Ground State of an adjoint SU(2) gauge theory from Neural Newtork assisted Variational Montecarlo

David Ward[1]

[1]Department of Physics
National Taiwan University

July 29, 2026

Table of Contents

1. Introduction
2. Hamiltonian Formalism and adjoint $SU(2)$
   - adjoint gauge-Higgs Hamiltonian
   - adjoint pure gauge Hamiltonian
3. Neural Wavefunctions and Ansatze
4. Simulation Results
   - 2+1d CNN Comparison
   - 3+1d Jastrow
5. Conclusion

Motivation

- Hamiltonian formalism offers us several attractive features: real time dynamics, evades sign problems, work directly with operators through the wavefunction; however, the vacuum groundstate and exact forms for the wavefuntion are sometimes difficult to determine.
- Nevertheless we have been able to make some progress, previous work has been able to determine the eigenvalue spectrum for a pure gauge theory using various montecarlo techniques.
  [Greensite – Nucl.Phys.B 158 (1979)]
  [Chin et al. – PhysRevD.31.3201 (1985)]
  [Greensite – Phys.Lett.B 191 (1987)]
  [Greensite – arXiv:1310.6706 (2013)]
- Machine learning offers us the ability implement neural wavefunctions which can make use of wavefunction ansatze to capture low lying eigenvalues in the energy spectrum for a given theory.

Motivation cont.

- Recent work by Spriggs et al. implemented a Lattice Gauge Neural Wavefunction using an equivariant approach to the gauge field. [arXiv:2509.12323]
- The results of this work have been able to capture low lying energies and thus shows promise for studying the vacuum groundstate.
- Previous work focused on the adjoint representation of the $SU(2)$ gauge-Higgs has shown a rich phase structure allowing us to study several confining properties of the vacuum.
- To this end, we simulated the adjoint gauge field in $SU(2)$ as part of a larger project on the adjoint gauge-Higgs vacuum wavefunction.

Why adjoint $SU(2)$?

- Adjoint $SU(2)$ has an explicit Z_2 symmetry which allows for a more rich study of the phase diagram of the gauge-Higgs theory.
  [Catumba et al. – PhysRevResearch.6.043172 (2024)]
  [Sarkar – arXiv:2210.09855 (2022)]
- These works in $D = 2 + 1$ and $D = 3 + 1$ have considered measurements of the Polyakov loop, plaquettes and typical probes for the confinement transition.
- These works did not make measurements of the mass for monopoles or look at center vortex operators; in addition to looking at these operators across phases we would like to understand the real time dynamics of these operators.

Hamiltonian

Our total target Hamiltonian splits into three sectors:

$$\mathcal{H} = \beta_E \mathcal{H}_G + \gamma_E \mathcal{H}_H + \mathcal{H}_C.$$

Making the explicit choice to fix to Coulomb gauge we can write the Hamiltonian in the following way,

$$\mathcal{H} = \sum_x \frac{1}{2}\mathrm{Tr}[E_L^i(x)E_L^i(x)] + \frac{\beta}{2}\sum_{ij} 1 - \mathrm{Tr}_A[U_i(x)U_j(x+i)U_i^\dagger(x+j)U_j^\dagger(x)]$$
$$+ \frac{1}{2}\pi_H\pi_H^\dagger + \gamma\sum_i \mathrm{Tr}[\varphi^\dagger(x)U_i(x)\varphi(x+i)] + V_{\mu,\lambda}(\varphi).$$

The last term H_C can be dropped in favor of an explicit Gauss' Law condition applied to each site on the lattice:

$$D_i E_L^{ia}(x) = q^a(x).$$

Adjoint SU(2) Pure Gauge

The electric field operator can be parameterized in terms of motion through the SU(2) manifold by way of the isomorphism with S_3,

$$\mathrm{Tr}[E_L^i(x)E_L^i(x)] \to -2\nabla_A^2.$$

The kinetic terms includes a normalization with respect to the metric coefficents from the projection to S_3 for the adjoint representation of the gauge field,

$$\nabla_A^2 \to \frac{1}{2}\nabla_{S_3}^2.$$

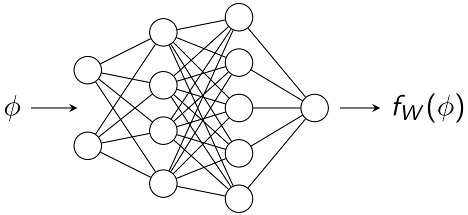
The gauge field sector now may be fully described in the basis constructed from links,

$$\mathcal{H}_G = \sum_x -\nabla_{S_3}^2 + \beta \sum_{ij} 1 - \mathrm{Tr}_A[U_i(x)U_j(x+i)U_i^\dagger(x+j)U_j^\dagger(x)],$$

$$D_i E_L^{ia}(x) = 0.$$

Neural Wavefunction

We can train a neural network to study gauge field configurations and assign weights to regions of the Hilbert space for a theory.



This makes using neural networks a great candidate for a variational method approach using monte carlo.
For a generic wavefunction in the Hilbert space we assume,

$$|\Psi\rangle = \Phi\,|\Psi_0\rangle\,,$$

where Φ is some fluctuation operator.

Neural Network and Variational Montecarlo

If we can create a parameterized form of the wavefunction through an ansatz e.g. Jastrow,

$$|\Psi\rangle = \sum_i f_\alpha(\psi_i) |\psi_i\rangle ,$$

then projection onto the groundstate is also controlled by the parameter.

$$\langle\Psi_0|\Psi\rangle = \sum_i \langle\Psi_0|f_\alpha(\psi_i)|\psi_i\rangle = f_\alpha(\Psi_0).$$

On the lattice Ψ is built from lattice configurations

$$|\Psi_\alpha\rangle = \sum_i f_\alpha(U_i) |U_i\rangle ,$$

using this form of the wavefuntion we can extract a parameterized energy

$$E_\alpha = \langle\Psi_\alpha|\mathcal{H}|\Psi_\alpha\rangle ,$$

and using automatic differentiation of the energy we can adjust the parameter to minimize the energy of the trial wavefunction.

Two Point Jastrow and Neural Wavefunction Ansatzae

In our simulations we will use two types of ansatze, the first being a modified Jastrow containing a two point correlation:

$$\Psi(x) = \prod_{x} e^{\alpha_\mu W_{\mu,\nu}(x)} \prod_{x'} e^{\beta^{\mu,\nu,\mu',\nu'}_{d(x,x')} W_{\mu',\nu'}(x) W_{\mu,\nu}(x')}.$$

Where

$$W_{\mu,\nu}(x) = \frac{1}{2}\mathrm{Tr}[P_{\mu,\nu}(x)],$$

and in the limit of $\beta \to 0$ we recover the typical Jastrow ansatz.
The second ansatz is the Lattice Gauge Neural Wavefunction(LGNWF), and is built from two layers:

- Plaquette Equivaraint
- Node Equivariant

The plaquette equivariant layer is constructed the weighted product of all plaquettes which contain the site x:

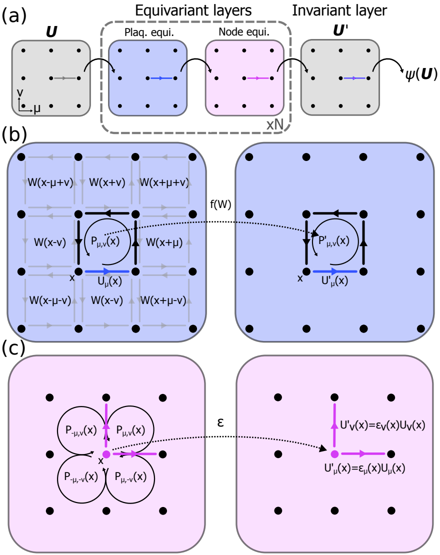
$$\epsilon(x) = e^{i\sum \gamma_i C^i_{\mu,\nu}(x)}$$

$$U_\mu(x) \to U'(x) = \epsilon(x) U_\mu(x)$$

The node equivariant layer is constructed from modification of the eigenvalues of the plaquette

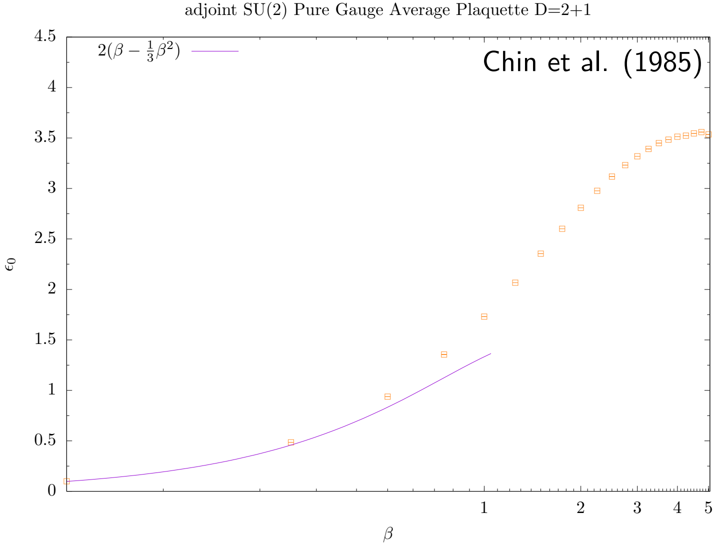
$$P_{\mu,\nu}(x) = V e^{i\lambda} V^\dagger \to$$
$$P'_{\mu,\nu}(x) = V f(\lambda, W) V^\dagger$$

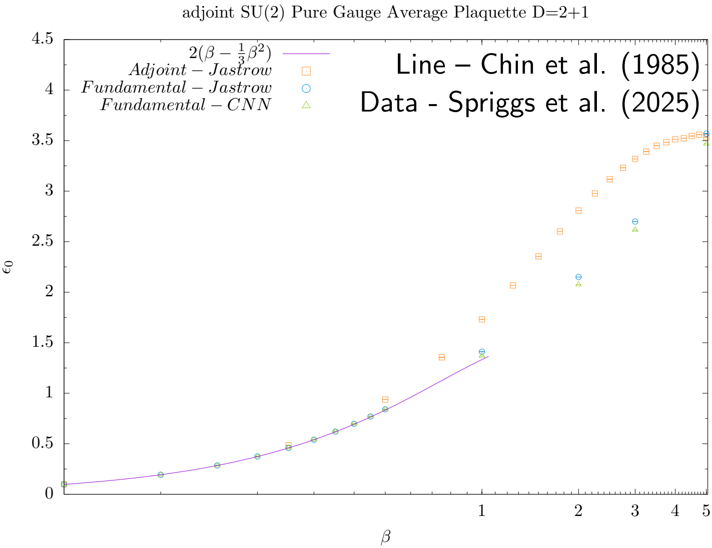
$$U_\mu(x) \to P'_{\mu,\nu}(x) P^\dagger_{\mu,\nu}(x) U_\mu(x)$$



Comparison of Ansatzae in $D = 2 + 1$

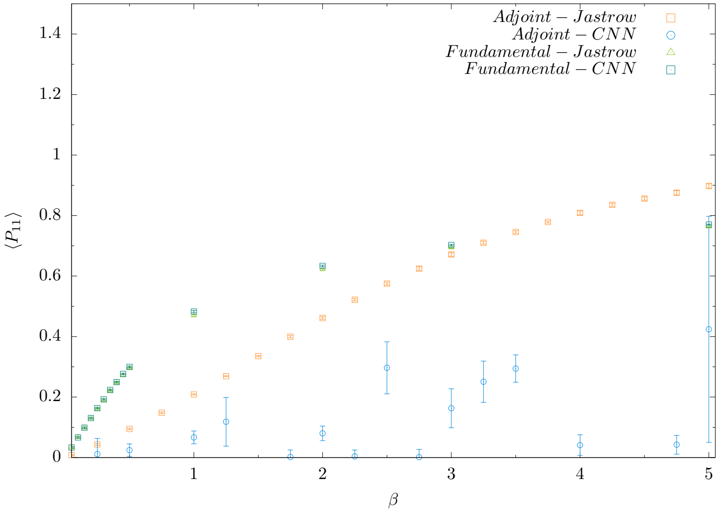
- For $D = 2 + 1$ we expect similar behaviors to the fundamental representation and so compared our results with those of Spriggs et al.
- We compared the groundstate energies, and plaquette behaviors between both the CNN and Jastrow ansatzae for our simulations.
- Due to a bug in our implementation of the Convolutional Neural Network part of the Equivariant Ansatz we were not able to compare the Jastrow and CNN in our simulation of the SU(2).


adjoint SU(2) Pure Gauge Average Plaquette D=2+1
Chin et al. (1985)
$2(\beta - \frac{1}{3}\beta^2)$
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
ϵ_0
1
2
3
4
5
β



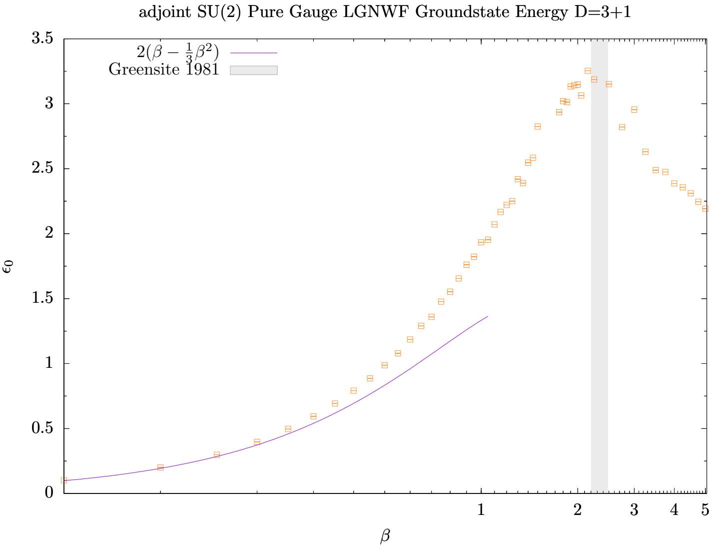
adjoint SU(2) Pure Gauge Average Plaquette D=2+1
$2(\beta - \frac{1}{3}\beta^2)$
$Adjoint - Jastrow$
$Fundamental - Jastrow$
$Fundamental - CNN$
Line – Chin et al. (1985)
Data - Spriggs et al. (2025)
ϵ_0
β
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
1
2
3
4
5

adjoint SU(2) Pure Gauge Average Plaquette D=2+1

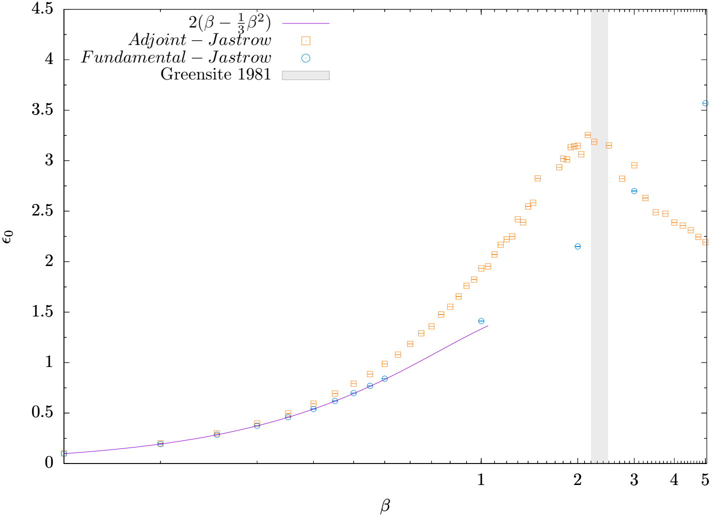


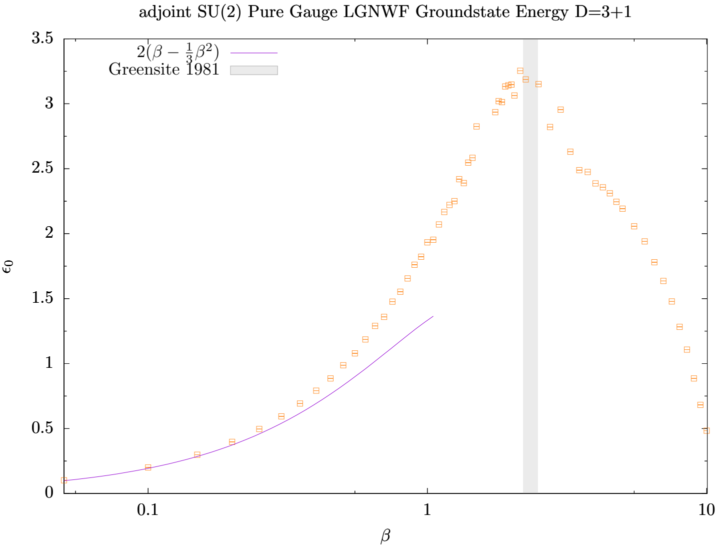
Adjoint $SU(2)$ in $D = 3 + 1$

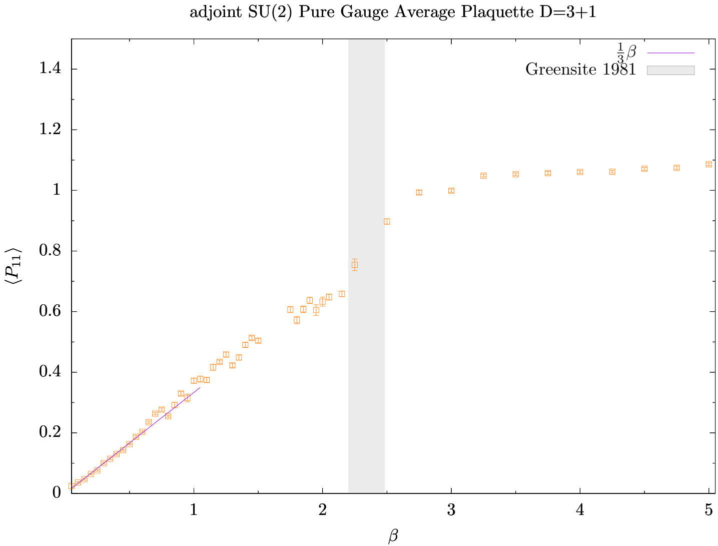
- In $D = 3 + 1$ we ran both CNN and Jastrow ansatzae but we only will present results from the Jastrow ansatz.
- Additionally, there is a well known 1st order phase transition for the $SO(3)$ pure gauge theory which occurs between $\beta \sim 2.2 - 2.48$. [Greensite – PhysRevLett.47.9 (1981)]
- This transition occurs due to breaking of the Z_2 center symmetry which is explicilty preserved in the adjoint representation.

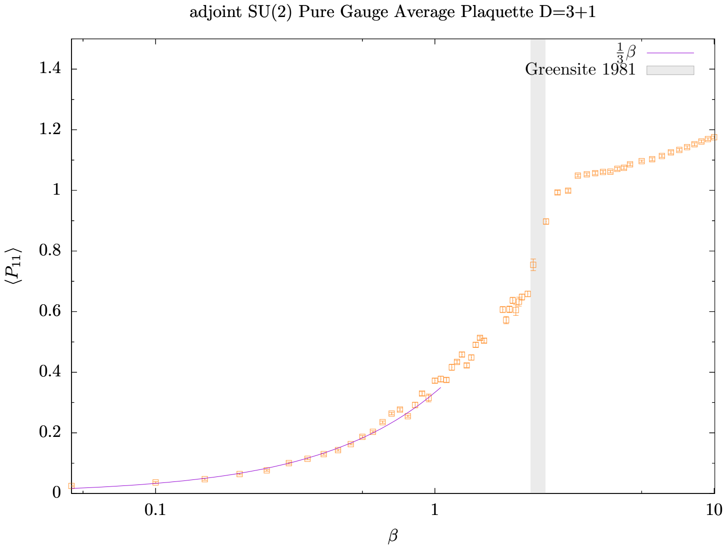

adjoint SU(2) Pure Gauge LGNWF Groundstate Energy D=3+1
$2(\beta - \frac{1}{3}\beta^2)$
Greensite 1981
ϵ_0
β
3.5
3
2.5
2
1.5
1
0.5
0
1
2
3
4
5

adjoint SU(2) Pure Gauge Average Plaquette D=3+1




adjoint SU(2) Pure Gauge LGNWF Groundstate Energy D=3+1
$2(\beta - \frac{1}{3}\beta^2)$
Greensite 1981
3.5
3
2.5
2
1.5
1
0.5
0
ϵ_0
0.1
1
10
β


adjoint SU(2) Pure Gauge Average Plaquette D=3+1
$\frac{1}{3}\beta$
Greensite 1981
1.4
1.2
1
0.8
0.6
0.4
0.2
0
$\langle P_{11} \rangle$
1
2
3
4
5
β

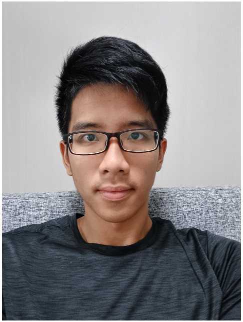

adjoint SU(2) Pure Gauge Average Plaquette D=3+1
$\frac{1}{3}\beta$
Greensite 1981
$\langle P_{11} \rangle$
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0.1
1
10
β

Conclusions

- Applying Neural Neworks as part of a variational technique has been shown to be an effective technique to capture low lying engenvalues in the pure gauge case.
- Extension of the theory to a full gauge-Higgs theory in $SU(2)$ is ongoing and looks promising, but will require increases in the efficiency of computing the combined gauge-scalar wavefunction.
- Our simulation of the adjoint $SU(2)$ theory for the pure gauge theory using the two body Jastrow was able to reproduce the first order phase transition visible from the plaquettes as well as the groundstate energies.

Acknowledgements

This work was done with support from NSTC of Taiwan under Grant No. 113-2112-M-002-033-MY3, 114-2811-M-002-169-MY3



- I would also like to acknowledge the contributions made to the project by my mentee Shen-Rong Liao

Thank you for your attention

- For scalar fields in the adjoint rep. we can parameterize the matrix by an auxiliary field function

$$\varphi(x) = e^{i\sigma^a \widetilde{\phi}^a(x)}.$$

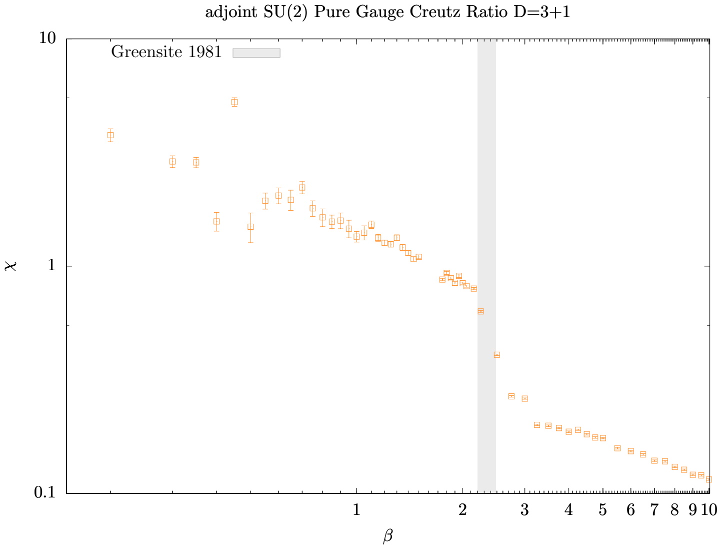
- The kinetic Higgs terms can, therefore, be parameterized in terms of motion through the SU(2) manifold by way of the isomorphism with S_3 akin to a gauge field in the fundamental representation,

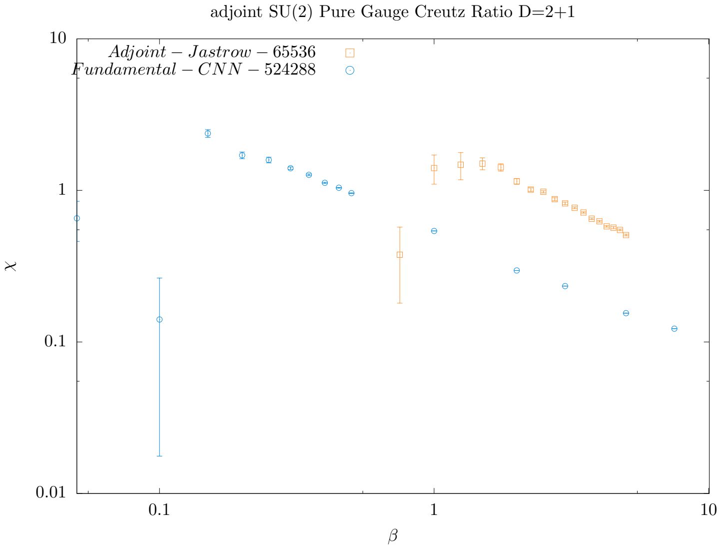
$$\pi_H \pi_H^\dagger(x) \to -4\nabla_F^2$$

- Normalization with respect to the metric coefficents for the fundamental representation projected to S_3 require,

$$\nabla_F^2 \to \frac{1}{4}\nabla_{S_3}^2.$$

$$\mathcal{H}_H = \sum_x \frac{1}{2}\nabla_{S_3}^2 + \gamma \sum_i \mathrm{Tr}[\varphi^\dagger(x) U_i(x) \varphi(x+i)] + V_{\mu,\lambda}(\varphi).$$


adjoint SU(2) Pure Gauge Creutz Ratio D=3+1
Greensite 1981
10
1
0.1
χ
1
2
3
4
5
6
7
8
9
10
β


adjoint SU(2) Pure Gauge Creutz Ratio D=2+1
$Adjoint - Jastrow - 65536$
$Fundamental - CNN - 524288$
10
1
0.1
0.01
0.1
1
10
χ
β