An algorithm for thimble regularization of lattice field theories and possibly (not?) only for that

F. Di Renzo University of Parma and INFN

in collaboration with G. Eruzzi and M. Brambilla

and M. Cristoforetti, A. Mukherjee, C. Schmidt, L. Scorzato, C. Torrero

LATTICE 2014

Columbia University, New York, 27-06-2014

Thimble regularization of field theories (M. Cristoforetti, F. Di Renzo, L. Scorzato - Phys.Rev.D86 2012) is still a fairly new attempt at the solution of the infamous sign problem.

While it is conceptually simple and elegant, it can be highly non trivial when it comes to algorithmic issues. Staying on the thimble is the relevant issue. I will present a few (almost trivial) results on a very basic model that are our starting point in the quest of a new algorithm, having in mind that we can maybe learn something useful for a broader field of applications. We are still quite far from success...

Agenda

- What is all this thimble story about (a quick primer)
- A few algorithmic solutions that are already available
- An interesting success for a toy model: ideal sampling on the thimble
- Can we step forward from that?
- Conclusions, perspectives, speculations

What is all this thimble story about (a quick primer)

A Summa of fundamental results from Morse Theory

$$dx^{n}g(x)e^{f(x)} dx^{n}g(x)e^{f(x)} dx^{n}g(x)e^{f(x)}$$

 $\Im t$

Everything comes with the right (real) dimension!

Thimbles in practice ...

$$S(\phi) = \frac{1}{2}\sigma\phi^2 + \frac{1}{4}\lambda\phi^4 \quad \text{with } \phi \in \mathbb{R}, \ \lambda \in \mathbb{R}^+ \text{ and } \sigma = \sigma_R + i \sigma_I \in \mathbb{C}.$$
$$\langle \phi^n \rangle = \frac{1}{Z} \int_{\mathbb{R}} \mathrm{d}\phi \ \phi^n \ e^{-S(\phi)} \quad Z = \int_{\mathbb{R}} \mathrm{d}\phi \ e^{-S(\phi)}$$

This is a prototypal toy model for a field theory displaying a sign problem. Since the (quite old) paper by J. Ambjorn (Phys.Lett.B 1985) it has been actracting attention.



A rich scenario ...



A continuity argument tells you the correct combinations of weights!

$$Z\left[\sigma_R \to 0^+\right] = Z\left[\sigma_R \to 0^-\right] = Z\left[\sigma_R = 0\right]$$

G. Eruzzi's talk on Tue



A number of issues!

- How many thimbles contribute? What are the weights? ... but for field theories ...
- It could be that we can take into account one single thimble
- Why? Thermodynamic limit plus universality

(same degrees of freedom, symmetries and symmetry representations,

PT, naive continuum limit as the original formulation...)



computable as in arXiv 1403.5637, accepted on PRD

and thus can be taken into account by reweighting.

In the following we will argue as if only one thimble should be taken into account (e.g. $\sigma_R > 0$ in our toy model) and will also (in the end) assume that the residual phase should be accounted for by reweighting once we have a suitable stochastic process on the thimble.



A few algorithmic solutions that we already know



On the thimble by very definition!

Noise should be extracted on the thimble!

But in our toy model we know very well the relevant direction: only 1 dimension, and thus the tangent space amounts simply to the only direction you always know, i.e. that of the gradient of the action!



Algorithm 1 A deneralization of angevin D_x , Rev. D86 2012 Langevin is the natural cardidate! $Z_0 \int_{\mathcal{J}_0}^{e^-iS_I} \int_{\mathcal{J}_0}^{e^-iS_I} a_x$



On the thimble by very definition!

Noise should be tangent to the thimble!



 $\Leftrightarrow |\partial S_R, \eta| = 0$

But since at the critical point we know the tangent space, the problem is that of trasporting a vector along our gradient flow

$$\mathcal{L}_{\partial S_R}(\eta) = 0 \qquad \Leftrightarrow [\partial S_R, \eta] = 0$$

$$0 = [\partial S_R, \eta(\tau)]_k = \sum_j \partial_j S_R \partial_j \eta_k(\tau) - \sum_j \eta_j(\tau) \partial_j \partial_j$$

Algorithm 1 How accurate must one be with Langevin?



In principle everything is there, in particular $\begin{aligned} \mathcal{L}_{\partial S_R}(\eta) &= 0 & \Leftrightarrow [\partial S_R, \eta] = 0 \\ \mathcal{L}_{\partial S_R}(\eta) &= 0 & \Leftrightarrow [\partial S_R, \eta] = 0 \end{aligned}$ $0 = [\partial S_R, \eta(\tau)]_k = \sum_j \partial_j S_R \partial_j \eta_k(\tau) - \sum_j \eta_j(\tau) \partial_j \partial_k S_R \\ \Leftrightarrow \overbrace{d\tau}^d \eta_j(\tau) = \sum_k \eta_k(\tau) \partial_k \partial_j S_R, \Biggr\}$

 $\Leftrightarrow [\partial S_R, \eta] = 0$



A question relevant in practice is: how much shall we go down towards the critical point?

And of course the other issue is finite integration step!

A different approach: HMC! See Fujii et al JHEP 1310 2013 Algorithm 1bis A very crude approach ... Gaussian approximation



- · · · · · · ·

Gaussian approximation: there are cases in which it works!





The Bose gas on the thimble (Cristoforetti, Di Renzo, Mukherjee, Scorzato Phys. Rev. D 88 2013)



-

Gaussian approximation: there are cases in which it works!



Gaussian approxim.: there are cases in which it does not work!

$$Z_N^{N_f}(m) = \int d\Phi d\Psi \det^{N_f} \left(D(\mu) + m \right) \exp\left(-N \cdot \operatorname{Tr}[\Psi^{\dagger}\Psi + \Phi^{\dagger}\Phi] \right)$$

$$D(\mu) + m = \begin{pmatrix} m & i\cosh(\mu)\Phi + \sinh(\mu)\Psi\\ i\cosh(\mu)\Phi^{\dagger} + \sinh(\mu)\Psi^{\dagger} & m \end{pmatrix}$$

See A. Mollgaard, K. Splittorff, Phys. Rev D88 2013



-

Algorithm 2 Metropolis

Originally proposed for U1 one-plaquette model (Cristoforetti, Mukherjee, Scorzato PRD Rapid 88 2013)

$$S[\phi] = S[\phi_0] + S_G[\eta] + \mathcal{O}(|\eta|^3)$$

 $S_G = \frac{1}{2} \sum_k \lambda_k \eta_k^2$

$$\phi_i = \phi_i^0 + \sum_k w_{ki} \eta_k$$

The idea of a Metropolis stems from the fact that near the critical point the action is gaussian and so ...

 $Hw_k=\lambda_k\bar{w}_k$ - Consider the theory that is purely gaussian and in the proximity of the critical point is "undistinguishable" from the complete theory

- Extract a point for that theory and ... $\frac{\mathrm{d}\phi_i(r)}{\mathrm{d}\phi_i(r)} = \frac{1}{\delta S}$ - ... take a SD towards the $\overline{r} \overline{sritical}$ point down to the region where the two theories are equivalent and ... - ... take a SA from that point for the complete theory $\phi_i(n+1) = \phi_i(n) + \delta r \overline{\frac{\delta S}{\delta \phi_i}}$

- Accept/reject with

$$P_{\rm acc} = \min\left\{1, e^{-\left[S^R(\phi') - S^R(\phi)\right] + \left[S_G(\eta') - S_G(\eta)\right]}\right\}$$

Algorithm 2 Metropolis

It works very well for our toy model (the algorithm has a technical parameter on which it depends!)





An interesting success for the toy model from which we try to step forward

Algorithm 3 Ideal sampling on the thimble!

Notice that for a 0-dim toy model an algorithm can be implemented that performs ideal sampling

The idea is simple: on each SD curve there is a 1-1 correspondence configurationaction, i.e. on a single SD if you extract a value for the action you extract a configuration ... and here we have only 2 SD curves!

Leaving out the residual phase that can be accounted for by reweighting ...



Let's try to step forward ...

$$Z = \int_{\mathcal{J}_{\sigma}} dz_1 \wedge \ldots \wedge z_n \, \mathrm{e}^{-S} = \sum_{\mathrm{charts} \, c} \int_{\Gamma_c} \prod_i^n dy_i^c \, \det(J) \, \mathrm{e}^{-S} = \sum_{\mathrm{charts} \, c} \int_{\Gamma_c} \prod_i^n dy_i^c \, \mathrm{e}^{i\Phi} \, \mathrm{e}^{-S}$$

Let's now discharge both charts and residual phase (we can go back to them ...)

$$Z = \int_{\Gamma} \prod_{i}^{n} dy_{i} e^{-S} = \int \prod_{i}^{n} dn_{i} \,\delta(|\vec{n}|^{2} - 1) \, Z_{\hat{n}} = \int \prod_{i}^{n} dn_{i} \,\delta(|\vec{n}|^{2} - 1) \int dt \,\Delta_{\hat{n}}(t) e^{-S(\hat{n},t)}$$

where a point on the thimble is singled out by giving the direction along which one leaves the critical point and the time one integrates Steepest Ascent for: (\hat{n}, t) .

Now a new probability is defined that we can (in principle) exactly sample

$$P_{\hat{n}}(t) = Z_{\hat{n}}^{-1} \Delta_{\hat{n}}(t) e^{-S(\hat{n},t)} \to F_{\hat{n}}(t) = Z_{\hat{n}}^{-1} \int_{-\infty}^{t} dt' \Delta_{\hat{n}}(t') e^{-S(\hat{n},t')} \to F_{\hat{n}}^{-1}(\xi)$$

But let's pause for a moment:

- I have not yet told you how we got the previous expressions;
- One could suspect all this is not at all trivial!
- Notice that the y coordinates are known once you transport a basis.

There is a tremendous amount of information in ...

... what we wrote down, which basically looks like a Faddeev-Popov trick to stay on the thimble. As a matter of fact we have inserted 1, written in an integral form ...

$$1 = \Delta_{\hat{n}}(t) \int \prod_{k}^{j} dn_{j} \,\delta(|\vec{n}|^{2} - 1) \int dt \,\prod_{i}^{n} \delta(y_{i} - y_{i}(\hat{n}, t))$$

The determinat is telling you how the action and the manifod itself are sensitive to variations with respect to directions and evolution time.

... but we can try to mimic the previous formula ...

... if we simply want to regard our sampling along a single ascent as a proposal ...

$$\tilde{P}_{\hat{n}}(t) = \tilde{Z}_{\hat{n}}^{-1} \,\tilde{\Delta}_{\hat{n}}(t) \,\mathrm{e}^{-S(\hat{n},t)} \to \tilde{F}_{\hat{n}}(t) = \tilde{Z}_{\hat{n}}^{-1} \int_{-\infty}^{t} dt' \,\tilde{\Delta}_{\hat{n}}(t') \,\mathrm{e}^{-S(\hat{n},t')} \to \tilde{F}_{\hat{n}}^{-1}(\xi)$$

... and a new version is ...

- Pick up a direction and a point (\hat{n},t) via $\,t=\tilde{F}_{\hat{n}}^{-1}(\xi)\,$

- Accept an evolution step $(\hat{n},t)
ightarrow (\hat{n}',t')$ with probability

$$\min\{1, \frac{\mathrm{e}^{-S(\hat{n}',t')}}{\mathrm{e}^{-S(\hat{n},t)}} \frac{\tilde{Z}_{\hat{n}}^{-1} \tilde{\Delta}_{\hat{n}}(t) \,\mathrm{e}^{-S(\hat{n},t)}}{\tilde{Z}_{\hat{n}'}^{-1} \tilde{\Delta}_{\hat{n}'}(t') \,\mathrm{e}^{-S(\hat{n}',t')}}\}$$

Conclusions, perspective, speculations ...

- Staying on a thimble can be an algorithmic challenge. We are trying to develop a new algorithm which tries to step forward from the ideal sampling which is viable in the case of O-dim toy models.

- Basically the main virtue we are looking for is: sampling configurations doing the thing which is under very good control in order to stay on the thimble, i.e. solving for the steepest ascent paths that define the thimble.

- We are brave enough to go for the "real thing" for fairly simple models (matrix models).

- Variants have to be devised in order to allievate the tremendous computational effort which is needed for ideal sampling. We have some ideas. These are under study; (if we succeed...) they would be good stochastic processes on the thimble. The quest is a close relative to that for the density of states.

- If you think about a little bit, it could even be (or maybe we would like to think ...) that there is a chance to extend the method to non-thimble applications. After all, over there we do know what the correct manifolds are.