Stochastic Reconfiguration

June 19, 2023

The trial wave function $|\Psi_T\rangle$ is parametrized in terms of a set of variational parameters $\mathbf{p} = \{p_1, \ldots, p_n\}$. Their optimal values are found minimizing the energy expectation value

$$E_T = \frac{\langle\Psi_T|H|\Psi_T\rangle}{\langle\Psi_T|\Psi_T\rangle} \geq E_0\,, \tag{1}$$

where E_0 is the true ground-state of the system. Standard steepest descent methods are proven to be highly inefficient when the Metropolis algorithm is employed to evaluate E_T and exhibit an impractically slow convergence. An alternative algorithm has been devised by noting that iterating the equation

$$|\Psi(\tau+\delta\tau)\rangle = (1-\mathcal{H}\delta\tau)|\Psi(\tau)\rangle \tag{2}$$

would converge to the ground-state of the system. We can take $|\Psi(\tau)\rangle = |\Psi_{p_0}\rangle$ where $|\Psi_{p_0}\rangle = |\Psi_T(\mathbf{p}=\mathbf{p}_0)\rangle$ is the trial wave function evaluated for a given set of parameters $\mathbf{p}_0$. We then approximate $|\Psi(\tau+\delta\tau)\rangle$ with a linear combination of the original wave function and its derivatives with respect to the variational parameters

$$(1-\mathcal{H}\delta\tau)|\Psi_{p_0}\rangle = \Delta p_0|\Psi_{p_0}\rangle + \sum_n \Delta p_n \mathcal{O}^n|\Psi_{p_0}\rangle\,, \tag{3}$$

where we defined

$$\mathcal{O}^n|\Psi_{p_0}\rangle = \frac{\partial}{\partial p_n}|\Psi_T\rangle\Big|_{\mathbf{p}=\mathbf{p_0}} \tag{4}$$

To obtain the coefficients Δp_0 and Δp_n it is convenient to multiply Eq. (**??**) by $\langle\Psi_{p_0}|\mathcal{O}^m$ and divide by $\langle\Psi_{p_0}|\Psi_{p_0}\rangle$

$$\frac{\langle\Psi_{p_0}|\mathcal{O}^m(1-\mathcal{H}\delta\tau)|\Psi_{p_0}\rangle}{\langle\Psi_{p_0}|\Psi_{p_0}\rangle} = \Delta p_0\frac{\langle\Psi_{p_0}|\mathcal{O}^m|\Psi_{p_0}\rangle}{\langle\Psi_{p_0}|\Psi_{p_0}\rangle} + \sum_n \Delta p_n\frac{\langle\Psi_{p_0}|\mathcal{O}^m\mathcal{O}^n|\Psi_{p_0}\rangle}{\langle\Psi_{p_0}|\Psi_{p_0}\rangle} \tag{5}$$

We separately consider the cases corresponding to $m = 0$ and $m \neq 0$ and introduce the notation $\langle \mathcal{O} \rangle = \langle \Psi_{p_0} | \mathcal{O} | \Psi_{p_0} \rangle / \langle \Psi_{p_0} | \Psi_{p_0} \rangle$:

$$\langle 1 - \mathcal{H}\delta\tau \rangle = \Delta p_0 + \sum_n \Delta p_n \langle \mathcal{O}^n \rangle \qquad m = 0 \tag{6}$$

$$\langle \mathcal{O}^m \rangle - \langle \mathcal{O}^m \mathcal{H} \rangle \delta\tau = \Delta p_0 \langle \mathcal{O}^m \rangle + \sum_n \Delta p_n \langle \mathcal{O}^m \mathcal{O}^n \rangle \qquad m \neq 0 \tag{7}$$

We isolate Δp_0 from the first line and substitute its expression into in the second one

$$\Delta p_0 = 1 - \langle \mathcal{H} \rangle \delta\tau - \sum_n \Delta p_n \langle \mathcal{O}^n \rangle$$

$$[\langle \mathcal{H} \rangle \langle \mathcal{O}^m \rangle - \langle \mathcal{O}^m \mathcal{H} \rangle] \delta\tau = \sum_n \Delta p_n \Big[\langle \mathcal{O}^m \mathcal{O}^n \rangle - \langle \mathcal{O}^m \rangle \langle \mathcal{O}^n \rangle \Big] \tag{8}$$

Introducing

$$\mathcal{S}^{mn} = \langle \mathcal{O}^m \mathcal{O}^n \rangle - \langle \mathcal{O}^m \rangle \langle \mathcal{O}^n \rangle \tag{9}$$

$$\mathcal{F}^m = [\langle \mathcal{H} \rangle \langle \mathcal{O}^m \rangle - \langle \mathcal{O}^m \mathcal{H} \rangle] \delta\tau \tag{10}$$

the above equation can be cast in the form

$$\sum \Delta p_n \mathcal{S}^{mn} = \mathcal{F}^m \tag{11}$$

We usually employ the Cholesky decomposition algorithm to solve the above equation for Δp_n and use them to obtain Δp_0 from the first line of Eq. (**??**). Finally, we rescale Δp_n as

$$\widetilde{\Delta} p_n = \frac{\Delta p_n}{\Delta p_0} \tag{12}$$

so that, using the fact that the wave-function normalization is immaterial, we can identify $(1 - \mathcal{H}\delta\tau)|\Psi_{p_0}\rangle$ with the Taylor expansion of the trial wave function evaluated at $\mathbf{p}_0 + \widetilde{\Delta}\mathbf{p}$.

1 Checks on linear order expansion with the overlap

Another test on the convergence of the algorithm can be carried out considering the normalized overlap square

$$\frac{|\langle \Psi_{p+\Delta p} | \Psi_p \rangle|^2}{\langle \Psi_{p+\Delta p} | \Psi_{p+\Delta p} \rangle \langle \Psi_p | \Psi_p \rangle} = 1 + \mathcal{O}(\Delta p^2) \tag{13}$$

that can be rewritten as follows

$$\begin{aligned} \frac{|\langle \Psi_{p+\Delta p} | \Psi_p \rangle|^2}{\langle \Psi_{p+\Delta p} | \Psi_{p+\Delta p} \rangle \langle \Psi_p | \Psi_p \rangle} &= \frac{\int dx \Psi_{p+\Delta p}(x) \Psi_p(x)}{\int dx |\Psi_{p+\Delta p}(x)|^2} \times \frac{\int dx \Psi_{p+\Delta p}(x) \Psi_p(x)}{\int dx |\Psi_p(x)|^2} \\ &= \frac{\int dx |\Psi_p(x)|^2 \frac{\Psi_{p+\Delta p}(x)}{\Psi_p(x)}}{\int dx |\Psi_p(x)|^2 \frac{|\Psi_{p+\Delta p}(x)|^2}{|\Psi_p(x)|^2}} \times \frac{\int dx |\Psi_p(x)|^2 \frac{\Psi_{p+\Delta p}(x)}{\Psi_p(x)}}{\int dx |\Psi_p(x)|^2} \,. \end{aligned} \tag{14}$$

The integrals appearing in last line can be estimated sampling configurations from $|\Psi_p(x)|^2$, using the appropriate re-weighting factors.