Contour Optimisation for the Sign Problem with an Efficient Evaluation of the Jacobian

Michael Kroyter

מכון טכנולוגי חולון
Holon Institute of Technology

Work in progress with *Naomi Solomensky*

(and 1805.04941, 2103.08948 with *Francis Bursa*)

University of Bern

23-January-2025

Outline

Introduction

Introduction

Many ways for defining contours that reduce the sign problem: Thimbles, flows, explicit expressions, machine learning.

Given N real integration variables, complexify the variables and look for a *real N dim* submanifold of the *complex N dim* space. The condition $Im(S) = C$ defines a $2N - 1$ dim object.

A lot of freedom in choosing a contour for $N > 1$.

Strategy – a Simple and Practical Approach

Requirements from an integration contour

- *Uniform deformations*:
  Easy to implement and effective in reducing the phase.
- $Im(S_{eff}) \approx C$:
  Exponential time is acceptable for small enough exponents.
- *Low computational cost*:
  In particular in the evaluation of the Jacobian.

We examine the method for the Bose Gas with chemical potential.

The mean phase factor, $\left\langle e^{i\,\mathrm{Im}(S_{eff})} \right\rangle_{\mathrm{Re}(S_{eff})}$, is used as a measure for the severity of the sign problem.

Contour Deformations for the d-Dimensional Bose Gas with $\mu \neq 0$

The Bose Gas at Finite Chemical Potential

Consider Bose gas with μ: A theory of a single complex scalar. Define $\alpha \equiv \frac{1}{2d+m^2}$. In lattice units the theory can be written as:

$$S = \frac{1}{\lambda\alpha^2}\sum_{\vec{r}}\Big(\Phi^*_{\vec{r}}\Phi_{\vec{r}} + (\Phi^*_{\vec{r}}\Phi_{\vec{r}})^2 - \alpha\sum_{\nu=0}^{d-1}\left(\Phi^*_{\vec{r}}\Phi_{\vec{r}+\hat{\nu}}e^{-\mu\delta_{\nu,0}} + \Phi^*_{\vec{r}+\hat{\nu}}\Phi_{\vec{r}}e^{\mu\delta_{\nu,0}}\right)\Big)$$

For an undeformed contour there is a sign problem:

$$\mathrm{Im}(S) = \frac{2\sinh\mu}{\lambda\alpha}\sum_{\vec{r}}\mathrm{Im}\left(\Phi^*_{\vec{r}}\Phi_{\vec{r}+\hat{0}}\right)$$

Complexify the (complex) field as:

$$\Phi_{\vec{r}} \to \Phi_{\vec{r}} = \phi_{\vec{r}} + i\psi_{\vec{r}}\,, \qquad \Phi^*_{\vec{r}} \to \bar{\Phi}_{\vec{r}} \equiv \phi^*_{\vec{r}} + i\psi^*_{\vec{r}}$$

$\bar{\Phi}_{\vec{r}}$ is the complex conjugate of $\Phi_{\vec{r}}$ only for the undeformed contour: $\psi_{\vec{r}} = 0$.

Defining a Contour

We look for a deformed contour by specifying $\psi_{\vec{r}} = \psi_{\vec{r}}\{\phi_{\vec{s}}\}$.

For a general deformation:

$$\begin{aligned}\mathrm{Im}(S) = \frac{2}{\lambda\alpha^2}\,\mathrm{Re}\sum_{\vec{r}}\Big(&\phi^*_{\vec{r}}\psi_{\vec{r}}\big(1+2|\phi_{\vec{r}}|^2-2|\psi_{\vec{r}}|^2\big)\\ &-\alpha\cosh\mu\big(\phi^*_{\vec{r}}\psi_{\vec{r}+\hat{0}}+\psi_{\vec{r}}\phi^*_{\vec{r}+\hat{0}}\big)-\alpha\sum_{\nu=1}^{d-1}\big(\phi^*_{\vec{r}}\psi_{\vec{r}+\hat{\nu}}+\psi_{\vec{r}}\phi^*_{\vec{r}+\hat{\nu}}\big)\\ &-i\alpha\sinh\mu\phi^*_{\vec{r}}\phi_{\vec{r}+\hat{0}}+i\sinh\mu\alpha\psi^*_{\vec{r}}\psi_{\vec{r}+\hat{0}}\Big)\end{aligned}$$

A very simple solution for $\mathrm{Im}(S)=0$ exists: $\psi_{\vec{r}} = \pm i\phi_{\vec{r}}$.

So are we done here?

Defining a Contour

We look for a deformed contour by specifying $\psi_{\vec{r}} = \psi_{\vec{r}}\{\phi_{\vec{s}}\}$.

For a general deformation:

$$\begin{aligned}\mathrm{Im}(S) = \frac{2}{\lambda\alpha^2}\,\mathrm{Re}\sum_{\vec{r}}\Big(&\phi^*_{\vec{r}}\psi_{\vec{r}}\big(1+2|\phi_{\vec{r}}|^2-2|\psi_{\vec{r}}|^2\big)\\ &-\alpha\cosh\mu\big(\phi^*_{\vec{r}}\psi_{\vec{r}+\hat{0}}+\psi_{\vec{r}}\phi^*_{\vec{r}+\hat{0}}\big)-\alpha\sum_{\nu=1}^{d-1}\big(\phi^*_{\vec{r}}\psi_{\vec{r}+\hat{\nu}}+\psi_{\vec{r}}\phi^*_{\vec{r}+\hat{\nu}}\big)\\ &-i\alpha\sinh\mu\phi^*_{\vec{r}}\phi_{\vec{r}+\hat{0}}+i\sinh\mu\alpha\psi^*_{\vec{r}}\psi_{\vec{r}+\hat{0}}\Big)\end{aligned}$$

A very simple solution for $\mathrm{Im}(S)=0$ exists: $\psi_{\vec{r}} = \pm i\phi_{\vec{r}}$.

So are we done here?

No. Now either $\Phi_{\vec{r}} = 0$ or $\bar{\Phi}_{\vec{r}} = 0$. Hence $S = 0$, not only $\mathrm{Im}(S)$.

Wrong boundary conditions.

Defining a Contour Using an Expansion

The term inducing the sign problem is proportional to the *small parameter* $0 \leq \alpha \leq \frac{1}{2d}$.

Hence, we define $\psi_{\vec{r}} = \sum_{n=1}^{\infty} \alpha^n \psi_{\vec{r}}^{(n)}$.

Expanding $\mathrm{Im}(S)$ to lowest order in α gives:

$$\mathrm{Re}\left(\sum_{\vec{r}} \phi_{\vec{r}}^{*}\Big(\big(1+2|\phi_{\vec{r}}|^2\big)\psi_{\vec{r}}^{(1)} - i\sinh\mu\,\phi_{\vec{r}+\hat{0}}\Big)\right) = 0$$

Many continuous solutions exist. A simple choice:

$$\psi_{\vec{r}}^{(1)} = i\sinh\mu\frac{\phi_{\vec{r}+\hat{0}}}{1+2|\phi_{\vec{r}}|^2} + i\phi_{\vec{r}} f_{\vec{r}}$$

with $f_{\vec{r}}$ an arbitrary real function of the fields.

Generalizing the Contour

The obtained expressions is certainly useful for small values of αe^{μ}.

How to improve it?

- Go to the next order in the α expansion.
  Certainly helps but an expansion is necessarily limited.
- Generalize to an ansatz:

$$\psi^{(1)}_{\vec{r}} = i \sinh \mu \frac{a_1 \phi_{\vec{r}} + a_2 \phi_{\vec{r}+\hat{0}}}{1 + b_1 |\phi_{\vec{r}}|^2 + b_2 |\phi_{\vec{r}+\hat{0}}|^2}$$

  a_1 gives a particular choice of $f_{\vec{r}}$.
  $b_2 > 0$ regularizes the contour.
- Do both, go to higher order and generalize.
  The expansion essentially guides us towards ansätze that could be useful also for large expansion parameter.

The Jacobian for the First Order Contours for $d = 1$

ψ_r depends only on ϕ_r and ϕ_{r+1}. Hence, the Jacobian has the form

$$J = \det \begin{pmatrix} A_1 & B_1 & 0 & 0 & \dots \\ 0 & A_2 & B_2 & 0 & \dots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & A_{L-1} & B_{L-1} \\ B_L & 0 & \dots & 0 & A_L \end{pmatrix}$$

A_r and B_r are known 2×2 matrices.

Without the B_L the Jacobian would have been given by a product of L determinants of size 2×2. Thus, the update of the Jacobian would have been $O(1)$ per site, or $O(L)$ per sweep.

Otherwise, it seems that *the cost is $O(L^4)$ per sweep*.

Evaluation of the Jacobian

The Jacobian matrix is not upper diagonal but it's not too bad.
Performing elementary block row operations leads to:

$$J = \det\left(\mathbf{1} - (-1)^L A_1^{-1} B_1 \cdot \ldots \cdot A_L^{-1} B_L\right) \det(A_1) \cdot \ldots \cdot \det(A_L)$$

A product of determinants of 2×2 matrices.
But the first determinant itself contains a product of $2L$ matrices.
The cost per sweep is $O(L^2)$.

Using cyclicity this can be made into $O(L)$, but then we have to invert the matrices $A_r^{-1} B_r$, but the B_r might not be invertible.

Evaluation of the Jacobian

The Jacobian matrix is not upper diagonal but it's not too bad.
Performing elementary block row operations leads to:

$$J = \det\left(\mathbf{1} - (-1)^L A_1^{-1} B_1 \cdot \ldots \cdot A_L^{-1} B_L\right) \det(A_1) \cdot \ldots \cdot \det(A_L)$$

A product of determinants of 2×2 matrices.
But the first determinant itself contains a product of $2L$ matrices.
The cost per sweep is $O(L^2)$.

Using cyclicity this can be made into $O(L)$, but then we have to invert the matrices $A_r^{-1} B_r$, but the B_r might not be invertible.

But going beyond first order for $d > 1$ would not be possible.
Need another approach.

Another Approach for Fast Evaluation of the Jacobian

Need fast algorithm for more general contours, for arbitrary d.

Revise the expressions for the evaluation of the "last lattice point".
Several such options were considered.

Another Approach for Fast Evaluation of the Jacobian

Need fast algorithm for more general contours, for arbitrary d.

Revise the expressions for the evaluation of the "last lattice point".
Several such options were considered.

Remaining problems

- Now the expressions are not uniform.
  Why is it a problem?
- As μ is increased the problem becomes more significant.
- For $d = 1$ this concerns a single lattice point.
  For general d there are L^{d-1} such points.
  This is enough to cause a significant sign problem.

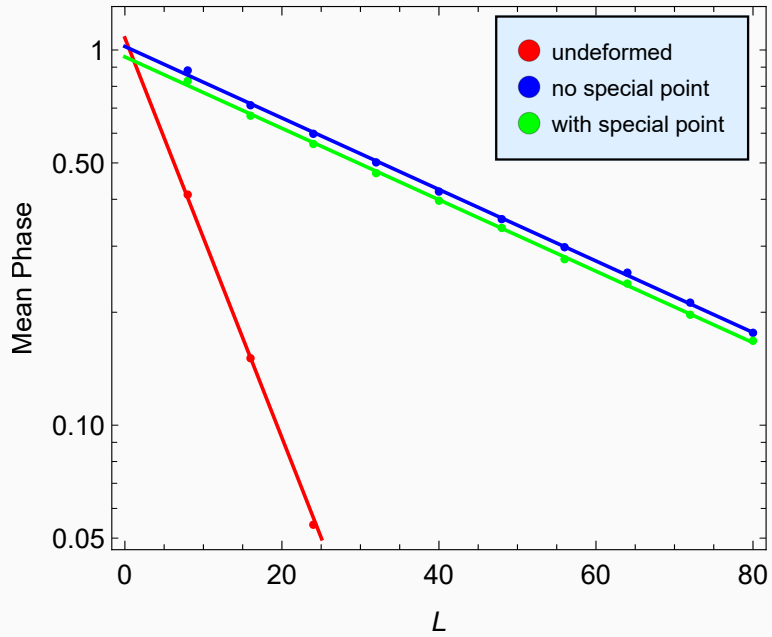
On the Bright Side

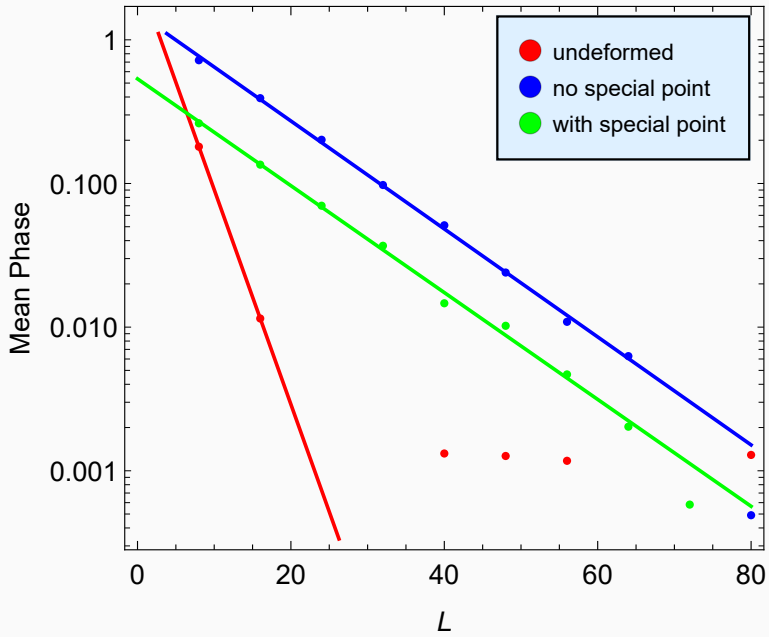
Not all is bad with this approach.

The problem is with the mean phase factor for a given lattice size.

It is important to examine what happens as we change this size.

The mean phase factor as a function of lattice size for $\lambda = 1$, $m = 1$, with $\mu = 1$ (left) and $\mu = 1.5$ (right) on a logarithmic scale





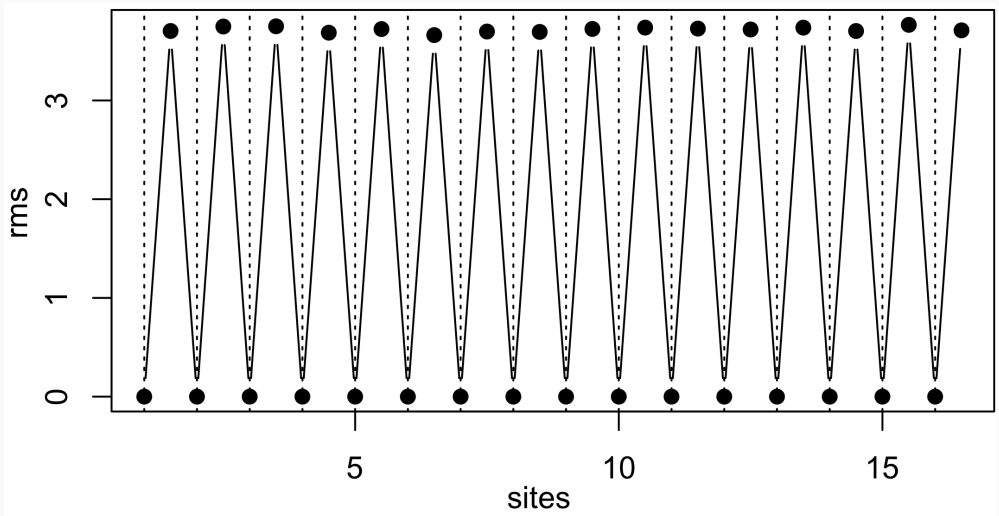
The Importance of Being Uniform

Undeformed Contour

The phase can get contributions from local terms (point r) and from nearest neighbour terms (point $r + \frac{1}{2}$).

RMS of the phase factor for the undeformed contour

$d = 1$, $L = 16$, $\lambda = 1$, $m = 1$ ($\alpha = \frac{1}{3}$, $\mu = 1$)



Only "half-integer points" contribute, as can be seen from S.

For uncorrelated contributions, the RMS of $\mathrm{Im}(S)$ would have been about $\sqrt{16}$ times the value at a single point.

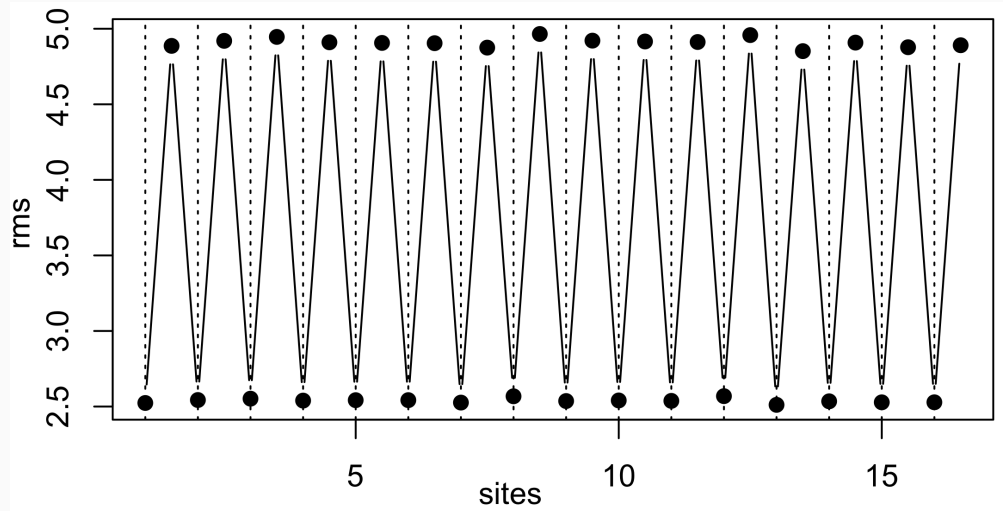
In fact, it is significantly lower.

Deformed Contour

Finding a good contour is as much about inducing proper correlations among the various contributions, as it is about reducing individual contributions.

RMS of the phase factor for the first order ansatz contour

$d = 1$, $L = 16$, $\lambda = 1$, $m = 1$ ($\alpha = \frac{1}{3}$), $\mu = 1$, no special point



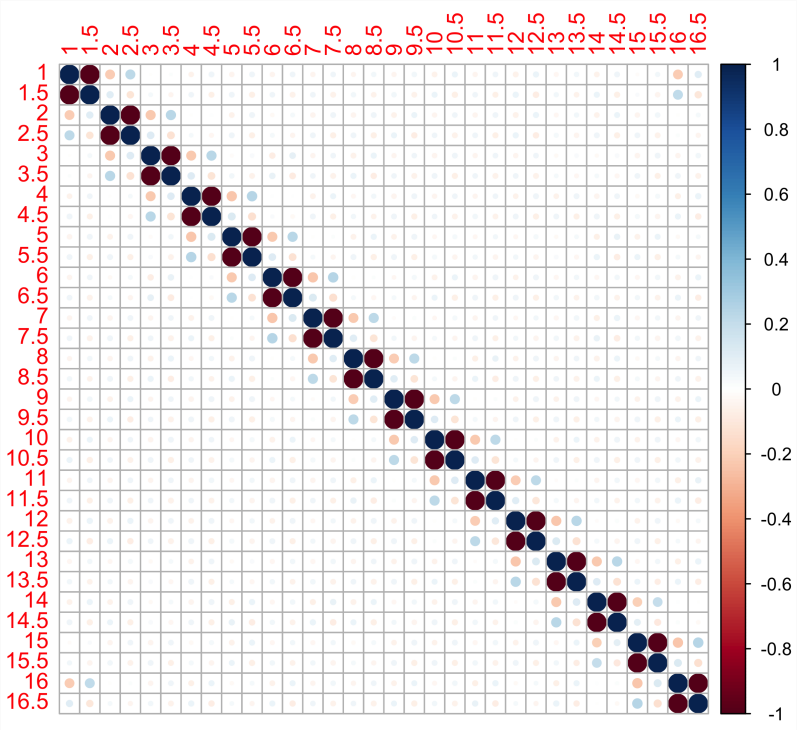
Now integer points also contribute and contributions are larger.

Still Im(S) is much reduced.

Correlations of Different Contributions to $\mathrm{Im}(S)$

Correlations among contributions to $\mathrm{Im}(S)$ from integer and half-integer lattice sites

$d = 1,\ L = 16,\ \lambda = 1,\ m = 1,\ \mu = 1$, no special point



Special treatment of a single point breaks this picture.

In $d > 1$ the situation is even worse.

Uniform Contours with Fast Evaluation of the Jacobian

Rewrite the action

Uniformity does not mean identical deformations at all sites.

Define parity on sites that alternates with the time coordinate.

Two natural options:

- Time coordinate parity: $|\vec{r}| = |(r_0, \ldots, r_{d-1})| \equiv (-1)^{r_0}$
- Chessboard parity: $|\vec{r}| = |(r_0, \ldots, r_{d-1})| \equiv (-1)^{r_0 + r_1 + \ldots + r_{d-1}}$

Declare that on even sites the fields are not deformed.

Rewrite the condition for canceling the lowest order term of $\mathrm{Im}(S)$:

$$\mathrm{Re} \sum_{\vec{r}\,\mathrm{odd}} \phi^*_{\vec{r}} \Big(\psi^{(1)}_{\vec{r}} \big(1 + 2|\phi_{\vec{r}}|^2\big) - i \sinh \mu \big(\phi_{\vec{r}+\hat{0}} - \phi_{\vec{r}-\hat{0}} \big) \Big)$$

A simple solution exists (for $\vec{r}$ odd):

$$\psi^{(1)}_{\vec{r}} = i \sinh \mu \frac{\phi_{\vec{r}+\hat{0}} - \phi_{\vec{r}-\hat{0}}}{1 + 2|\phi_{\vec{r}}|^2} + i \phi_{\vec{r}} f_{\vec{r}}$$

Ansätze and the Jacobian

Generalizing to an ansatz is straighforward:

$$\psi_{\vec{r}} = i\alpha \sinh \mu \, \frac{a_1(\phi_{\vec{r}+\hat{0}} - \phi_{\vec{r}-\hat{0}})}{1 + b_1|\phi_{\vec{r}}|^2 + b_2|\phi_{\vec{r}+\hat{0}} - \phi_{\vec{r}-\hat{0}}|^2}$$

For the evaluation of the Jacobian we put the even lattice points after the odd ones.

Upper triangular Jacobian matrix. Thus:

$$J = \prod_{\vec{r}\,\text{odd}} J_{\vec{r}}$$

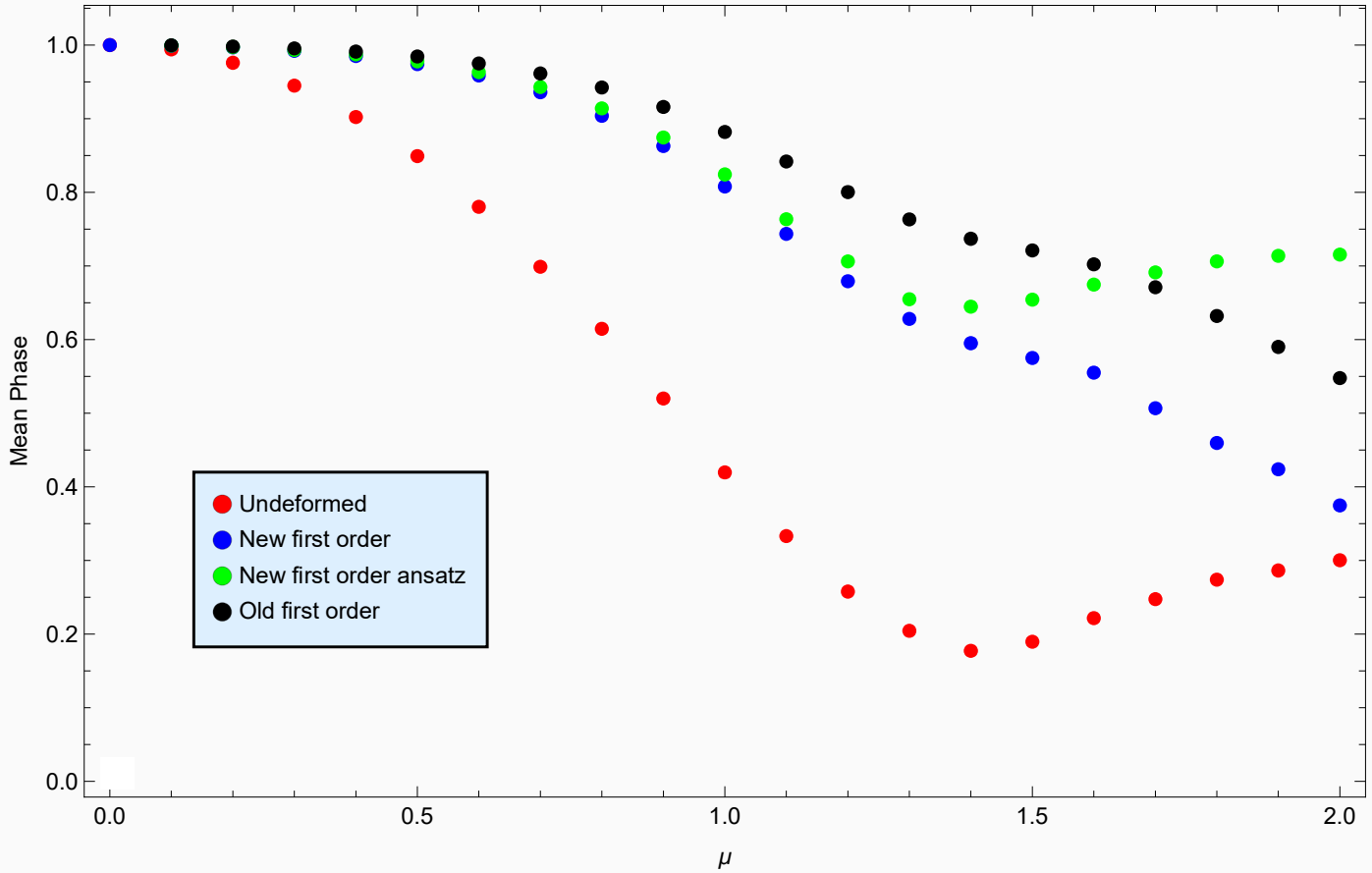
$$J_{\vec{r}} = 1 - \frac{2ib_1 \,\text{Re}\left(\phi^*_{\vec{r}}\psi_{\vec{r}}\right)}{1 + b_1|\phi_{\vec{r}}|^2 + b_2|\phi_{\vec{r}+\hat{0}} - \phi_{\vec{r}-\hat{0}}|^2}$$

At each site we have to evaluate a simple local expression.

Comparing Contours

The mean phase factor as a function of μ for for different contours

$$d = 1,\ \lambda = 1,\ L = 8,\ m = 1$$



Higher Order Ansätze

Plugging $\psi^{(1)}$ into the action we get for the second order:

$$\mathrm{Re}\left(\sum_{\vec{r}} \phi^*_{\vec{r}} \psi^{(2)}_{\vec{r}} \left(1+2|\phi_{\vec{r}}|^2\right) - 2i \sinh\mu \sum_{\vec{r}\,\mathrm{odd}} \frac{1}{1+2|\phi_{\vec{r}}|^2} \cdot \right.$$
$$\left. \cdot \left(\cosh\mu \phi_{\vec{r}+\hat{0}} \phi^*_{\vec{r}-\hat{0}} + \sum_{\nu=1}^{d-1} \phi_{\vec{r}+\hat{\nu}} \phi^*_{\vec{r}-\hat{\nu}} \right) \right) = 0$$

Second order terms at odd lattice sites are not enough.

Instead, we rewrite again, and choose $\psi_{\vec{r}} \neq 0$ only if $\vec{r} \equiv_4 0$.

Need chessboard (quaternary) parity now.

$$\psi^{(2)}_{\vec{r}} = \frac{i \sinh(2\mu)}{1+2|\phi_{\vec{r}}|^2} \left(\frac{\phi_{\vec{r}+2\hat{0}}}{1+2|\phi_{\vec{r}+\hat{0}}|^2} - \frac{\phi_{\vec{r}-2\hat{0}}}{1+2|\phi_{\vec{r}-\hat{0}}|^2} \right)$$
$$+ \frac{2i \sinh\mu}{1+2|\phi_{\vec{r}}|^2} \sum_{\nu=1}^{d-1} \left(\frac{\phi_{\vec{r}+2\hat{\nu}}}{1+2|\phi_{\vec{r}+\hat{\nu}}|^2} - \frac{\phi_{\vec{r}-2\hat{\nu}}}{1+2|\phi_{\vec{r}-\hat{\nu}}|^2} \right)$$

Problems with the Jacobian Again

The obtained expression exactly works for second order but the Jacobian Matrix is again not simple.
Modify the explicit expression to a similarly looking ansatz,

$$\psi^{(2)}_{\vec{r}} = \frac{i}{\left(1 + b_3|\phi_{\vec{r}}|^2 + b_4 \sum_{\nu=0}^{d-1} |\phi_{\vec{r}+2\hat{\nu}} - \phi_{\vec{r}-2\hat{\nu}}|^2\right)^2} \cdot$$
$$\cdot\left(a_2 \sinh(2\mu)\left(\phi_{\vec{r}+2\hat{0}} - \phi_{\vec{r}-2\hat{0}}\right) + a_3 \sinh\mu \sum_{\nu=1}^{d-1}\left(\phi_{\vec{r}+2\hat{\nu}} - \phi_{\vec{r}-2\hat{\nu}}\right)\right)$$

More terms can be added to account for correlations between next to nearest neighbours with odd parity.

Summary

Summary

Contour deformation is a viable approach for the sign problem.

Address the source of the sign problem (nearest $\hat{0}$ neighbours).

An efficient computational method with a determinant that can be evaluated with complexity $O(V)$ for arbitrary d.

Interpreting the expansion

The expansion is in powers of $\alpha = \frac{1}{2d+m^2}$:

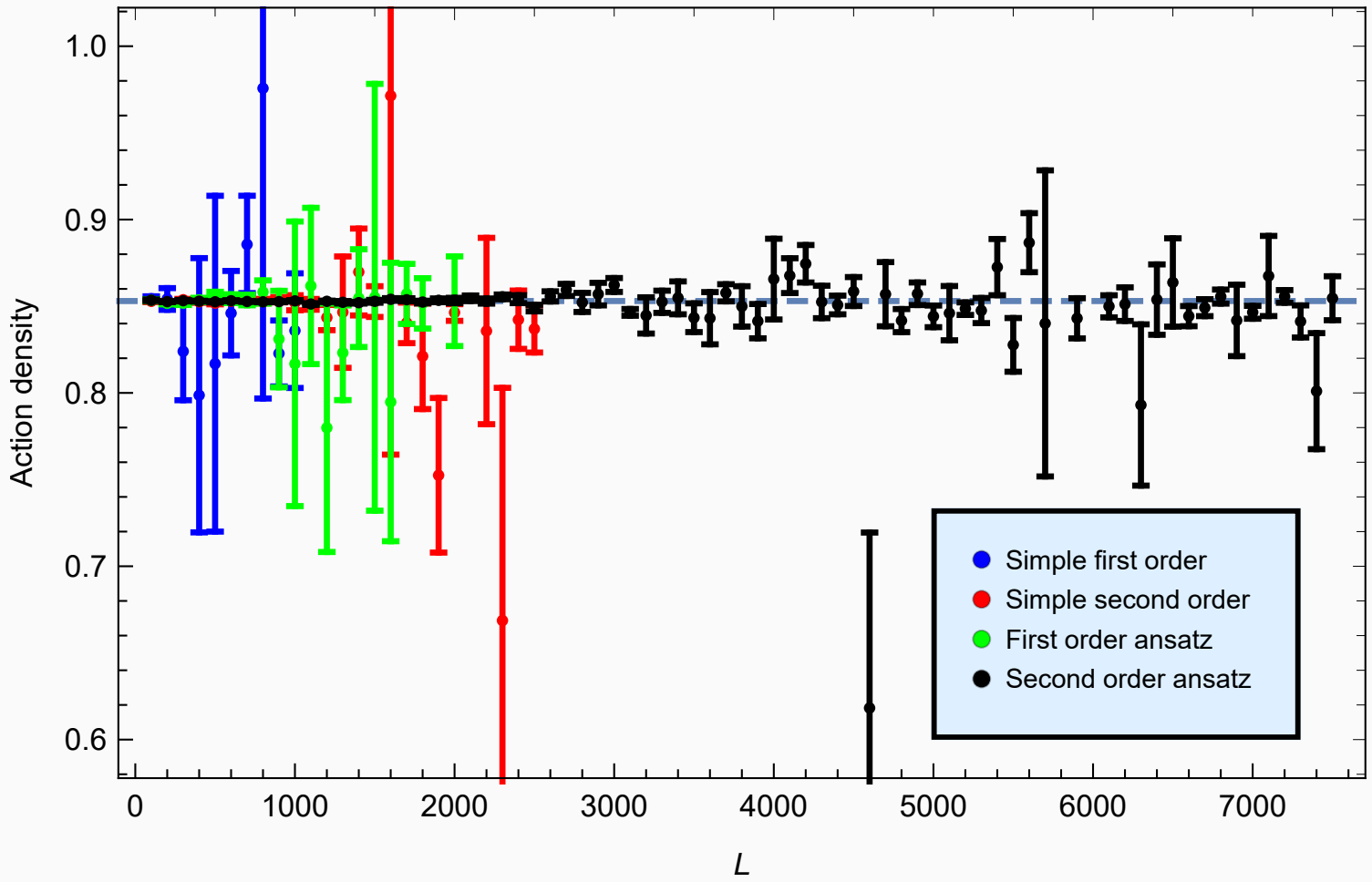
- Expansion around infinite mass.
- Expansion around infinite d.
- Expansion in order of neighbours.

THANK YOU

Action Density as a Function of Lattice Size

The action density, $\frac{\langle S \rangle}{L}$, as a function of lattice size L

Various contours with special point for $\lambda = 1$, $m = 1$, $\mu = 1$



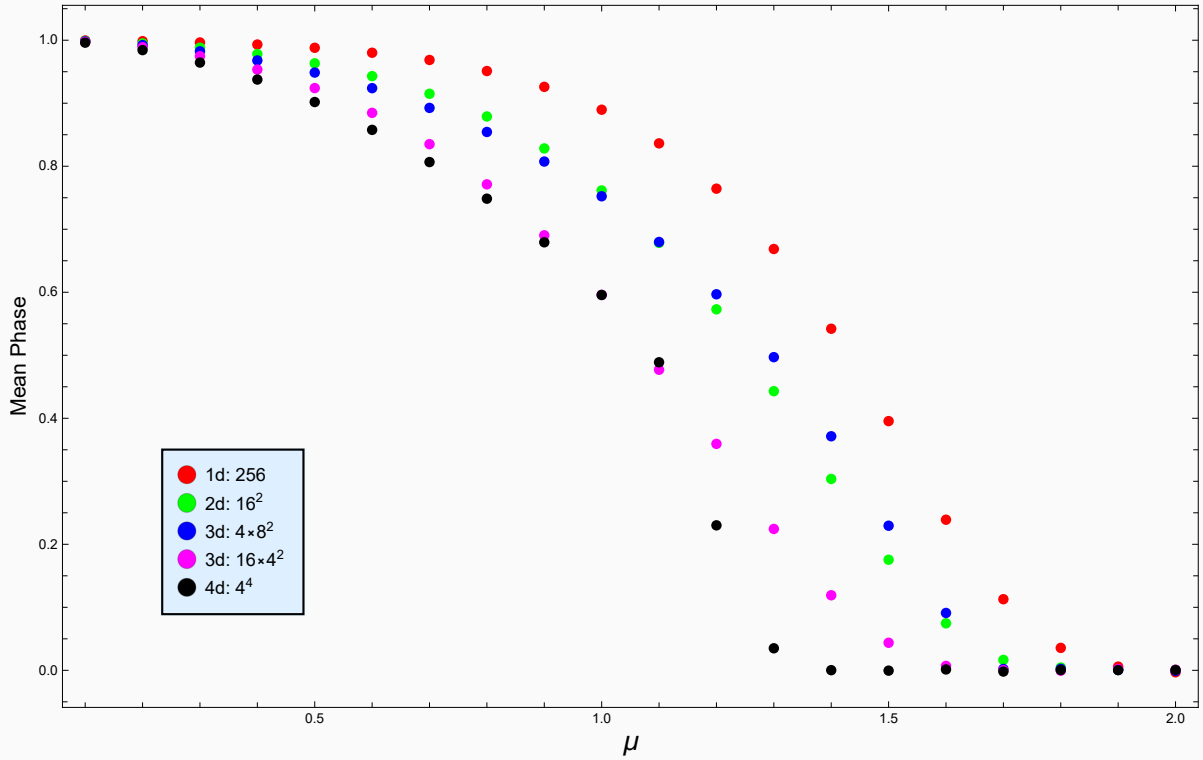
Comparing Different Dimensions – Fixed α

Simple 1^{st} order contours, 256 lattice sites, $\lambda = 1$

Fixed value of $\alpha = \frac{1}{9}$ ($m = 1$ for $d = 4$, $m = \sqrt{7}$ for $d = 1$)

The mean phase factor is larger for smaller d

Results depend also on the geometry



Comparing Different Dimensions – Fixed m

Simple 1^{st} order contours, 256 lattice sites, $\lambda = 1$

Fixed value of $m = 1$

The mean phase factor is smaller for smaller d