Variational Monte Carlo with Proximal Policy Optimization

Requirement

```
pyTorch >= 1.8.0 + cu111
```

Neural-network quantum state (NQS)

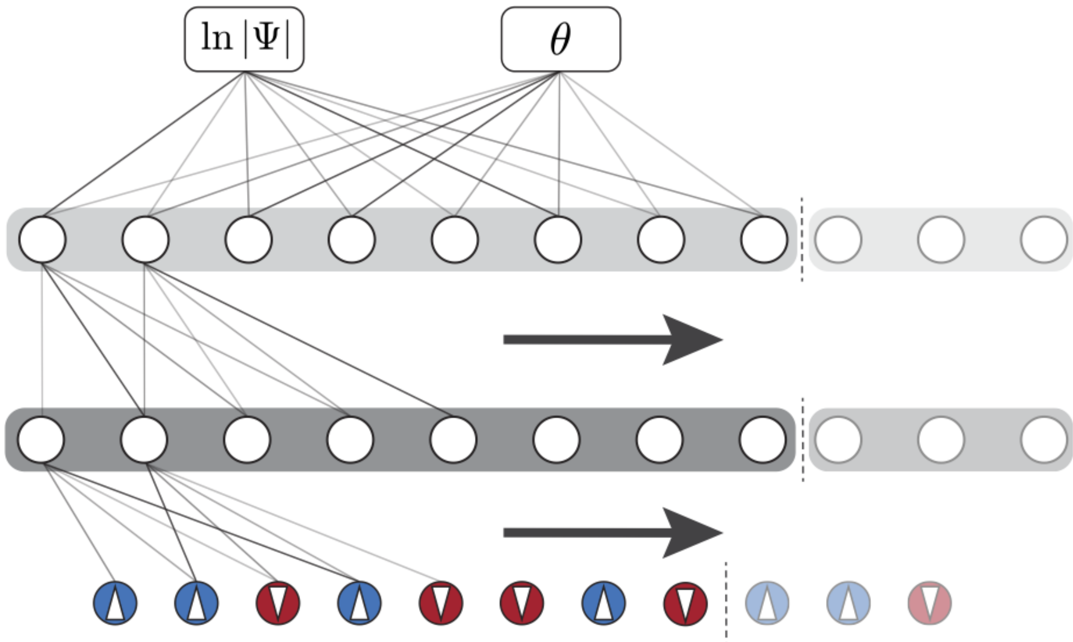
```
algos.core
```

Two types of convolution neural network are implemented as NQS.

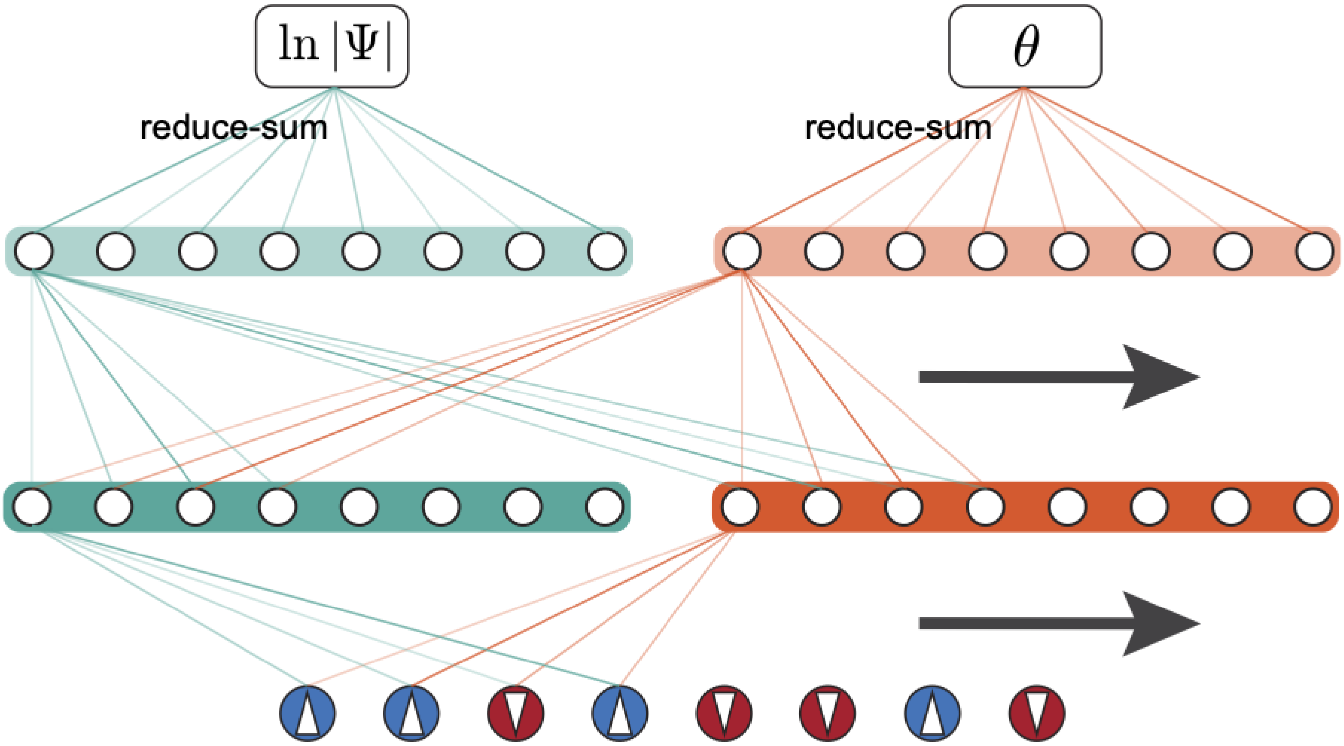
Inputs: spin configurations.

Outputs:
$\log|\Psi| \text{ and } \theta$

- pesudocomplex CNN with only real parameters.



- complex CNN contains two sublayers in each layer, one stands for real part and the other denotes the imag part of the complex layer.



Optimization: Natural gradient descent

Optimization object,

$$\text{minimize}_w E_w(\Psi(s;w)) \quad \text{s.t.} D_{FS}^2(\Psi_{\text{old}}(s;w_{\text{old}}), \Psi(s;w)) \leq \delta.$$

where
E_w
is the energy estimated via importance sampling and
D_{FS}
is the Fubini-Study distance.

PPO algorthim

```
algos.pesudocomplex_ppo or algos.complex_ppo
```

NGD can be approximately solved by PPO-clip and PPO-clip updates wavefunctions via,

$$w_{k+1} = \text{argmin}_w \mathbb{E}[L(w_k, w)]$$

with the loss function,

$$L(w_k, w) = \max\left(\frac{|\Psi_w|^2}{|\Psi_{w_k}|^2} E_{w \text{ or } w_k}, \text{clip}\left(\frac{|\Psi_w|^2}{|\Psi_{w_k}|^2}, 1-\epsilon, 1+\epsilon\right) E_{w \text{ or } w_k}\right),$$

where
E_w

is estimated by the current wavefunction
Ψ_w
and
E_{w_k}
is estimated by the old wavefunction
Ψ_{w_k}
.

Stochastic reconfiguration (TDVP)

```
algos.pesudocomplex_sr or algos.complex_sr
```

The optimization object of NGD is also equivalent to,

$$\text{minimize}_{\Delta w}\{E_w + \nabla_w E_w \Delta w\} \;\; \text{s.t.} \;\; \frac{1}{2}\Delta w^\dagger \mathbf{S} \Delta \omega < \delta,$$

due to,

$$D_{FS}^2 \approx \sum_{ij} dw_i^* dw_j [\langle \mathcal{O}_i^* \mathcal{O}_j \rangle - \langle \mathcal{O}_i^* \rangle \langle \mathcal{O}_j \rangle],$$

where
$S_{ij} = [\langle \mathcal{O}_i^* \mathcal{O}_j \rangle - \langle \mathcal{O}_i^* \rangle \langle \mathcal{O}_j \rangle]$
is the stochastic reconfiguration matrix and
$\mathcal{O}_i = \partial_w \log \Psi$
.
Such a contitional minimal problem is hence equivalent to:

$$\text{minimize}_{\Delta w}\{E_w + \nabla_w E_w + \lambda(\frac{1}{2}\Delta w^\dagger \mathbf{S} \Delta w - \epsilon)\},$$

with a Lagrange multiplier
λ
. Its minimal satisfies,

$$\mathbf{S}\Delta w = -\alpha \nabla_w E_w,$$

which is the exact TDVP equation.