Learning the Spectrum with Neural Quantum States

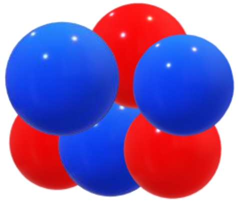
J. Rozalén Sarmiento, A. Rios

ICCUB | Institut de Ciències del Cosmos
UNIVERSITAT DE BARCELONA

EXCELENCIA MARÍA DE MAEZTU 2020–2024

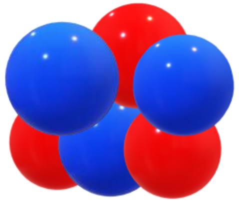
UNIVERSITAT DE BARCELONA

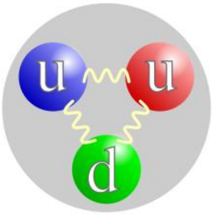
Ab Initio Nuclear structure in a (nut)shell

1. Select your nucleus 

Ab Initio Nuclear structure in a (nut)shell

1. Select your nucleus

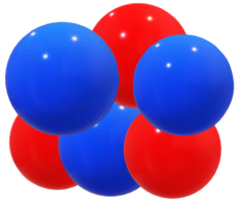


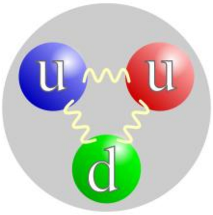


2. Select your Hamiltonian $\mathcal{L}_{\mathrm{QCD}}$ → $\hat{H}_{\mathrm{nuc}} = \hat{T} + \hat{V}$

Ab Initio Nuclear structure in a (nut)shell

1. Select your nucleus





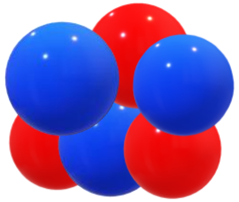
2. Select your Hamiltonian

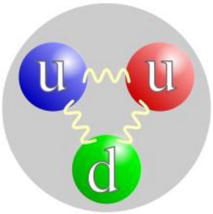
$$\mathcal{L}_{\mathrm{QCD}} \Longrightarrow \hat{H}_{\mathrm{nuc}} = \hat{T} + \hat{V}$$

3. Select your method

Ab Initio Nuclear structure in a (nut)shell

1. Select your nucleus





2. Select your Hamiltonian

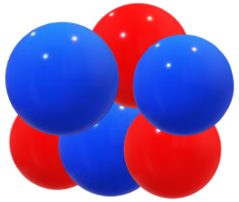
$$\mathcal{L}_{\mathrm{QCD}} \longrightarrow \hat{H}_{\mathrm{nuc}} = \hat{T} + \hat{V}$$

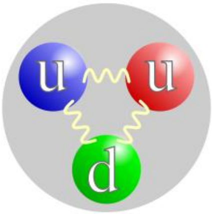
3. Select your method

4. Predict your wave function $\longrightarrow$ $\psi_{\mathrm{nucleus}}(\mathbf{x};\sigma,\tau)$

Ab Initio Nuclear structure in a (nut)shell

1. Select your nucleus





2. Select your Hamiltonian

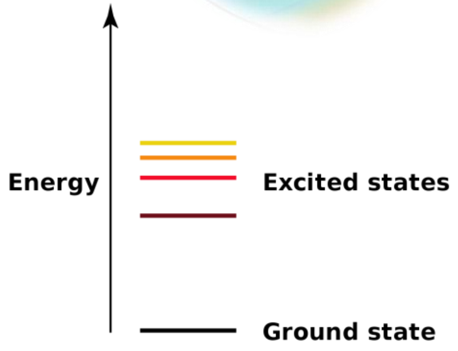
$$\mathcal{L}_{\mathrm{QCD}} \longrightarrow \hat{H}_{\mathrm{nuc}} = \hat{T} + \hat{V}$$

3. Select your method

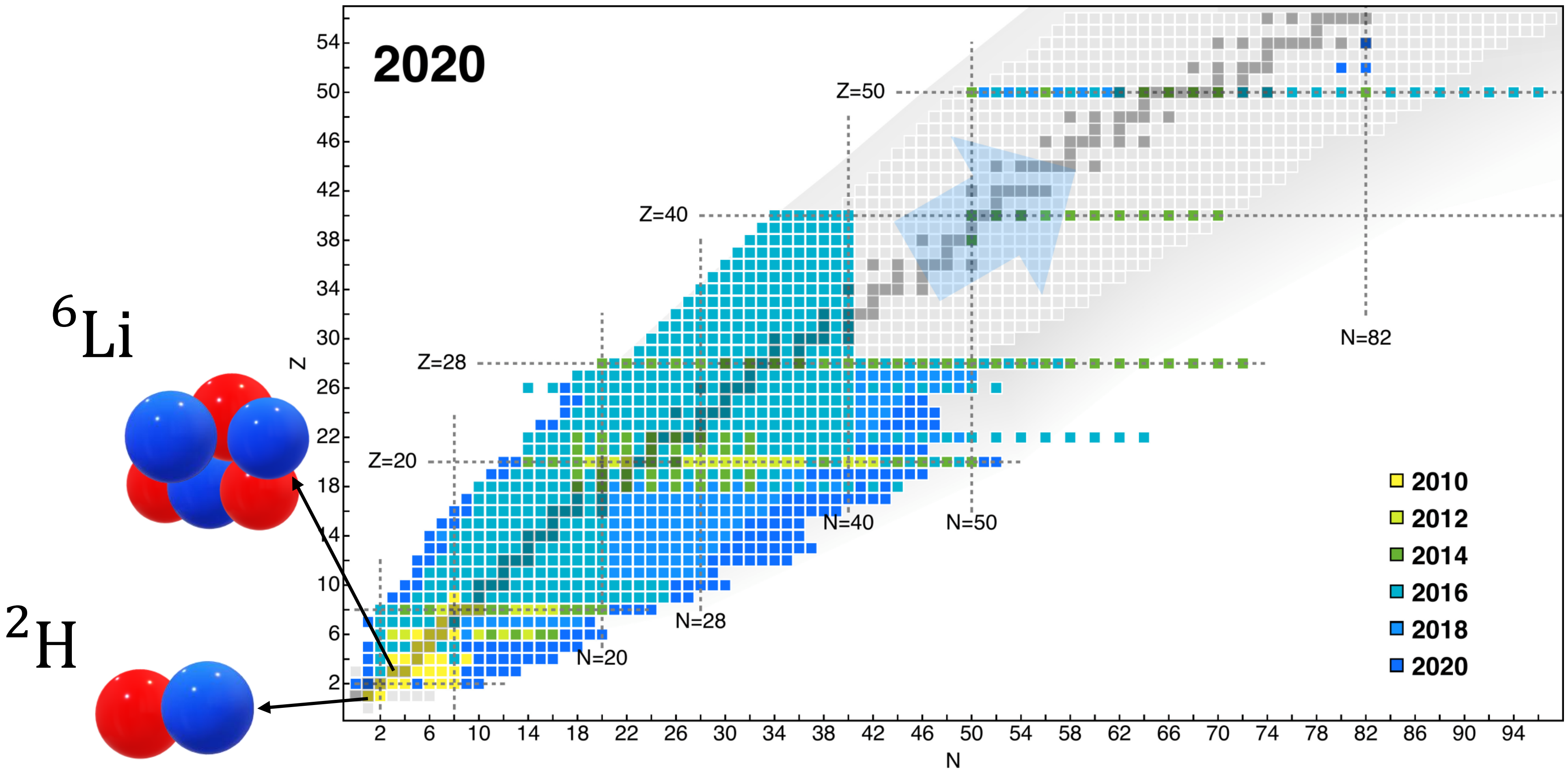
shape?

4. Predict your wave function

$$\longrightarrow \psi_{\mathrm{nucleus}}(\mathbf{x}; \sigma, \tau)$$

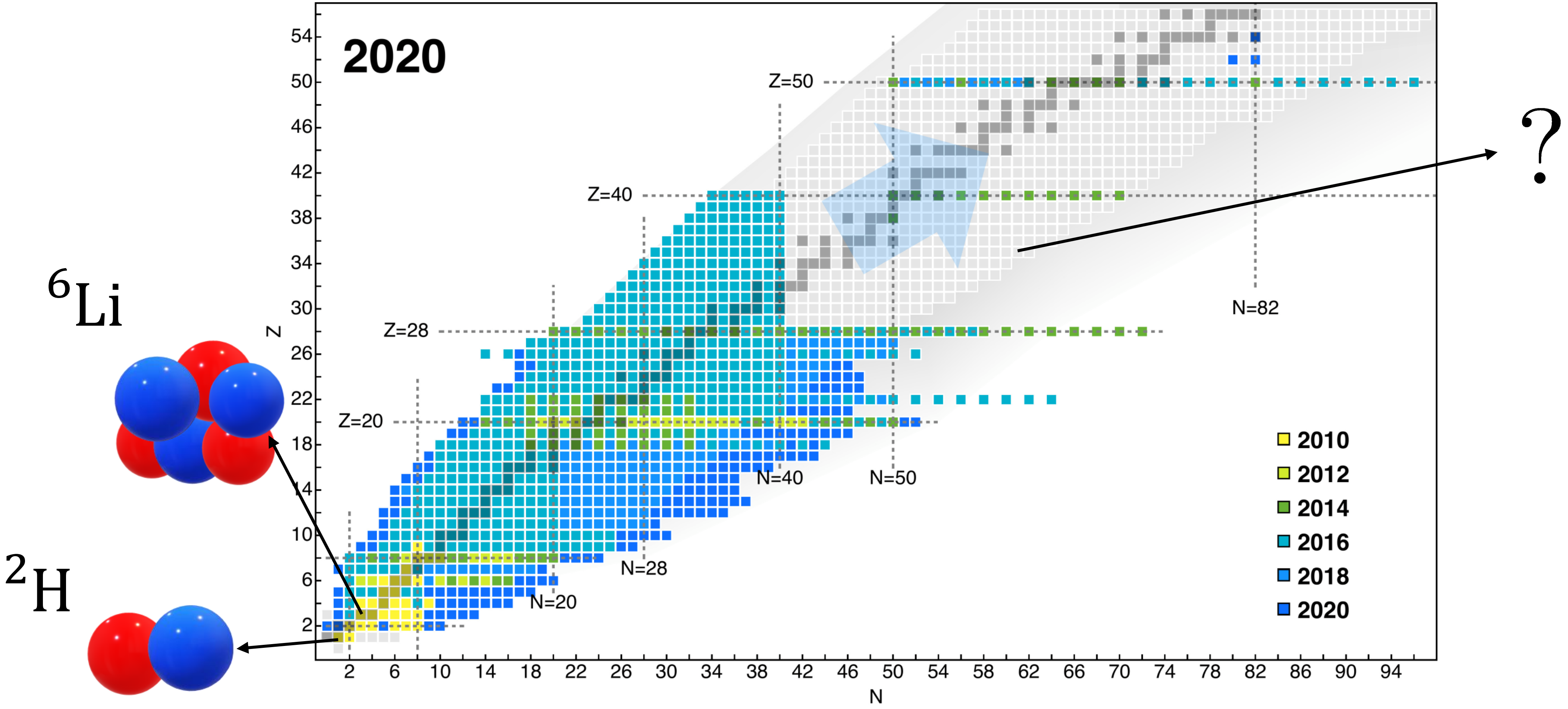


Ab Initio Nuclear Structure



H. Hergert, Front. Phys. **07** (2020)

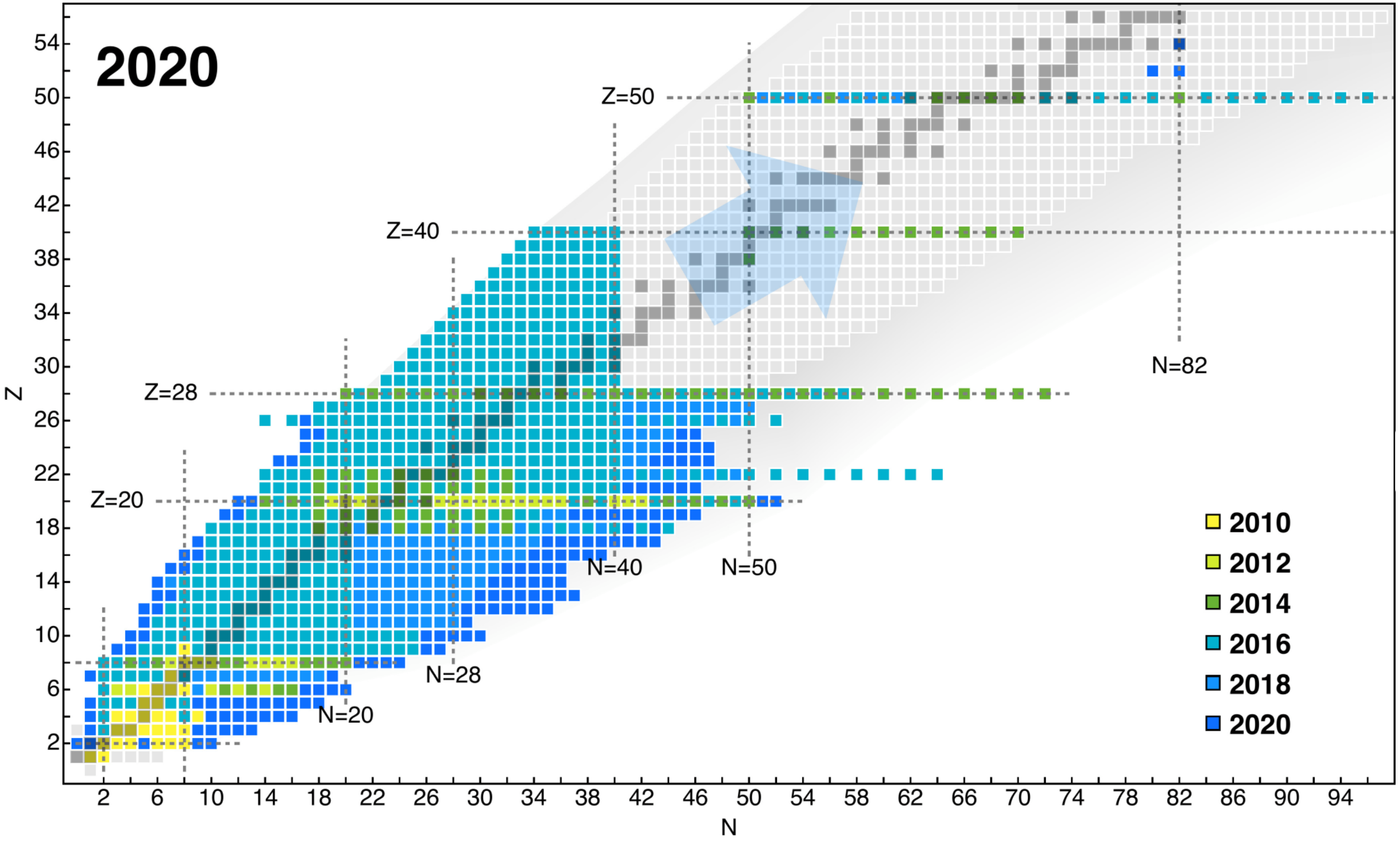
Ab Initio Nuclear Structure



H. Hergert, Front. Phys. **07** (2020)

Ab Initio Nuclear Structure

- IT-NCSM
- MR-IMSRG
- VS-IMSRG
- ADC

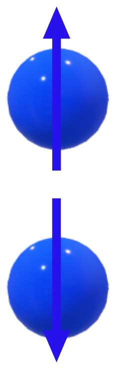


- Lattice EFT
- CCSD
- Λ-CCSD

H. Hergert, Front. Phys. **07** (2020)

Exponential Scaling Problem

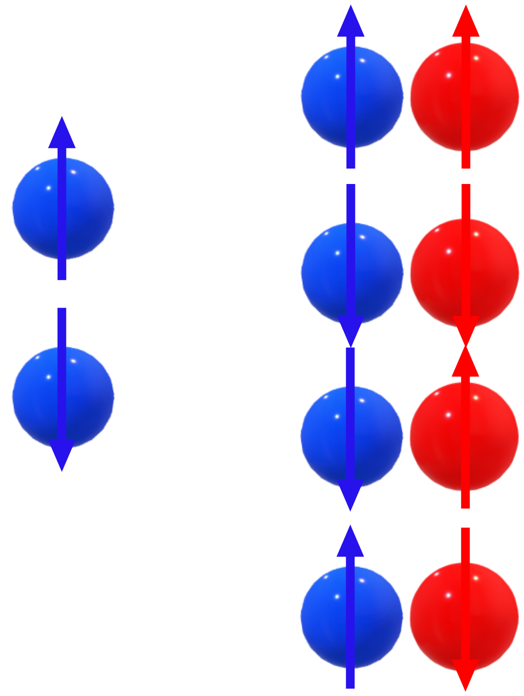
$N=1$



$2^N=2$

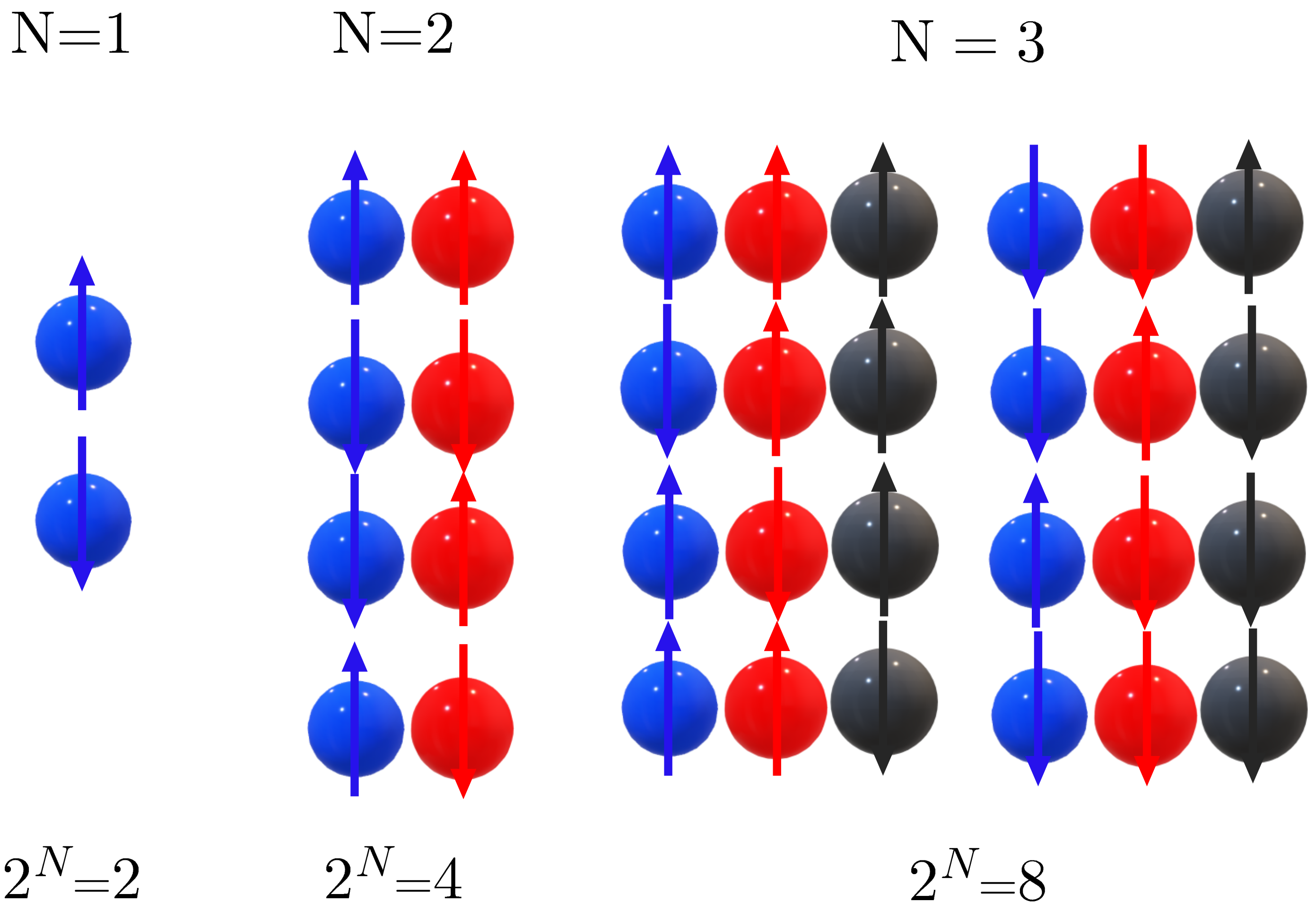
Exponential Scaling Problem

N=1 N=2



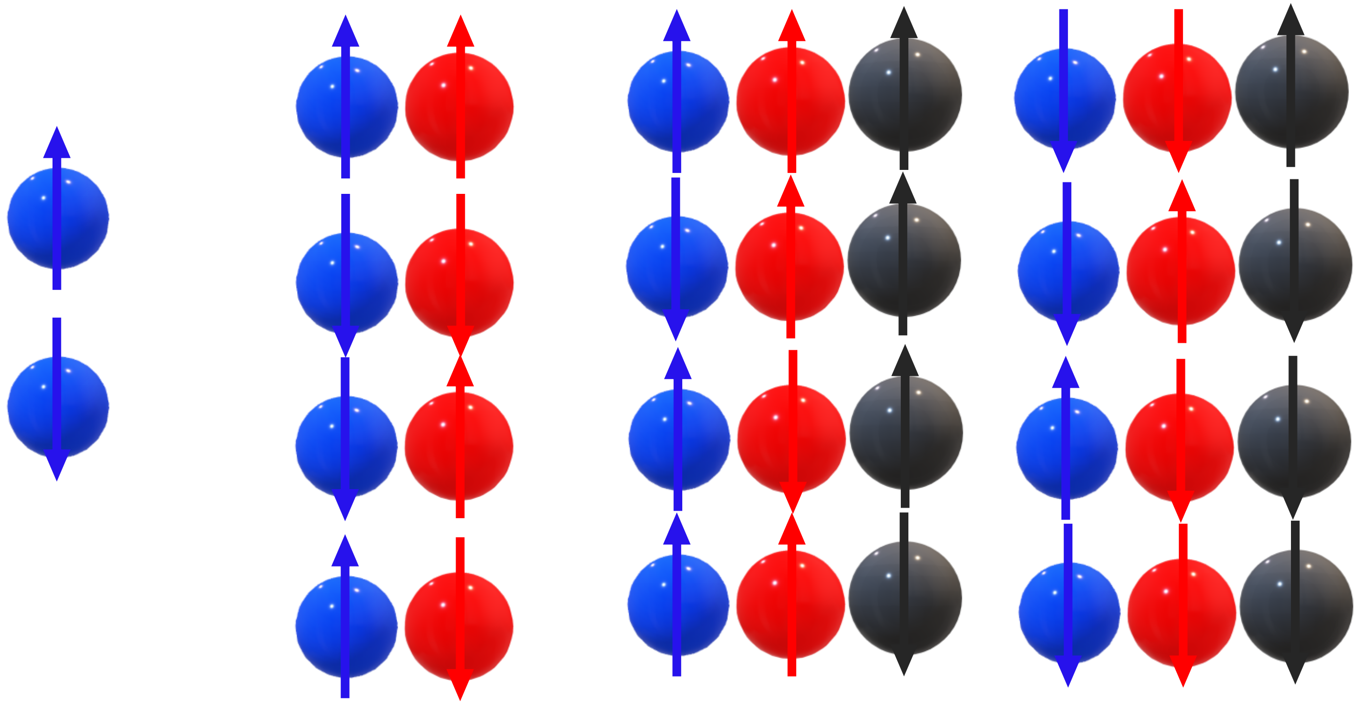
2^N=2 2^N=4

Exponential Scaling Problem



Exponential Scaling Problem

N=1 N=2 N = 3 N = 80 (Zr)

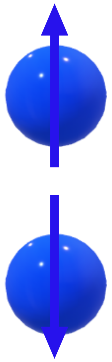


2^N=2 2^N=4 2^N=8 2^N=1,208,925,819,614,629,174,706,176

($1{,}2 \times 10^{24}$)

Exponential Scaling Problem

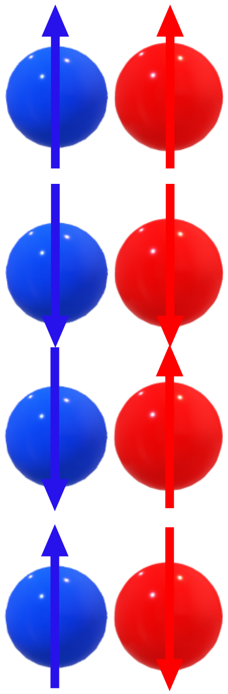
N=1



2^N=2

=16B

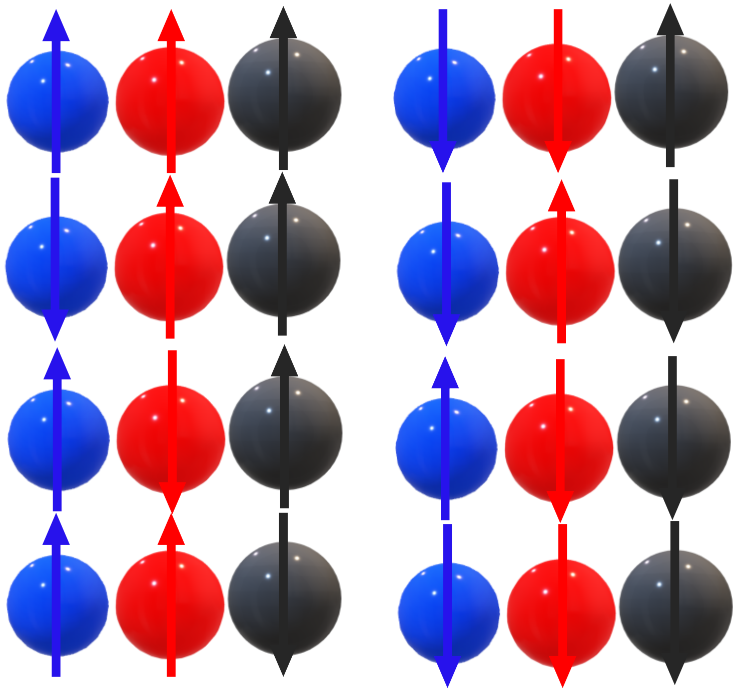
N=2



2^N=4

=32B

N = 3



2^N=8

=64B

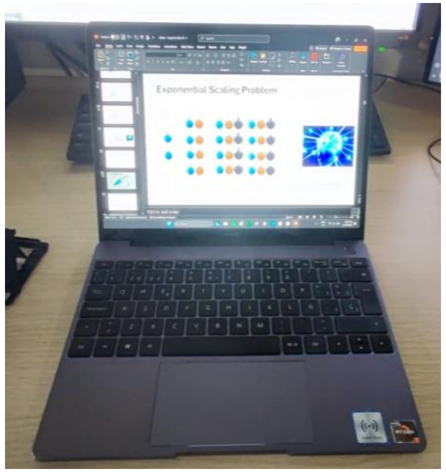
N = 80 (Zr)

2^N=1,208,925,819,614,629,174,706,176

$\sim 10^{25}$B

Exponential Scaling Problem

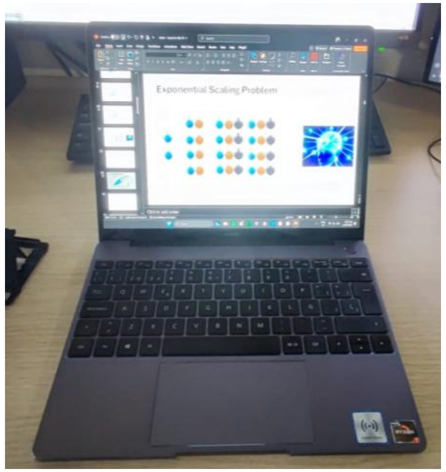
My laptop



~512 GB ~10^{11}B

Exponential Scaling Problem

My laptop



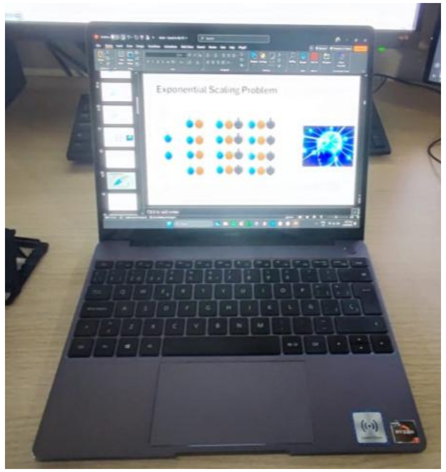
~512 GB ~10^{11}B

My rich friend's laptop

~8 TB ~ 10^{12}B

Exponential Scaling Problem

My laptop



~512 GB ~10^{11}B

My rich friend's laptop

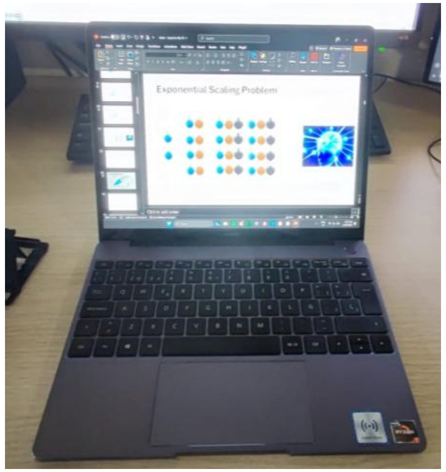
~8 TB ~ 10^{12}B

El Capitán

~5 PB $\approx 10^{15}\ \mathrm{B}$

Exponential Scaling Problem

My laptop



~512 GB ~10^{11}B

My rich friend's laptop

~8 TB ~ 10^{12}B

El Capitán

~5 PB ≈ 10^{15} B

What we need:  $\sim 10^{25}$B What we have:  $\sim 10^{15}$B

The Solution: AI

How can AI help us?

Neural Quantum States

- **Variational principle** $$\frac{\langle \psi_\theta | \hat{H} | \psi_\theta \rangle}{\langle \psi_\theta | \psi_\theta \rangle} \geq E_{\mathrm{GS}}$$ ➡ Find θ that minimises $$\frac{\langle \psi_\theta | \hat{H} | \psi_\theta \rangle}{\langle \psi_\theta | \psi_\theta \rangle}$$

G. Carleo and M. Troyer, Science (2017)

Neural Quantum States

- **Variational principle** $\frac{\langle \psi_\theta | \hat{H} | \psi_\theta \rangle}{\langle \psi_\theta | \psi_\theta \rangle} \geq E_{\text{GS}}$ ➡ Find θ that minimises $\frac{\langle \psi_\theta | \hat{H} | \psi_\theta \rangle}{\langle \psi_\theta | \psi_\theta \rangle}$

- **Variational Monte Carlo**

$$|\psi_\theta\rangle = \int \psi_\theta(X) dX = \int \psi_\theta(\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_N}) d\mathbf{x_1} d\mathbf{x_2} \ldots d\mathbf{x_N}$$

$$\frac{\langle \psi_\theta | \hat{H} | \psi_\theta \rangle}{\langle \psi_\theta | \psi_\theta \rangle} = \mathbb{E}_{X \sim p_\theta} \left[\frac{\langle X | \hat{H} | \psi_\theta \rangle}{\langle X | \psi_\theta \rangle} \right] \simeq \frac{1}{N_s} \sum_{i=1}^{N_s} \frac{\langle X_i | \hat{H} | \psi_\theta \rangle}{\langle X_i | \psi_\theta \rangle}$$

"Local energy"

G. Carleo and M. Troyer, Science (2017)

Neural Quantum States

- **Variational principle** $\frac{\langle\psi_\theta|\hat{H}|\psi_\theta\rangle}{\langle\psi_\theta|\psi_\theta\rangle} \geq E_{\text{GS}}$ ➡ Find θ that minimises $\frac{\langle\psi_\theta|\hat{H}|\psi_\theta\rangle}{\langle\psi_\theta|\psi_\theta\rangle}$

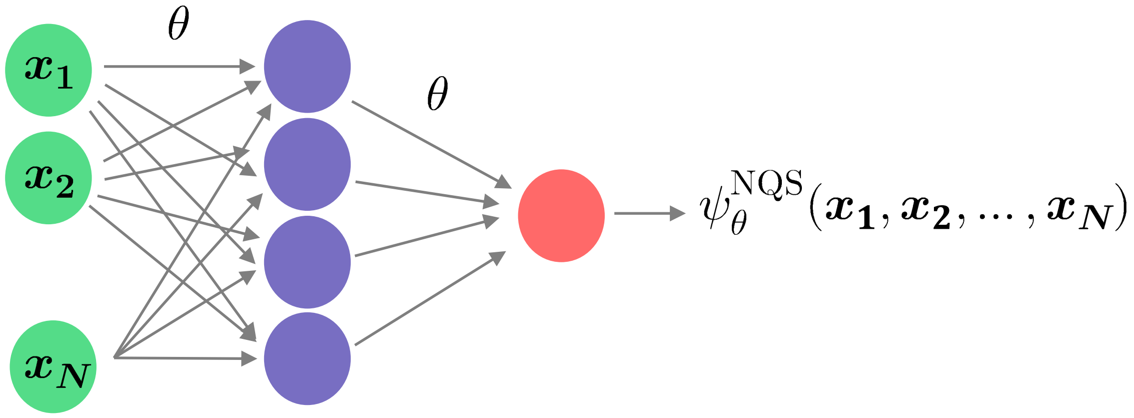
- **Variational Monte Carlo**

$$|\psi_\theta\rangle = \int \psi_\theta(X)dX = \int \psi_\theta(\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_N})d\mathbf{x_1}d\mathbf{x_2}\ldots d\mathbf{x_N}$$

$$\frac{\langle\psi_\theta|\hat{H}|\psi_\theta\rangle}{\langle\psi_\theta|\psi_\theta\rangle} = \mathbb{E}_{X\sim p_\theta}\left[\frac{\langle X|\hat{H}|\psi_\theta\rangle}{\langle X|\psi_\theta\rangle}\right] \simeq \frac{1}{N_s}\sum_{i=1}^{N_s}\frac{\langle X_i|\hat{H}|\psi_\theta\rangle}{\langle X_i|\psi_\theta\rangle}$$

"Local energy"

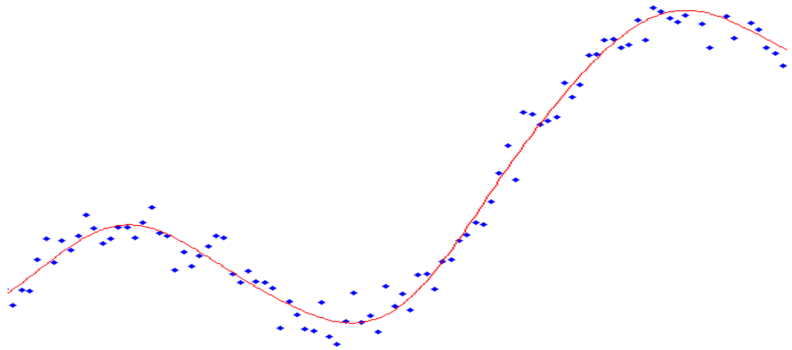
- **NQS Ansatz:** $\psi_{\text{NQS}}(\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_N}; \boldsymbol{\theta})$



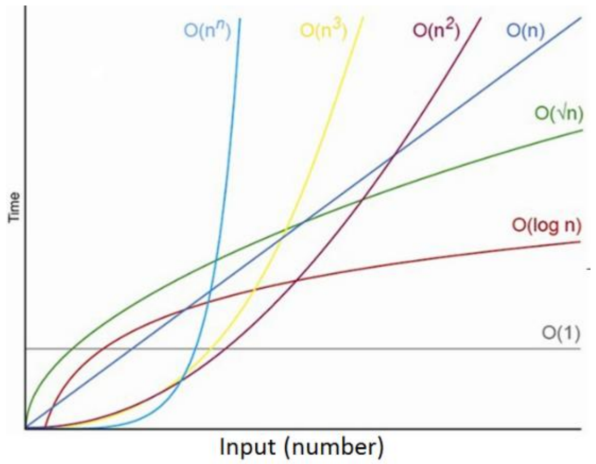
G. Carleo and M. Troyer, Science (2017)

Why Neural Networks?

- **NNs have "∞ power"**: a neural network can approximate any continuous function

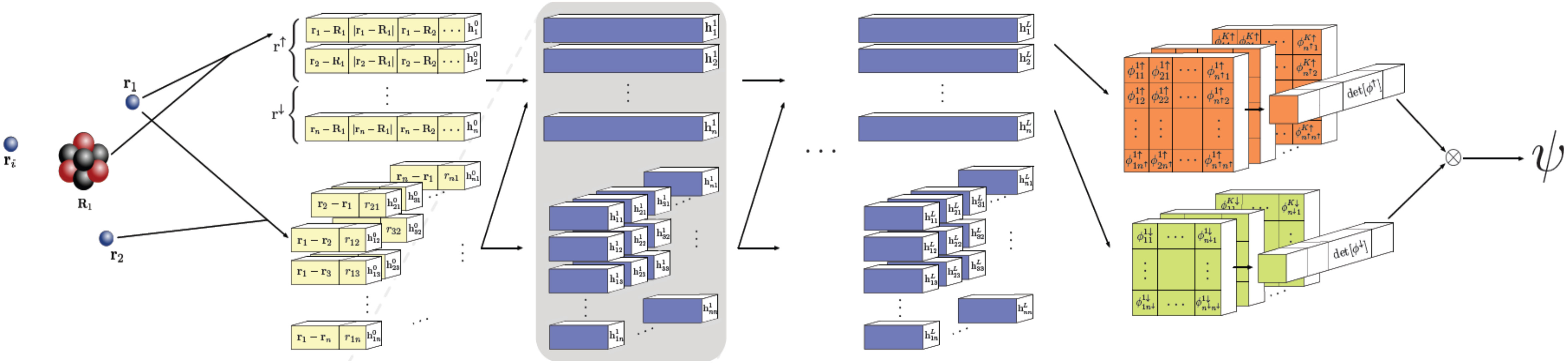


- **Space complexity:** polynomial scaling of memory resources… possibly!



NQS showcases

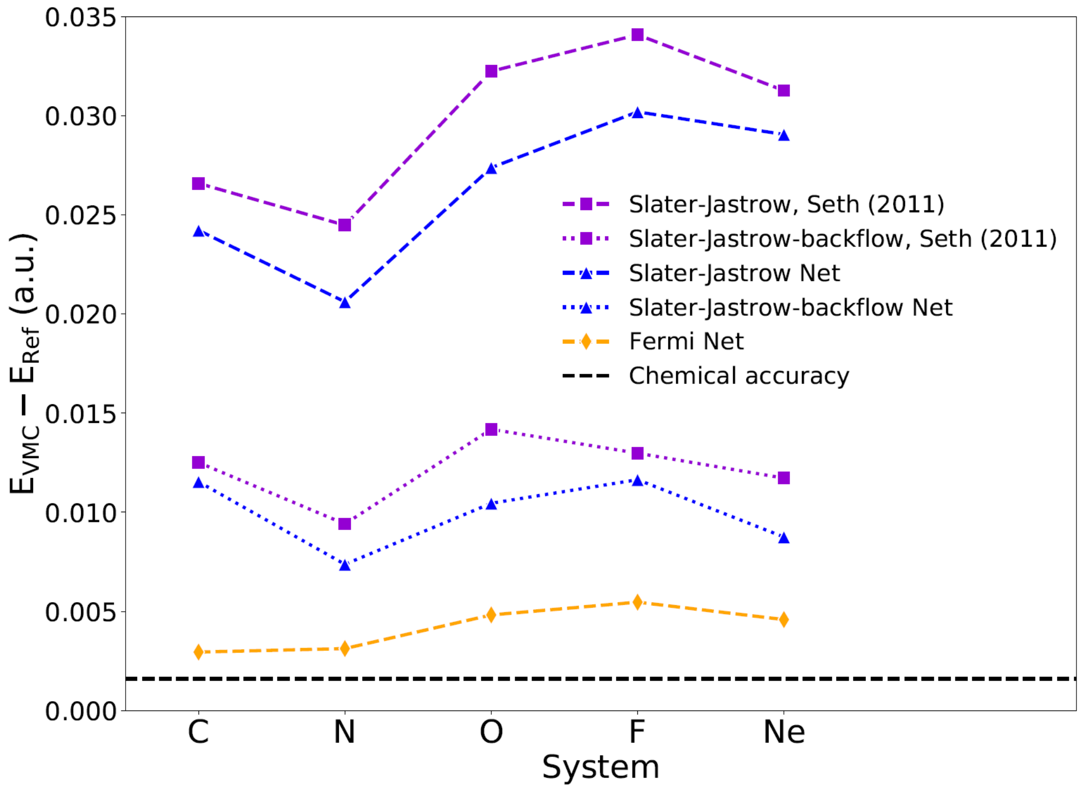
FermiNet (Quantum Chemistry)

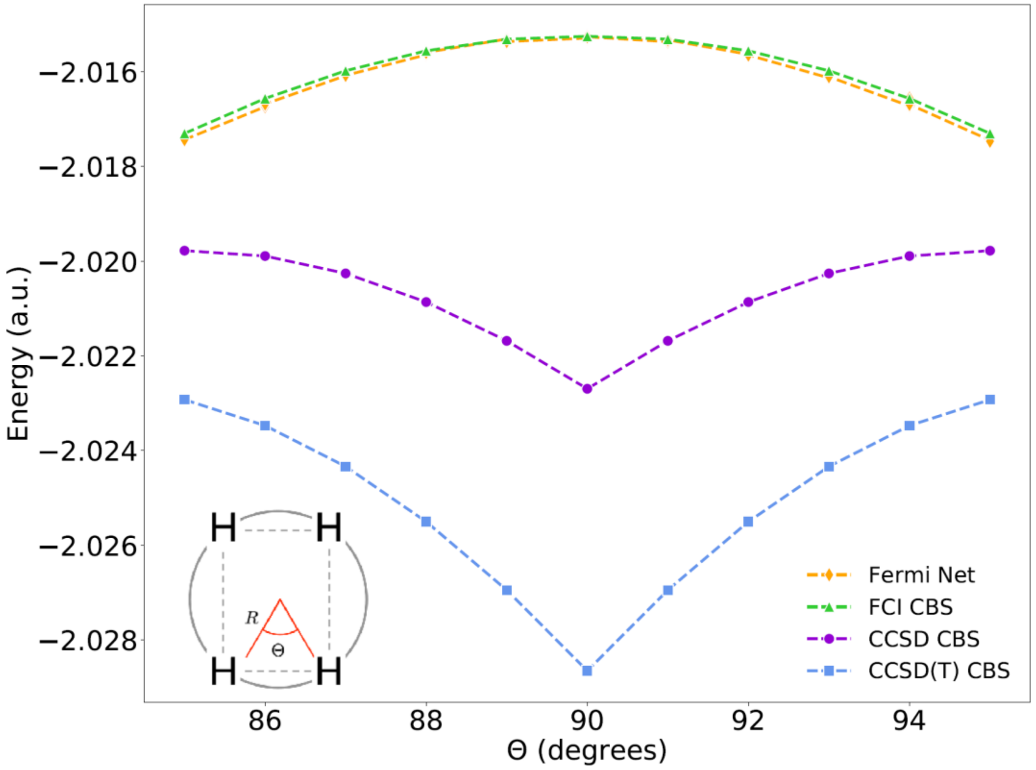


DeepMind

D. Pfau, J. S. Spencer *et al*, 2020, *Phys. Rev. Res.* **2**.

FermiNet (Quantum Chemistry)





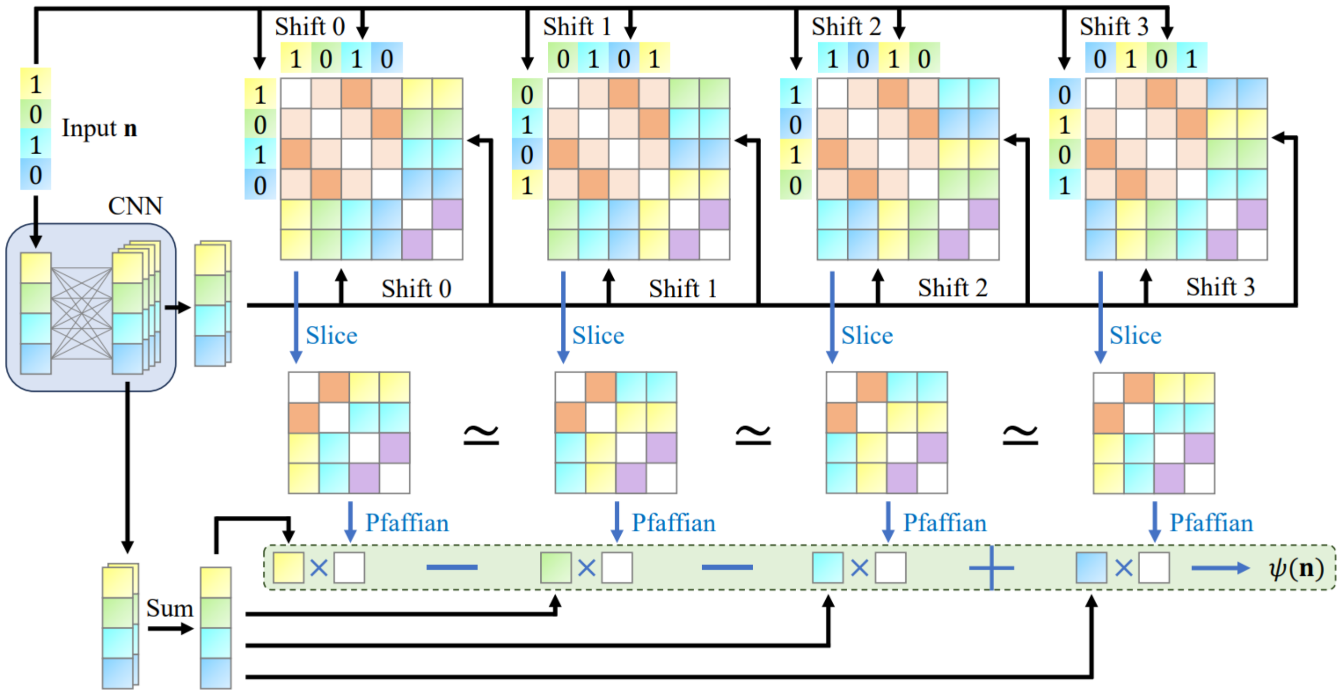
D. Pfau, J. S. Spencer *et al*, 2020, *Phys. Rev. Res.* **2**.

Pfaffian-based Networks (Superconductivity)

- **System:** Hubbard model

$$\hat{\mathcal{H}} = -t \sum_{\langle ij \rangle,\sigma} \hat{c}^\dagger_{i,\sigma}\hat{c}_{j,\sigma} + U \sum_i \hat{n}_{i,\uparrow}\hat{n}_{i,\downarrow}$$

- **NQS Ansatz:**



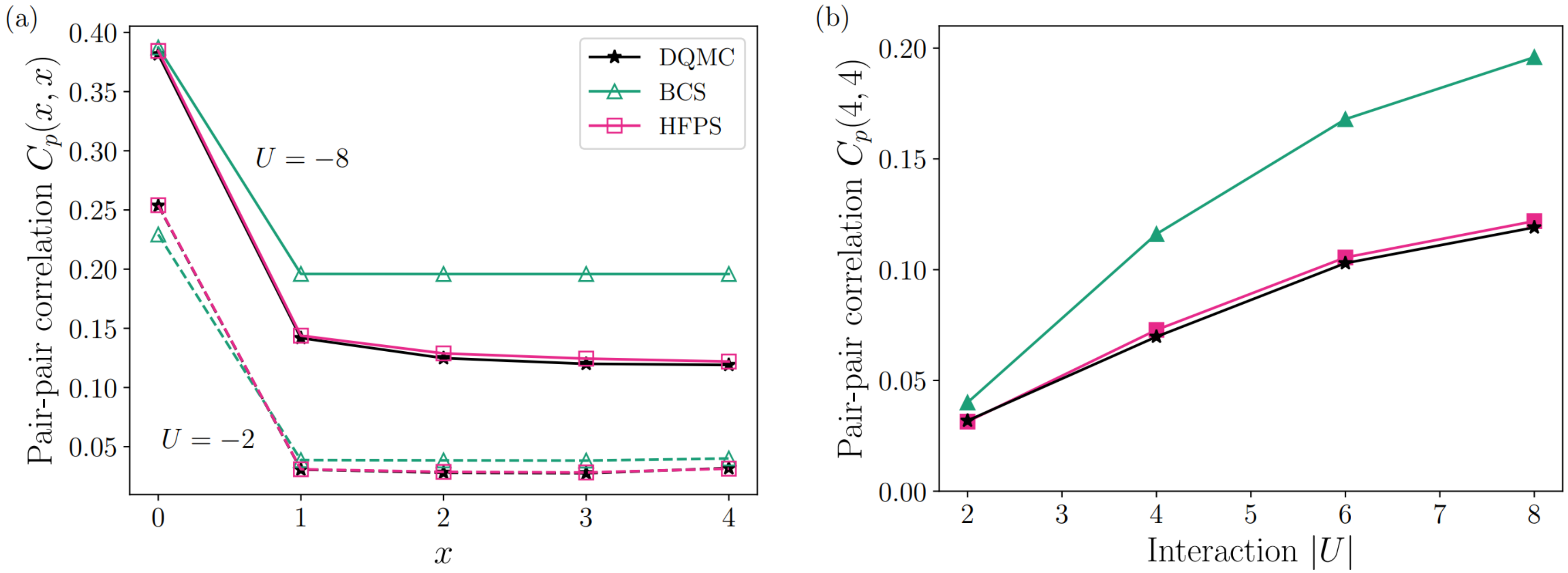
A. Chen *et al, arXiv:2507.10705 (2025)*

Pfaffian-based Networks (Superconductivity)

- **System:** Hubbard model

$$\hat{\mathcal{H}} = -t \sum_{\langle ij \rangle, \sigma} \hat{c}^{\dagger}_{i,\sigma} \hat{c}_{j,\sigma} + U \sum_{i} \hat{n}_{i,\uparrow} \hat{n}_{i,\downarrow}$$

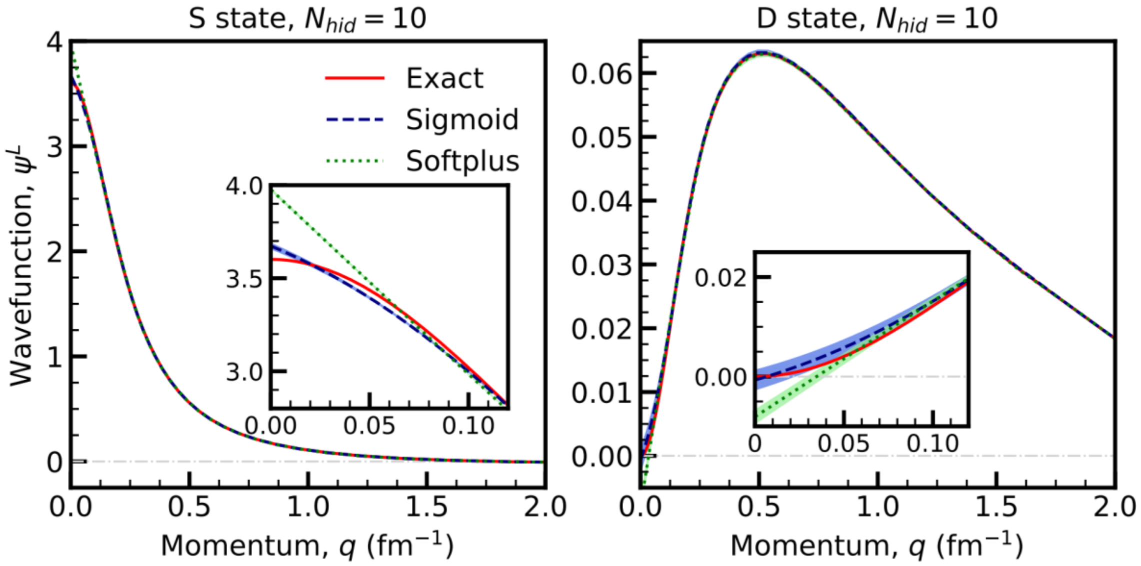
- **Performance:**

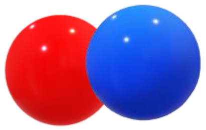


A. Chen *et al, arXiv:2507.10705 (2025)*

NQS for Nuclei

2020

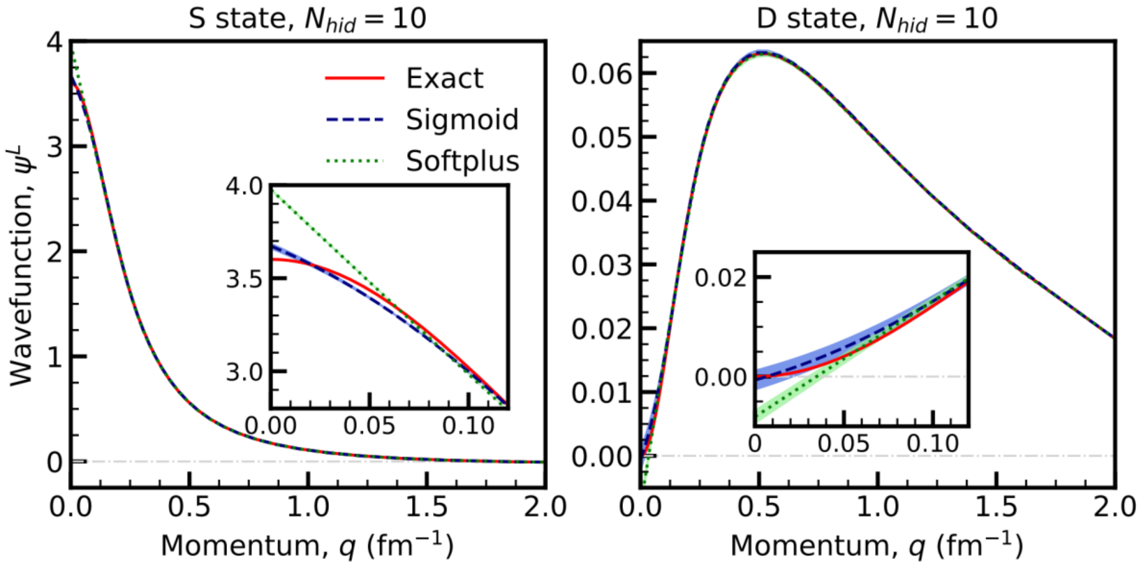


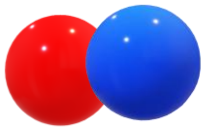


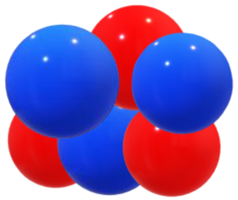
J. Keeble & A. Rios, Phys. Lett. B **809** (2020)

NQS for Nuclei

2020 → ...





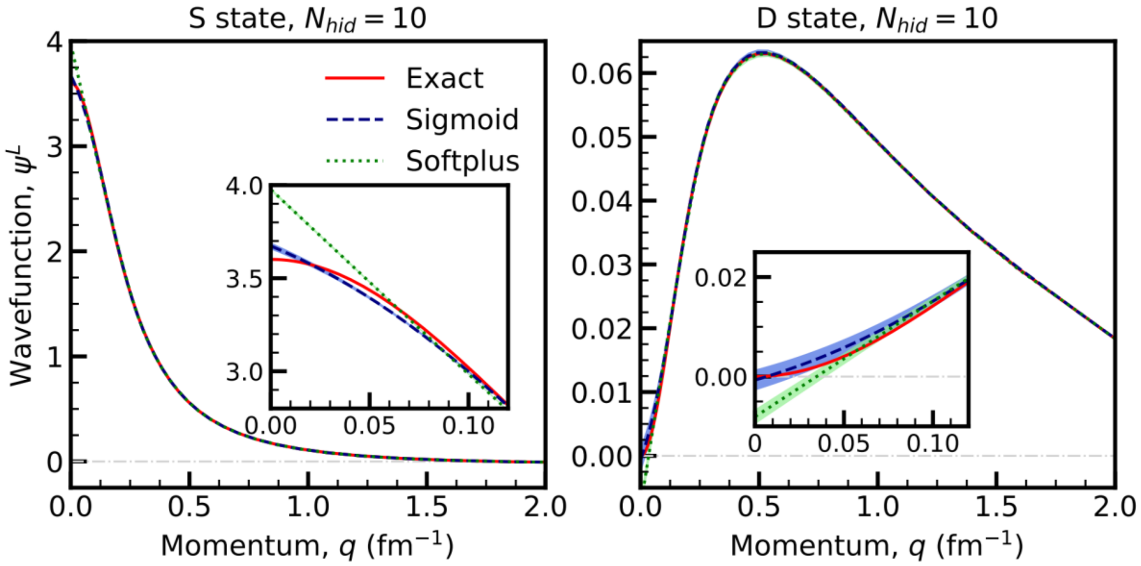


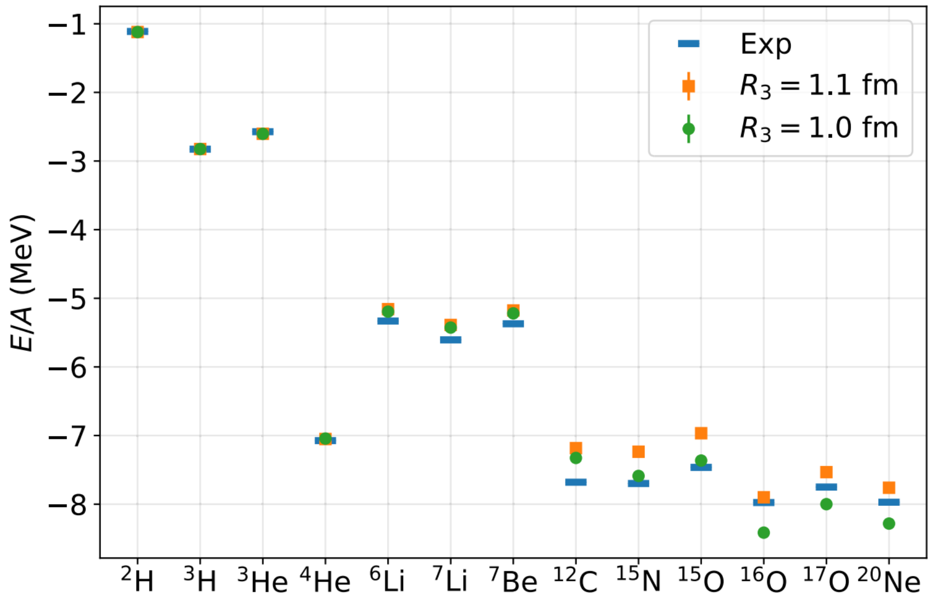
J. Keeble & A. Rios, Phys. Lett. B **809** (2020)

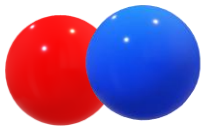
A. Gnech et al, Few-Body Syst. **63** (2022)

NQS for Nuclei

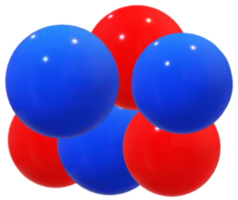
2020 → ... → **2024**



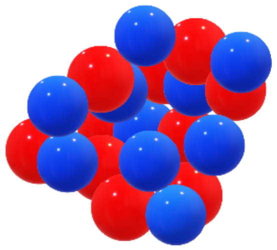




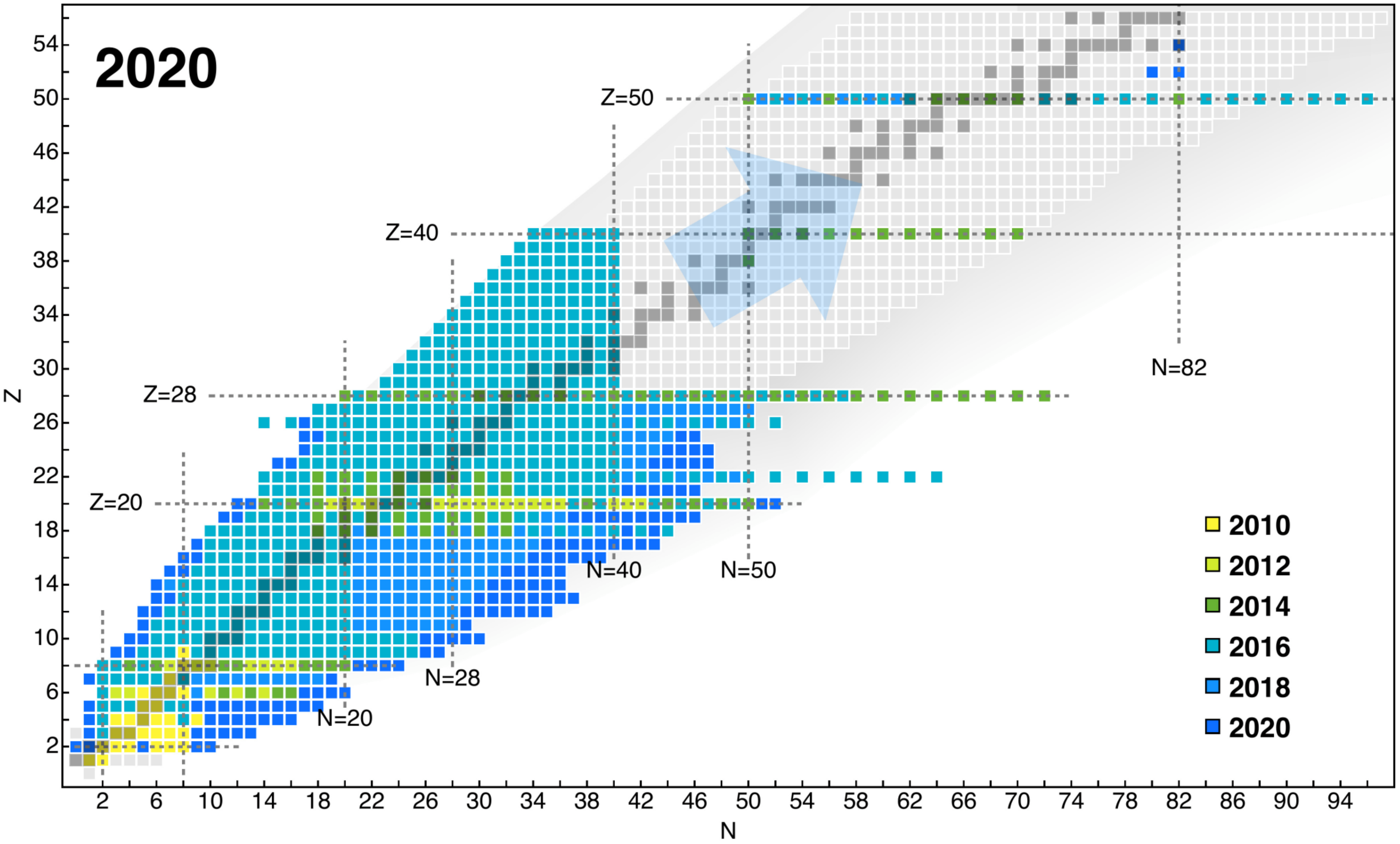
J. Keeble & A. Rios, Phys. Lett. B **809** (2020)



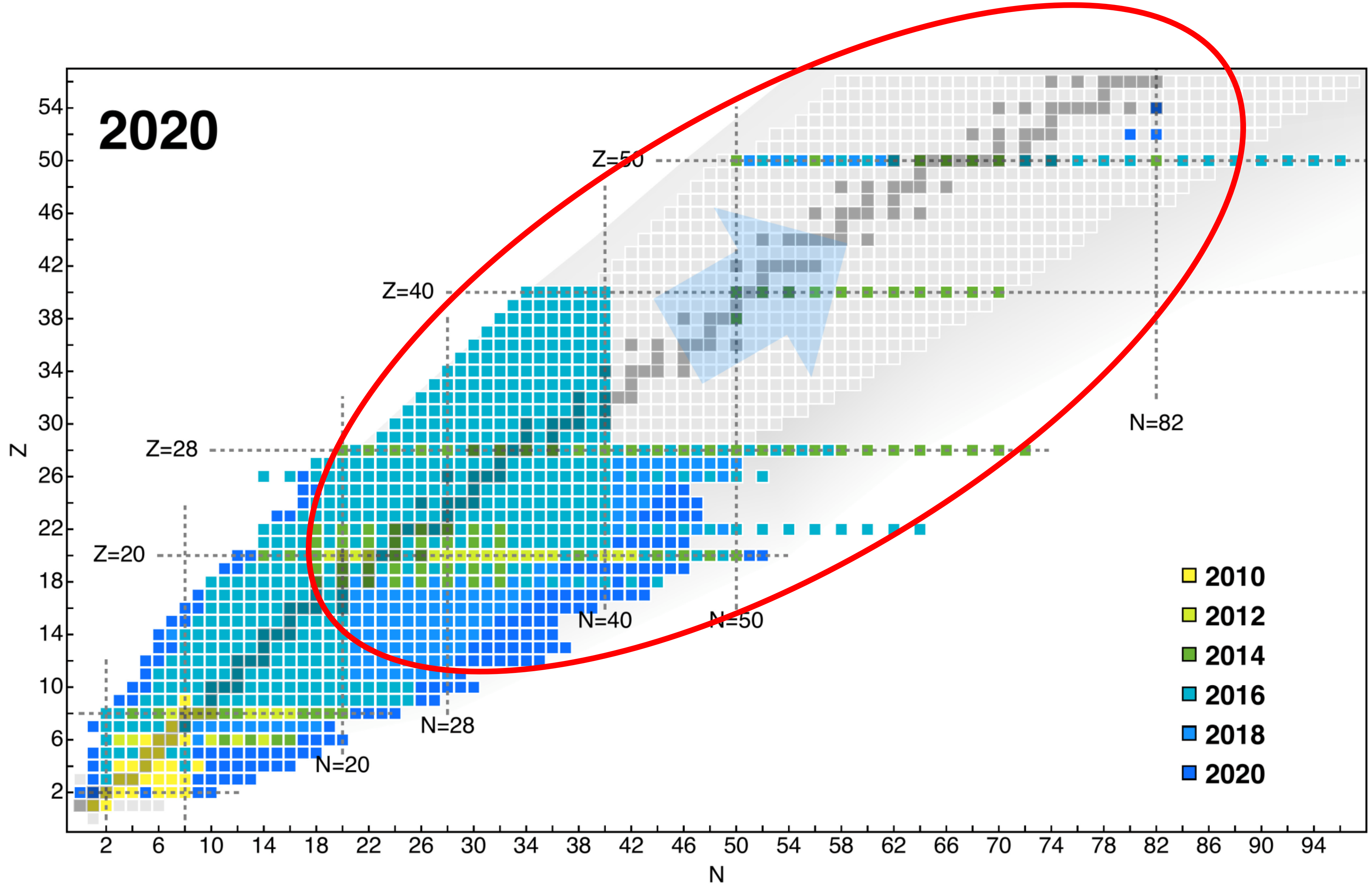
A. Gnech et al, Few-Body Syst. **63** (2022)

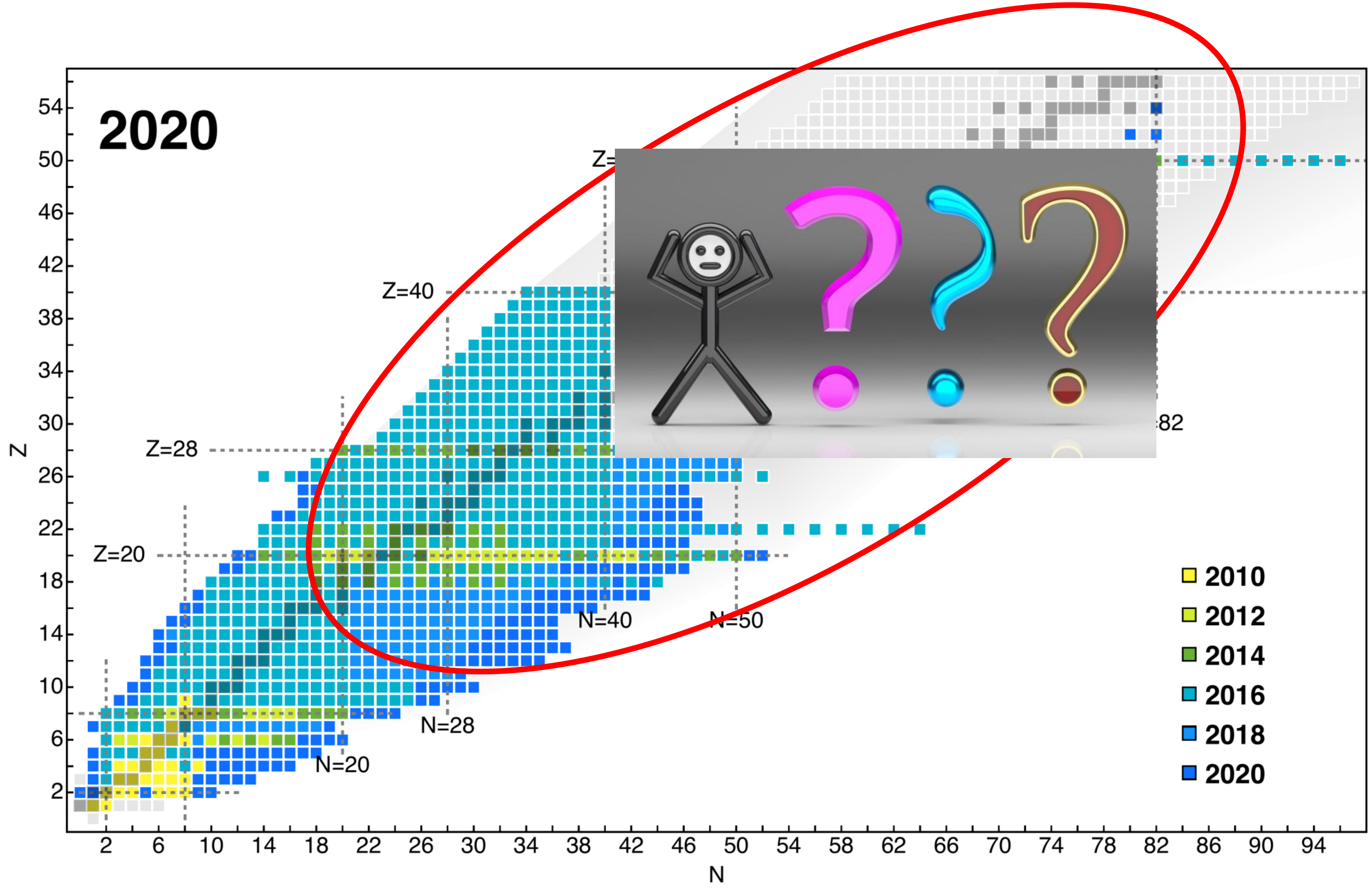


A. Gnech et al, PRL **133** (2024)



2020
Z=50
Z=40
Z=28
Z=20
N=20
N=28
N=40
N=50
N=82
2010
2012
2014
2016
2018
2020
Z
N

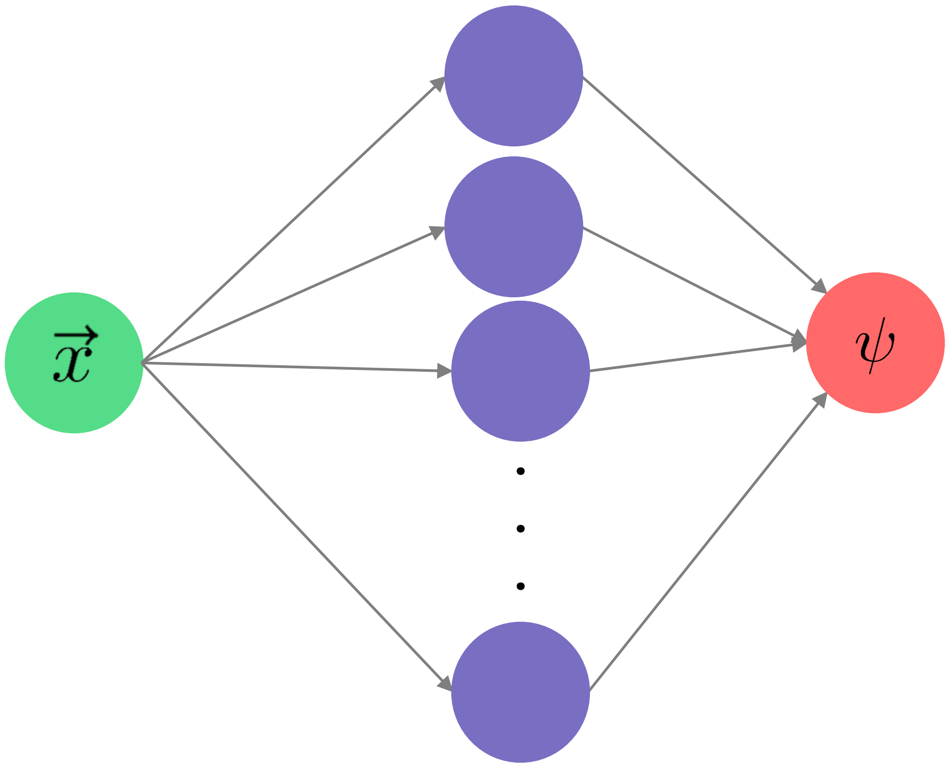

2020
Z=50
Z=40
Z=28
Z=20
N=82
N=40
N=50
N=28
N=20
2010
2012
2014
2016
2018
2020
Z
N



2020
Z=40
Z=28
Z=20
N=20
N=28
N=40
N=50
82
Z
N
2010
2012
2014
2016
2018
2020

1. Neural Network Architecture

Physical properties in NNs



- **Particle exchange symmetry**

$$\psi(x_1, x_2, \ldots, x_N) = \pm\psi(x_2, x_1, \ldots, x_N)$$

- **Spherical symmetry**

$$\psi(\vec{x}) = \psi(R\vec{x})$$

- **Time-reversal symmetry**

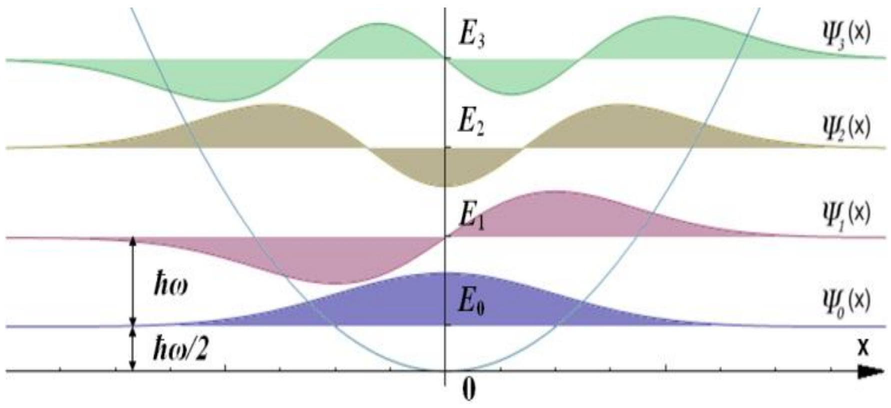
$$\psi(\vec{x}, \sigma) = \psi(T(\vec{x}, \sigma))$$

Physical properties in NNs

- **1D Harmonic Oscillator**

$$\hat{H} = \frac{1}{2}\nabla^2 + \frac{1}{2}mw^2x^2$$

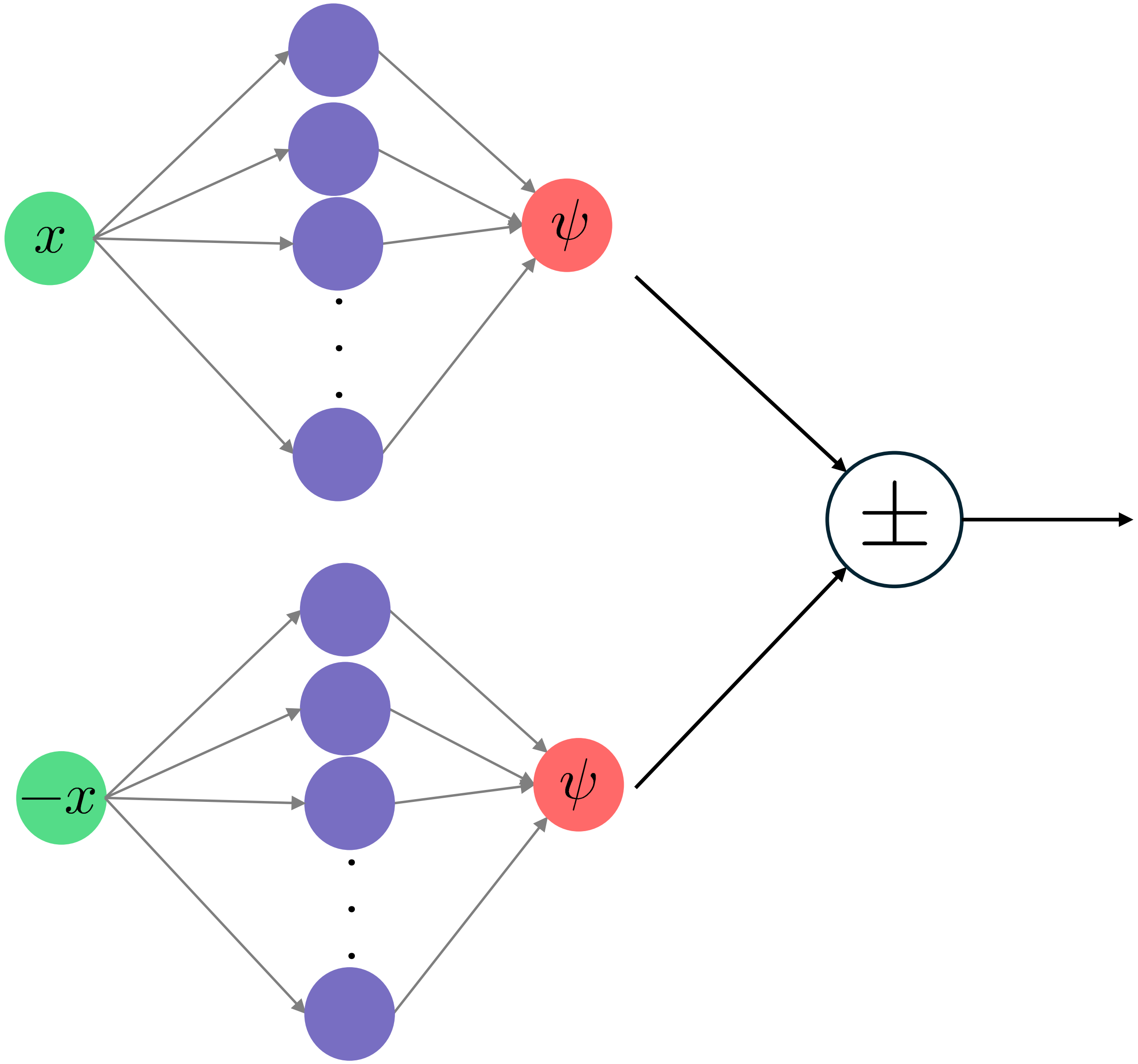
$$\Rightarrow [\hat{H},\ \hat{P}] = 0,\ \hat{P}\quad \text{parity}$$



- **Reflection symmetry**

$$\begin{cases} \psi_+(x) := \psi(x) + \psi(-x) \\ \psi_-(x) := \psi(x) - \psi(-x) \end{cases}$$

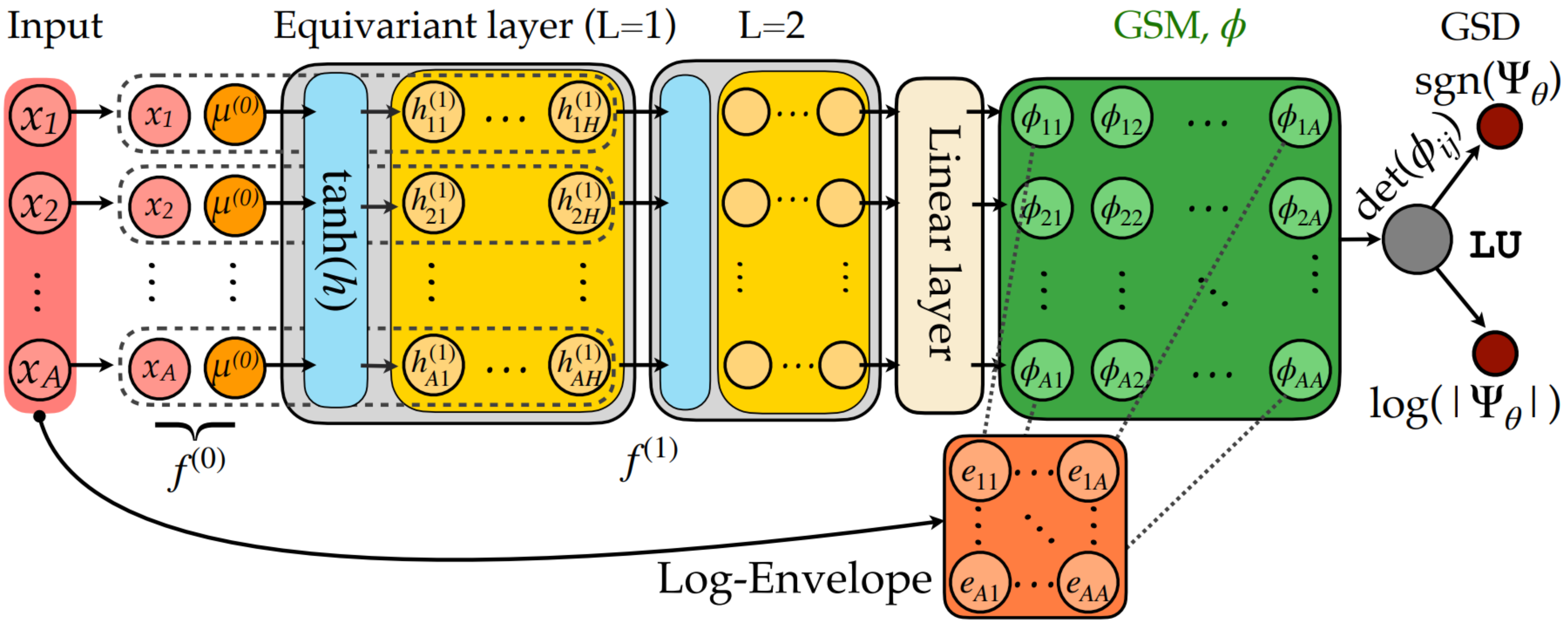
Reflection Symmetry



$$\psi_{\pm}(x) := \psi(x) \pm \psi(-x)$$

Particle Exchange Symmetry

$$\psi_{\text{NQS}}(x_1, x_2, \ldots, x_N) = \varphi_{\text{EQUIV}} \circ \det\, \phi_{\text{GSM}}(x_1, x_2, \ldots, x_N)$$



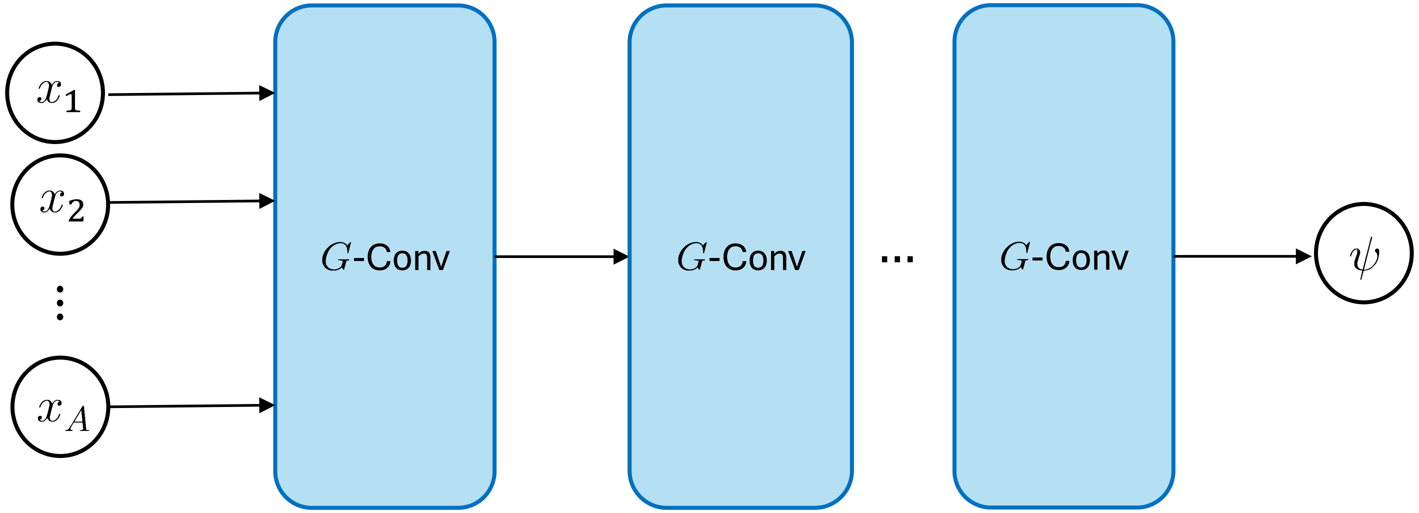
J.W.T Keeble, M. Drissi , A. Rojo-Francàs, B. Juliá-Díaz & A. Rios, Phys. Rev. A. **108** (2023)

Towards a universal recipe…

- **Continous groups:** what about Lie groups? $SU(2)$ $\mathrm{SO}(3)$
- **Mixed groups:** what about product groups? $SU(2) \times S_N$

Towards a universal recipe...

- **Continous groups:** what about Lie groups? $SU(2)$ $SO(3)$
- **Mixed groups:** what about product groups? $SU(2) \times S_N$

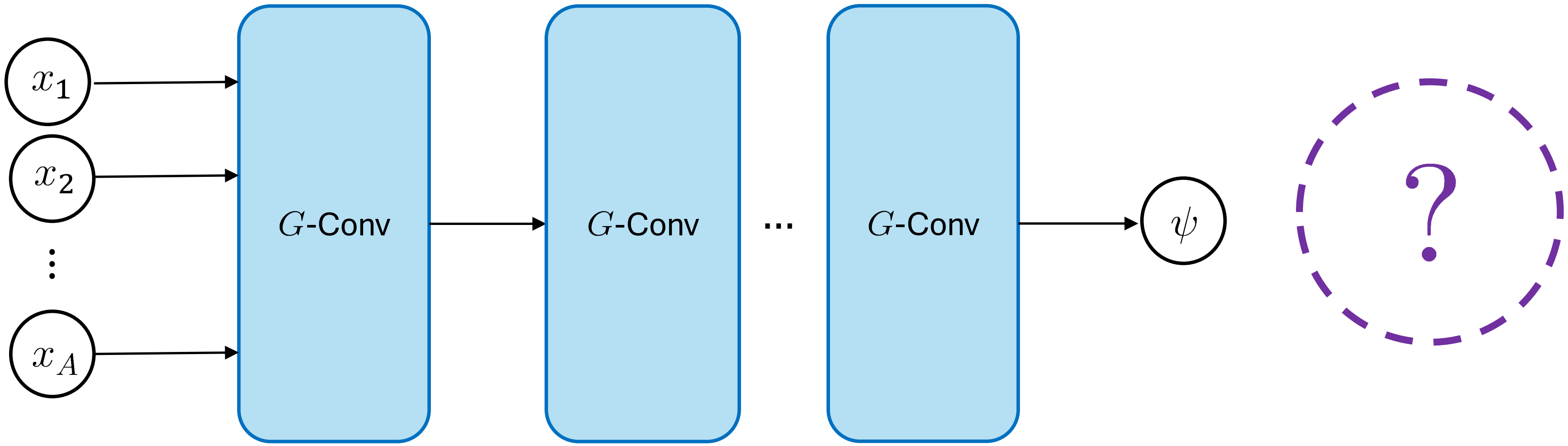


R. Kondor & S. Trivedi, arXiv:1802.03690

Towards a universal recipe...

- **Continous groups:** what about Lie groups? $SU(2)$ $SO(3)$

- **Mixed groups:** what about product groups? $SU(2) \times S_N$

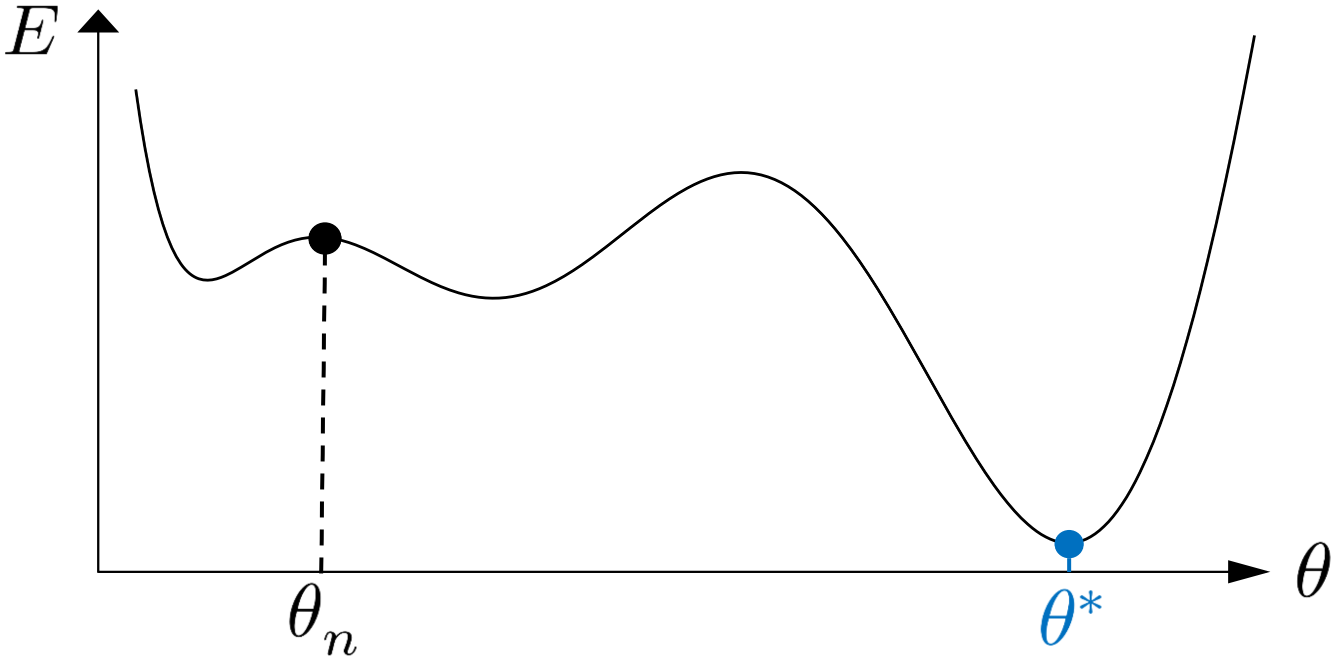


R. Kondor & S. Trivedi, arXiv:1802.03690

2. Optimisation Strategy

Gradient Descent

- **Local Linear Model:** $M_n(\delta) = L^T \delta + C$
- **Update:** $\theta_{n+1} = \theta_n + \delta_n$ where $\delta_n \equiv \mathrm{argmin}_{\delta \in T_n} M_n(\delta)$

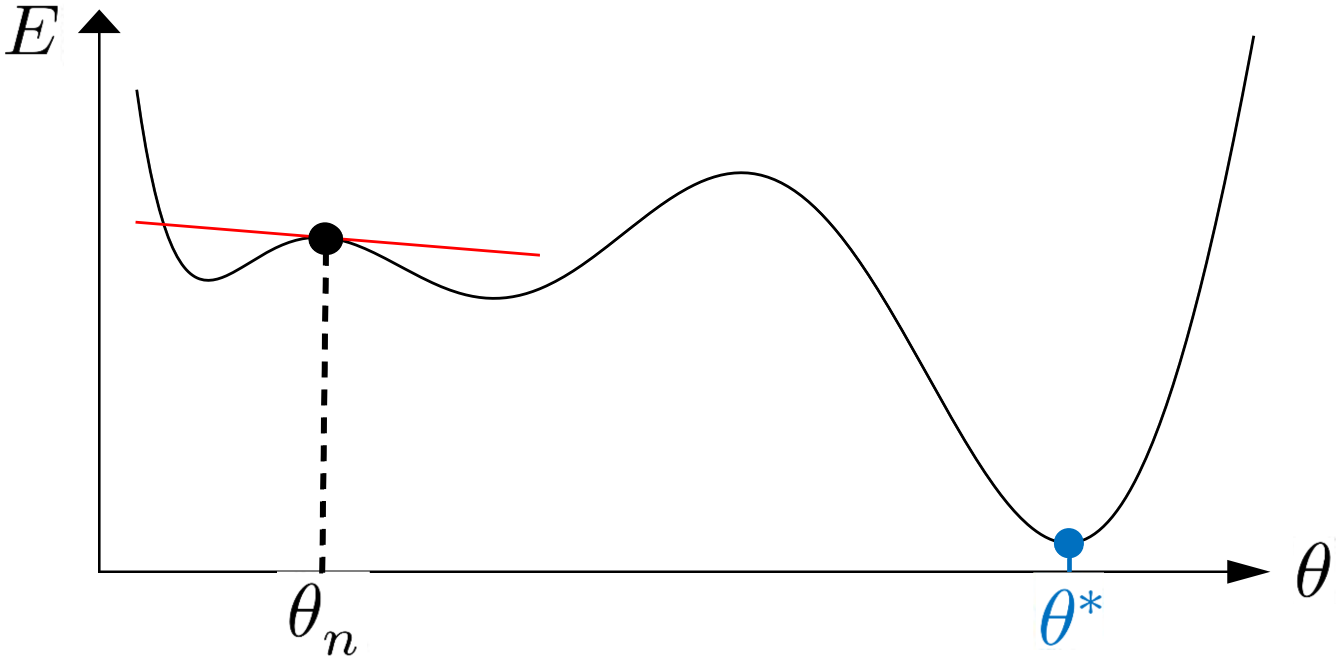


Gradient Descent

- **Local Linear Model:** $M_n(\delta) = L^T\delta + C$
- **Update:** $\theta_{n+1} = \theta_n + \delta_n$ where $\delta_n \equiv \text{argmin}_{\delta \in T_n} M_n(\delta)$

$$\delta_n = -\alpha\nabla E(\theta_n)$$

$$\theta_{n+1} = \theta_n - \alpha\nabla E(\theta_n)$$

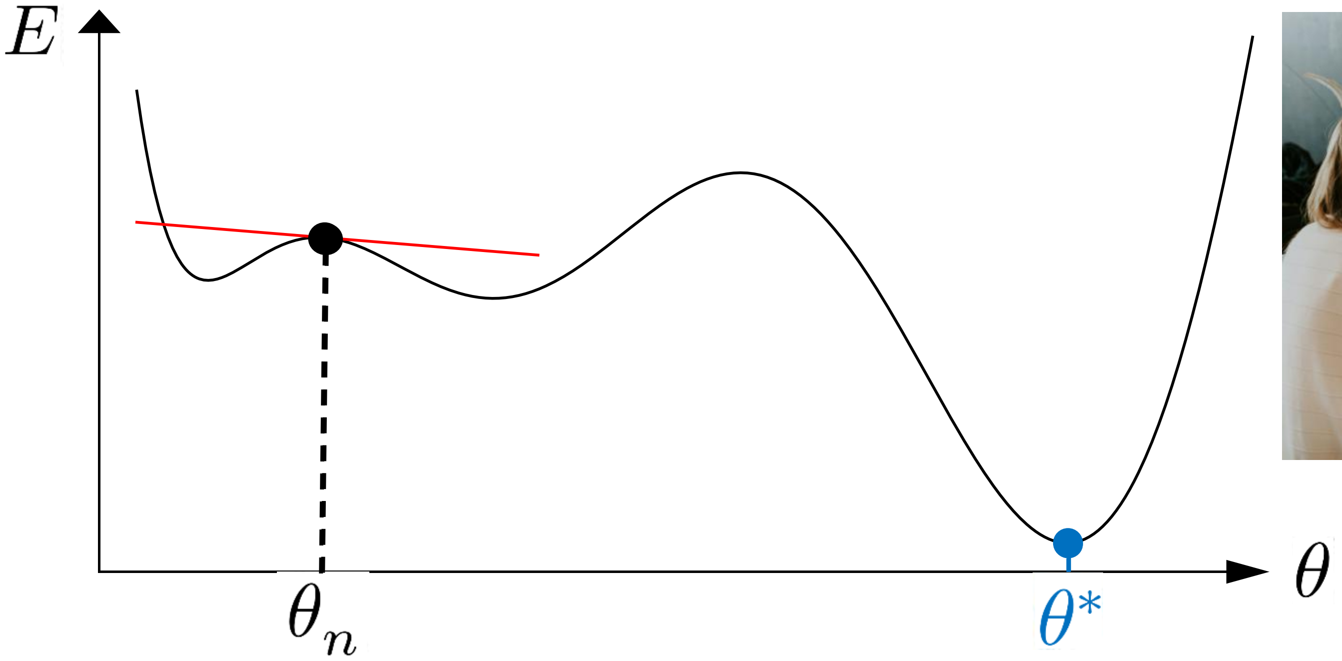


Gradient Descent

- **Local Linear Model:** $M_n(\delta) = L^T\delta + C$
- **Update:** $\theta_{n+1} = \theta_n + \delta_n$ where $\delta_n \equiv \text{argmin}_{\delta \in T_n} M_n(\delta)$

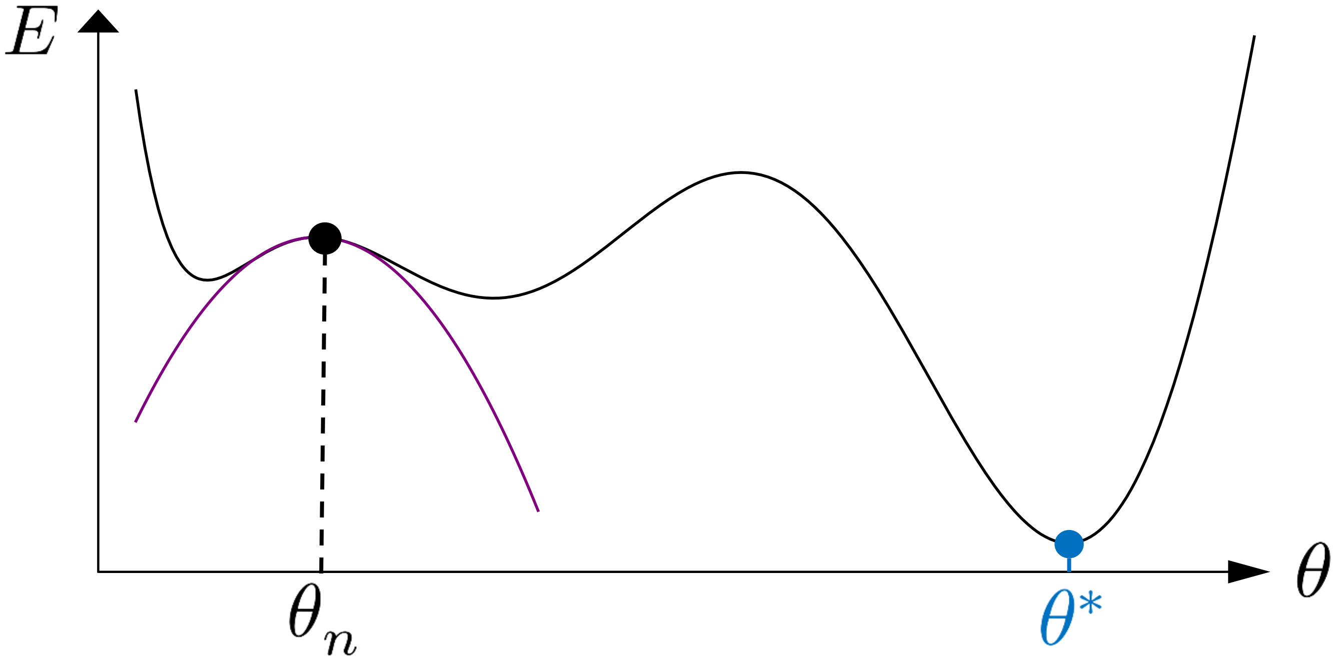
$$\delta_n = -\alpha \nabla E(\theta_n)$$

$$\theta_{n+1} = \theta_n - \alpha \nabla E(\theta_n)$$



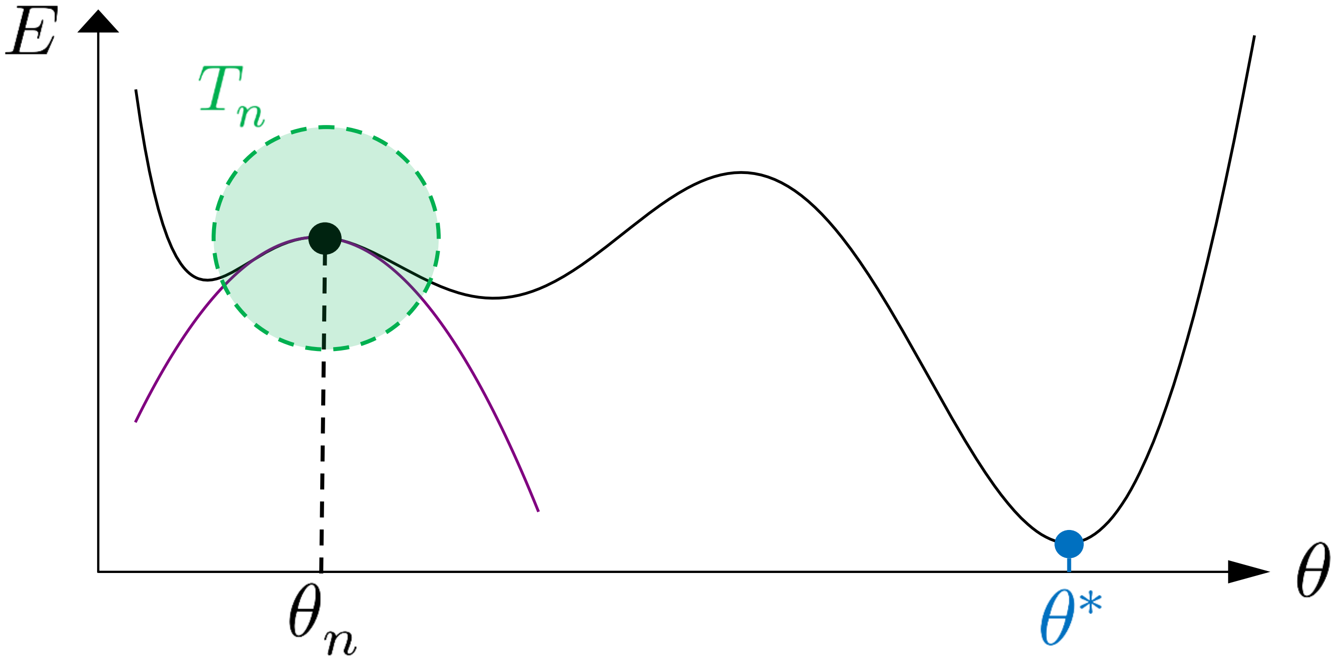
Second-order optimisation

- **Local Quadratic Model:** $M_n(\delta) = \frac{1}{2}\delta^T Q \delta + L^T \delta + C$
- **Trust region:** $T_n = \{\delta \colon \delta^T R_n \delta \leq r^2\}$
- **Update:** $\theta_{n+1} = \theta_n + \delta_n$ where $\delta_n \equiv \mathrm{argmin}_{\delta \in T_n} M_n(\delta)$



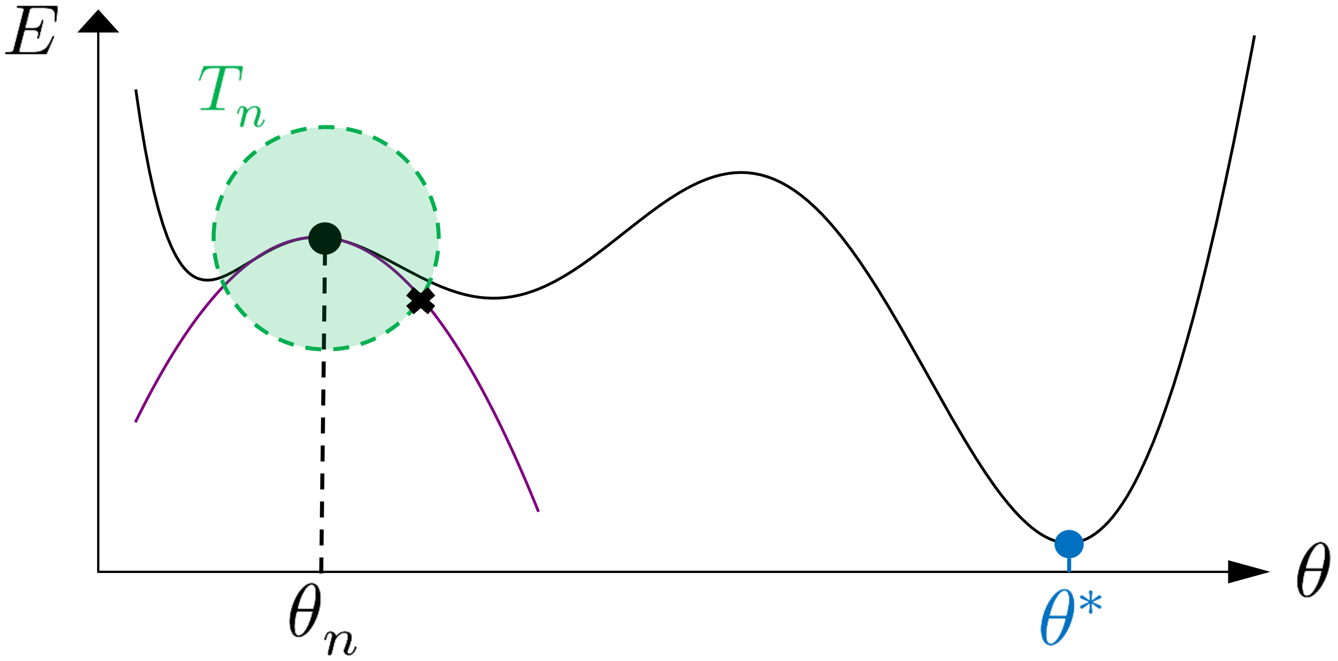
Second-order optimisation

- **Local Quadratic Model:** $M_n(\delta) = \frac{1}{2}\delta^T Q \delta + L^T \delta + C$
- **Trust region:** $T_n = \{\delta \colon \delta^T R_n \delta \leq r^2\}$
- **Update:** $\theta_{n+1} = \theta_n + \delta_n$ where $\delta_n \equiv \text{argmin}_{\delta \in T_n} M_n(\delta)$



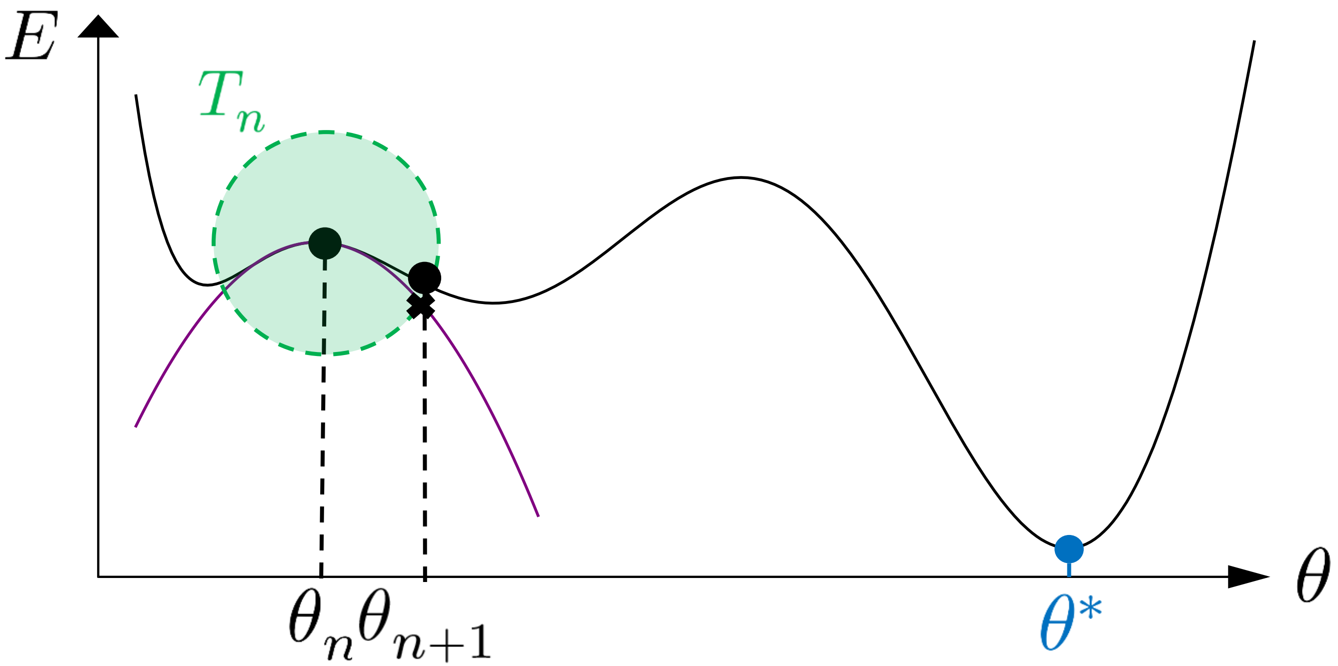
Second-order optimisation

- **Local Quadratic Model:** $M_n(\delta) = \frac{1}{2}\delta^T Q \delta + L^T \delta + C$
- **Trust region:** $T_n = \{\delta : \delta^T R_n \delta \leq r^2\}$
- **Update:** $\theta_{n+1} = \theta_n + \delta_n$ where $\delta_n \equiv \mathrm{argmin}_{\delta \in T_n} M_n(\delta)$



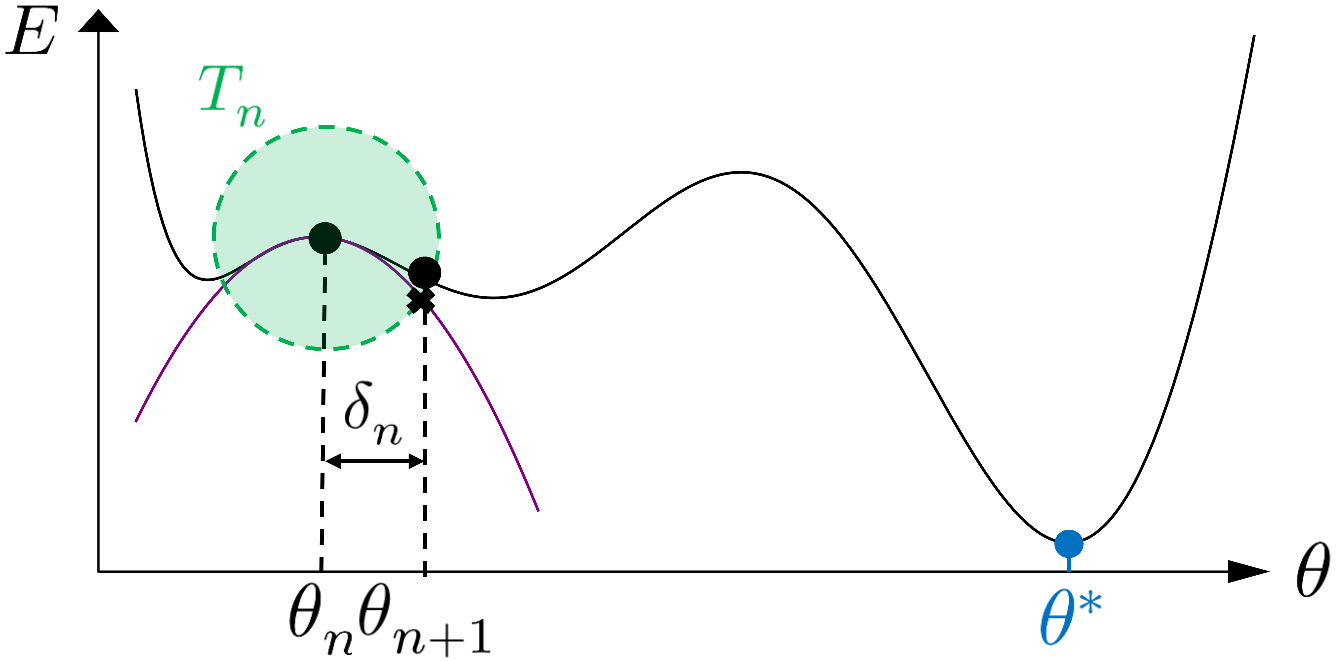
Second-order optimisation

- **Local Quadratic Model:** $M_n(\delta) = \frac{1}{2}\delta^T Q \delta + L^T \delta + C$
- **Trust region:** $T_n = \{\delta \colon \delta^T R_n \delta \leq r^2\}$
- **Update:** $\theta_{n+1} = \theta_n + \delta_n$ where $\delta_n \equiv \operatorname{argmin}_{\delta \in T_n} M_n(\delta)$



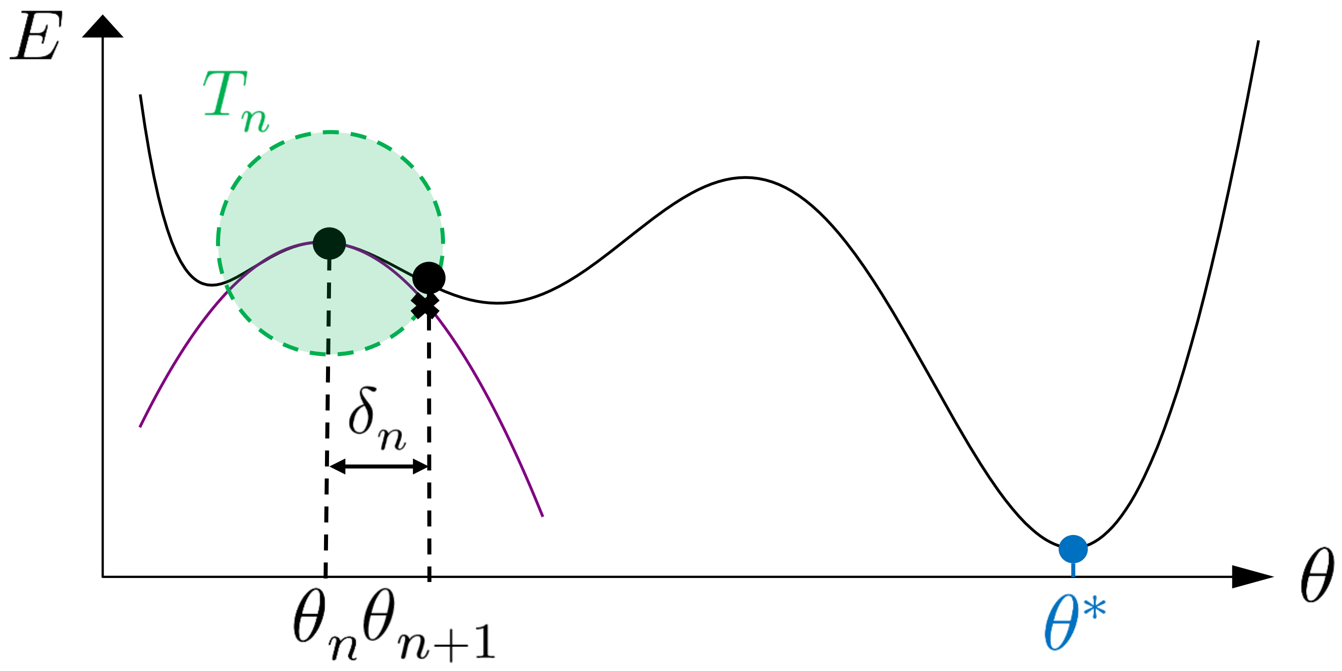
Second-order optimisation

- **Local Quadratic Model:** $M_n(\delta) = \frac{1}{2}\delta^T Q \delta + L^T \delta + C$
- **Trust region:** $T_n = \{\delta : \delta^T R_n \delta \leq r^2\}$
- **Update:** $\theta_{n+1} = \theta_n + \delta_n$ where $\delta_n \equiv \text{argmin}_{\delta \in T_n} M_n(\delta)$



Second-order optimisation

- **Local Quadratic Model:** $M_n(\delta) = \frac{1}{2}\delta^T Q \delta + L^T \delta + C$
- **Trust region:** $T_n = \{\delta : \delta^T R_n \delta \leq r^2\}$
- **Update:** $\theta_{n+1} = \theta_n + \delta_n$ where $\delta_n = -(Q_n + \lambda_n R_n)^{-1} \nabla E(\theta_n)$

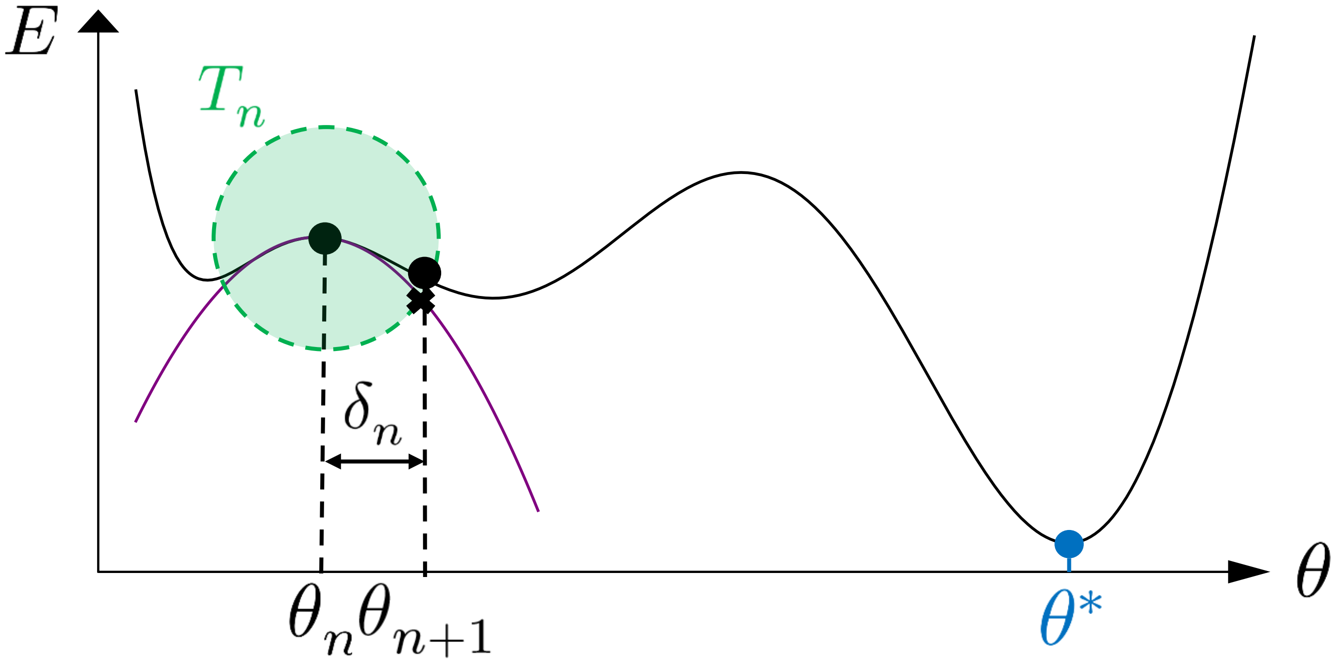


Second-order optimisation

- **Local Quadratic Model:** $M_n(\delta) = \frac{1}{2}\delta^T Q \delta + L^T \delta + C$

$$L = \nabla_\theta E$$

- **Trust region:** $T_n = \{\delta : \delta^T R_n \delta \leq r^2\}$
- **Update:** $\theta_{n+1} = \theta_n + \delta_n$ where $\delta_n = -(Q_n + \lambda_n R_n)^{-1} \nabla E(\theta_n)$

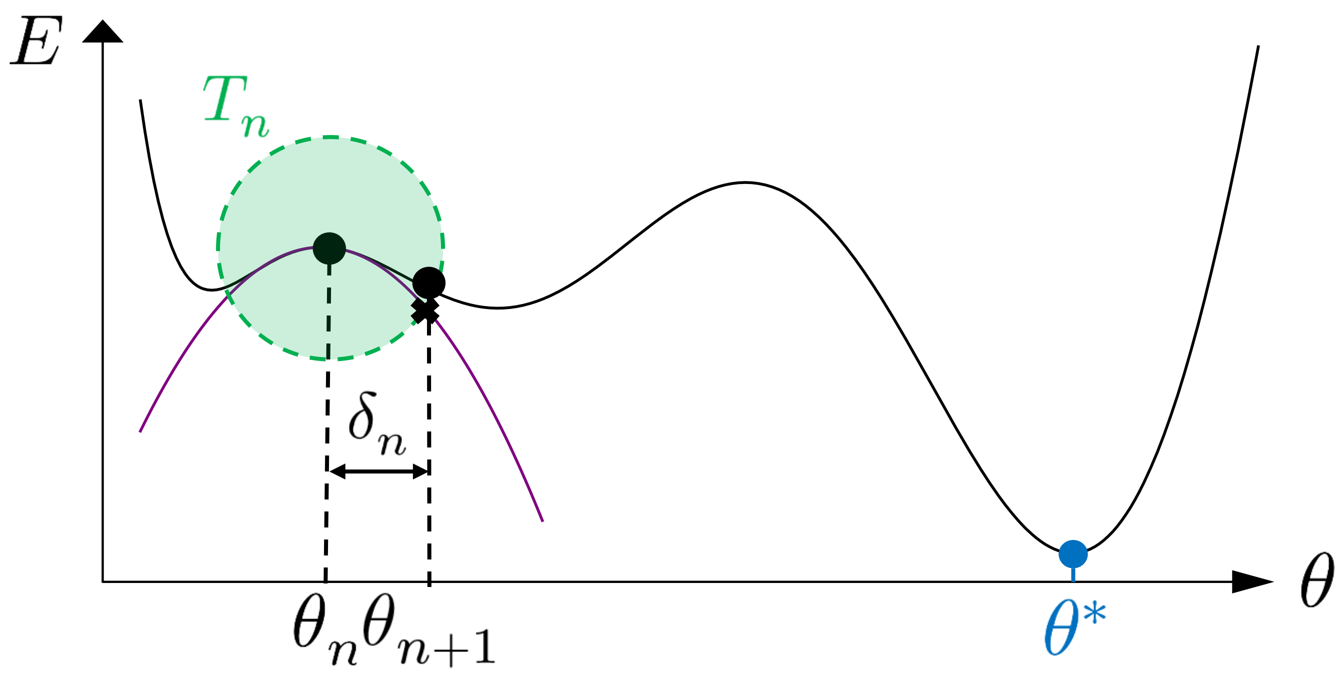


Second-order optimisation

- **Local Quadratic Model:** $M_n(\delta) = \frac{1}{2}\delta^T Q \delta + L^T \delta + C$
- **Trust region:** $T_n = \{\delta : \delta^T R_n \delta \le r^2\}$
- **Update:** $\theta_{n+1} = \theta_n + \delta_n$ where $\delta_n = -\left(Q_n + \lambda_n R_n\right)^{-1} \nabla E(\theta_n)$

$L = \nabla_\theta E$

$Q?$



Stochastic Reconfiguration

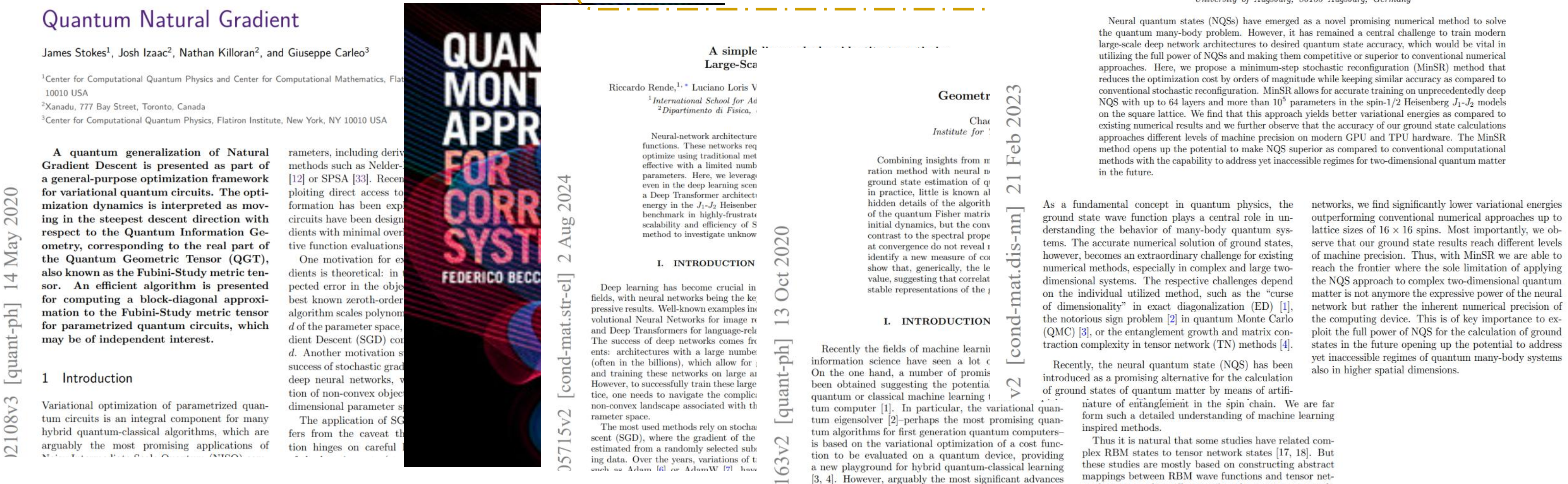
"Quantum Geometric Tensor"

- **Local Quadratic Model:** $M_n(\delta) = \frac{1}{2}\delta^T Q \delta + L^T \delta + C, \quad Q_{ij}(\theta) = \left\langle \frac{\partial \psi_\theta}{\partial \theta_i}, \frac{\partial \psi_\theta}{\partial \theta_j} \right\rangle - \left\langle \frac{\partial \psi_\theta}{\partial \theta_i}, \psi_\theta \right\rangle \left\langle \psi_\theta, \frac{\partial \psi_\theta}{\partial \theta_j} \right\rangle$
- **Trust region:** $T_n = \{\delta : \delta^T R_n \delta \leq r^2\}, \quad R_n = 1$
- **Update:** $\theta_{n+1} = \theta_n + \delta_n$ where $\delta_n = -(Q_n + \lambda_n R_n)^{-1} \nabla E(\theta_n)$

Stochastic Reconfiguration

"Quantum Geometric Tensor"

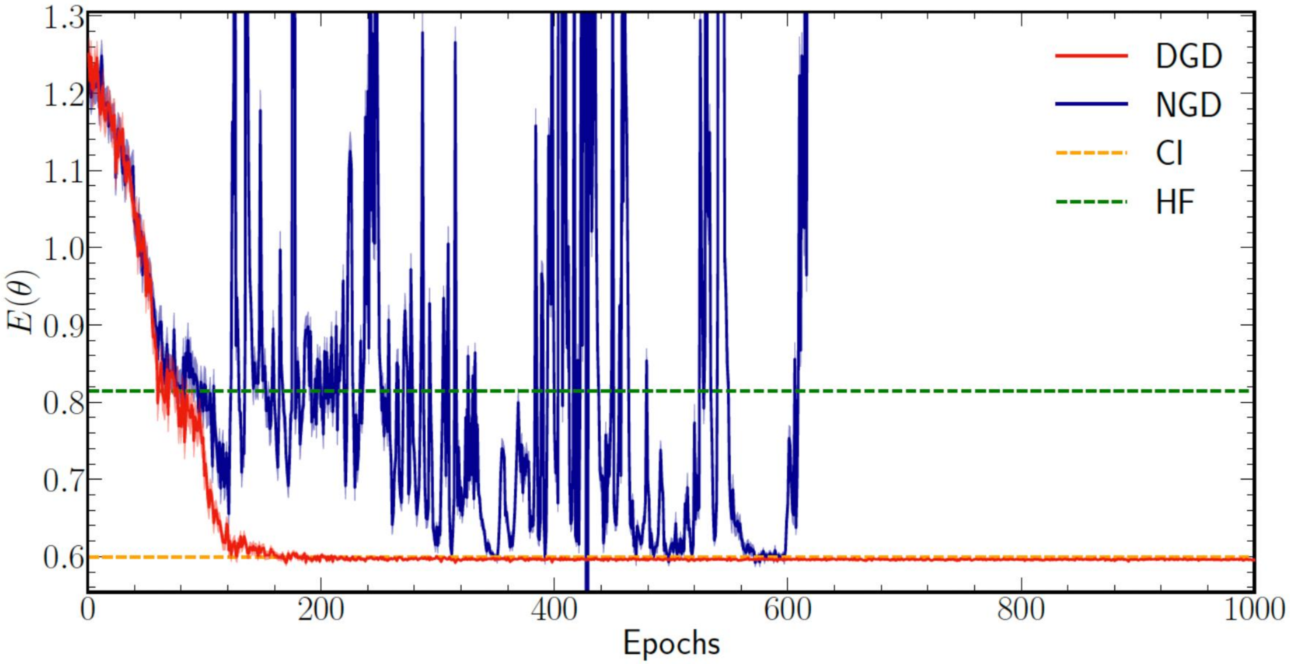
- **Local Quadratic Model:** $M_n(\delta) = \frac{1}{2}\delta^T Q \delta + L^T \delta + C,$ $\quad Q_{ij}(\theta) = \left\langle \frac{\partial \psi_\theta}{\partial \theta_i}, \frac{\partial \psi_\theta}{\partial \theta_j} \right\rangle - \left\langle \frac{\partial \psi_\theta}{\partial \theta_i}, \psi_\theta \right\rangle \left\langle \psi_\theta, \frac{\partial \psi_\theta}{\partial \theta_j} \right\rangle$
- **Trust region:** $T_n = \{\delta : \delta^T R_n \delta \leq r^2\}, \quad R_n = 1$
- **Update:** $\theta_{n+1} = \theta_n + \delta_n$ where $\delta_n = -(Q_n + \lambda_n R_n$



Decisional Gradient Descent

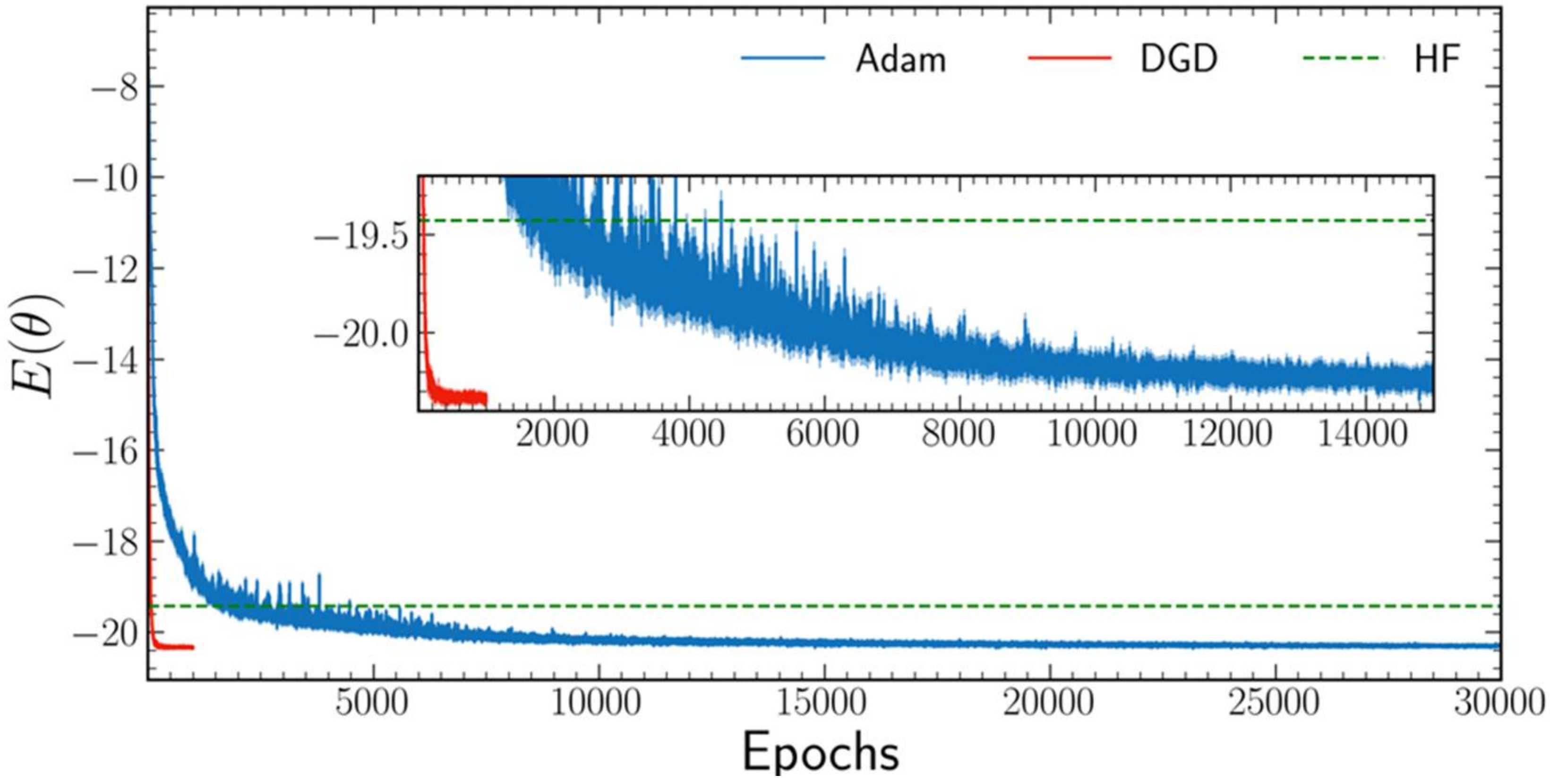
- **System:** chain of **1D spin-polarized fermions** $H = -\frac{1}{2}\sum_{i=1}^{A}\nabla_i^2 + \frac{1}{2}\sum_{i=1}^{A} x_i^2 + \frac{V_0}{\sqrt{2\pi}\sigma_0}\sum_{i<j} e^{-\frac{(x_i - x_j)^2}{2\sigma_0^2}}$

- **Quadratic term:** $\mathrm{Q}_{\mathrm{VMC}}(\theta)_{\theta_i\theta_j} \equiv \frac{1}{4}\sum_{k=1}^{A} E_{X\sim p_\theta}\left[\partial_{\theta_i}\partial_{x_k}\ln p_\theta(X)\partial_{\theta_j}\partial_{x_k}\ln p_\theta(X)\right]$



M. Drissi, J.W.T Keeble, J. Rozalén Sarmiento & A. Rios, Phil. Trans. R. Soc. A. **382** (2024)

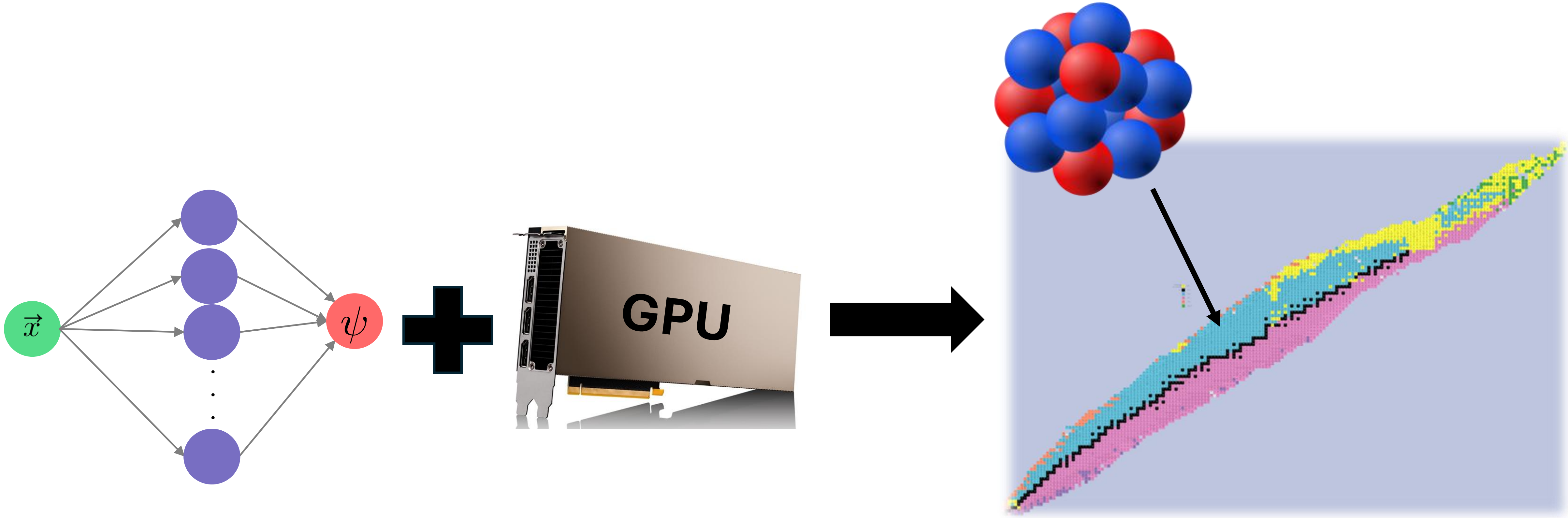
DGD vs. ADAM



M. Drissi, J.W.T Keeble, J. Rozalén Sarmiento & A. Rios, Phil. Trans. R. Soc. A. **382** (2024)

Conclusions

Takeaway



Thank you

J. Rozalén Sarmiento, A. Rios

javi.rozalen@icc.ub.edu

Institut de Ciències del Cosmos
UNIVERSITAT DE BARCELONA
EXCELENCIA MARÍA DE MAEZTU 2020-2024

UNIVERSITAT DE BARCELONA