Surrogate models of nuclear Density Functional Theory with autoencoders

Decoding nuclear fission

LLNL: M Verriere, N. Schunck, I. Kim, P. Marevic
CEA, France: R.D. Lasseri, D. Regnier, M. Frosini

May 24th, 2023

LLNL-PRES-849489

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

Lawrence Livermore National Laboratory

Nuclear fission is a key ingredient in a vast range of applications.

CHEMISTRY

The Quest for Superheavy Elements and the Island of Stability

A race is on to create the world's heaviest elements—and to explore the periodic table's "island of stability," where these elements exist for more than a moment

By Christoph E. Düllmann, Michael Block on March 1, 2018

Scientific American

FORBES > INNOVATION > SCIENCE

60 Years Of Starstuff: How Humanity Discovered Where Our Elements Come From

Starts With A Bang Contributor
Starts With A Bang Contributor Group

Follow

POST WRITTEN BY

Paul Halpern

Physicist Paul Halpern is author of The Quantum Labyrinth: How Richard Feynman and John Wheeler Revolutionized Time and Reality.

Forbes

Apr 18, 2017, 10:00am EDT

Nuclear Plant $16 Billion Over Budget Arrives for Atomic Revival

The first newly constructed US reactors in decades will make Vogtle the biggest energy generator in the country

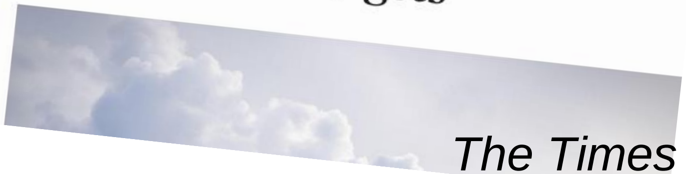
By Will Wade and Josh Saul
Graphics by Dave Merrill
May 15, 2023 at 5:00 AM PDT

Bloomberg

Scots back nuclear power to help meet net zero targets

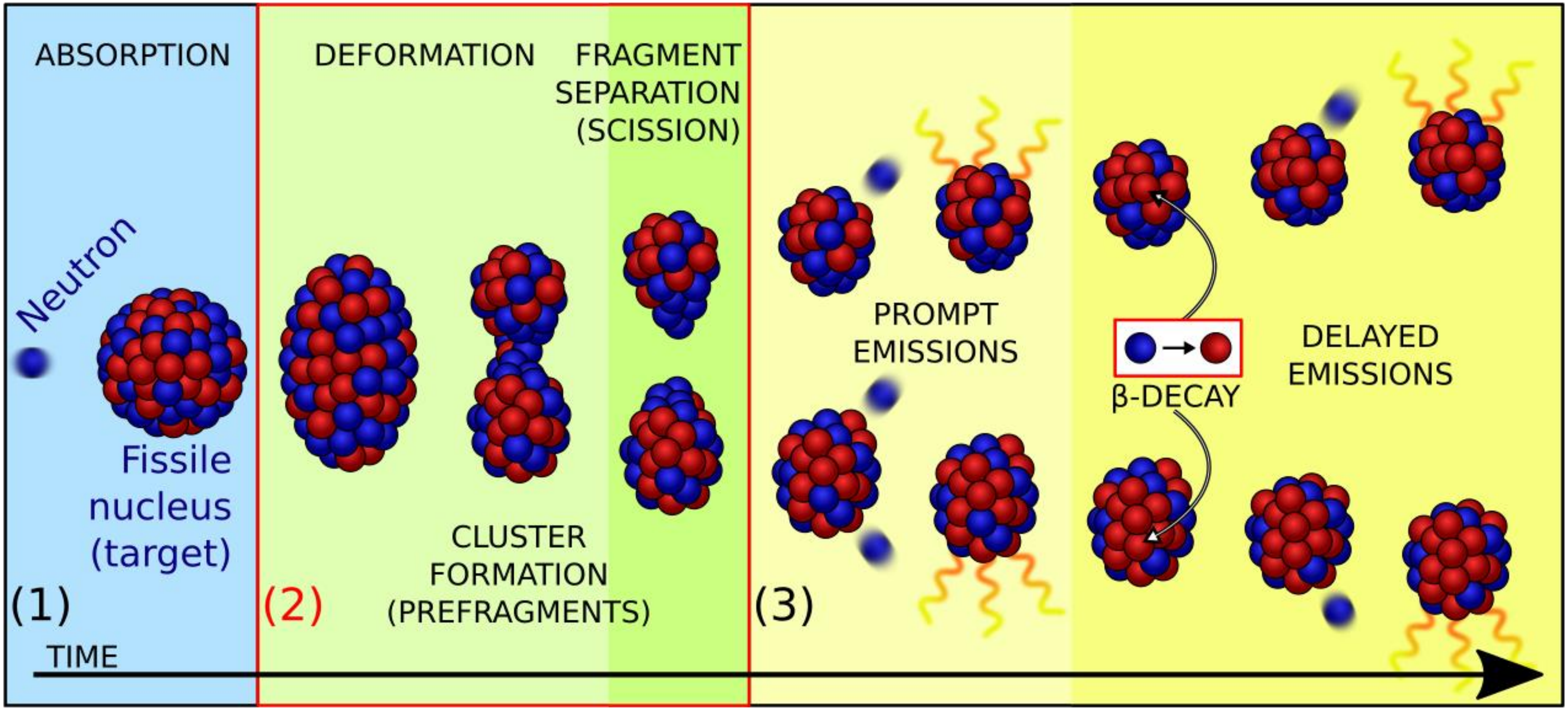
Greig Cameron

Tuesday May 16 2023, 12.01am BST, The Times

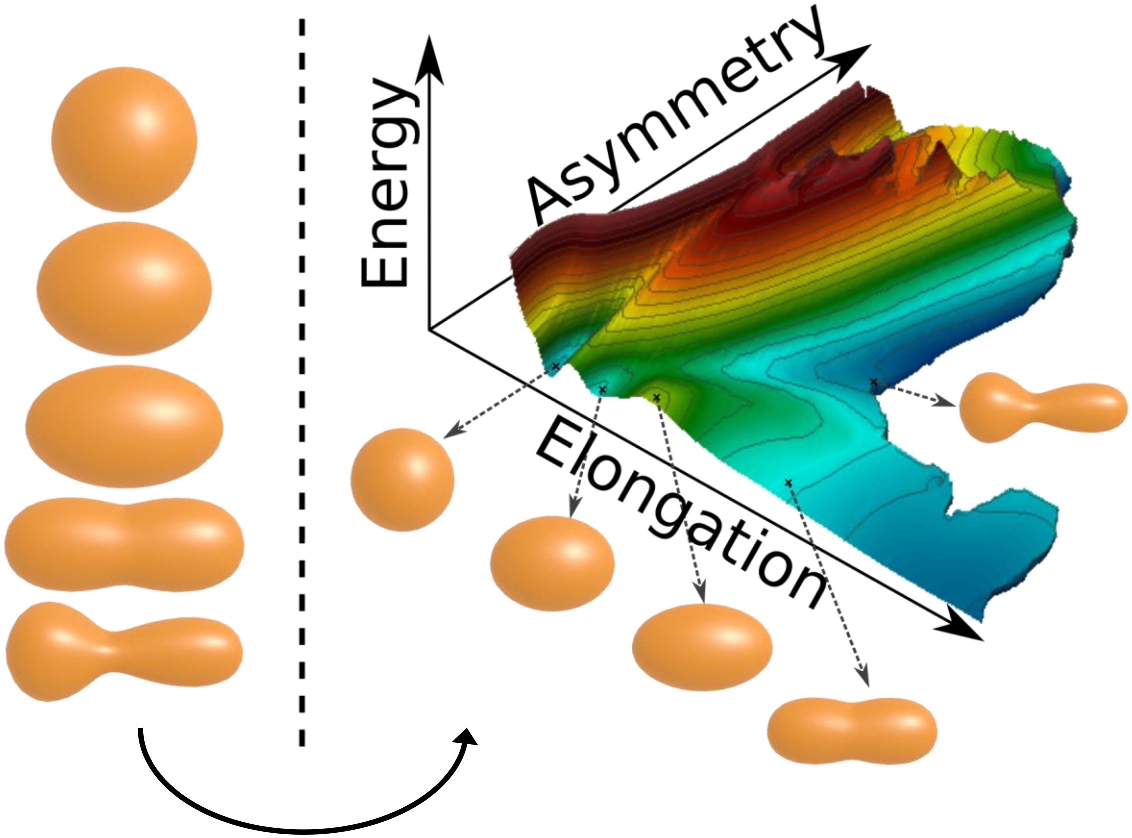


The Times

We aim at describing the various steps of nuclear fission.



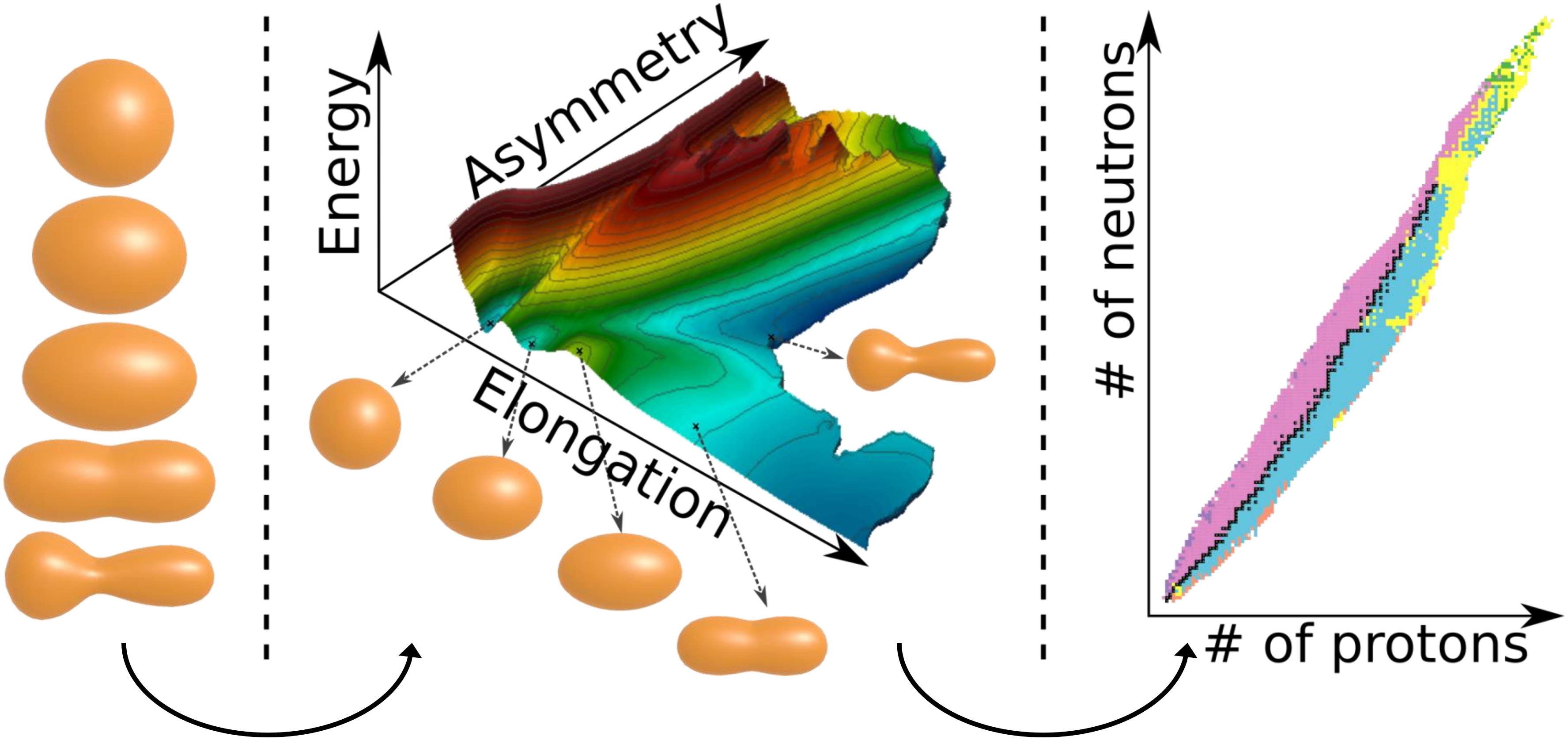
We use a fully microscopic approach to describe nuclei: the nuclear Density Functional Theory.



Minimization of the energy for each shape.
× 5,000 shapes (2D)
× 100,000 shapes (3D)

2D: **Potential Energy Surface (PES)**
ND: Potential Energy Landscape

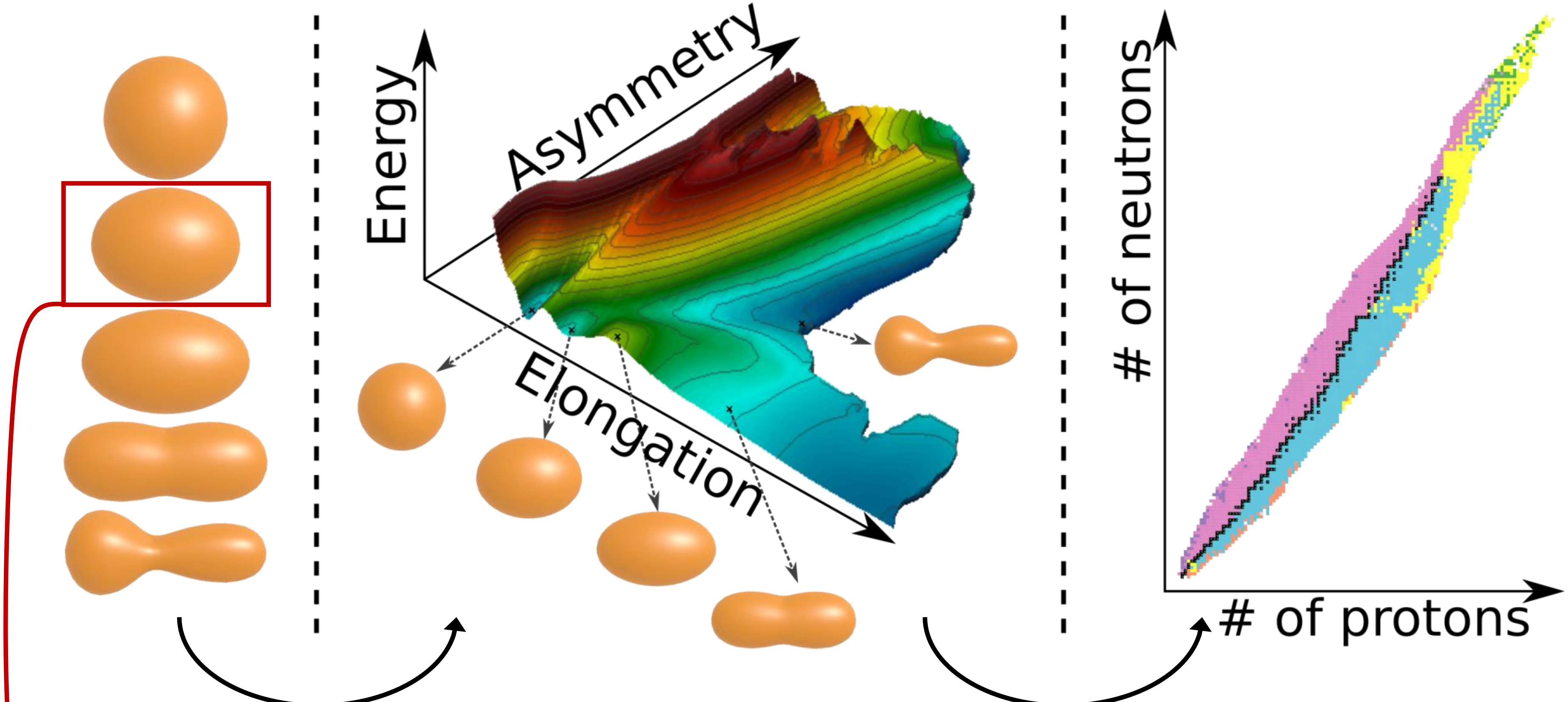
We use a fully microscopic approach to describe nuclei: the nuclear Density Functional Theory.



Minimization of the energy for each shape.
$\times$ 5,000 shapes (2D)
$\times$ 100,000 shapes (3D)

$\times$ 3,000 atomic nuclei of interest.

Fission calculations with nuclear DFT across all known nuclei remain out of reach.

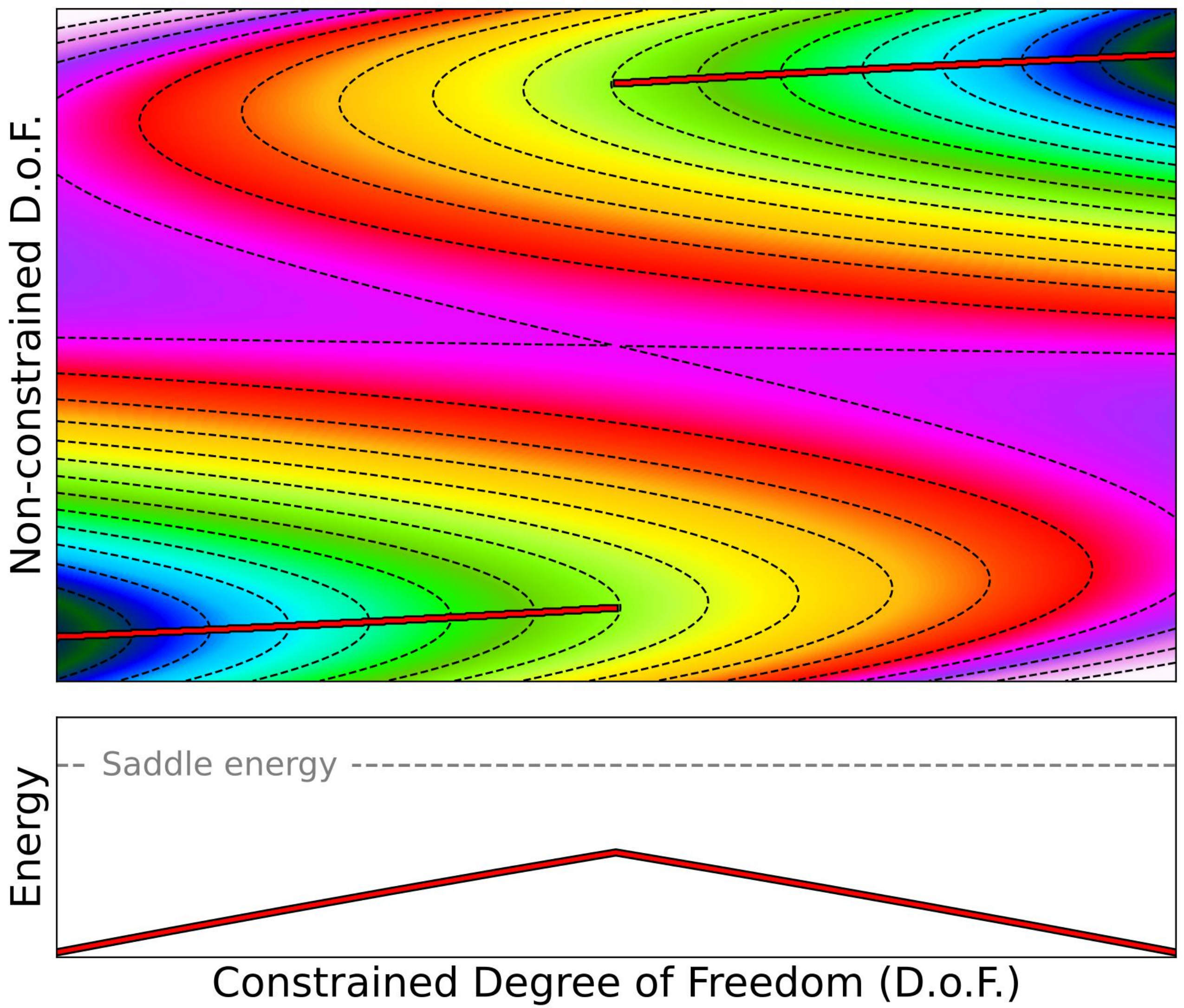


Minimization of the energy

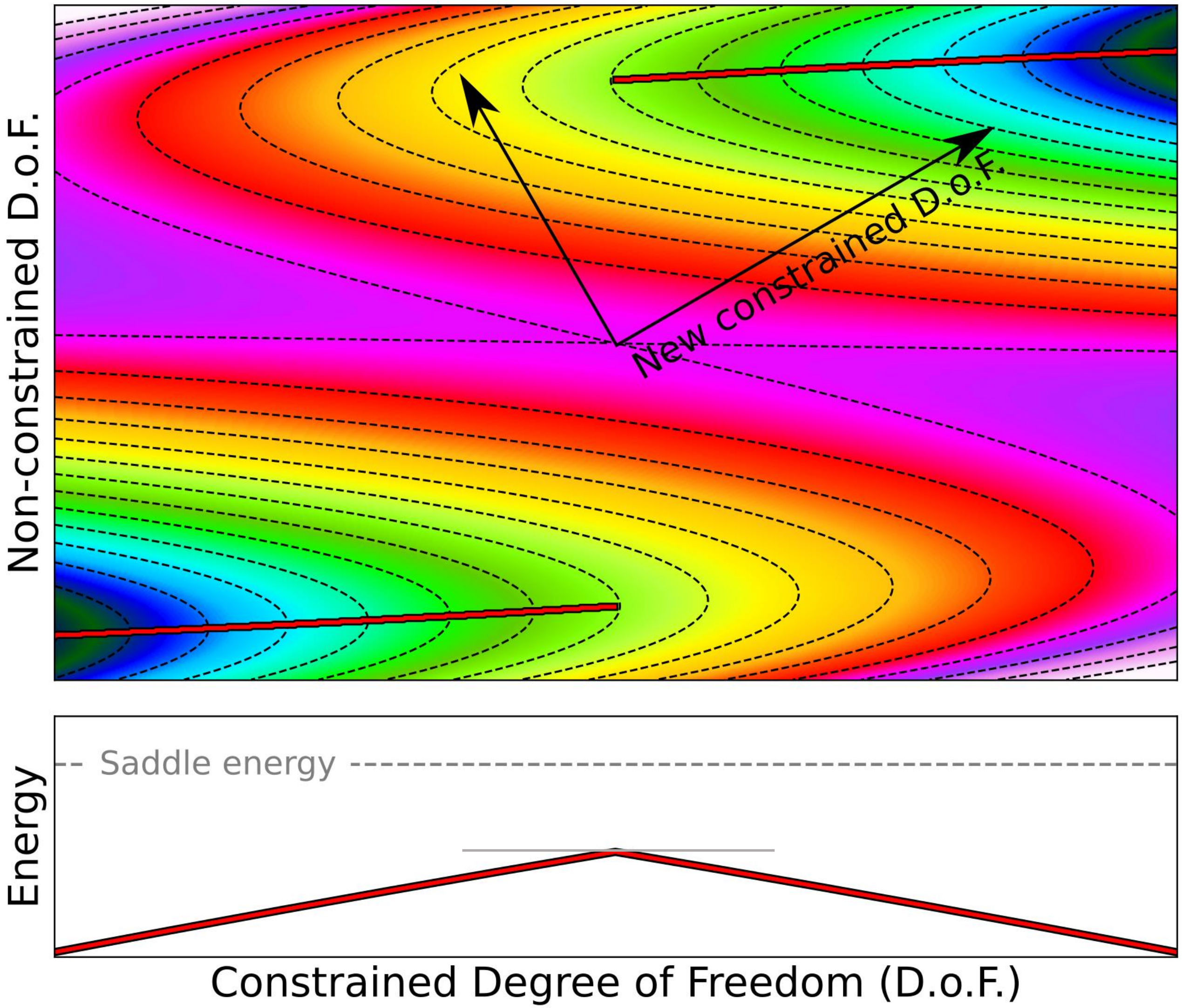
× 3,000 atomic nuclei

One shape	→	*One surface*	→	*Full calculation*
1 cpu.h		5 k cpu.h (2D)		**2,000 cpu.year (2D)**
		100 k cpu.h (3D)		**35,000 cpu.year (3D)**

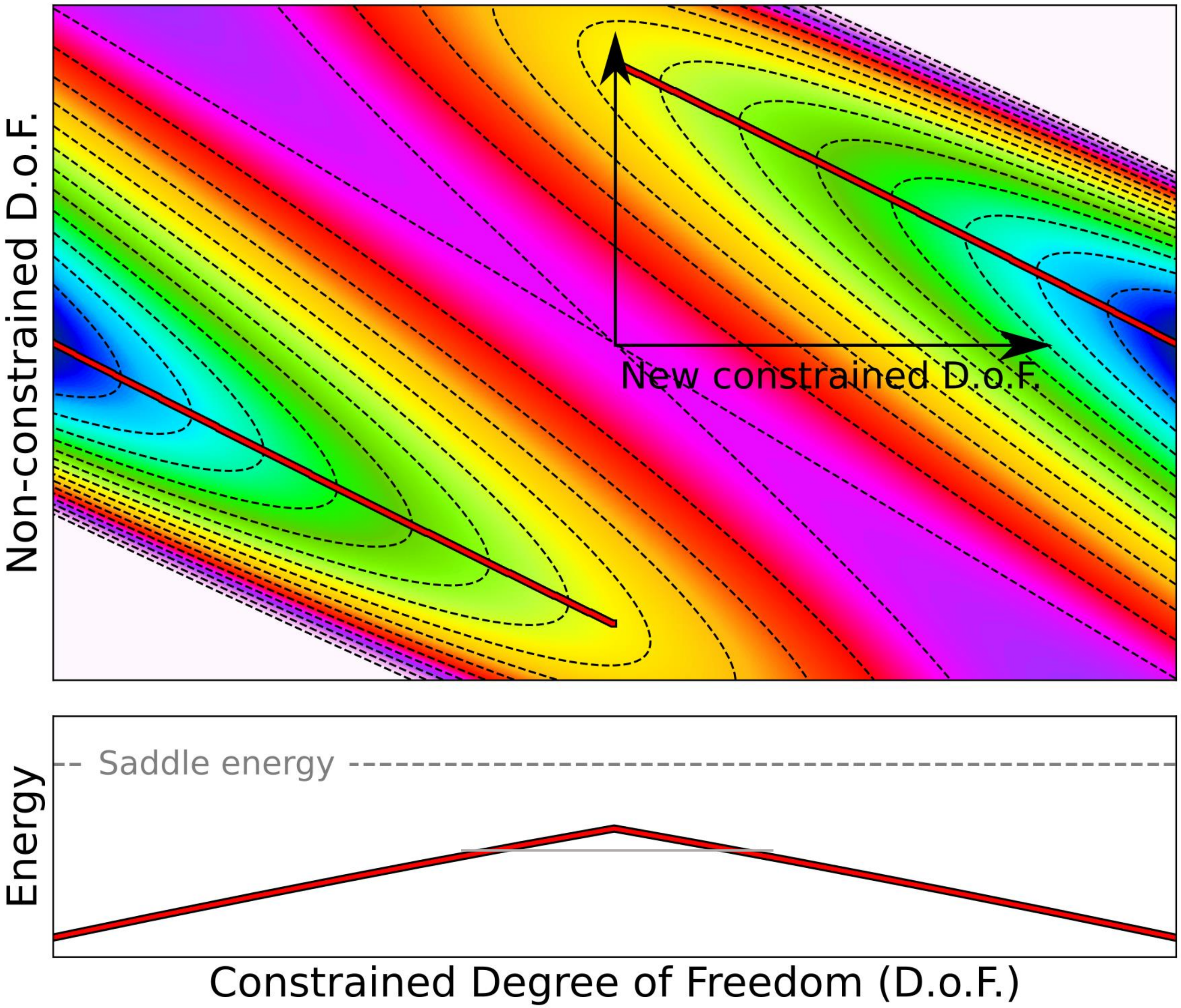
(1) Minimizing the energy leads to spurious connections between different channels.



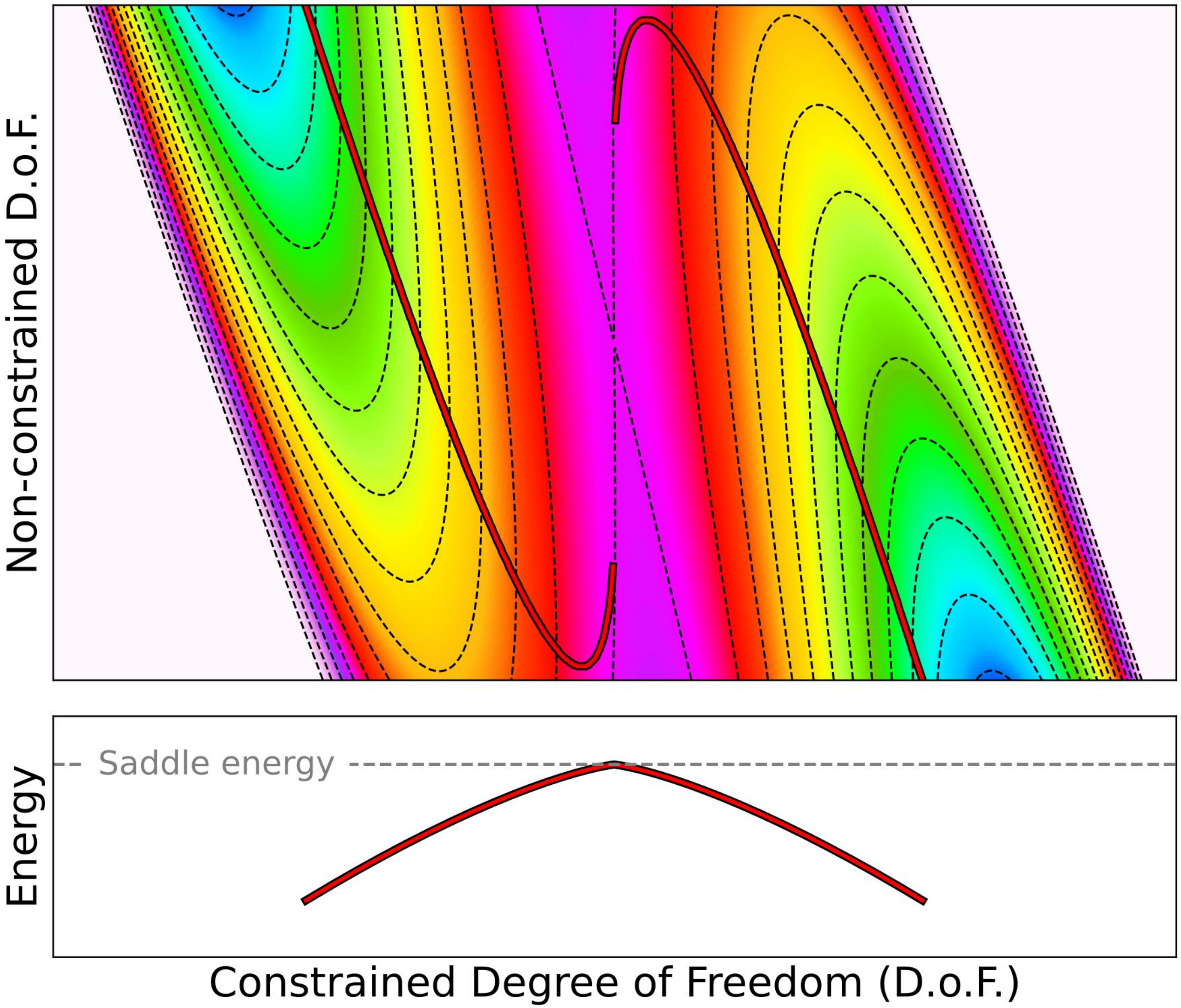
(2) The choice of constrained degrees of freedom is somewhat arbitrary.



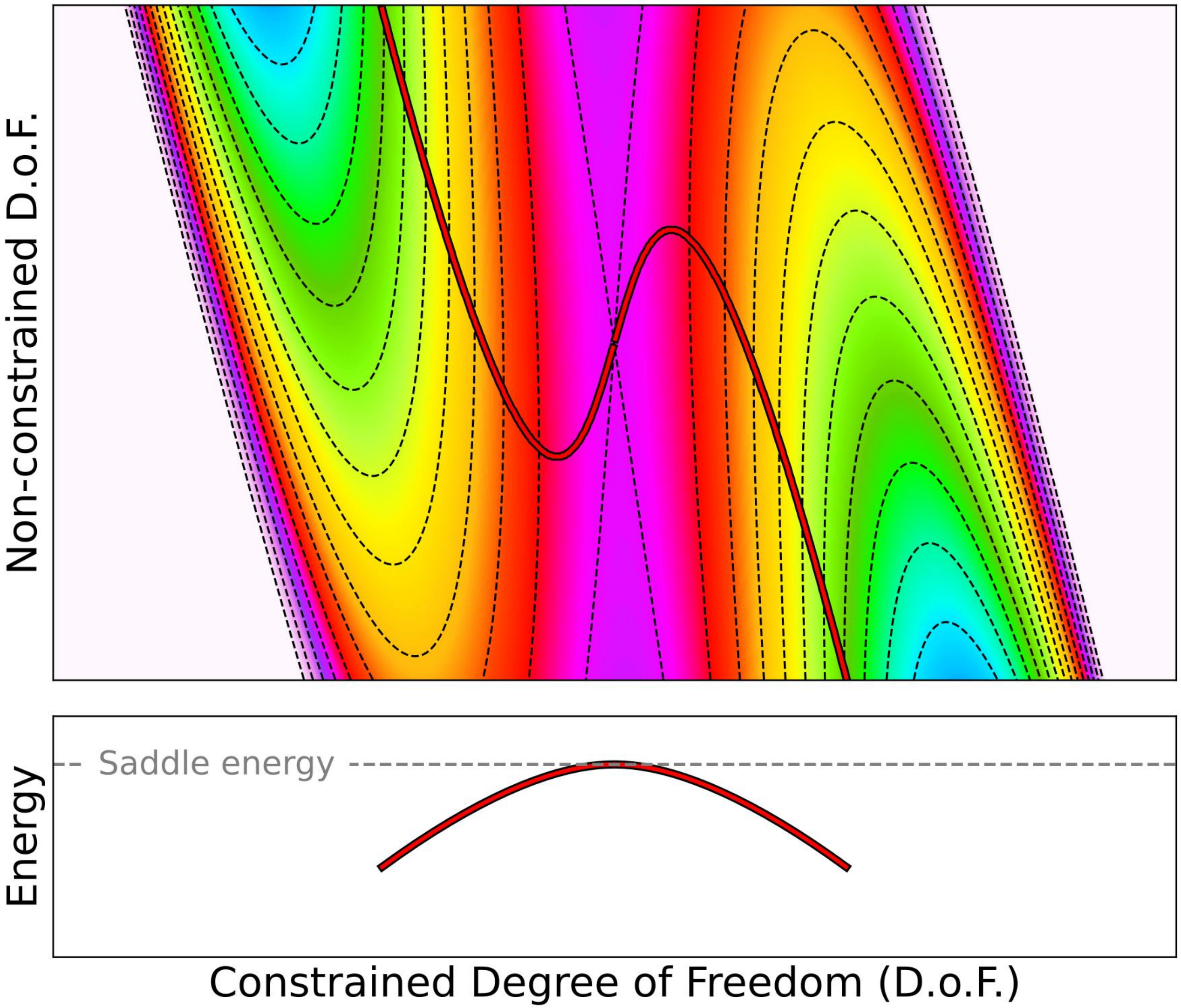
(2) The choice of constrained degrees of freedom is somewhat arbitrary.



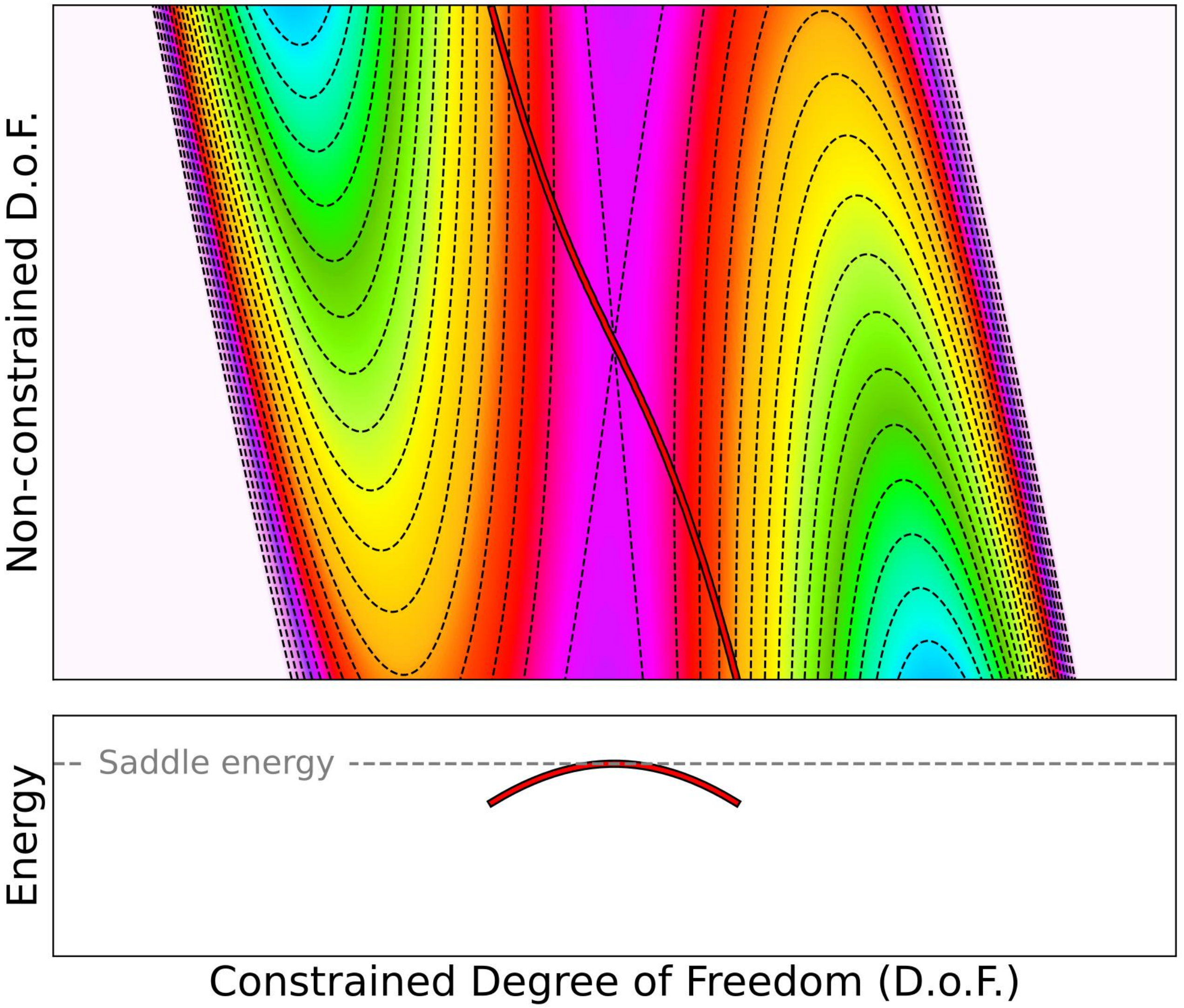
(2) The choice of constrained degrees of freedom is somewhat arbitrary.



(2) The choice of constrained degrees of freedom is somewhat arbitrary.



(3) Another choice of constrained degrees of freedom can remove the discontinuities.



Nuclear DFT still is a great theoretical contender for the description of fission.

- It is ***predictive almost everywhere*** across the nuclear chart, far from known nuclei.
- It is microscopic, and thus enables to ***connect the latest developments*** in nuclear interaction with the description of heavy nuclei.
- It is a ***very flexible framework*** that enables physicists to study a wide range of phenomena.

Can we find a computationally efficient surrogate model of nuclear DFT that preserves its most important features?

➔ We are exploring the use of machine learning to learn an efficient representation of DFT degrees of freedom.

We use autoencoder neural networks to build our surrogate model of nuclear DFT.

Required properties for the surrogate model:

1. It has to be ***computationally efficient***.

   We want, eventually, to tackle astrophysics simulations.

   Nuclear DFT states are described by millions of parameters, we want to be scalable.

2. It has to predict ***simply connected*** (“continuous”) manifolds.

   No missing saddle = no missing physics.

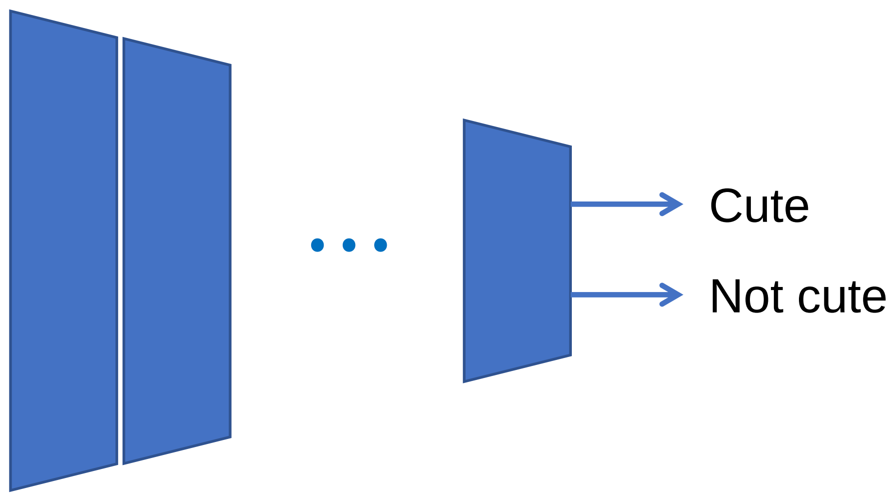
3. We can ***choose its dimension*** *D*.

   We want D=1 (potential energy line) or D=2 (potential energy surface).

4. It has to ***reproduce*** states far from discontinuities.

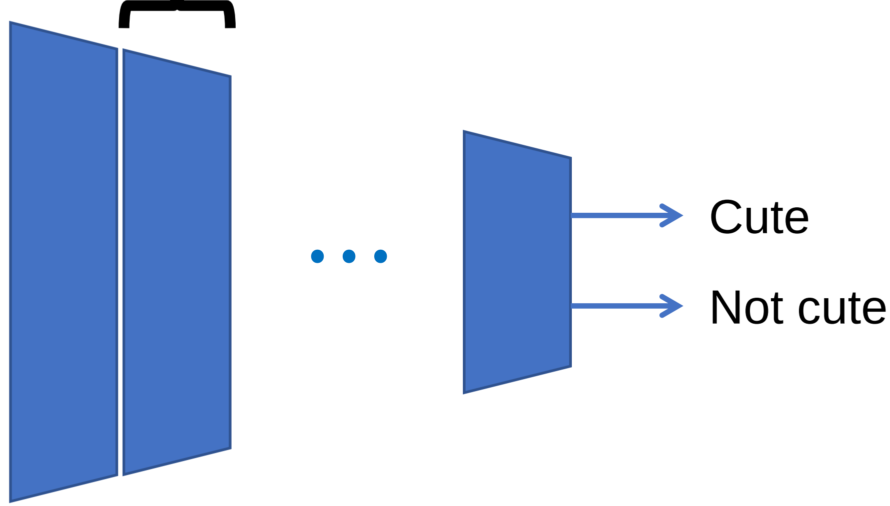
   The surrogate model has to reproduce the model where it works well.

(Feedforward) neural networks are the sequential application of neural layers.



In our case, neural layers are the composition of a linear map and a nonlinear activation function.

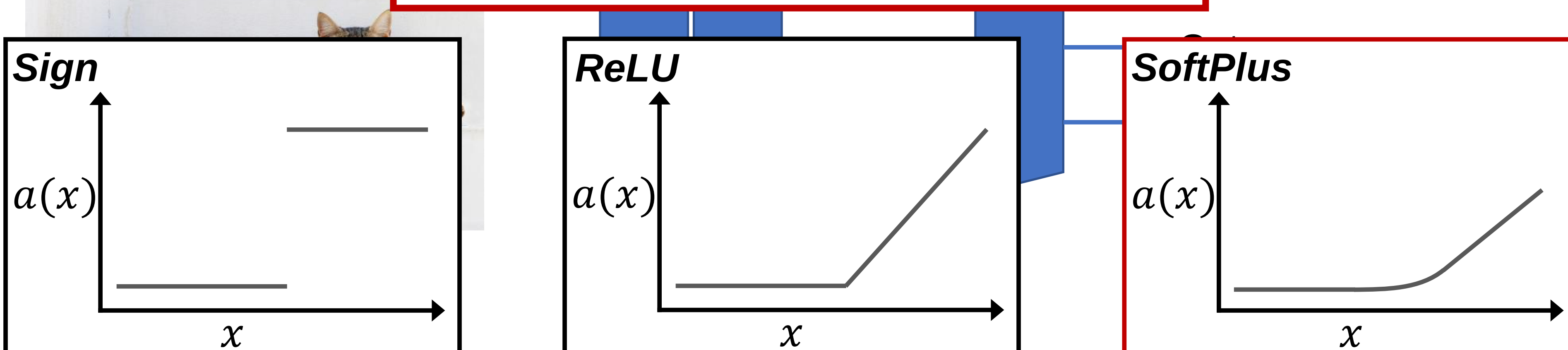
$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = a\left\{\begin{pmatrix} w_{11} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{mn} \end{pmatrix} \times \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}\right\}$$



The choice of activation function determines the smoothness of the neural network.

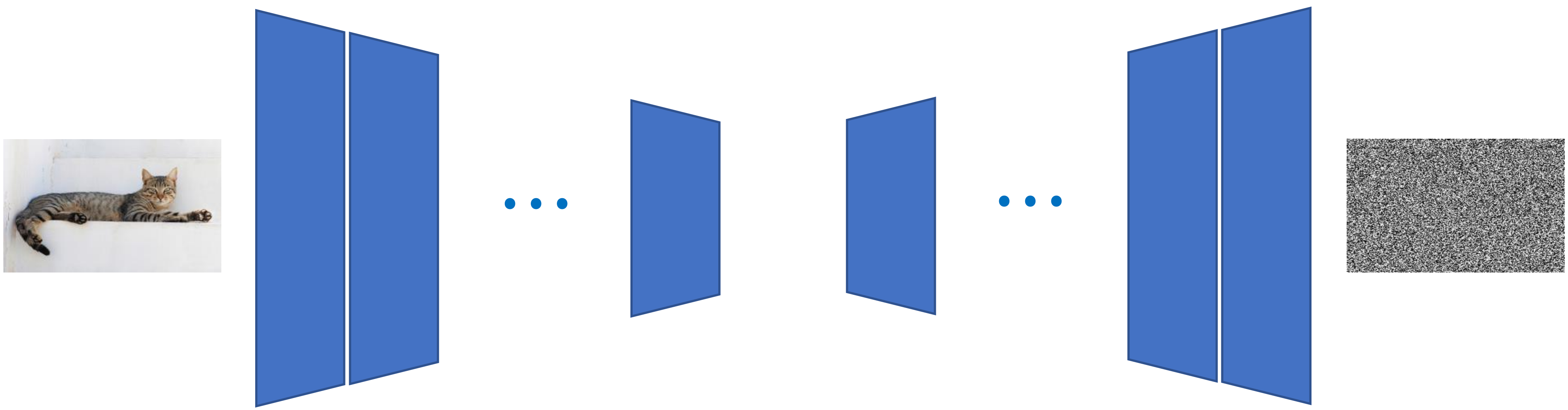
$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \underbrace{a\left\{\begin{pmatrix} w_{11} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{mn} \end{pmatrix} \times \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}\right\}}$$

- **Sign function**: discontinuous.
- **ReLU**: continuous.
- **SoftPlus, Tanh**: differentiable.



Autoencoders are analogous to zip/unzip.

Input (original file) | Encoder (zip) | Code ∈ Latent space (compressed file) | Decoder (unzip) | Output (uncompressed file)



Autoencoders are analogous to zip/unzip.

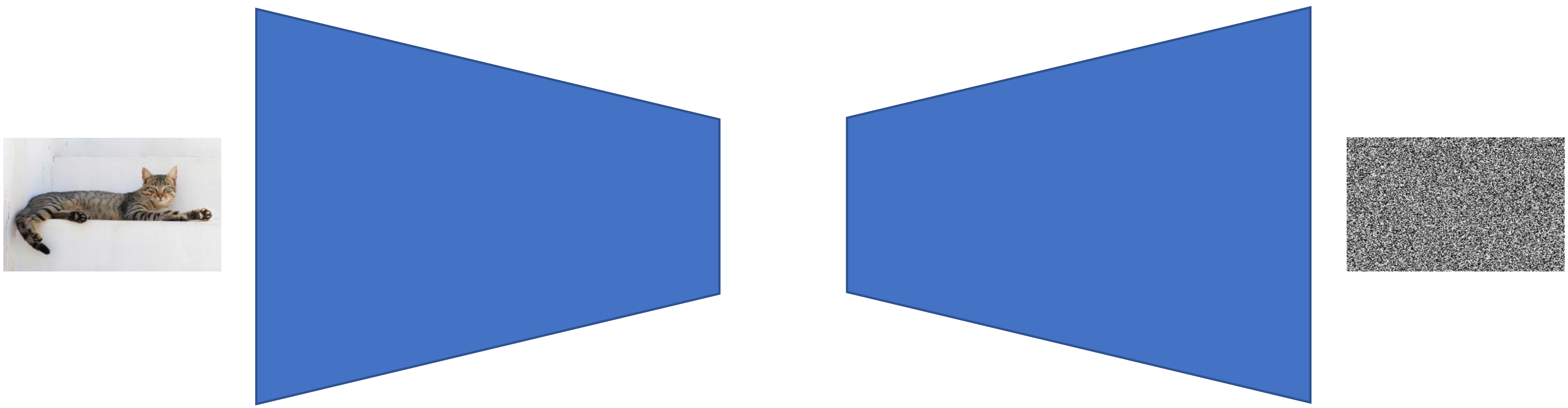
Input
(original file)

Encoder
(zip)

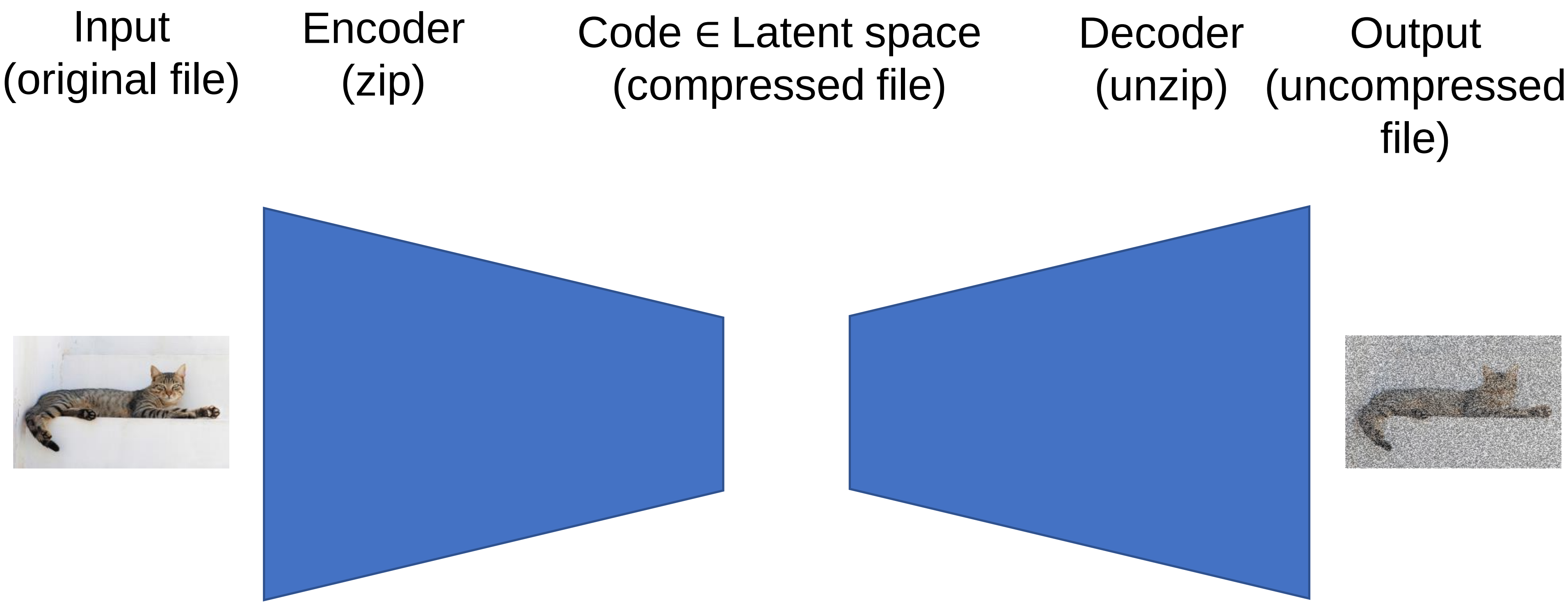
Code ∈ Latent space
(compressed file)

Decoder
(unzip)

Output
(uncompressed file)



Autoencoders are analogous to zip/unzip.



Autoencoders are analogous to zip/unzip.

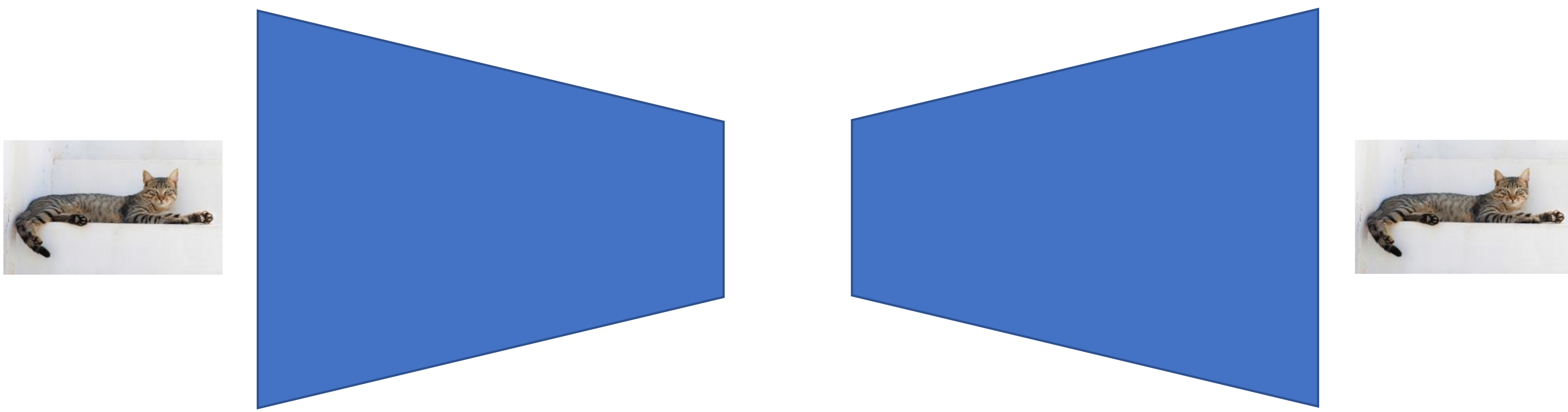
Input
(original file)

Encoder
(zip)

Code ∈ Latent space
(compressed file)

Decoder
(unzip)

Output
(uncompressed file)



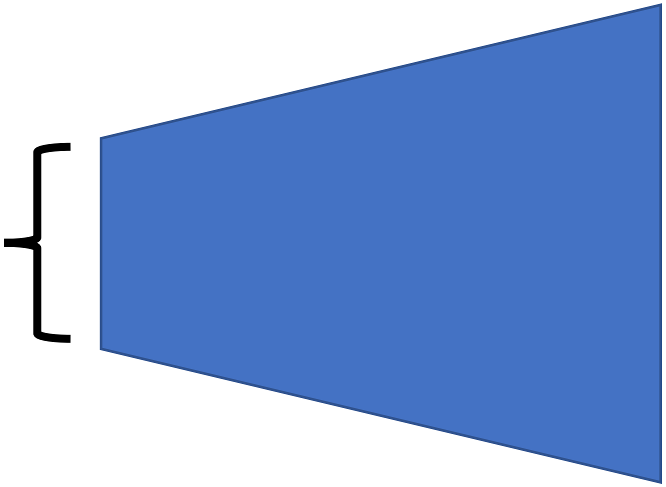
The latent space contains the new DoFs, and the decoder is a continuous surrogate model.

Code ∈ Latent space
(compressed file)

Decoder
(unzip)

Output
(uncompressed file)

New degrees of freedom

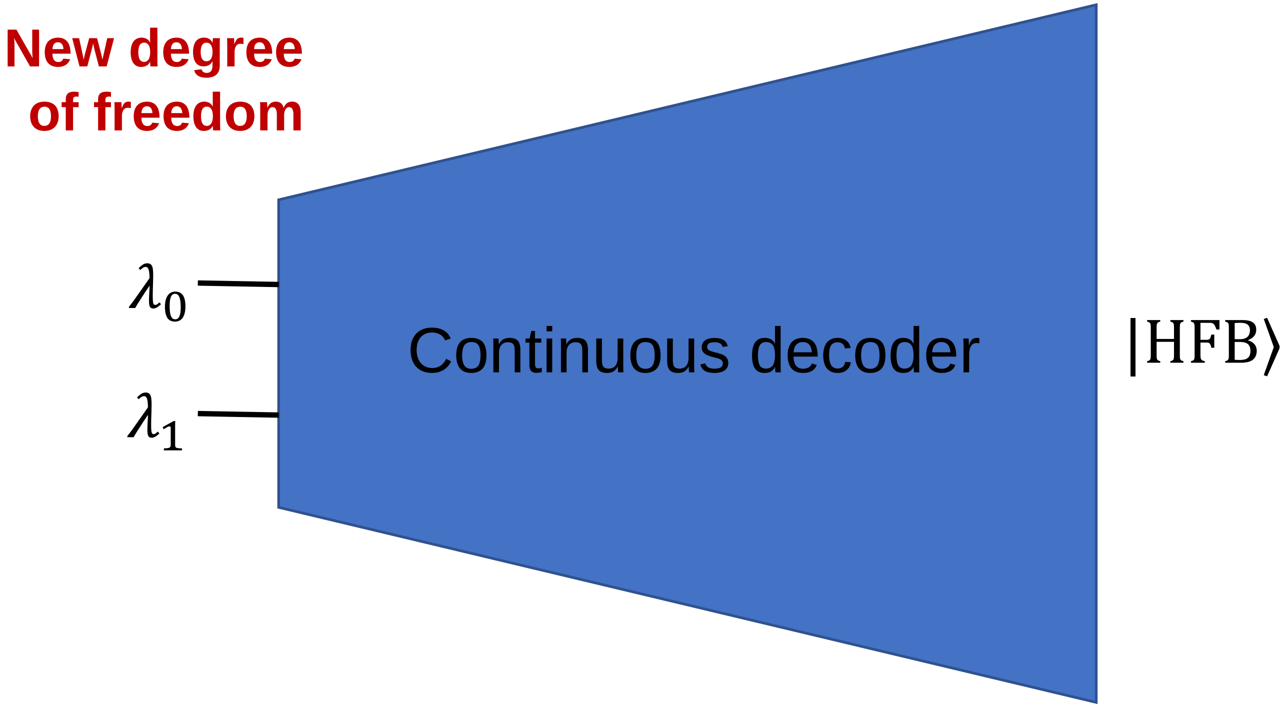


We have tackled the question in two different ways.

I. We have fitted a continuous variational autoencoder on the **orbitals** of a **2-D** Potential Energy Surface **with pairing**, a.k.a, *Hartree-Fock-Bogoliubov* (**HFB**) states.

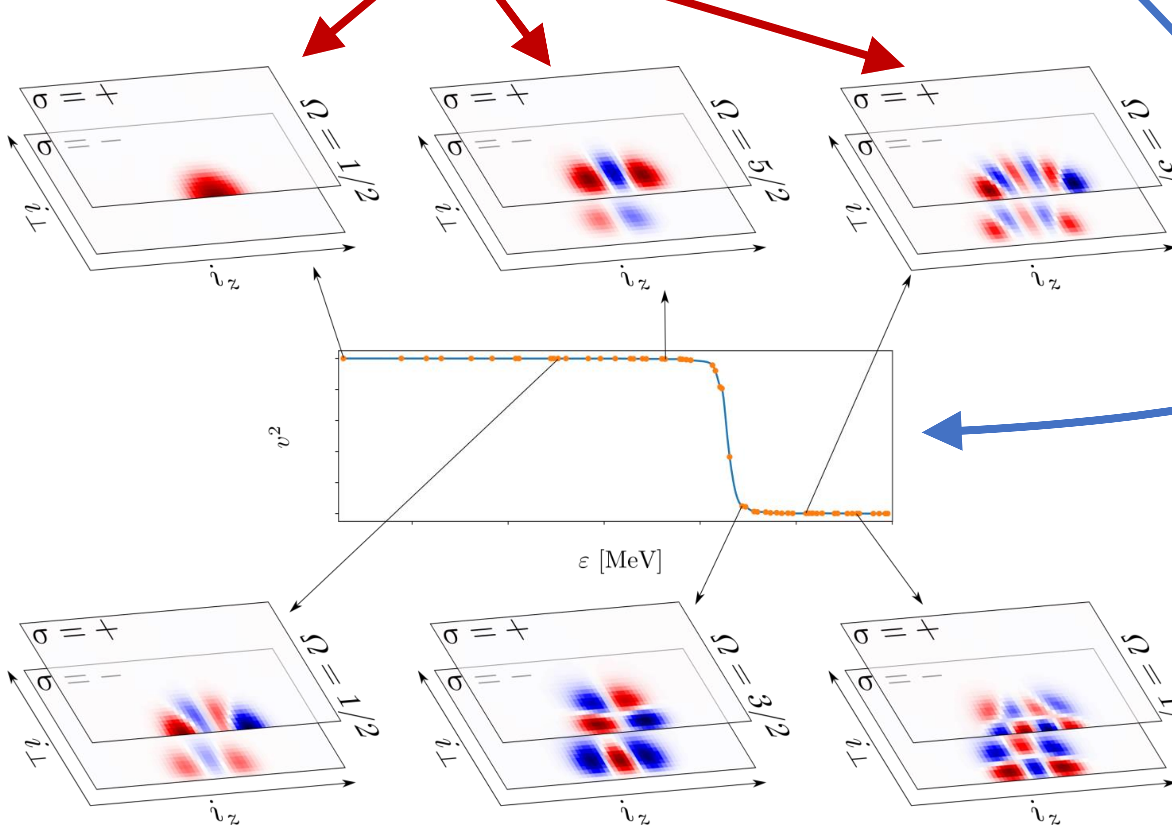
II. We have fitted a continuous variational autoencoder on a **1-D** Potential Energy Landscape **without pairing**, a.k.a., *Hartree-Fock* (**HF**) states.

I. We aim at compressing nuclear DFT states with pairing (HFB states) in a 2-D PES.

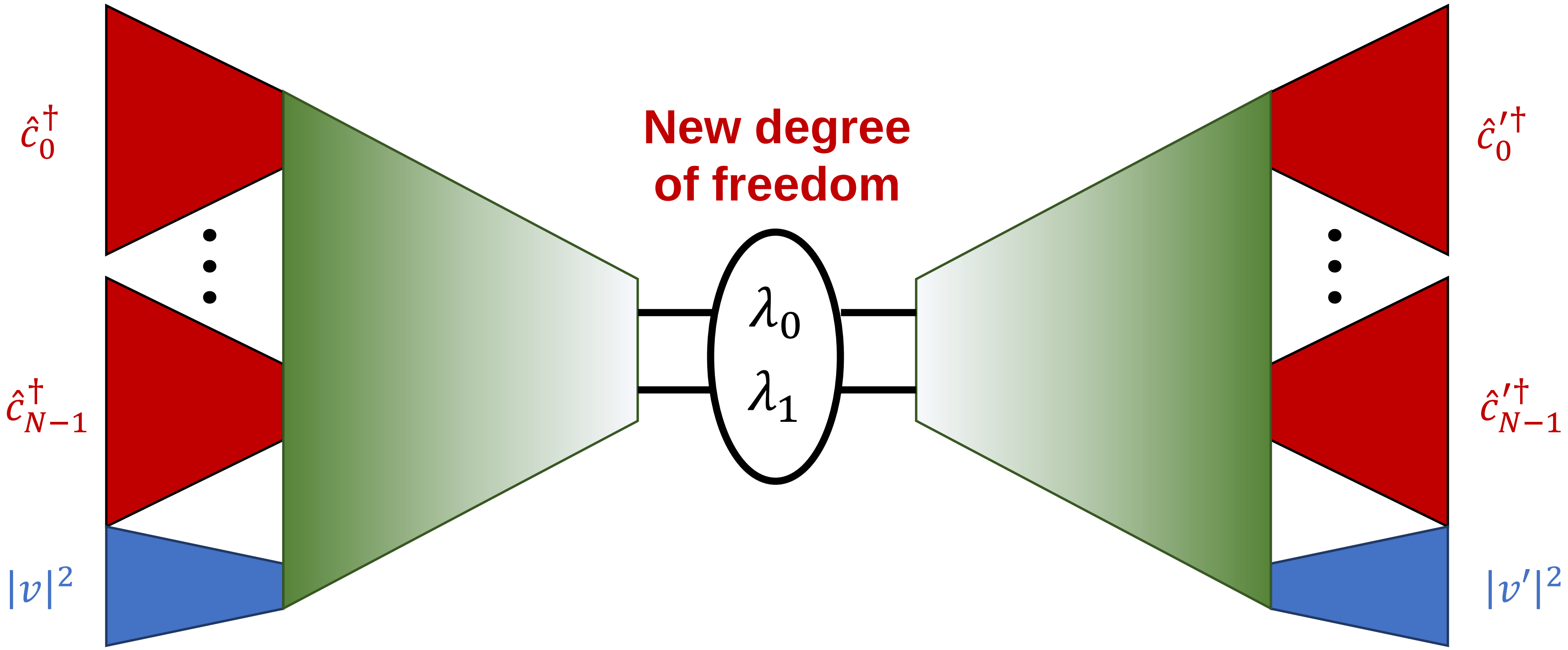


I. We use the Bloch-Messiah decomposition of HFB states (with pairing).

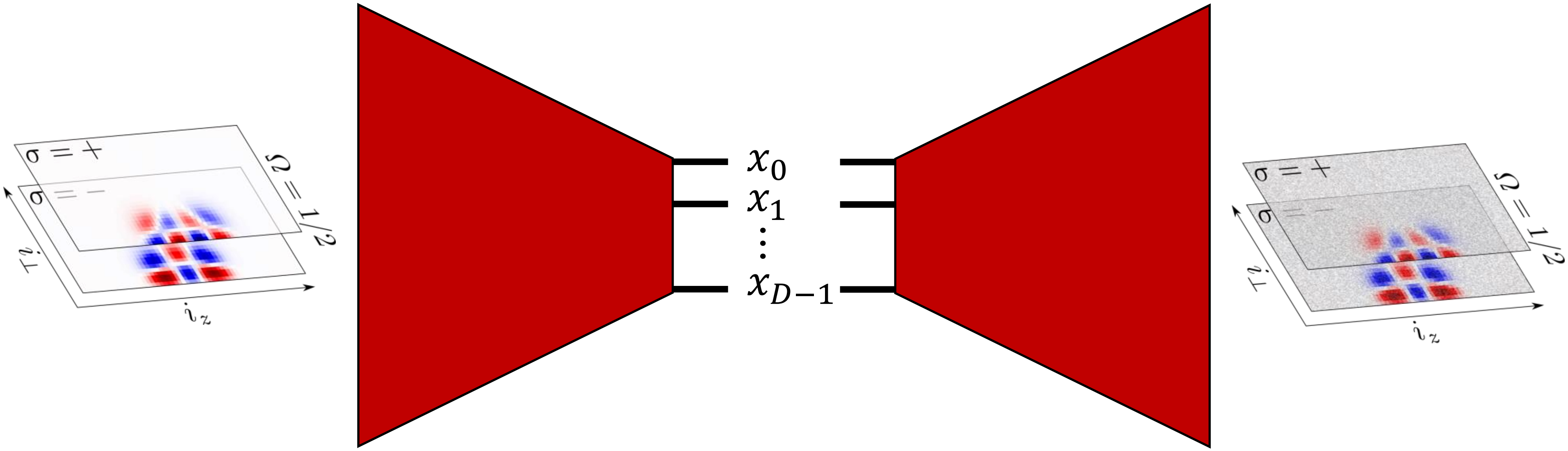
$$|\mathrm{HFB}\rangle \propto \prod_{i>0} \left(u_i + v_i \hat{c}_i^\dagger \hat{c}_{\bar{\imath}}^\dagger \right) |-\rangle, \qquad |u_i|^2 + |v_i|^2 = 1.$$



I. We can give the autoencoder a block structure and train each block separately.



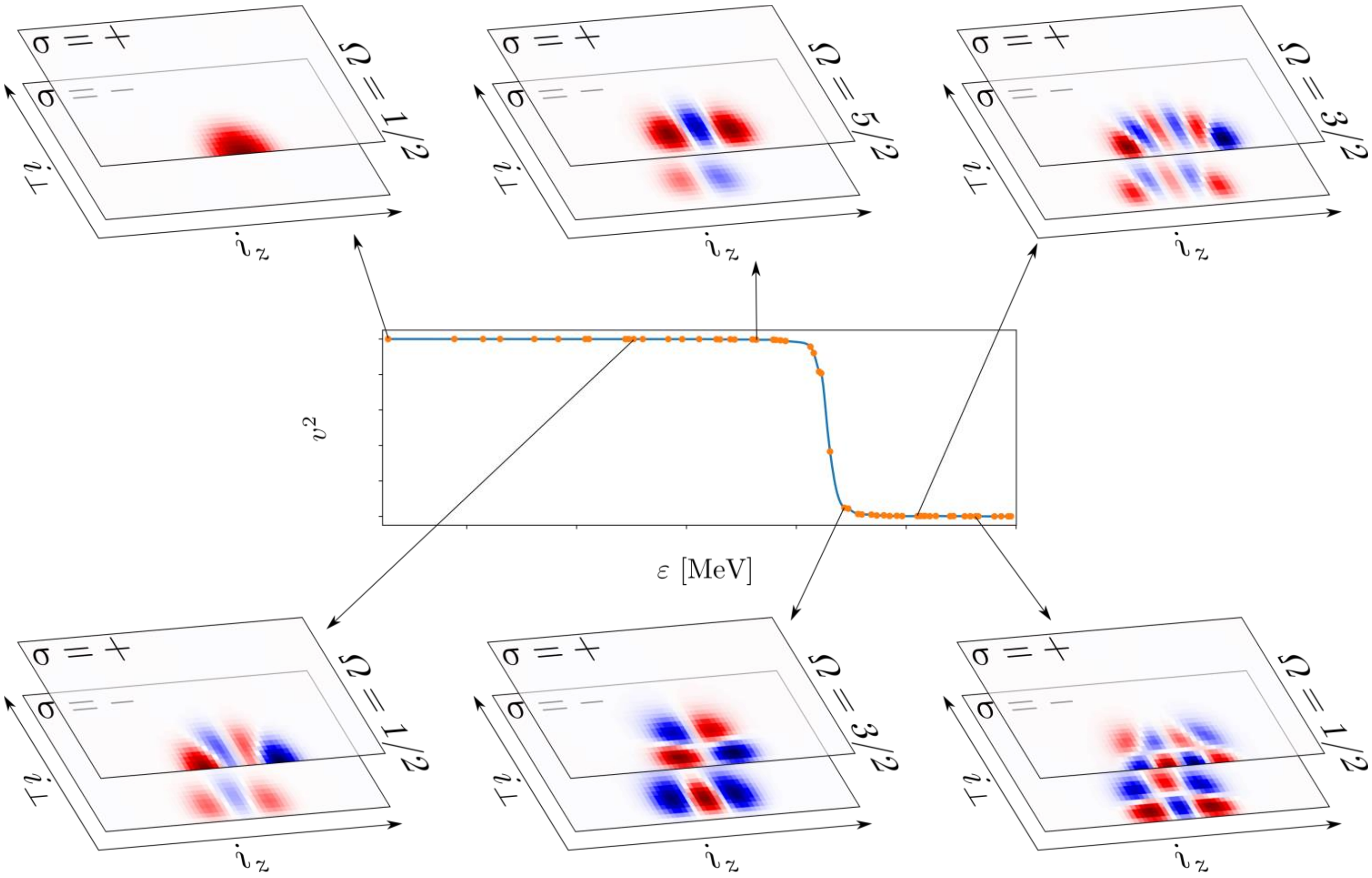
I. We trained the orbital block on all the orbitals of all the HFB states in the PES.



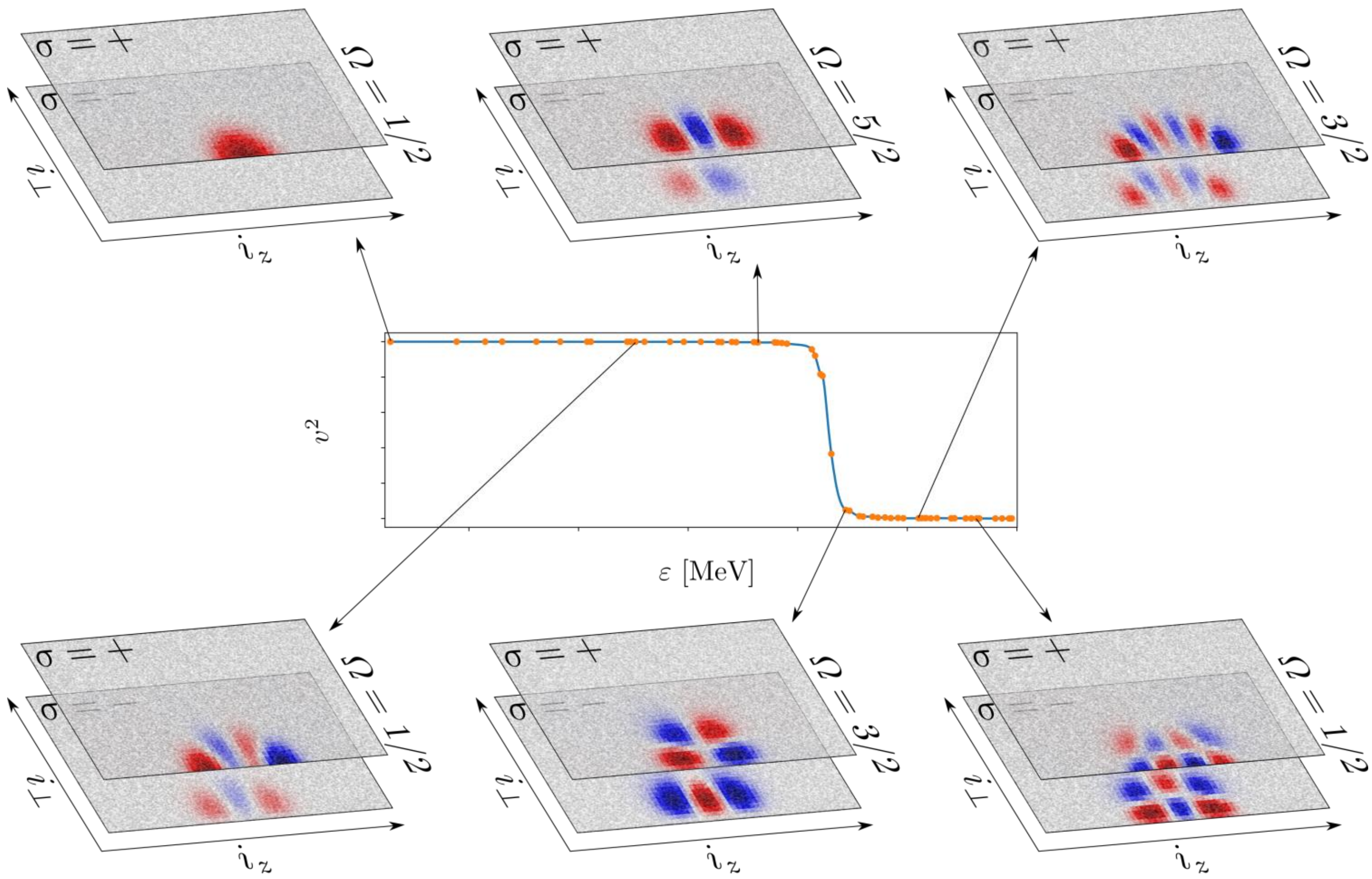
Rough estimation of the optimal code size d:

q_{20}, q_{30}, ε, Ω, τ, unknown others: $D \geq 5$

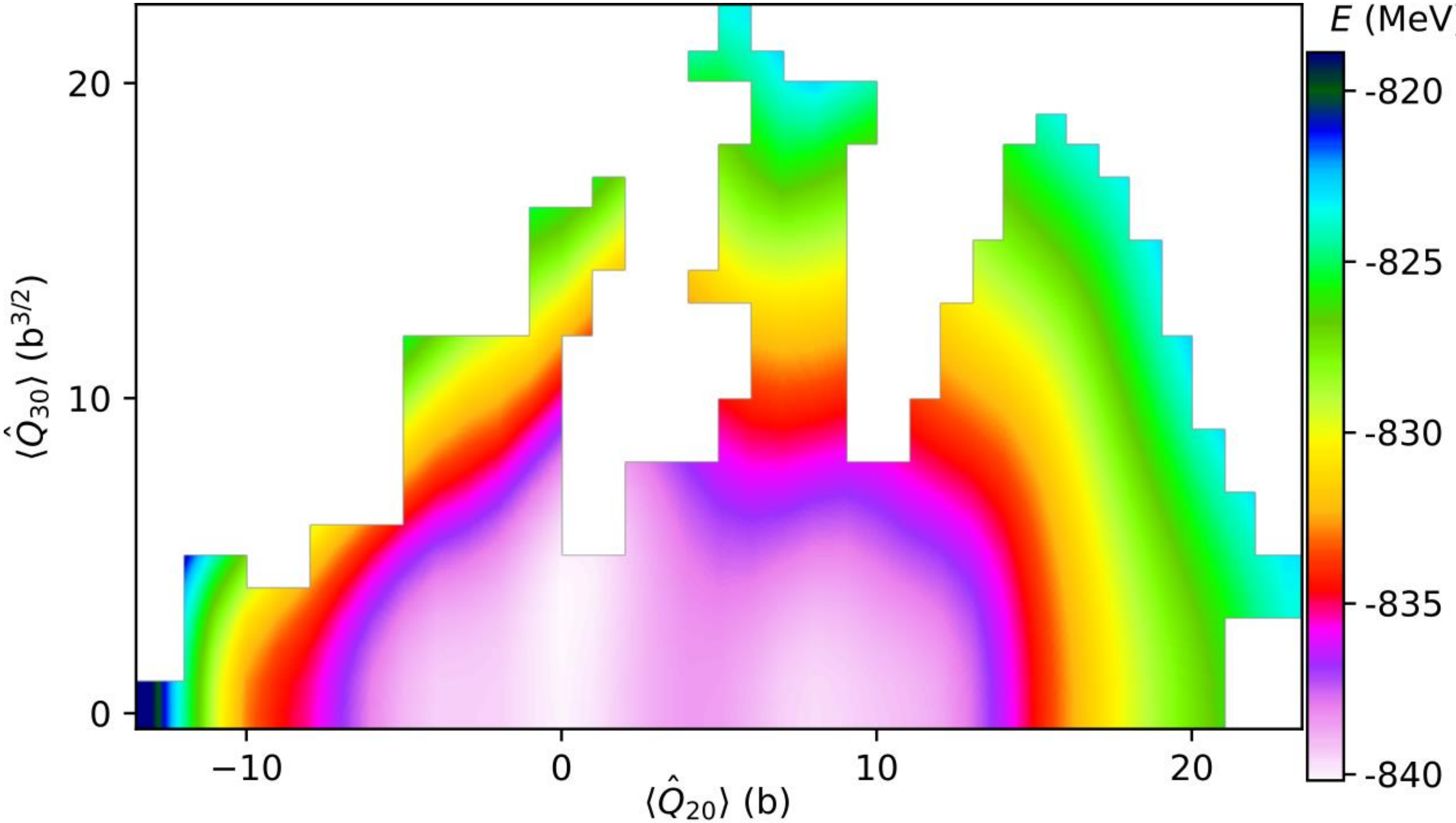
I. Quality testing: we replace the orbitals with the decoded ones and recompute the energy.



I. Quality testing: we replace the orbitals with the decoded ones and recompute the energy.

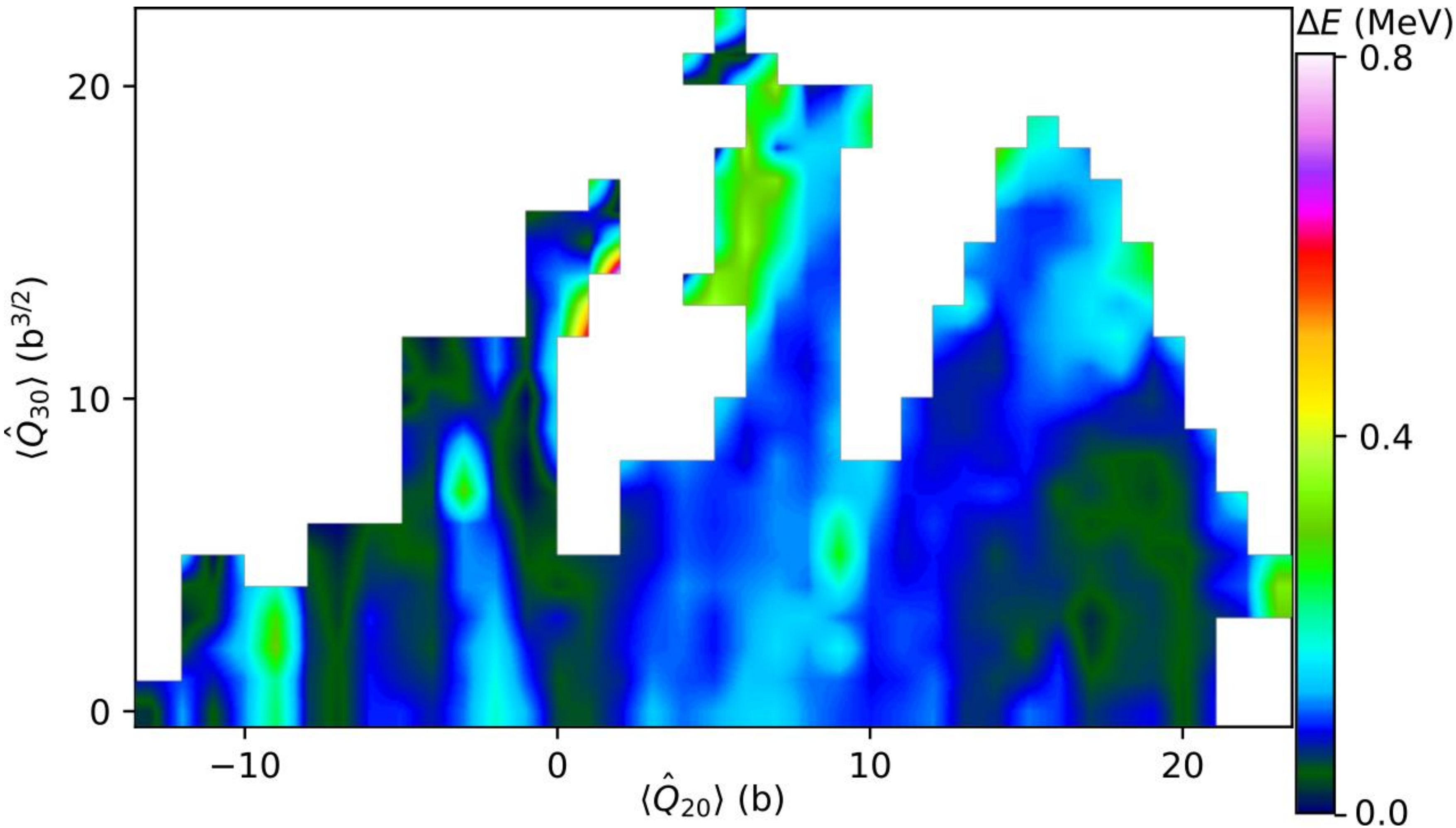


I. We compressed the orbitals for ^{98}Zr using a latent space $D = 20$.



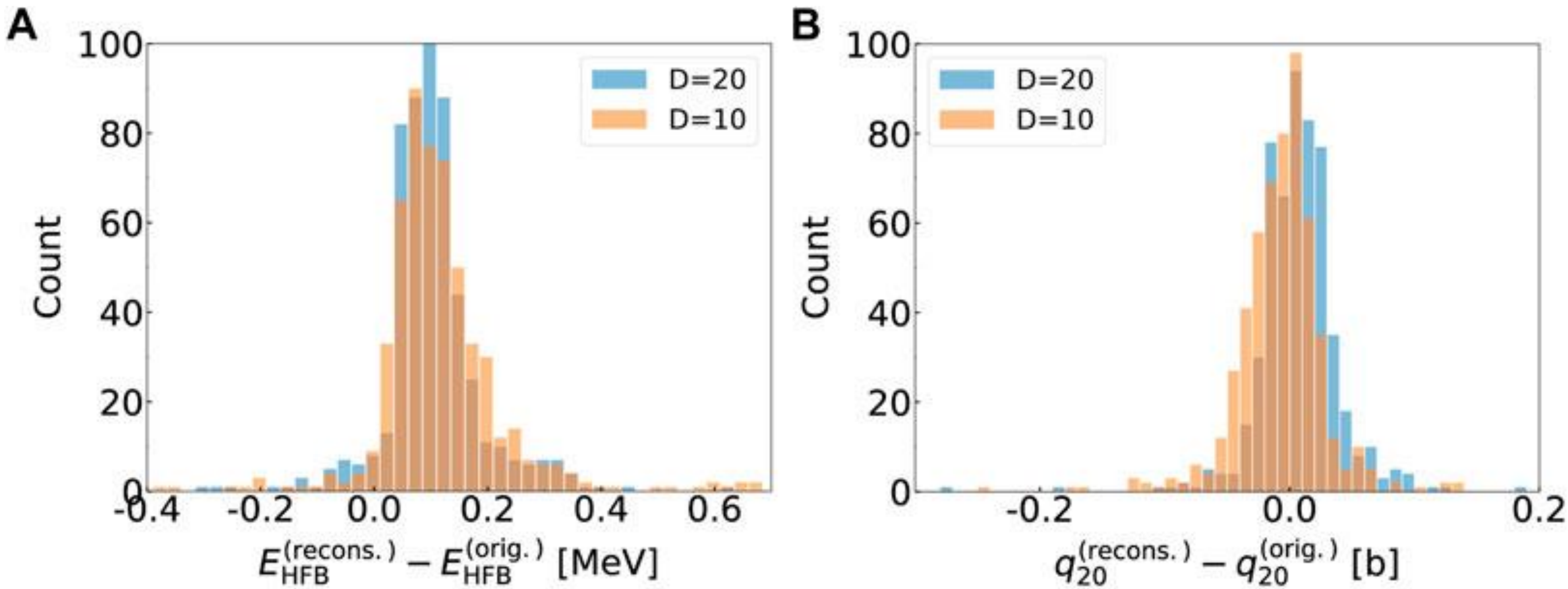
Verriere M, Schunck N, Kim I, Marević P, Quinlan K, Ngo MN, Regnier D and Lasseri R.D., *Front. Phys.* 10:1028370.
doi: 10.3389/fphy.2022.1028370

I. The error on the HFB energy is always below 1 MeV, and mostly below 0.2 MeV.



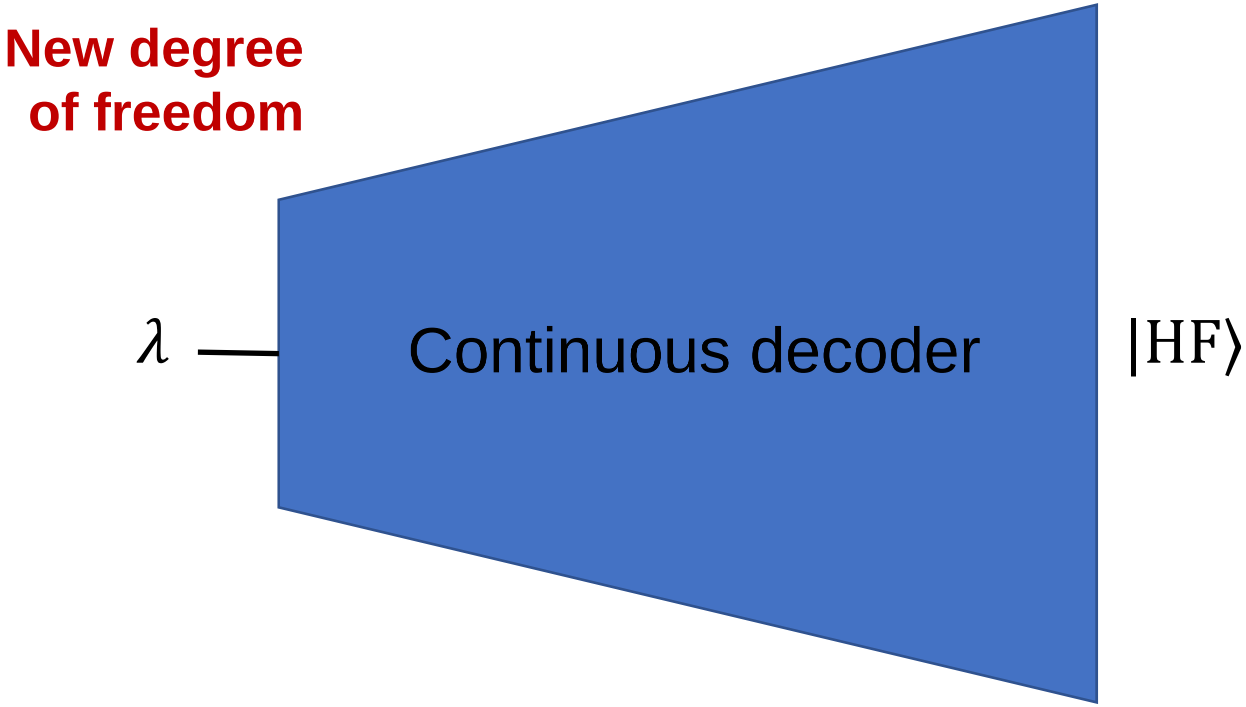
Verriere M, Schunck N, Kim I, Marević P, Quinlan K, Ngo MN, Regnier D and Lasseri R.D., *Front. Phys.* 10:1028370. doi: 10.3389/fphy.2022.1028370

I. We did the same work with $D = 10$, and we obtained a very similar error distribution.



Verriere M, Schunck N, Kim I, Marević P, Quinlan K, Ngo MN, Regnier D and Lasseri R.D., *Front. Phys.* 10:1028370.
doi: 10.3389/fphy.2022.1028370

II. We aim at compressing nuclear DFT states without pairing (HF states) in a 1-D PEL.



II. The Thouless theorem gives a one-to-one mapping between HF states and matrices.

1. We can decompose all the HF states $|\mathrm{HF}\rangle$ as

$$|\mathrm{HF}\rangle \propto \exp\left(\sum_{p<A,\, h<N_\mathrm{b}} Z_{ph}\, \hat{a}_h^\dagger \hat{a}_p\right)|\Phi_0\rangle.$$

2. We need to find a $|\Phi_0\rangle$ **not** orthogonal to any of the HF states. We use the **Karcher mean** of the training set.

3. Now, not only Z entirely represents the state $|\mathrm{HF}\rangle$, but there is also a **one-to-one correspondence** between the $A \times N_\mathrm{b}$ matrices and the HF states not orthogonal to $|\Phi_0\rangle$.

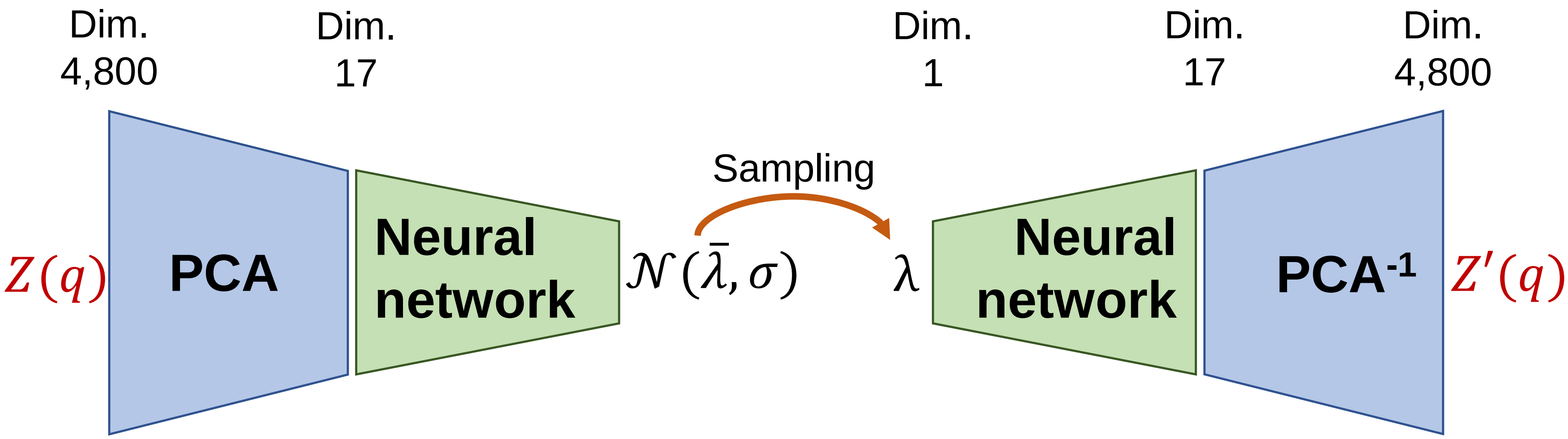
II. We consider a variational autoencoder with a 1-D latent space for the isotope ^{16}O.

4. We compute the 16×300 matrices $Z = Z(q)$ for each precomputed HF state $|\mathrm{HF}(q)\rangle$.

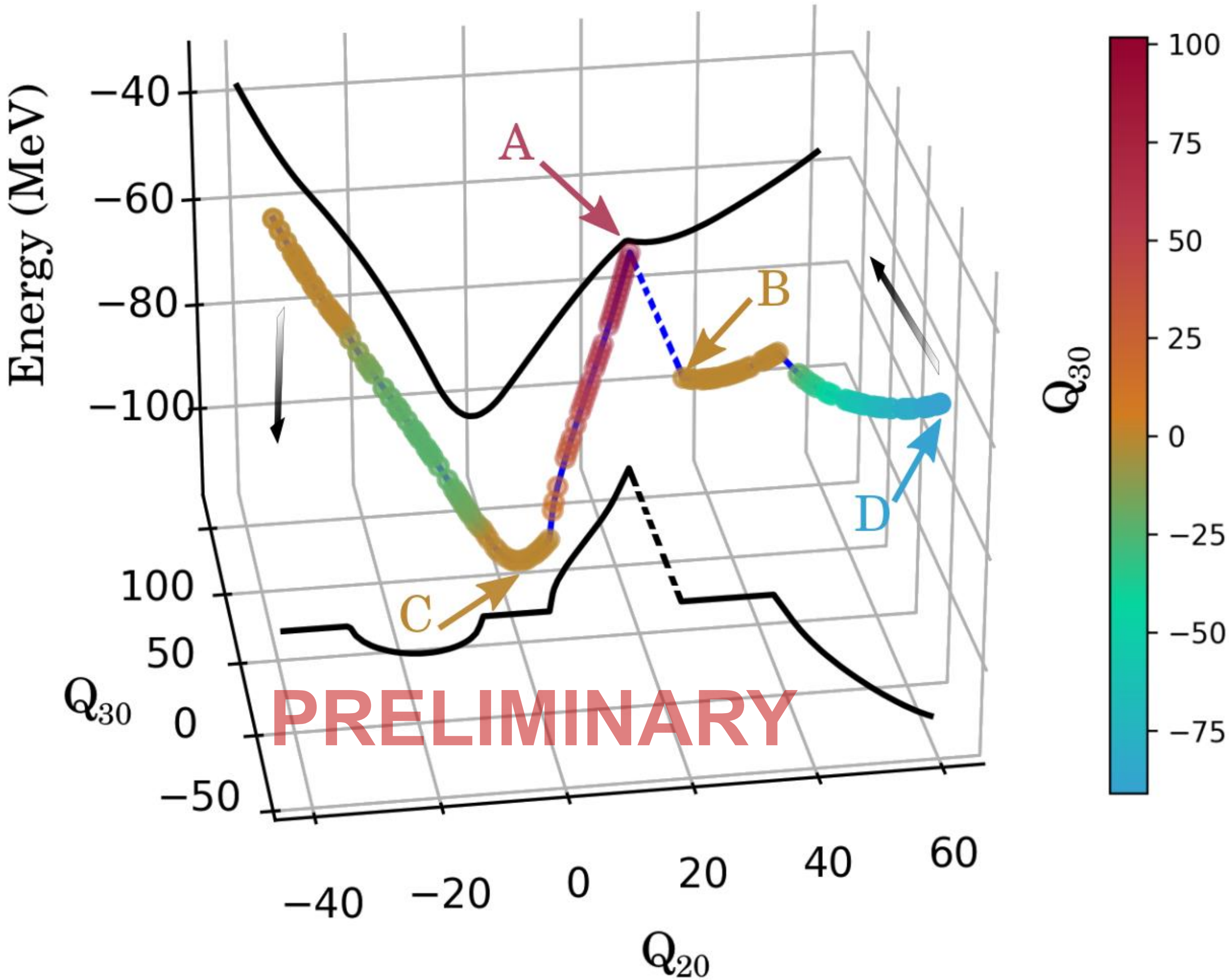
5. We randomly split all the $Z(q)$ into:

   training set: 70% validation set: 20% testing set: 10%

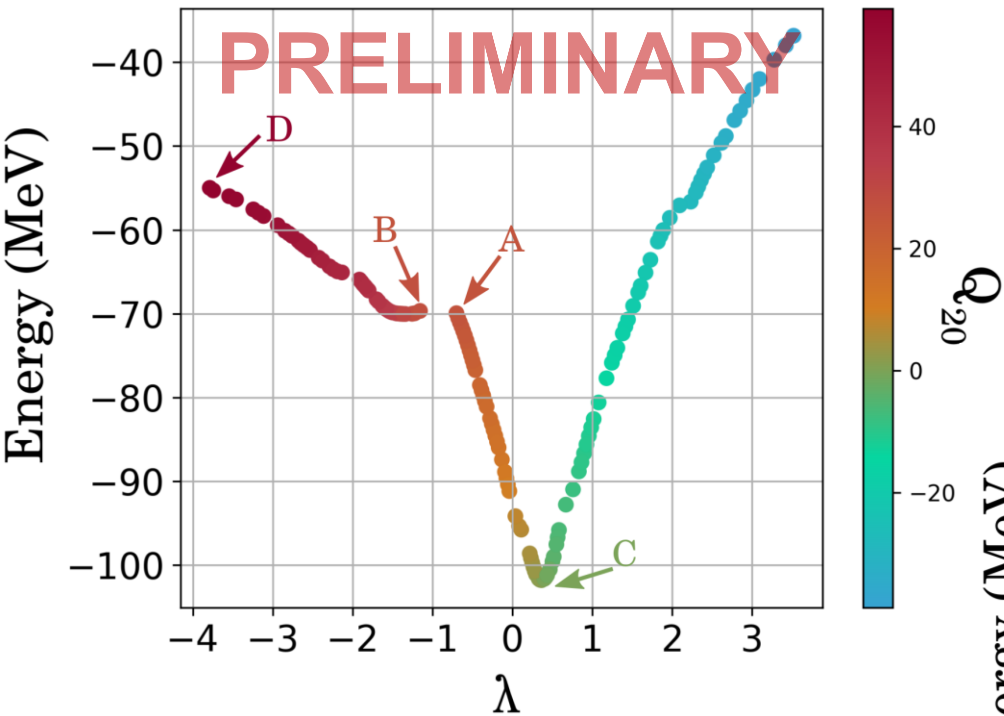
6. We use this architecture:

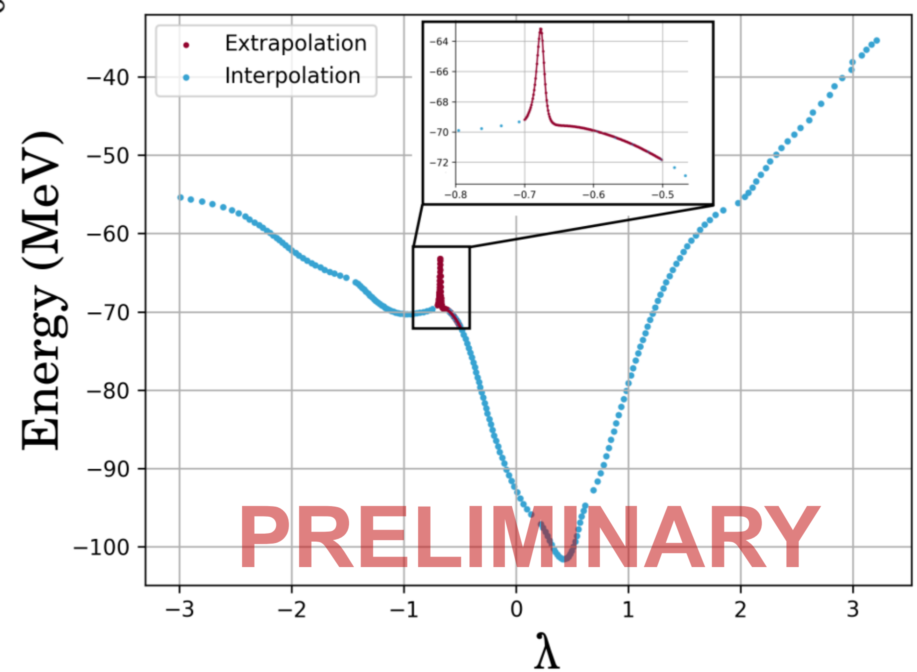


II. We applied this approach on a set of HF states with one constrained DoF for the nucleus ^{16}O.



II. We are able to find a physically relevant continuous degree of freedom in 1D.





R.D. Lasseri, D. Regnier, M. Frosini, M. Verriere, N. Schunck, A. Penon, **to be published** (2023)

We have built the first surrogate models of nuclear DFT with deep neural networks.

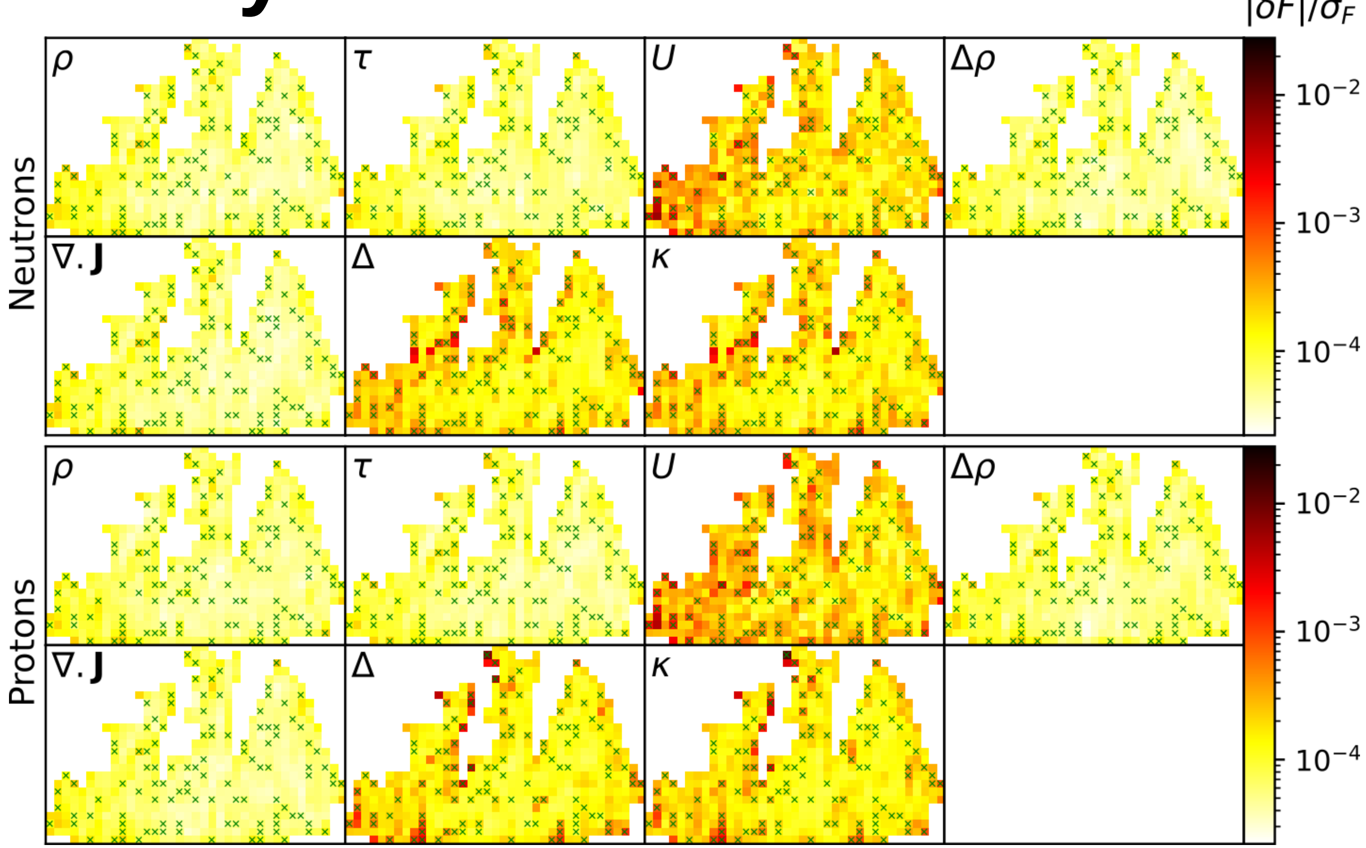
1. We gained a lot of insight on the structure of nuclear DFT states, with and without pairing, and how to build, train and use (variational) autoencoders.

2. We obtain **qualitatively good results** in the compression of orbitals, but **going further seems challenging**:
   - we need to follow the orbitals, but we want to only explicitly compute a few nuclear DFT states,
   - conical intersections might impose to use more dimensions than physical DoFs.

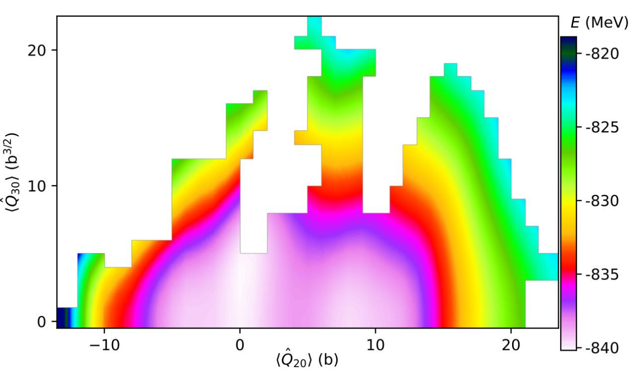
3. We are now exploring the use of deep neural networks for the fields and densities of nuclear DFT with pairing included…

… and we are able to compress HFB states with great accuracy.

PRELIMINARY



Mean absolute error for each normalized fields in the ^{98}Zr potential energy surface.



Lawrence Livermore
National Laboratory