice*, *Phys. Rev. D* **25** (1982) 2649–2657.
- [2] J. Bloch, T. Wettig, *Overlap Dirac Operator at Nonzero Chemical Potential and Random Matrix Theory*, *Phys. Rev. Lett.* **97** (2006), 012003 [[hep-lat/0604020](#)].
- [3] N. J. Higham, *Functions of Matrices: Theory and Computation*, SIAM, Philadelphia 2008.
- [4] G. H. Golub, C. F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore and London 1996.
- [5] J. Bloch, A. Frommer, B. Lang, T. Wettig, *An iterative method to compute the sign function of a non-Hermitian matrix and its application to the overlap Dirac operator at nonzero chemical potential*, *Comput.Phys.Commun.* **177** (2007), 933–943 [[0704.3486](#)].
- [6] J. Bloch, T. Brey, T. Wettig, *Comparing iterative methods to compute the overlap Dirac operator at nonzero chemical potential*, in proceedings of Lattice 2008, [PoS\(LATTICE 2008\)027](#).
- [7] J. Bloch, S. Heybrock, *A nested Krylov subspace method to compute the sign function of large complex matrices*, *Comput. Phys. Commun.* **182** (2011), 878–889 [[0912.4457](#)].
- [8] A. Griewank, A. Walther, *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*, SIAM, Philadelphia 2008.
- [9] R. Mathias, *A Chain Rule for Matrix Functions and Applications*, *SIAM J. Matrix Anal. & Appl.* **17** (1996), 610–620.
- [10] M. Lüscher, P. Weisz, *On-shell improved lattice gauge theories*, *Comm. Math. Phys.* **97** (1985), 59–77.