Path Integral-Enabled Methods within the Stochastic Representation of Wavefunctions

Liam Bernheimer

THE COHEN GROUP

School of Chemistry | Faculty of Exact Sciences

Tel Aviv University

TREX Symposium on Bridging Quantum Monte Carlo and High-Performance Simulations

Esch-sur-Alzette, Luxembourg

February 2024

جامعة تل ابيب

TEL AVIV UNIVERSITY אוניברסיטת תל אביב

Outline

Outline

The problem

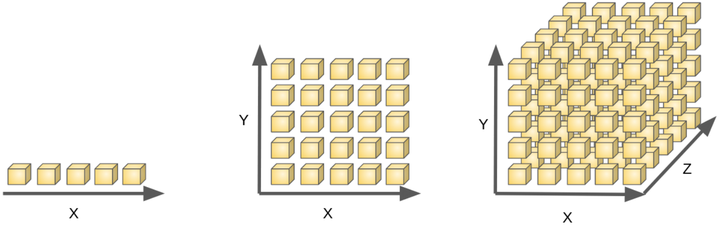
- Finding the ground state of many-body quantum systems is hard.

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V\left(\mathrm{R}\right)\right]\psi_n\left(\mathrm{R}\right) = E_n\psi_n\left(\mathrm{R}\right)$$

- Analytical solutions for systems beyond the Hydrogen atom are rare.

Curse of dimensionality

Straight forward numerically accurate approximations suffer from impractical scaling.



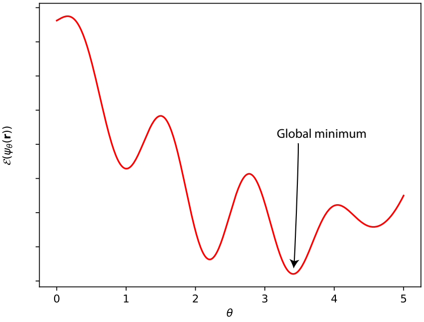
https://www.i2tutorials.com/what-do-you-mean-by-curse-of-dimensionality-what-are-the-different-ways-to-deal-with-it/

Convoluted approximations are required!

The variational principle

We can find for the ground state by searching for the global minimum of the energy. Lower is always better.

$$\mathcal{E}\left(\psi_{\boldsymbol{\theta}}(\mathrm{R})\right) = \frac{\left\langle \psi_{\boldsymbol{\theta}} \middle| \hat{H} \middle| \psi_{\boldsymbol{\theta}} \right\rangle}{\left\langle \psi_{\boldsymbol{\theta}} \middle| \psi_{\boldsymbol{\theta}} \right\rangle} = \sum_n \left| c_n \right|^2 E_n \geq E_0$$

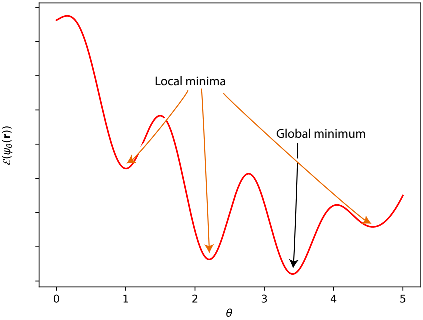


Optimization

Global optimization is also hard. Derivatives with respect to all parameters and spatial coordinates are needed.

$$\nabla^2 \psi_{\boldsymbol{\theta}}(\mathsf{R})$$

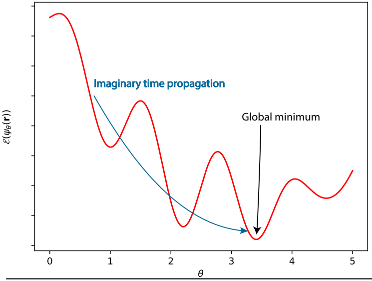
$$\boldsymbol{\theta} \to \boldsymbol{\theta} - \eta \nabla_{\boldsymbol{\theta}} \mathcal{E}\left(\psi_{\boldsymbol{\theta}}(\mathsf{R})\right)$$



Imaginary Time Propagation / Stochastic Reconfiguration[1]

Imaginary time propagation bypasses optimization issues but is computationally impractical for large ansatzes with existing methods.

$$\psi(\mathrm{R},\tau) = \sum_{n=0}^{\infty} c_n \varphi_n(\mathrm{R}) e^{-\frac{1}{\hbar}E_n\tau} \xrightarrow[E_0 \leq E_1 \leq \ldots \leq E_n]{\tau\to\infty} \psi(\mathrm{R},\tau) \propto \varphi_0(\mathrm{R})$$



$$\theta_\alpha \to \theta_\alpha - \eta \sum_\beta S^{-1}_{\alpha\beta} f_\beta$$

[1] S. Sorella, Phys. Rev. B 71, 241103 (2005).

Outline

Introduction

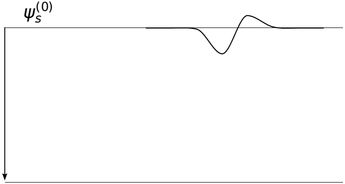
The Stochastic Representation & Path Integration

Methods
- Symmetry Enforcement
- Energy Estimation

Results

The Stochastic Representation[2]

Given a tractable ansatz $\psi_s^{(0)}$, we should be able to:

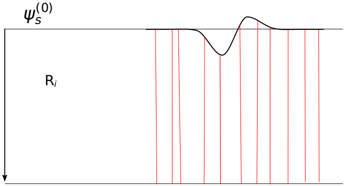


[2] H. Atanasova, L. Bernheimer, and G. Cohen, Nat Commun 14, 1 (2023).

The Stochastic Representation[2]

Given a tractable ansatz $\psi_s^{(0)}$, we should be able to:

- Select a set of sample coordinates R_i.

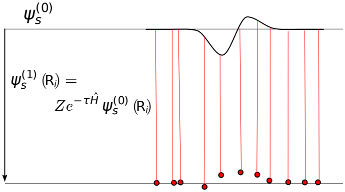


[2] H. Atanasova, L. Bernheimer, and G. Cohen, Nat Commun 14, 1 (2023).

The Stochastic Representation[2]

Given a tractable ansatz $\psi_s^{(0)}$, we should be able to:

- Select a set of sample coordinates R_i.
- Obtain the set of numbers $\psi_s^{(1)}(\mathsf{R}_i) \equiv e^{-\tau\hat{H}}\psi_s^{(0)}(\mathsf{R}_i)$, so we have pairs $\left\{\mathsf{R}_i, \psi_s^{(1)}(\mathsf{R}_i)\right\}$.

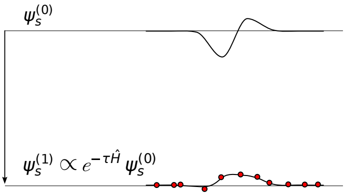


[2] H. Atanasova, L. Bernheimer, and G. Cohen, Nat Commun 14, 1 (2023).

The Stochastic Representation[2]

Given a tractable ansatz $\psi_s^{(0)}$, we should be able to:

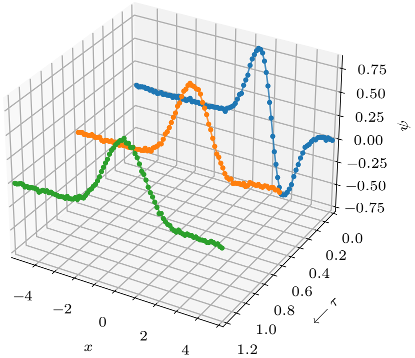
- Select a set of sample coordinates R_i.
- Obtain the set of numbers $\psi_s^{(1)}(R_i) \equiv e^{-\tau \hat{H}} \psi_s^{(0)}(R_i)$, so we have pairs $\left\{R_i, \psi_s^{(1)}(R_i)\right\}$.
- Interpolate over the samples to result in a new, propagated tractable ansatz $\psi_s^{(1)}(R)$.



[2] H. Atanasova, L. Bernheimer, and G. Cohen, Nat Commun 14, 1 (2023).

The Stochastic Representation

Repeating this will eventually result in the wavefunction at long imaginary time, i.e., the ground state. Consequently, the optimization problem transitions into a supervised regression task.



Previous Works

Similar ideas were devised before us:

- D. Kochkov and B. K. Clark, arXiv:1811.12423 (2018).
- I. L. Gutiérrez and C. B. Mendl, Quantum 6, 627 (2022).
- J. Gacon, J. Nys, R. Rossi, S. Woerner, and G. Carleo, arXiv:2303.12839 (2023).

Path Integration

- Simplest way: Euler method, $e^{-\Delta\tau\frac{\hat{H}}{\hbar}} \simeq 1 - \Delta\tau\frac{\hat{H}}{\hbar}$. Requires taking spatial derivatives, e.g. $-\frac{\hbar^2}{2m}\nabla_i^2\psi_s^{(0)}(\mathrm{R}_i)$.
- Another way: perform path integration. No derivatives needed!

$$e^{-\tau\frac{\hat{H}}{\hbar}}\psi(\mathrm{R}_0) = \lim_{N\to\infty}\left(\frac{mN}{2\pi\tau\hbar}\right)^{\frac{Nd}{2}} \times \int_{(\mathbb{R}^d)^N} \exp\left\{-\frac{1}{\hbar}\sum_{j=1}^{N}\varepsilon\left[\frac{m}{2}\left|\frac{\mathrm{R}_j - \mathrm{R}_{j-1}}{\varepsilon}\right|^2 + V(\mathrm{R}_{j-1})\right]\right\} \times \psi(\mathrm{R}_N)\, d\mathrm{R}_1 d\mathrm{R}_2 \cdots d\mathrm{R}_N$$

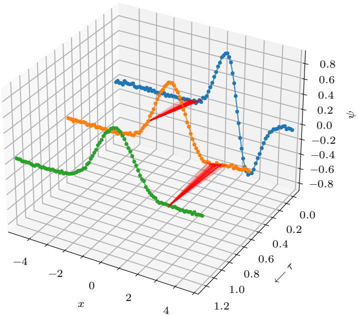
$$\varepsilon = \frac{\tau}{N}$$

Path Integration

- Path integration for long times is limited by sign problems, but shorter times can be evaluated approximately, $N = 1$ ($\Delta\tau \equiv \varepsilon$).

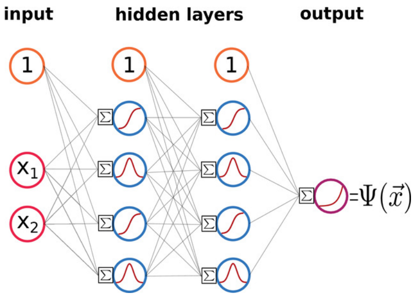
$$e^{-\Delta\tau \frac{\hat{H}}{\hbar}} \psi(\mathrm{R}_0) \simeq \left(\frac{m}{2\pi\Delta\tau\hbar}\right)^{\frac{d}{2}} \int_{\mathbb{R}^d} \exp\left\{-\frac{1}{\hbar} S_{\mathrm{E}}^{\mathrm{L}}(\mathrm{R}_0,\mathrm{R}_N,\Delta\tau)\right\} \times \psi(\mathrm{R}_N) d\mathrm{R}_N$$

- Can be carried out with Monte Carlo!



Regression

$$J(\theta) = \frac{1}{M}\sum_{i=1}^{M} \frac{\left(\psi_\theta\left(\bar{\pi}_i \mathrm{R}_i\right) - \sigma\left(\bar{\pi}_i\right) e^{-\Delta\tau\hat{H}/\hbar}\psi\left(\mathrm{R}_i\right)\right)^2}{\left|e^{-\Delta\tau\hat{H}/\hbar}\psi\left(\mathrm{R}_i\right)\right|^2 + \epsilon}.$$



https://onlinelibrary.wiley.com/doi/10.1002/adts.202000269

Outline

Outline

Exchange Symmetry of Identical Particles

The spatial wavefunction is (anti)symmetric to exchange of identical bosons (fermions).

- Bosons are relatively easy to treat due to the lack of nodes
- Fermions are hard - more sign problems

Fermionic Symmetry

Fermionic symmetry is usually enforced via Slater determinants,

$$\psi\left(\mathrm{r}_1, \mathrm{r}_2, \ldots, \mathrm{r}_N\right) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1\left(\mathrm{r}_1\right) & \chi_2\left(\mathrm{r}_1\right) & \cdots & \chi_N\left(\mathrm{r}_1\right) \\ \chi_1\left(\mathrm{r}_2\right) & \chi_2\left(\mathrm{r}_2\right) & \cdots & \chi_N\left(\mathrm{r}_2\right) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1\left(\mathrm{r}_N\right) & \chi_2\left(\mathrm{r}_N\right) & \cdots & \chi_N\left(\mathrm{r}_N\right) \end{vmatrix}$$

The time complexity for computing a determinant is of the order $\mathcal{O}\left(N^3\right)$ (or $\mathcal{O}\left(N^2\right)$ for Vandermonde ansatzes).

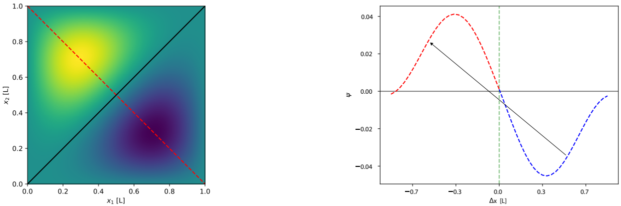
Inexpensive Symmetry Enforcement

To enforce symmetry of N particles, order their coordinates lexicographically and consider the sign change in the corresponding space subset. [3]

$\mathrm{r}_i < \mathrm{r}_j$ if $x_i < x_j$, or $x_i = x_j$ and $y_i < y_j$ or $x_i = x_j$ and so on.

$$\psi\left(\{\mathrm{r}_1, \ldots, \mathrm{r}_N\}\right) = \sigma\left(\bar{\pi}\right)\psi_{\frac{1}{N!}}\left(\left\{\mathrm{r}_{\bar{\pi}(1)}, \ldots, \mathrm{r}_{\bar{\pi}(N)}\right\}\right)$$

For example, 2 fermions in a 1D box:



[3] M. Hutter, ArXiv:2007.15298 [Quant-Ph] (2020).

Inexpensive Symmetry Enforcement

To enforce symmetry of N particles, order their coordinates lexicographically and consider the sign change in the corresponding space subset. [3]

$\mathrm{r}_i < \mathrm{r}_j$ if $x_i < x_j$, or $x_i = x_j$ and $y_i < y_j$ or $x_i = x_j$ and so on.

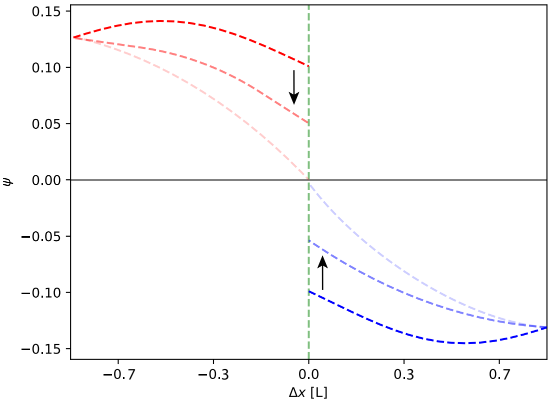
$$\psi\left(\{\mathrm{r}_1, \ldots, \mathrm{r}_N\}\right) = \sigma\left(\bar{\pi}\right) \psi_{\frac{1}{N!}}\left(\left\{\mathrm{r}_{\bar{\pi}(1)}, \ldots, \mathrm{r}_{\bar{\pi}(N)}\right\}\right)$$

Using quicksort lowers the symmetrization complexity to $\mathcal{O}\left(N \log N\right)$, with a small prefactor!

[3] M. Hutter, ArXiv:2007.15298 [Quant-Ph] (2020).

Inexpensive Symmetry Enforcement

Diverging derivatives are treatable only via the path integral approach!



Outline

Introduction

The Stochastic Representation & Path Integration

Methods

Symmetry Enforcement

Energy Estimation

Results

Energy estimation

A non-variational energy estimation can be extracted immediately.

- At long imaginary times we can write

$$e^{-\Delta\tau\frac{\hat{H}}{\hbar}}\varphi_0(\mathrm{R}) = e^{-\Delta\tau\frac{E_0}{\hbar}}\varphi_0(\mathrm{R})$$

$$E_0 = -\hbar\frac{\ln\left(\frac{e^{-\Delta\tau\frac{\hat{H}}{\hbar}}\varphi_0(\mathrm{R})}{\varphi_0(\mathrm{R})}\right)}{\Delta\tau}$$

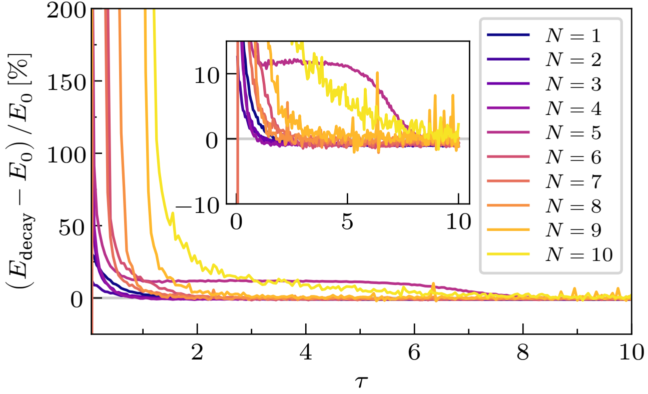
- At each step we check if the "decay" estimation has converged

$$E_{\text{decay}}(\psi) = -\hbar\frac{\ln\left(\frac{e^{-\Delta\tau\frac{\hat{H}}{\hbar}}\psi(\mathrm{R_i})}{\psi(\mathrm{R_i})}\right)}{\Delta\tau}$$

Outline

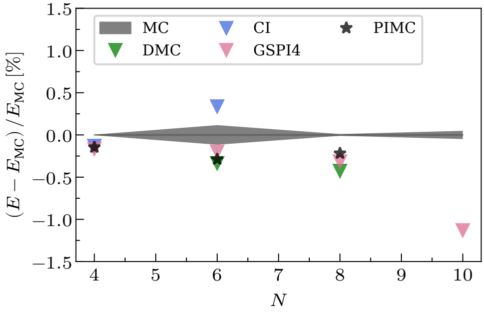
Non-interacting fermions in a 2D harmonic trap

$$\hat{H} = \frac{1}{2m}\nabla^2 + \frac{1}{2}m\omega^2 \sum_i \mathrm{r}_i^2$$



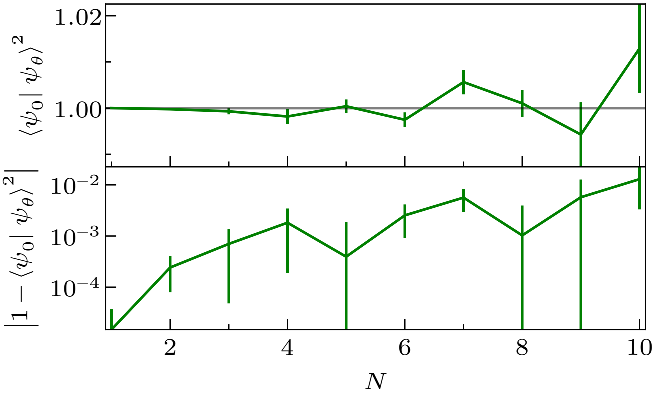
Spin polarized interacting fermions in a 2D harmonic trap

$$\hat{H} = \frac{1}{2m}\nabla^2 + \frac{1}{2}m\omega^2\sum_i r_i^2 + \sum_{i>j}\frac{\lambda}{r_{ij}}, \lambda = 8$$



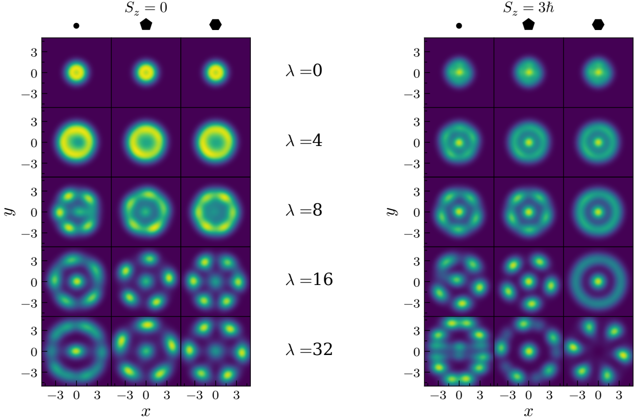
Fidelity

The quality of the regression decreases slower than exponentially.



Wigner crystallization for $N = 6$

We shed light on the different transitions to a Wigner molecule of the ground-state and spin-polarized cases.



Summary

- SRW is an **alternative to variational Monte Carlo** that enables more robust optimization with **scalable imaginary time propagation.**
- **Path integration** obviates the need for **spatial derivatives**, enabling utilization of non-differentiable or even discontinuous ansatzes.
- Non-differentiable ansatzes are **helpful in machine learning**, and enable **highly efficient (anti)symmetry enforcement** by lexicographic sorting of coordinates.

Thank You!

Bonus A - linear euclidean action

$$\mathrm{R}^{L}(t) = \left(1 - \frac{t}{\Delta\tau}\right)\mathrm{R}_0 + \frac{t}{\Delta\tau}\mathrm{R}_N$$

$$S_{\mathrm{E}}^{\mathrm{L}}(\mathrm{R}_0, \mathrm{R}_N, \Delta\tau) = \int_0^{\Delta\tau} \left[\frac{m}{2}\left|\frac{\mathrm{R}_N - \mathrm{R}_0}{\Delta\tau}\right|^2 + V\left(\mathrm{R}^{L}(t)\right)\right] dt$$
$$= \frac{m}{2\Delta\tau}\left|\mathrm{R}_N - \mathrm{R}_0\right|^2 + \int_0^{\Delta\tau} V\left(\mathrm{R}^{L}(t)\right) dt$$

$$e^{-\hat{H}\Delta\tau/\hbar}\psi(\mathrm{R}_0) \simeq \left(\frac{m}{2\pi\Delta\tau\hbar}\right)^{d/2}\frac{1}{n_s} \times \sum_{\mathrm{R}_i\sim\mathcal{N}_{\mu,\sigma^2}} \frac{\exp\left\{-\frac{1}{\hbar}S_{\mathrm{E}}^{\mathrm{L}}(\mathrm{R}_0, \mathrm{R}_i, \Delta\tau)\right\}}{\mathcal{N}_{\mu,\sigma^2}(\mathrm{R}_i)}\psi(\mathrm{R}_i)$$

Bonus B - derivative-free variational energy estimation

We can avoid the differentiation of the wavefunction by convoluting it with a Gaussian.

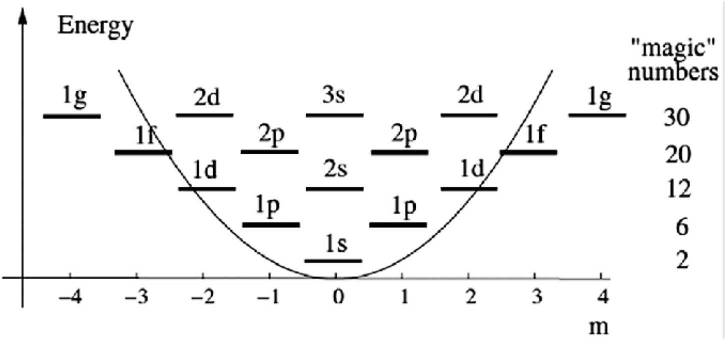
$$\psi(\mathrm{R}) \to \widetilde{\psi}(\mathrm{R}) = \psi(\mathrm{R}) * \mathcal{N}_{0,\sigma^2}(\mathrm{R})$$
$$= \int_{\mathbb{R}^d} \psi\left(\mathrm{R} - \mathrm{k}'\right) \mathcal{N}_{0,\sigma^2}\left(\mathrm{k}'\right) d\mathrm{k}'$$

$$E = \frac{\left\langle \widetilde{\psi} \middle| \hat{H} \middle| \widetilde{\psi} \right\rangle}{\left\langle \widetilde{\psi} \middle| \widetilde{\psi} \right\rangle} \simeq \frac{\sum_{\mathrm{R},\mathrm{k}',\mathrm{k}''} A(\mathrm{R}, \mathrm{k}', \mathrm{k}'')}{\sum_{\mathrm{R},\mathrm{k}',\mathrm{k}''} B(\mathrm{R}, \mathrm{k}', \mathrm{k}'')} \geq E_0$$

$$A\left(\mathrm{R}, \mathrm{k}', \mathrm{k}''\right) \equiv B\left(\mathrm{R}, \mathrm{k}', \mathrm{k}''\right) \times \frac{\mathrm{R}''^2 - \sigma^2 + \sigma^4 V(\mathrm{R})}{\sigma^4}$$

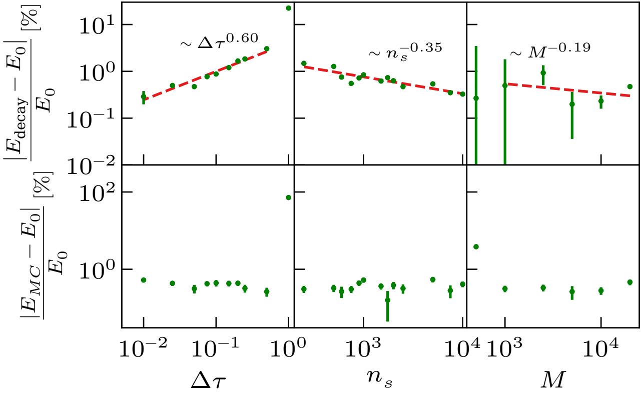
$$B\left(\mathrm{R}, \mathrm{k}', \mathrm{k}''\right) \equiv \frac{\psi^*\left(\mathrm{R} - \mathrm{k}'\right) \psi\left(\mathrm{R} - \mathrm{k}''\right)}{\left|\psi(\mathrm{R})\right|^2}$$

Bonus C - 2D harmonic oscillator energy levels



https://www.researchgate.net/figure/energy-levels-and-number-of-electrons-for-shell-closings-of-the-2d-harmonic-oscillator_fig7_306243994

Bonus D - Scaling results



Bonus E - Runtime scaling