Artificial Neural Network-Based Exchange and Correlations Functionals

João Paulo Almeida de Mendonça *(He/Him)*
Université Grenoble Alpes, CNRS, Grenoble INP
SIMaP, 38000 Grenoble, France
joao-paulo.almeida-de-mendonca@grenoble-inp.fr

What we will cover on this talk?

- Density Functional Theory (DFT)
  - What is DFT? What can we do with if?
  - Theoretical Background
  - Challenges in DFT
  - How to expand if?

- Artificial Neural Network (ANN)
  - What are ANNs? Why they are interesting do DFT?
  - Understanding Feedforward Neural Networks
  - How can it be a DFT Functional?
  - Challenges in Training

- ANN Functionals
  - Some examples on the literature and how they do it
  - My work and how we do it at our group

- Preparing for the practical section!

Please, help me by making questions!

This is a very interdisciplinary topic and no one is expected to know everything about it.

If not now, stop me at some moment on the conference and ask!

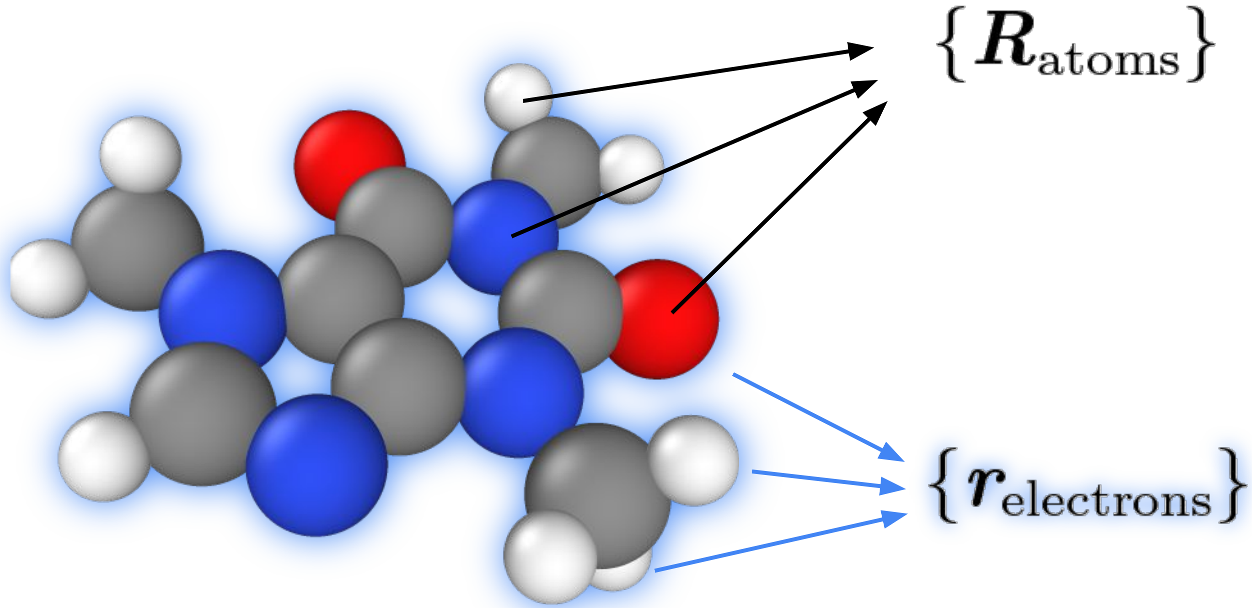
"We all know something. We all choose to ignore something. And that's why we always learn." (Paulo Freire)

What is Density Functional Theory (DFT)?

"A theory used to describe **many-fermion systems** in which **the energy is a functional of the density of fermions**. Density functional theory has been used extensively in the theory of **electrons in atoms, molecules, and solids** and in the theory of nucleons in nuclei."

*(From **A Dictionary of Physics** - Oxford University Press)*

How this is done?



$$H_{\text{elec.+atoms}} = \underbrace{-\sum_i \frac{1}{2}\nabla_i^2 - \sum_A \frac{1}{2M_A}\nabla_A^2}_{\text{Kinetic Energy}} \underbrace{- \sum_{i,A}\frac{Z_A}{r_{iA}} + \sum_{i>j}\frac{1}{r_{ij}} + \sum_{B>A}\frac{Z_A Z_B}{R_{AB}}}_{\text{Potential Energy}}$$

$$H_{\text{elec.+atoms}} = -\sum_i \frac{1}{2}\nabla_i^2 - \sum_A \frac{1}{2M_A}\nabla_A^2 - \sum_{i,A}\frac{Z_A}{r_{iA}} + \sum_{i>j}\frac{1}{r_{ij}} + \sum_{B>A}\frac{Z_A Z_B}{R_{AB}}$$

a proton is ~1840 times the electron mass

Born-Oppenheimer Approximation:

- The use an approximate ansatz for the complete many-body problem that **allows the wave functions of atomic nuclei and electrons to be treated separately.**
- This approximation is motivated by the mass difference between nuclei and electrons.
- The electronic Hamiltonian than can be written as the **kinetic energy of the electrons, the electron interaction with fixed nuclei, and the electron-electron interaction.**

$$\hat{\mathbf{H}} = -\sum_i \frac{1}{2}\nabla_i^2 - \sum_{i,A}\frac{Z_A}{r_{iA}} + \sum_{i>j}\frac{1}{r_{ij}}$$

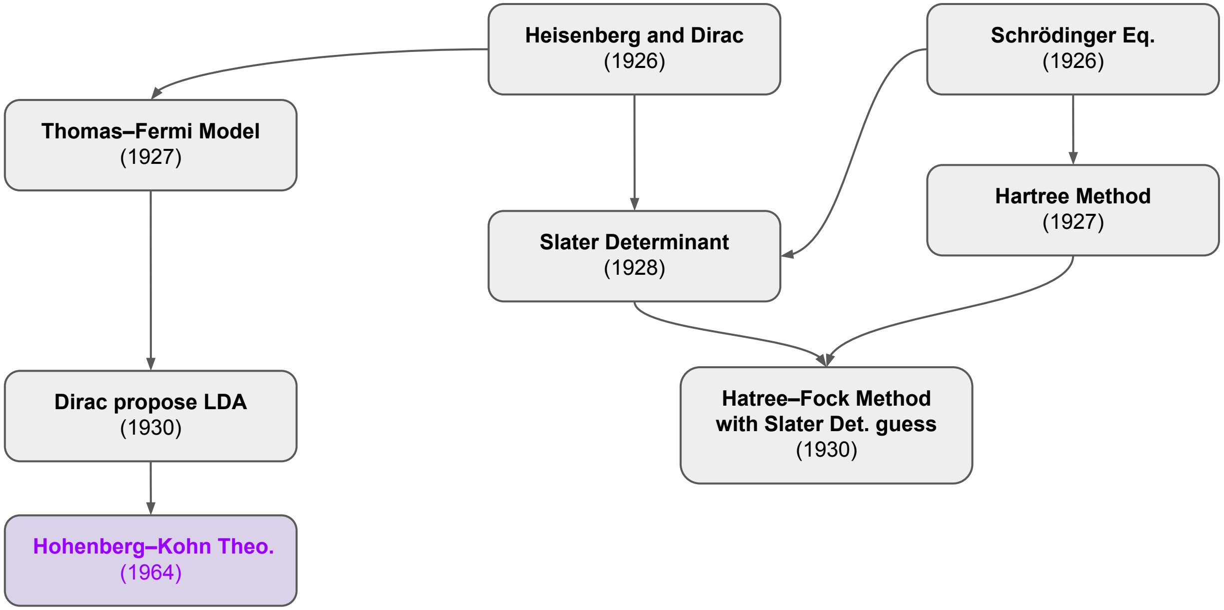
$$\left[-\sum_{i}^{N}\frac{1}{2}\nabla_i^2-\sum_{i,A}^{N,N_{\mathrm{atoms}}}\frac{Z_A}{r_{iA}}+\sum_{i>j}^{N,N}\frac{1}{r_{ij}}\right]\Psi(\mathbf{r}_1,...,\mathbf{r}_N)=E\Psi(\mathbf{r}_1,...,\mathbf{r}_N)$$

The normalized electronic wavefunction Ψ can be associated to a electronic density $n(\mathbf{r})$:

$$n(\mathbf{r})=N\int \mathrm{d}^3\mathbf{r}_2\cdots\int \mathrm{d}^3\mathbf{r}_N\,\Psi^*(\mathbf{r}_1,\ldots,\mathbf{r}_N)\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N)$$

This definition **can be inverted for the ground state**, so the density $n_0(\mathbf{r})$ unambiguously defines $\Psi_0[n_0(\mathbf{r})]$, and as consequence defines the ground state energy of the electronic system $E_0[n_0(\mathbf{r})]$.

Hince, ***Density Functional Theory***!

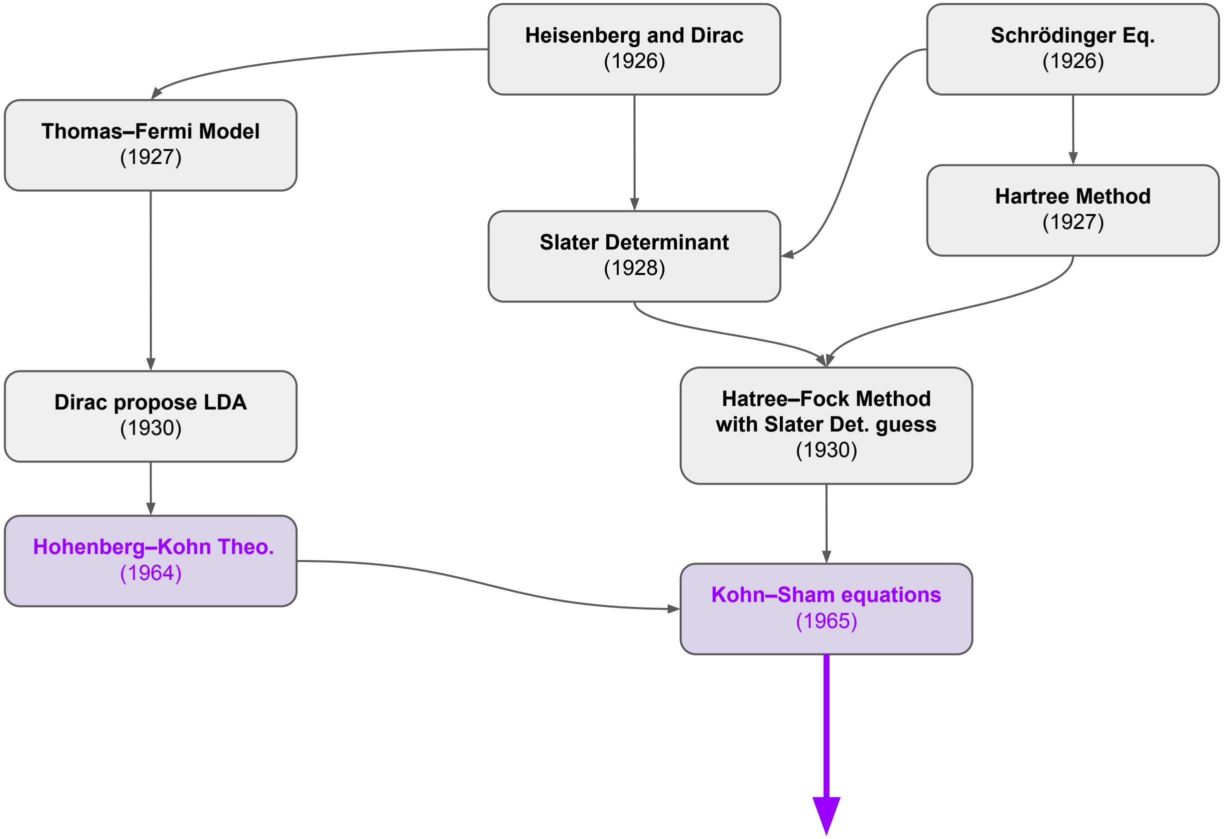


Theorem 1. The external potential v_{ext} (and hence the total energy E_{tot}), **is a unique functional of the electron density** $n(\mathbf{r})$.

Theorem 2. The functional that delivers the ground-state energy of the system ($E_{tot}[n(\mathbf{r})]$) **gives the lowest energy** E_0 **if and only if the input density is the true ground-state density**.

VERY GENERAL

Valid for all systems where $v_{ext}(\mathbf{r})$ defines pure ground state(s)!



Single non-interacting particle approximation for $n(\mathbf{r})$:

$$n(\mathbf{r}) = \sum_i^N \left|\varphi_i(\mathbf{r})\right|^2$$

The single particle eigenstate problem can be solved using the **Kohn–Sham equations**:

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V_{\text{ext}}(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}^3\mathbf{r}' + V_{\text{xc}}[n(\mathbf{r})]\right]\varphi_i(\mathbf{r}) = \varepsilon_i\varphi_i(\mathbf{r})$$

We can recover the energy in the many-electrons system:

$$E = \sum \varepsilon_i - \frac{1}{2}\iint \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}'d\mathbf{r} + E_{\text{xc}}[n(\mathbf{r})] - \int \frac{\delta E_{\text{xc}}[n(\mathbf{r})]}{\delta n(\mathbf{r})}n(\mathbf{r})d\mathbf{r}.$$

Kohn–Sham equation solved self-consistently

Guess $\rho(\mathbf{r})$

$$v_{\text{eff}}(\mathbf{r}) = v_{\text{ext}}(\mathbf{r}) + e^2 \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\, d\mathbf{r}' + \frac{\delta E_{\text{xc}}[\rho]}{\delta \rho(\mathbf{r})},$$

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + v_{\text{eff}}(\mathbf{r})\right)\varphi_i(\mathbf{r}) = \varepsilon_i \varphi_i(\mathbf{r}).$$

$$\rho(\mathbf{r}) = \sum_i^N |\varphi_i(\mathbf{r})|^2.$$

$\rho(\mathbf{r})$ converged?

no

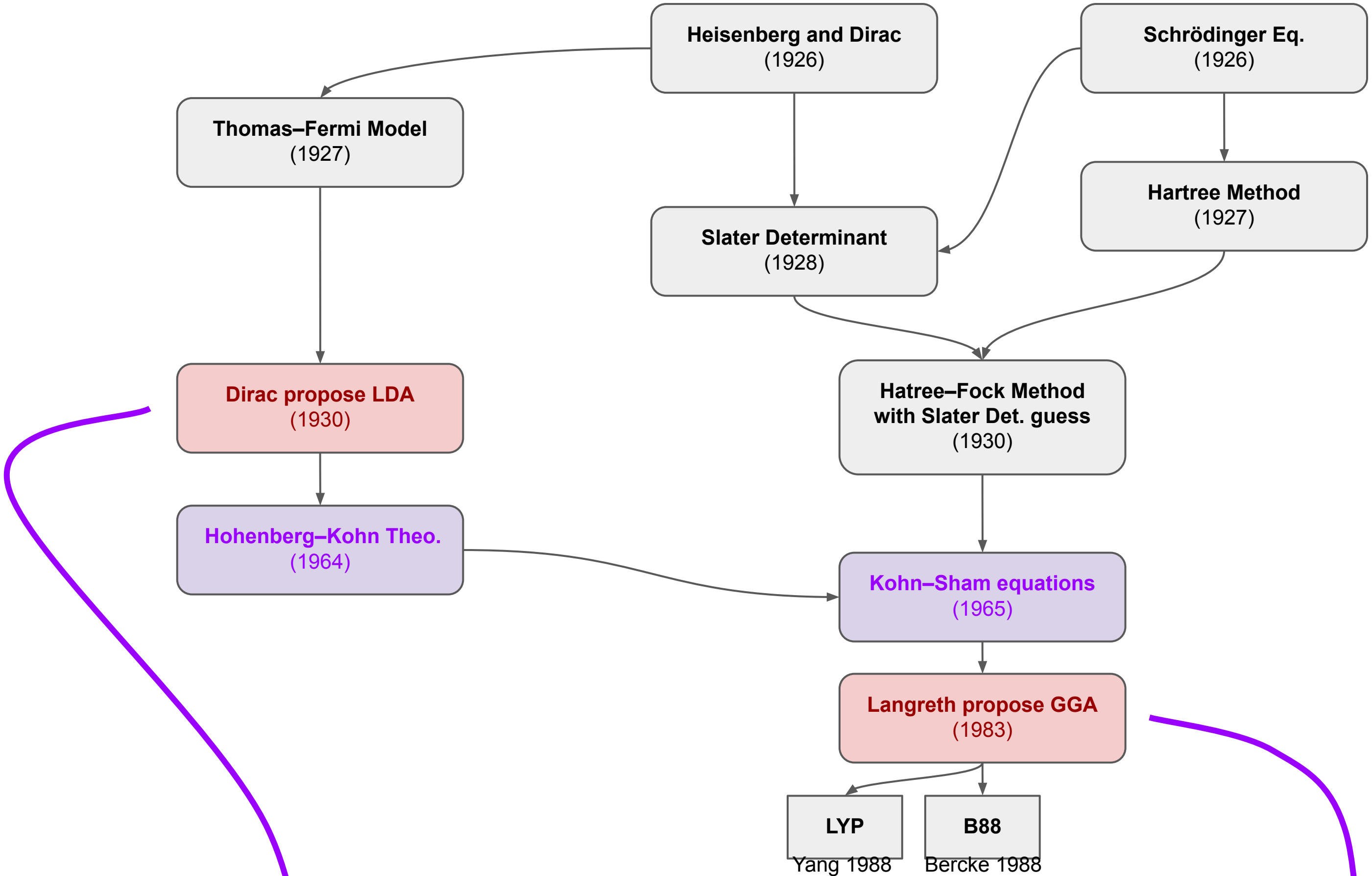
yes

$$E = \sum_i^N \varepsilon_i - E_{\text{H}}[\rho] + E_{\text{xc}}[\rho] - \int \frac{\delta E_{\text{xc}}[\rho]}{\delta \rho(\mathbf{r})} \rho(\mathbf{r})\, d\mathbf{r}.$$

$$E_{\text{xc}}[\rho(\mathbf{r})] = \int \rho(\mathbf{r}) \varepsilon_{\text{xc}}[\rho(\mathbf{r}), \nabla\rho(\mathbf{r})\ldots] d\mathbf{r}$$

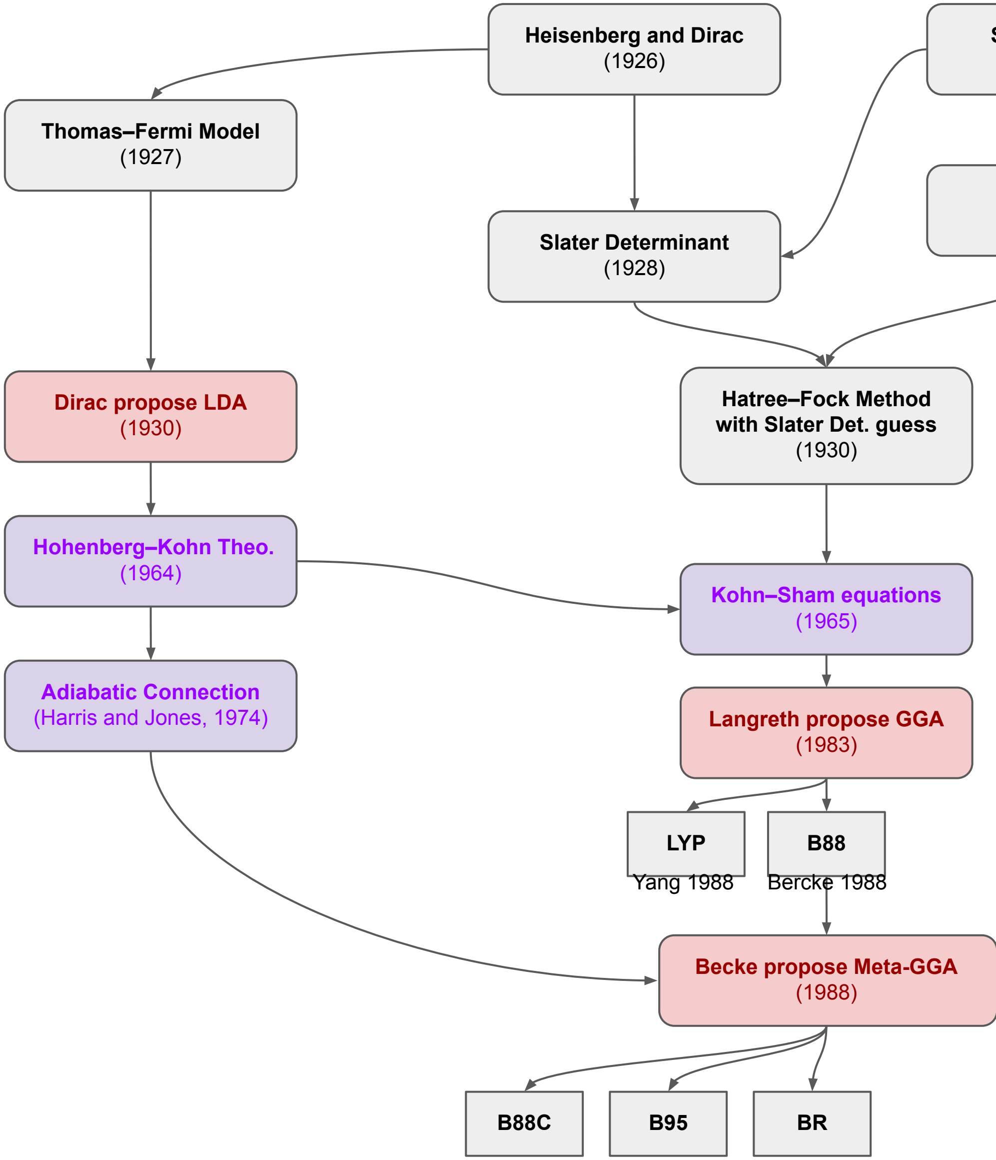
We know from the Hohenberg–Kohn Theorems that this $\boldsymbol{\varepsilon}_{\mathbf{xc}}$ **exist, but unfortunately it can not be derived from first principles.**

Building good approximations for this functional is on the core of the quality of DFT results and applicability.



$$E_{\mathrm{xc}}^{\mathrm{LDA}}[n] = \int n(\mathbf{r})\, \varepsilon_{\mathrm{xc}}[n]\, \mathrm{d}\mathbf{r}$$

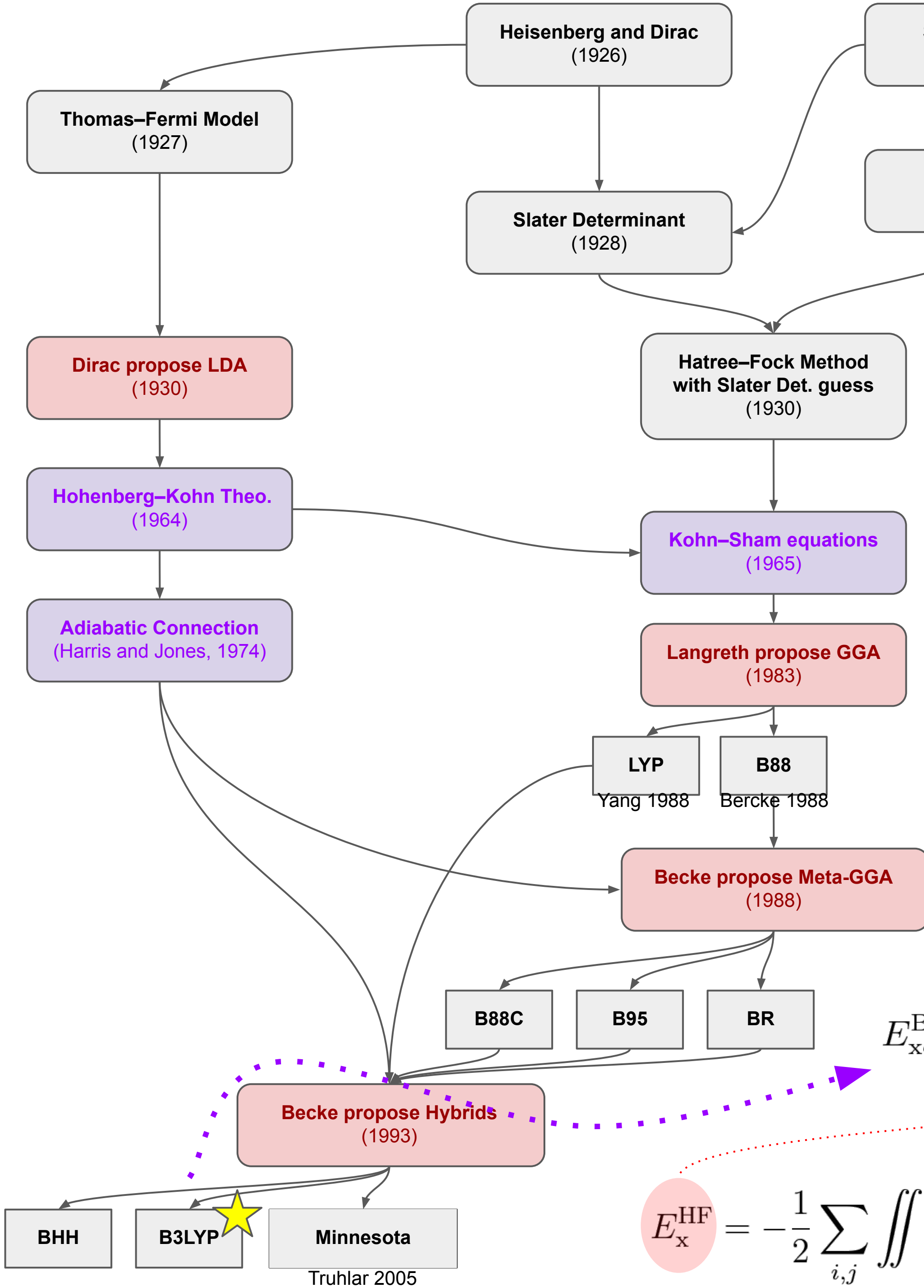
$$E_{\mathrm{xc}}^{\mathrm{GGA}}[n] = \int n(\mathbf{r})\, \varepsilon_{\mathrm{xc}}[n, \nabla n,]\, \mathrm{d}\mathbf{r}$$



Kinetic-energy density

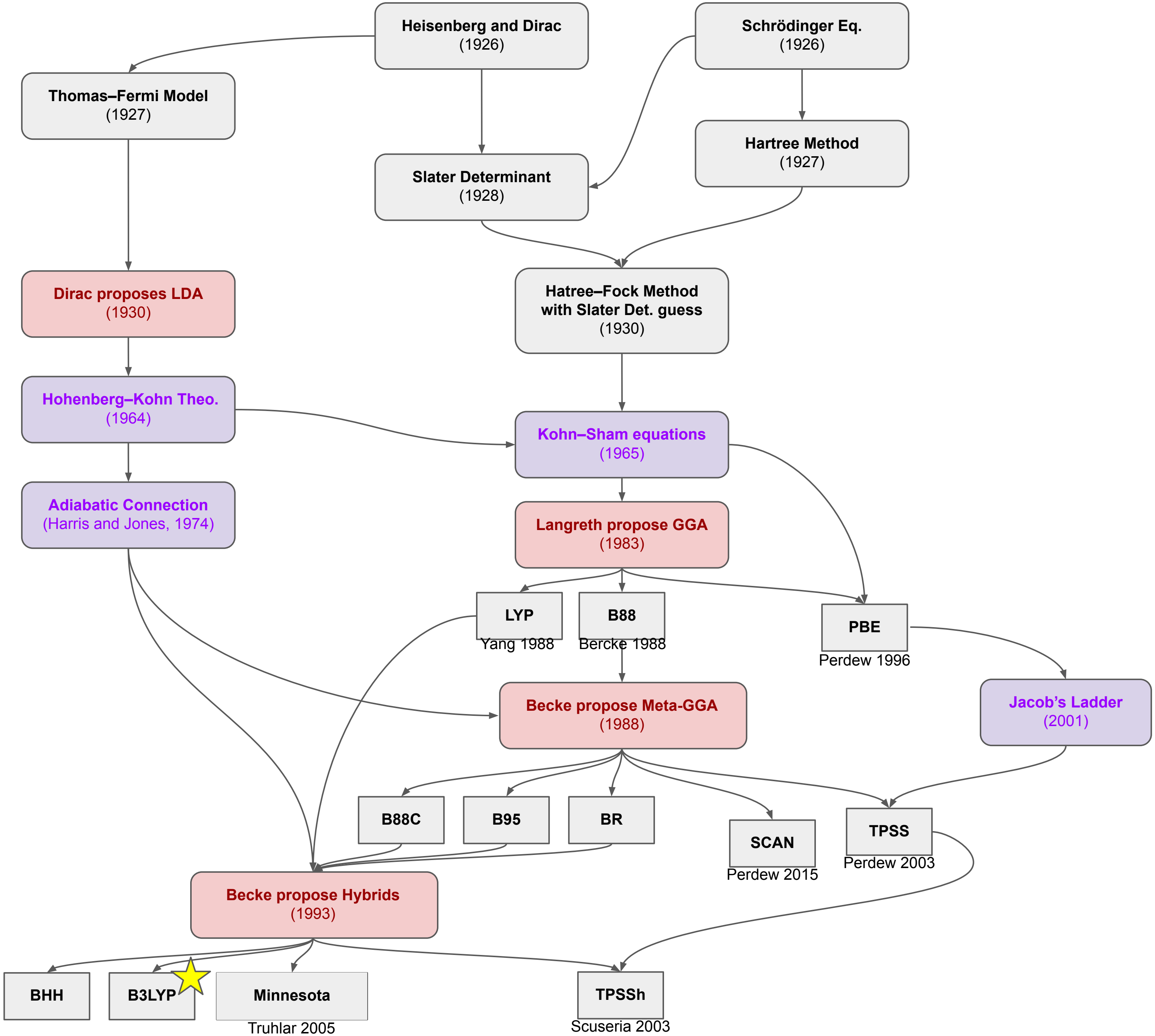
$$\tau = \frac{1}{2} \sum_{i}^{\text{occ}} |\nabla \varphi_i|^2$$

$$E_{\text{xc}}^{\text{meta-GGA}}[n] = \int n(\mathbf{r})\, \varepsilon_{\text{xc}}[n, \nabla n, \tau]\, \mathrm{d}\mathbf{r}$$



$$E_{\mathrm{xc}}^{\mathrm{B3LYP}} = (1-a)E_{\mathrm{x}}^{\mathrm{LDA}} + aE_{\mathrm{x}}^{\mathrm{HF}} + + b \vartriangle E_{\mathrm{x}}^{\mathrm{B88}} + (1-c)E_{\mathrm{c}}^{\mathrm{LDA}} + cE_{\mathrm{c}}^{\mathrm{LYP}}$$

$$E_{\mathrm{x}}^{\mathrm{HF}} = -\frac{1}{2}\sum_{i,j}\iint \varphi_i^*(\mathbf{r})\varphi_j^*(\mathbf{r}')\frac{1}{|\mathbf{r}-\mathbf{r}'|}\varphi_j(\mathbf{r})\varphi_i(\mathbf{r}')\,d\mathbf{r}\,d\mathbf{r}'$$



Heisenberg and Dirac
(1926)
Schrödinger Eq.
(1926)
Thomas–Fermi Model
(1927)
Slater Determinant
(1928)
Hartree Method
(1927)
Dirac proposes LDA
(1930)
Hatree–Fock Method
with Slater Det. guess
(1930)
Hohenberg–Kohn Theo.
(1964)
Kohn–Sham equations
(1965)
Adiabatic Connection
(Harris and Jones, 1974)
Langreth propose GGA
(1983)
LYP
Yang 1988
B88
Bercke 1988
PBE
Perdew 1996
Becke propose Meta-GGA
(1988)
Jacob's Ladder
(2001)
B88C
B95
BR
SCAN
Perdew 2015
TPSS
Perdew 2003
Becke propose Hybrids
(1993)
BHH
B3LYP
Minnesota
Truhlar 2005
TPSSh
Scuseria 2003

"And he[Jabob] dreamed, and behold a ladder set up on the earth, and the top of it reached to heaven; and behold the angels of God ascending and descending on it."

[Gen 28:12]

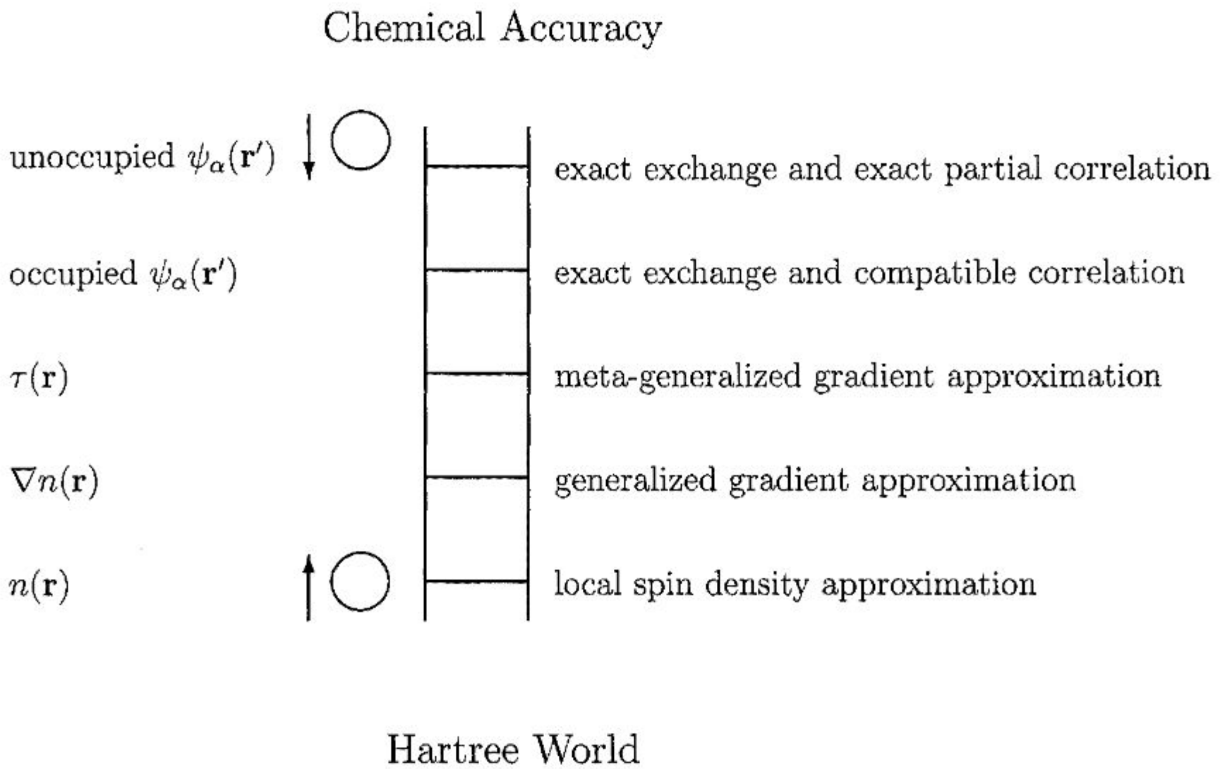
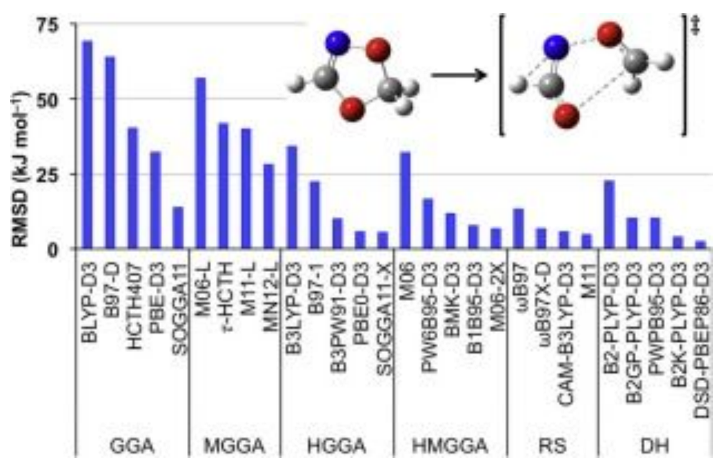


FIGURE 1. Jacob's ladder of density functional approximations. Any resemblance to the Tower of Babel is purely coincidental. Also shown are angels in the spherical approximation, ascending and descending. Users are free to choose the rungs appropriate to their accuracy requirements and computational resources. However, at present their safety can be guaranteed only on the two lowest rungs.

Adapted from ***Challenges for Density Functional Theory***
A. J. Cohen, P. Mori-Sánchez, and W. Yang
dx.doi.org/10.1021/cr200107z

So why are we searching for new functionals?

- Challenge 1: The Need To Improve the Description of Reaction Barriers and Dispersion/van der Waals Interactions
  - Due to the **local nature of the LDA or GGA functional form**, it is not possible for these functionals to accurately describe this **non-local phenomena**.
  - Non-local functionals build over Hartree-Fock exchange are also completely wrong, since they all exhibit long-range repulsive behavior. **The performance of most popular functionals on simple weakly bound dimers is extremely poor**.



10.1016/j.chemphys.2015.07.005m

Adapted from ***Challenges for Density Functional Theory***
A. J. Cohen, P. Mori-Sánchez, and W. Yang
dx.doi.org/10.1021/cr200107z

So why are we searching for new functionals?

- Challenge 2: Delocalization Error and Static Correlation Error
  - **a single electron can interact with itself**, known as self-interaction error.
  - **The exact functional on the Hohenberg–Kohn Theory does not have any self-interaction**, i.e., the exchange energy exactly cancels the Coulomb energy for one electron.
  - This and similar errors are at the heart of many failures with the currently used approximations.

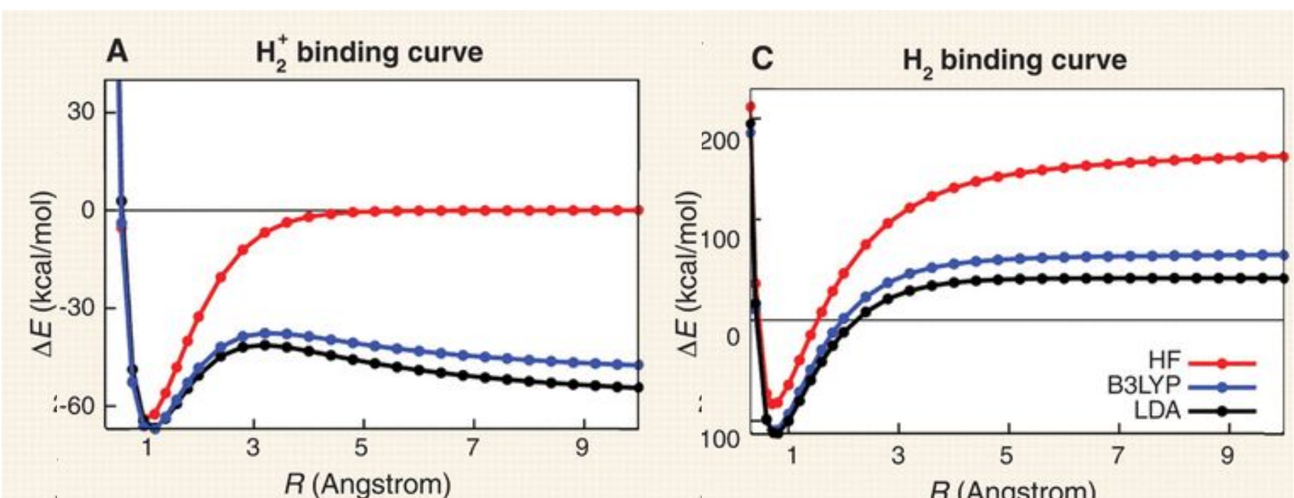
$$n(\mathbf{r}) = \left|\varphi_0(\mathbf{r})\right|^2$$

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V_{\text{ext}}(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}^3\mathbf{r}' + V_{\text{xc}}[n(\mathbf{r})]\right]\varphi_i(\mathbf{r}) = \varepsilon_i\varphi_i(\mathbf{r})$$

Adapted from ***Challenges for Density Functional Theory***
A. J. Cohen, P. Mori-Sánchez, and W. Yang
dx.doi.org/10.1021/cr200107z

So why are we searching for new functionals?

- Challenge 3: The Energy of Two Protons Separated by Infinity with One and Two Electrons: Strong Correlation
  - Except for multiconfigurational methods, most mean-field theories struggle to describe strongly correlated systems. This is evident from some very simple tests involving infinitely separated protons with varying numbers of electrons. **Currently, all functionals fail even for the simplest of these, infinitely stretched H_2^+ and infinitely stretched H_2.**
  - In order to satisfy exact fundamental conditions and not to suffer from systematic errors, the energy functionals must have the correct discontinuous behavior at integer numbers of electrons.



10.1126/science.1158722

How are people trying to push *DFA* forward?

- **Range separation and Local Hybrids:** This was motivated by localization errors on DFAs. The idea, originally from the groups of Savin and Gill, whas to separate the electron-electron interaction into **two parts, one long-range and the other short-range**.

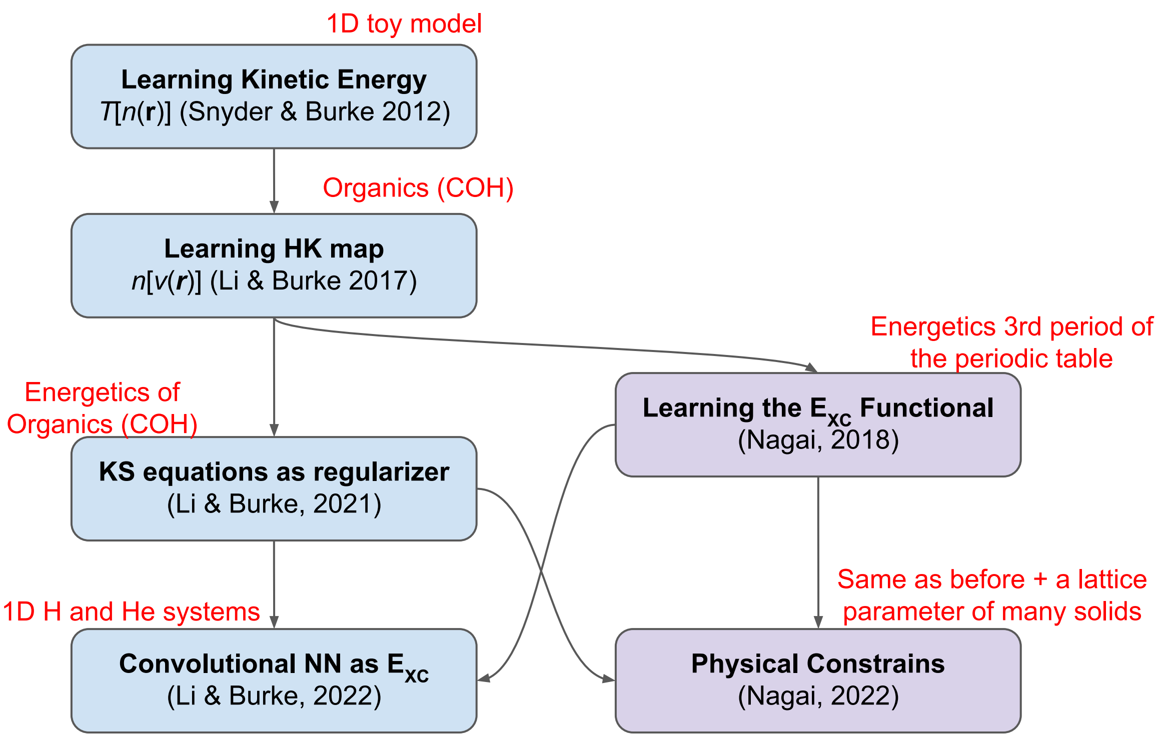
- **Objective Oriented Fitting:** Making educated guesses on the shape of the DFA functional, one can **fit parameters to suit specific research interests**. The efforts on this direction is what motivate the recent additions to the Minnesota family of functionals.

- **Generalized Adiabatic Connection:** Normal hybrids, like B3LYP, **use a linear approximation to the adiabatic connection**. Both Perdew, Burke (from PBE) and Yang (from LYP) have already expanded this approach using larger orders. Some functionals: ISI (2000), SPL (1999) and LB (2009).

- **Functional of Unoccupied Orbitals:** The idea here is to add a electron-electron interaction model that models excitations via the **inclusion of the unoccupied KS orbitals in the exchange and correlation functional**. Examples include LDA+U (2012) and the MP2 based B2PLYP (2006).

And where does Machine Learning appear?

- **We are in the *Objective Oriented Fitting*:** All work on machine learning functionals is on that category. Machine learning is a data driven approach, so its subject dependent by definition.



"... [ML approximations] achieves chemical accuracy using many more inputs, but requires far less insight into the underlying physics."

(K. Burke, 2012)

Why Neural Networks and not a simpler thing?

- We want to approximate $E_{xc}[n(\mathbf{r})]$, that we know **exist and is unique**.
- But we also know that **the one-to-one correspondence between n and E_{xc} contains non-analytic structures** (e.g., discontinuities and singularities).
- Neural Networks are shown to be good **universal approximators** even in the case of intricate functions.

Multilayer Feedforward Networks are Universal Approximators

KUR' HORNIK

Technische Universität Wien

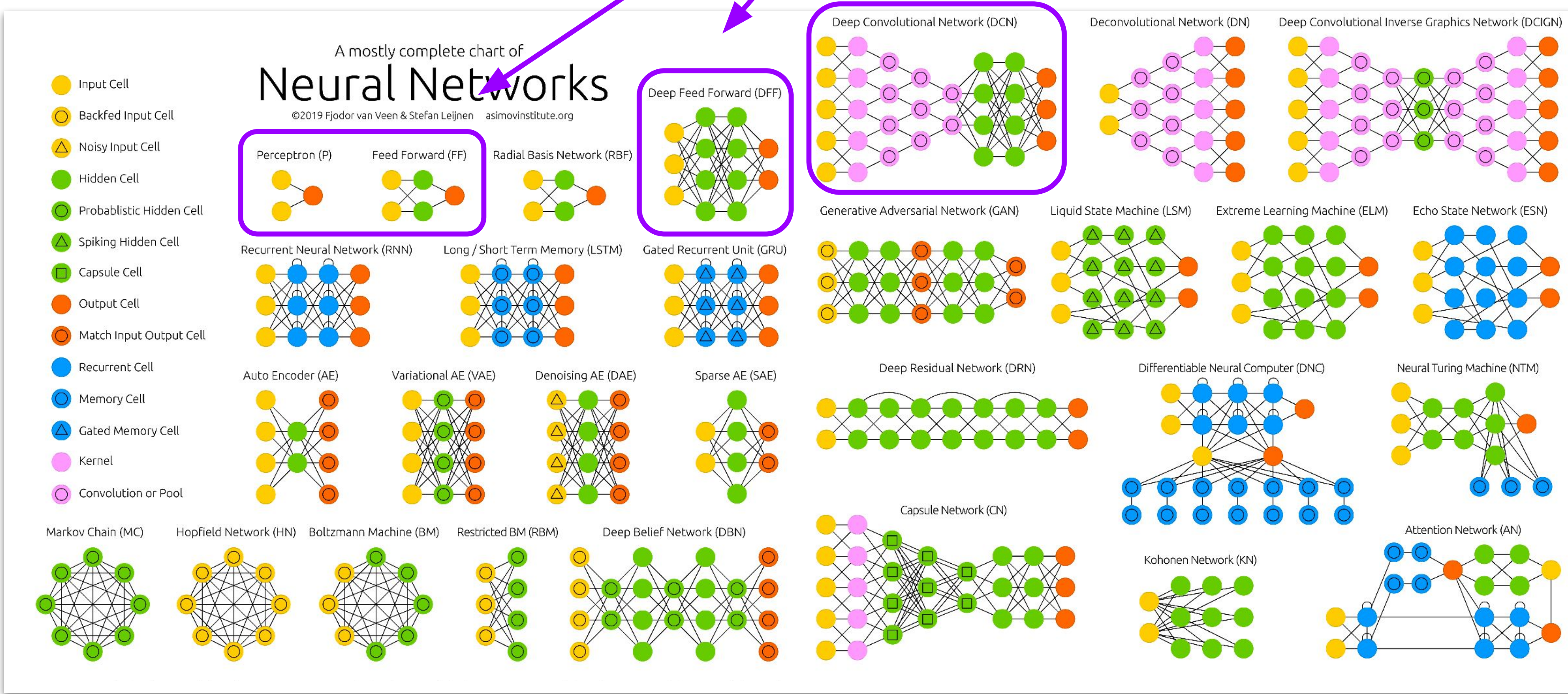
MAXWELL STINCHCOMBE AND HALBEK WHITE

University of California, San Diego

(*Received 16 September 1988; revised and accepted 9 March 1989*)

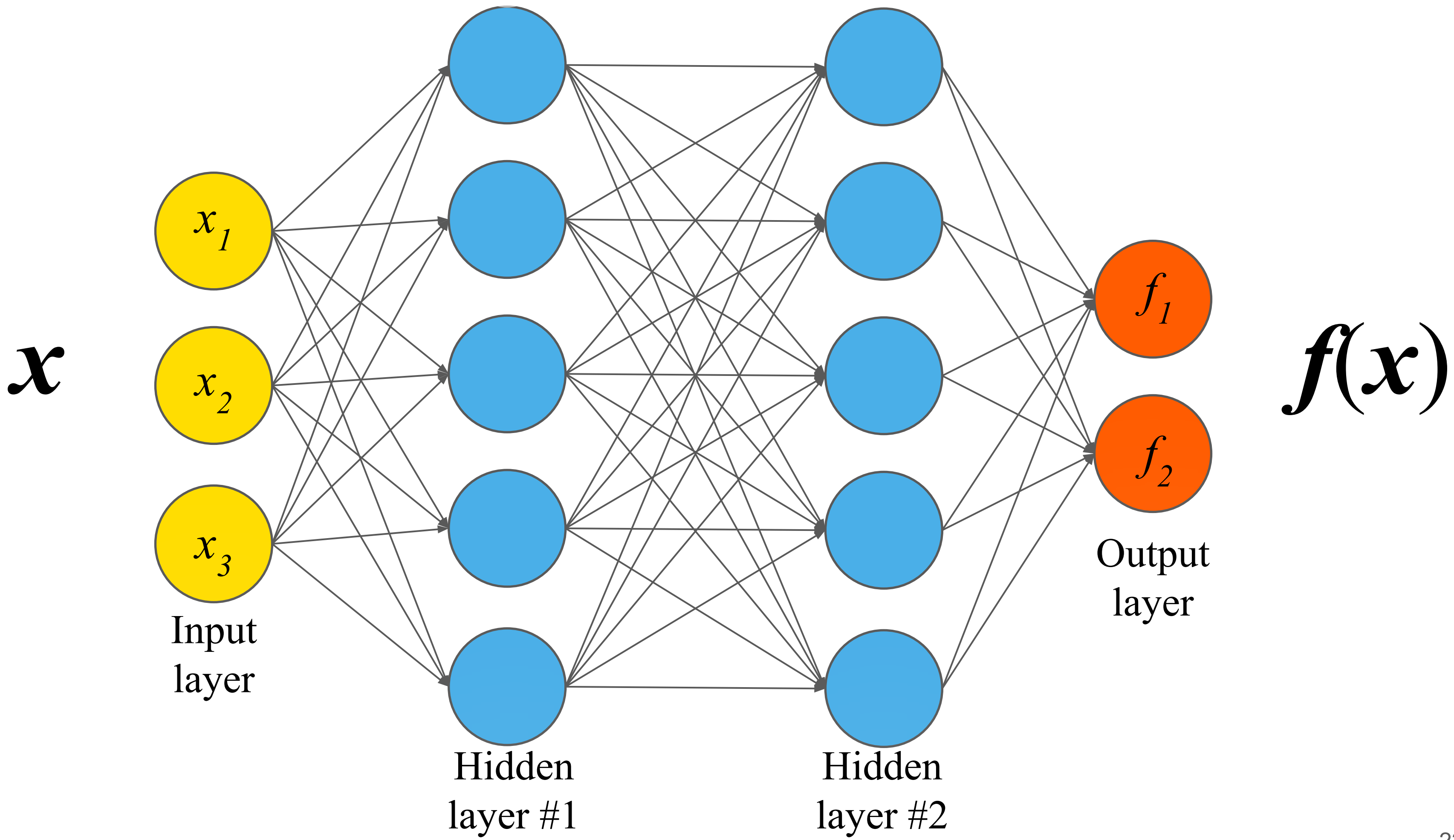
Abstract—*This paper rigorously establishes that standard multilayer feedforward networks with as few as one hidden layer using arbitrary squashing functions are capable of approximating any Borel measurable function from one finite dimensional space to another to any desired degree of accuracy, provided sufficiently many hidden units are available. In this sense, multilayer feedforward networks are a class of universal approximators.*

Already used on Machine Learning Functionals

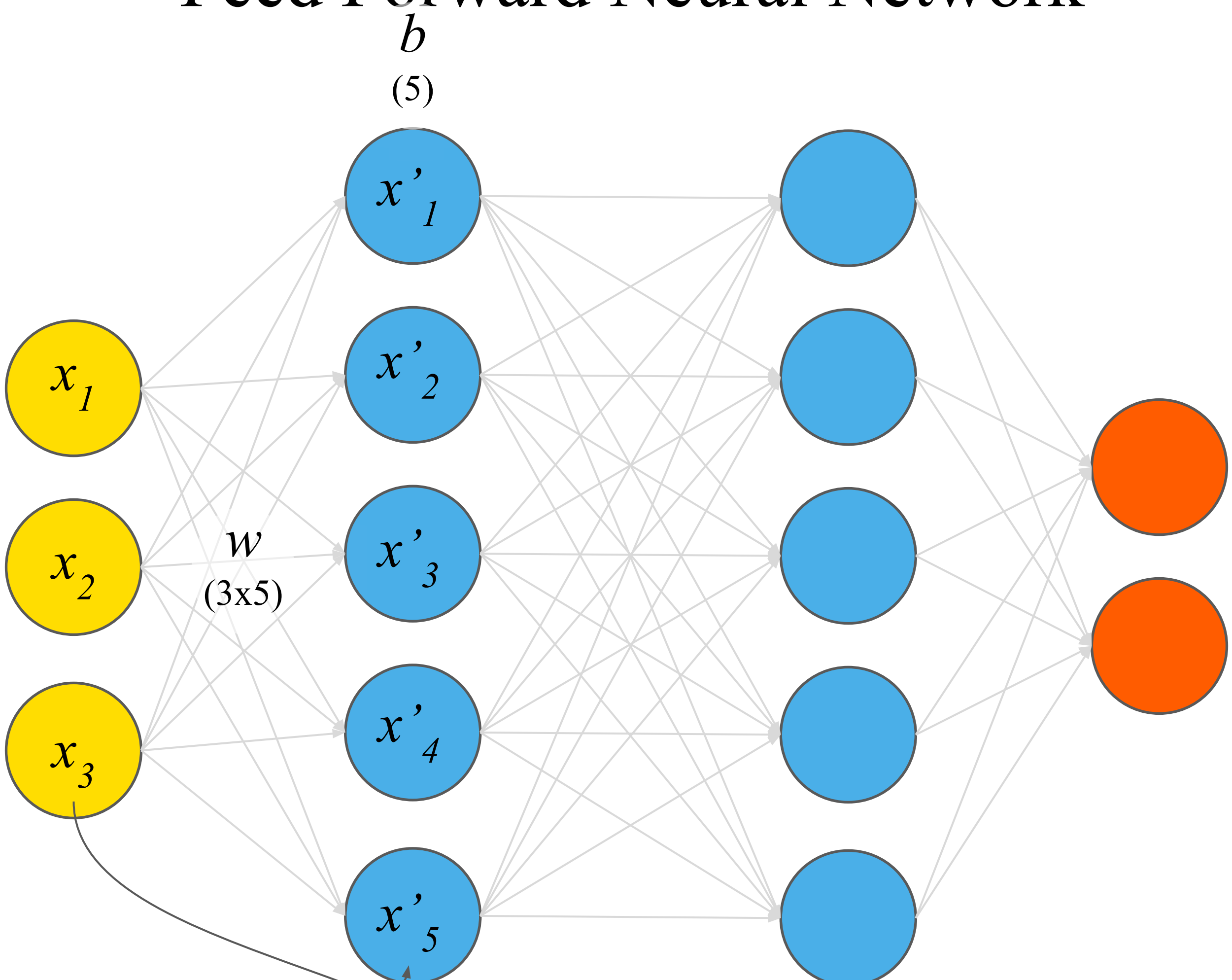


The Neural Network Zoo - *The Asimov Institute*

Feed Forward Neural Network



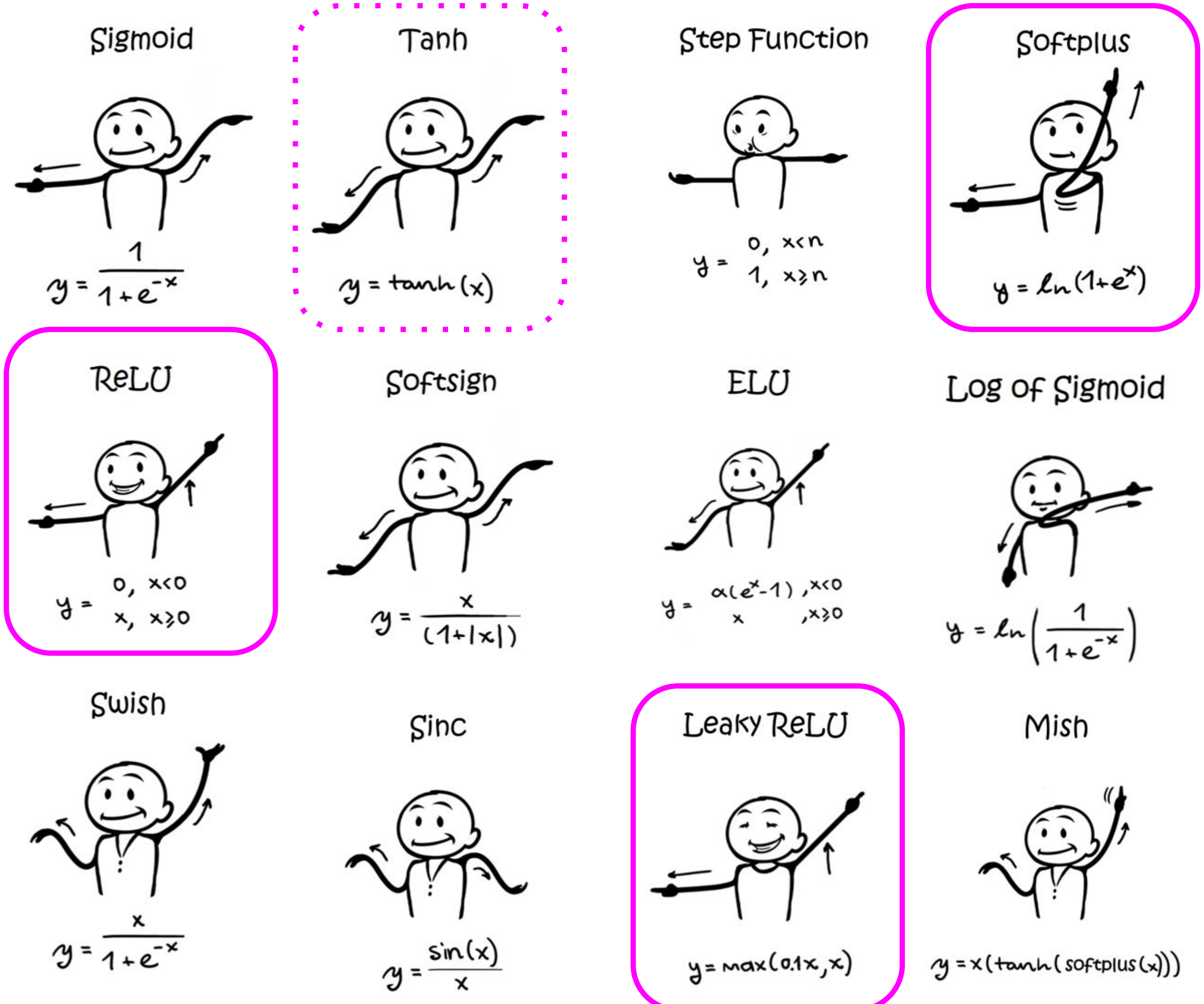
Feed Forward Neural Network



$$x'_i = f_{\text{activation}}(w_{ij} x_j + b_i)$$

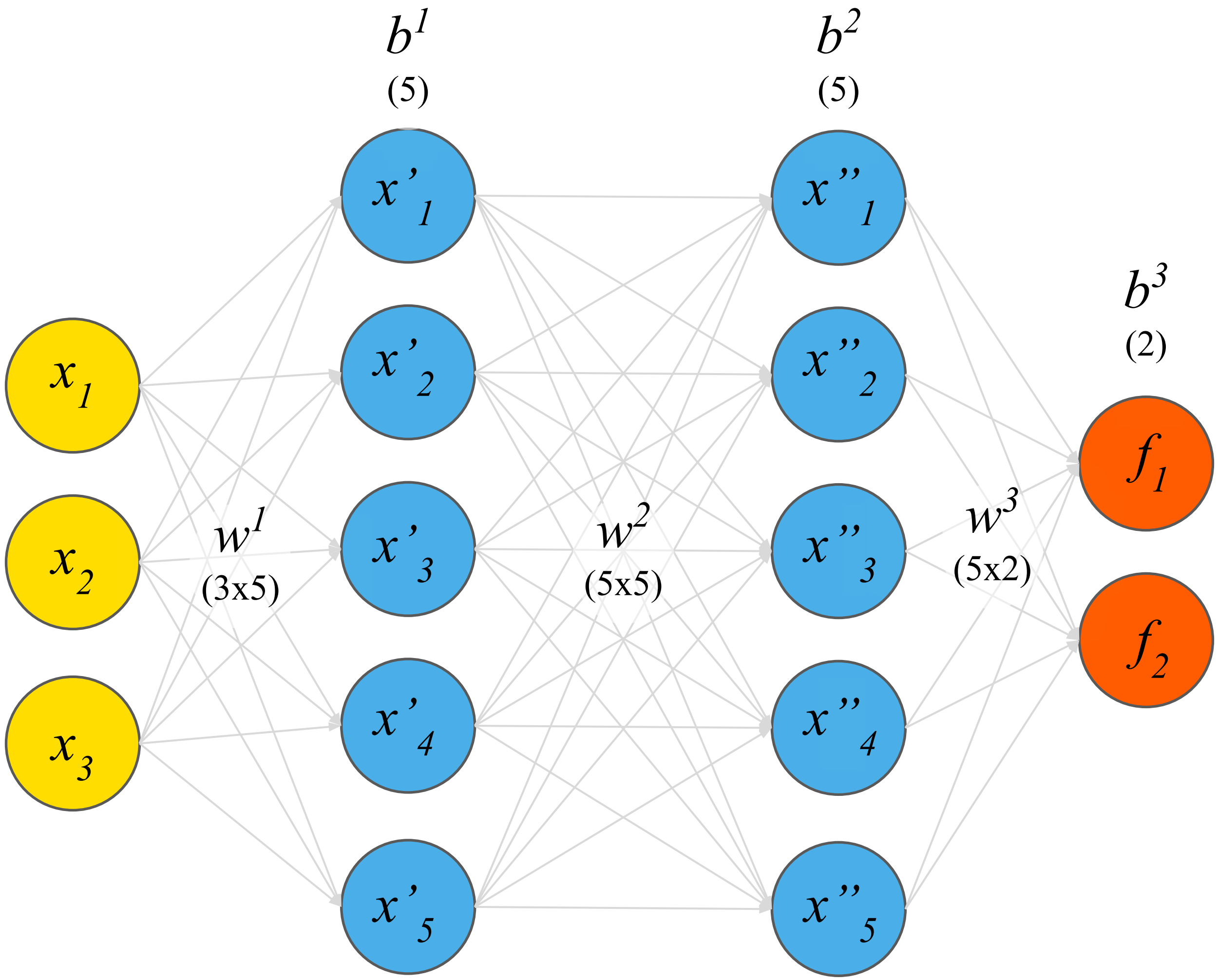
Inspired on how neurons work!

Already used on Machine Learning Functionals



Illustrations by Sefik Ilkin Serengil

Feed Forward Neural Network

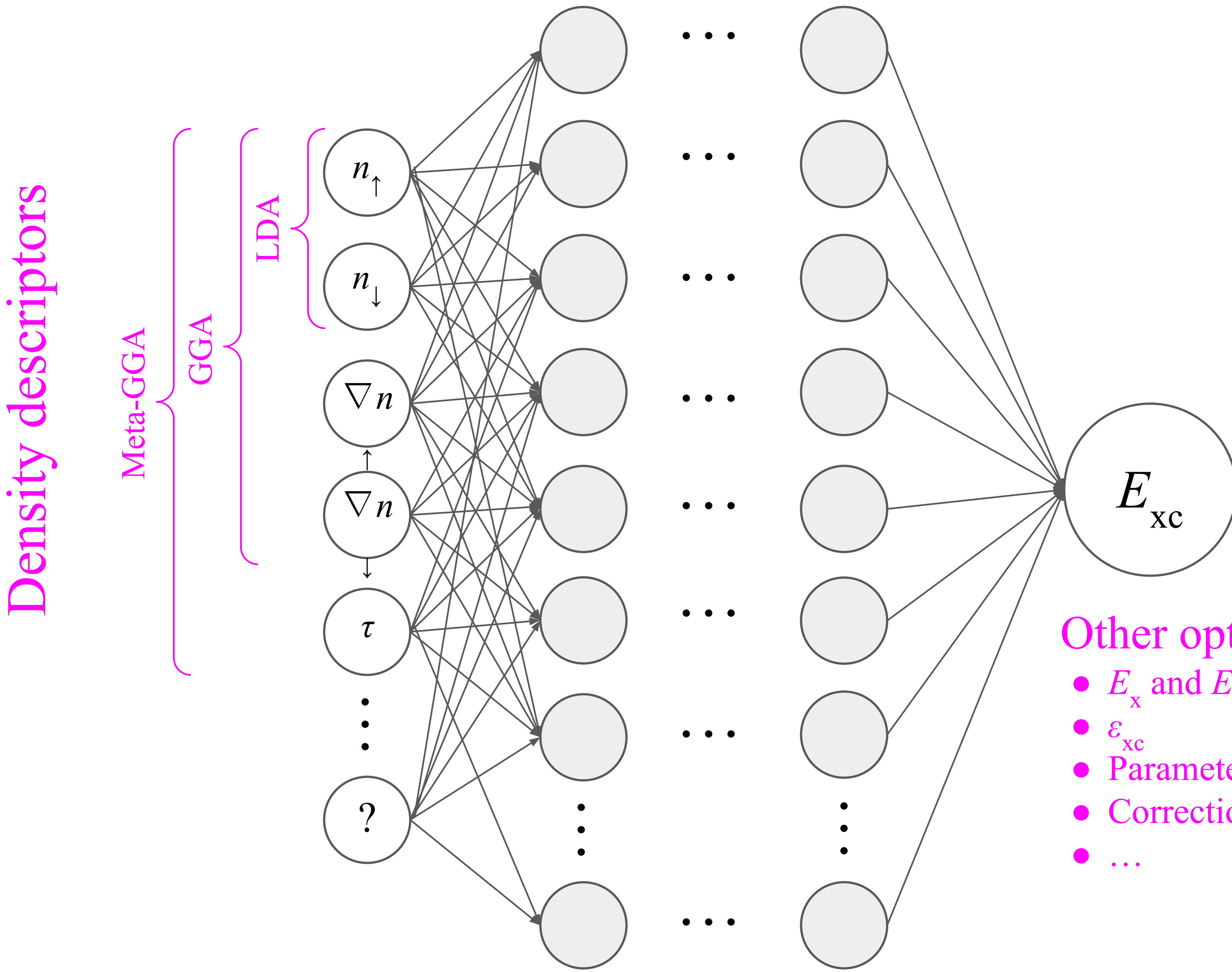


$$x'_i = f_{\text{activation}}(w^1_{ij} x_j + b^1_i)$$

$$f_i = f_{\text{activation}}(w^3_{ij} x'_j + b^3_i)$$

$$x''_i = f_{\text{activation}}(w^2_{ij} x'_j + b^2_i)$$

Artificial Neural Network as E_{xc} Functionals



Other options:

- E_x and E_c
- ε_{xc}
- Parameters
- Corrections
- …

How do we get values for weights and biases?

You will need to define a loss function (Δ_{err}).
It should be:

- *Be quick to compute,*
- *Represente the physico-chemical accuracy of the model*

Δ_{err} defines an order and equivalence relation in the solution space, having a global minimum on the perfect model.

TRAINING = OPTIMIZING Δ_{err}

Relevant Features

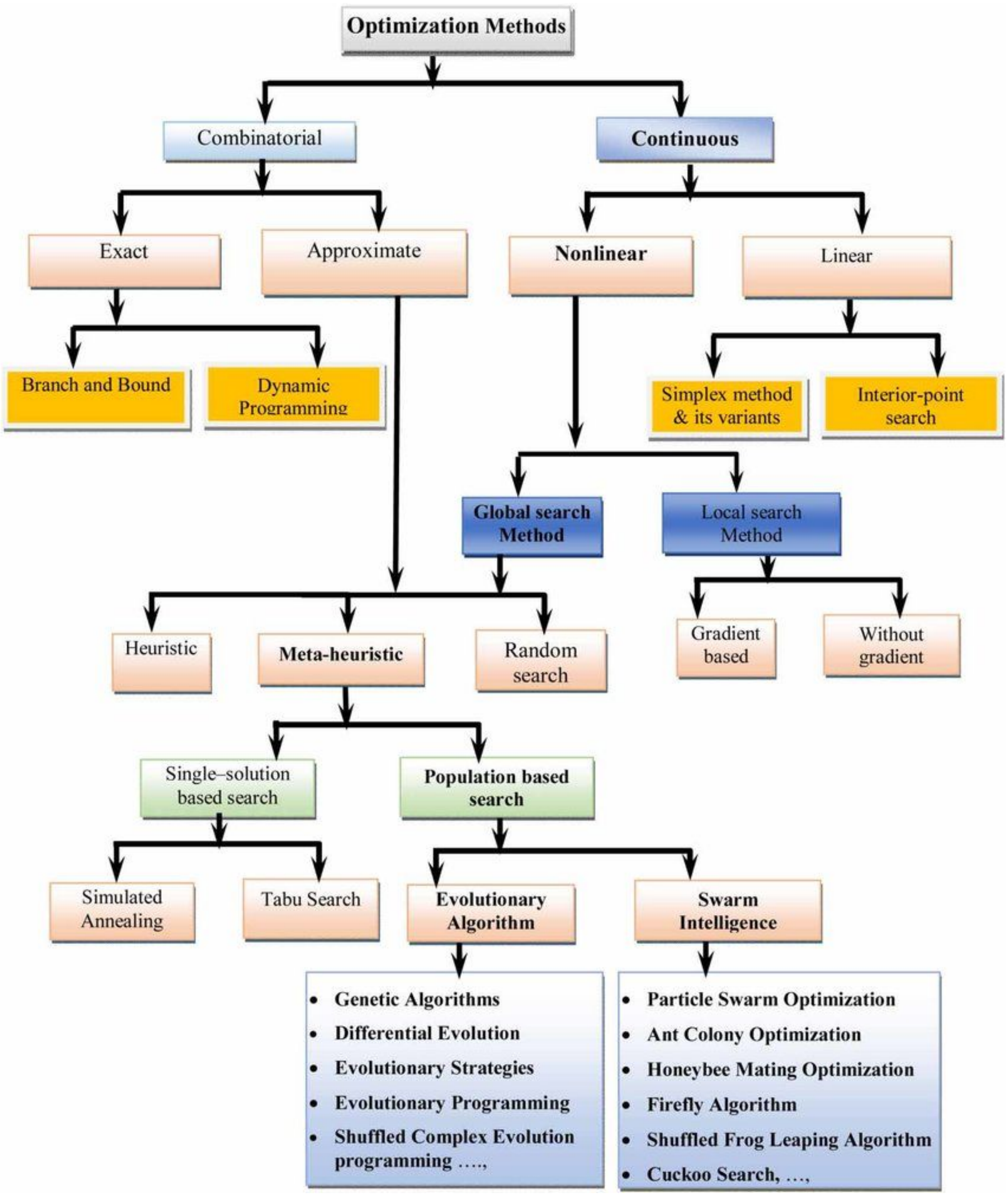
i.e., the quantities in which my training should be based on.

Shape of the Loss Function

i.e., the relations between the features and how they define a good functional.

Training Data Set

i.e., defining the data used on the calculation of the loss function.



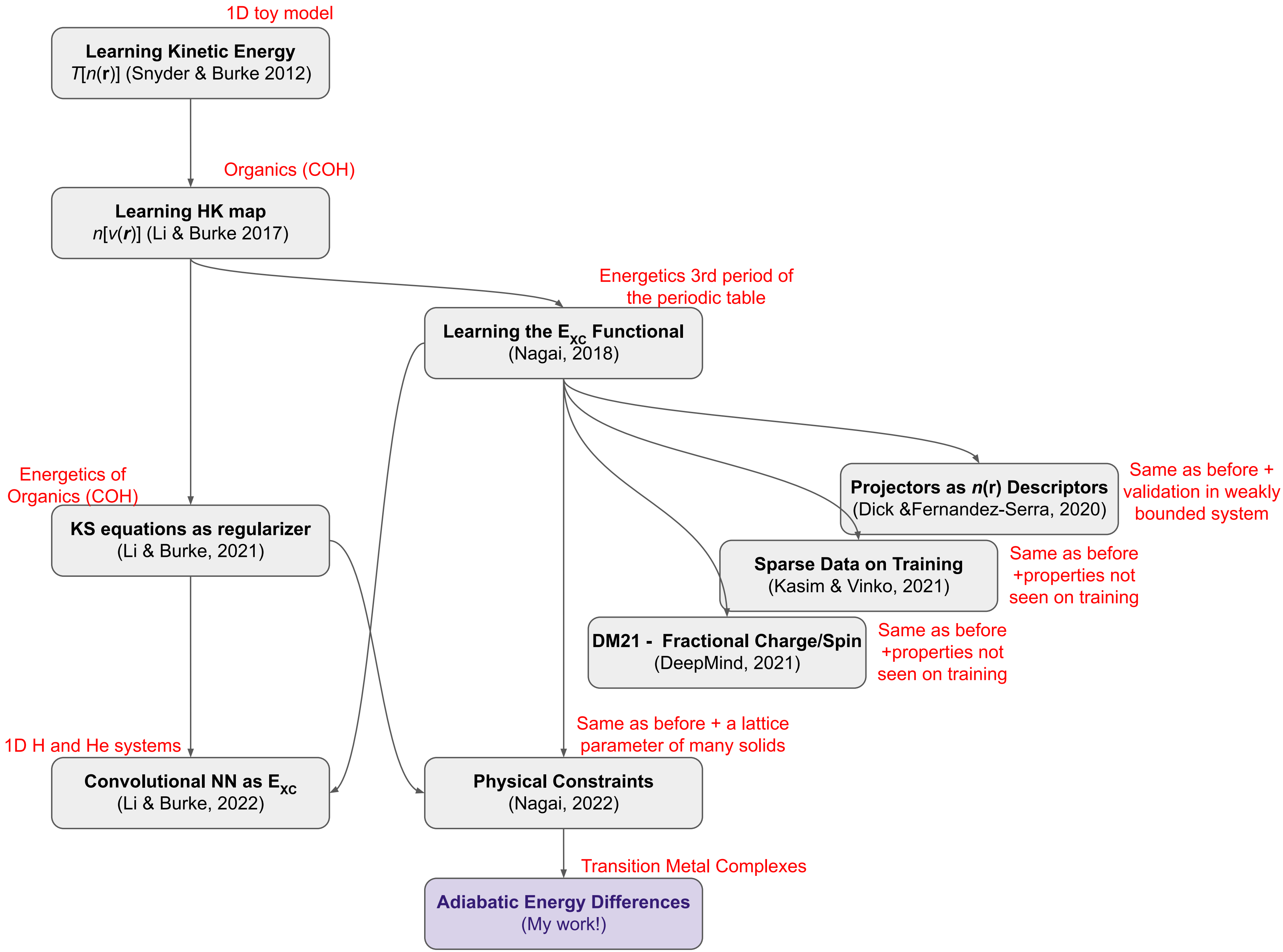
$$\Delta_{\text{err}} = c_1 \sum_i |AE_i^{\text{ref}} - AE_i^{\text{DFT}}| + c_2 \sum_j \frac{1}{N_e} \int |\rho_j^{\text{ref}} - \rho_j^{\text{DFT}}| dV$$

Combinatorial or continuous?

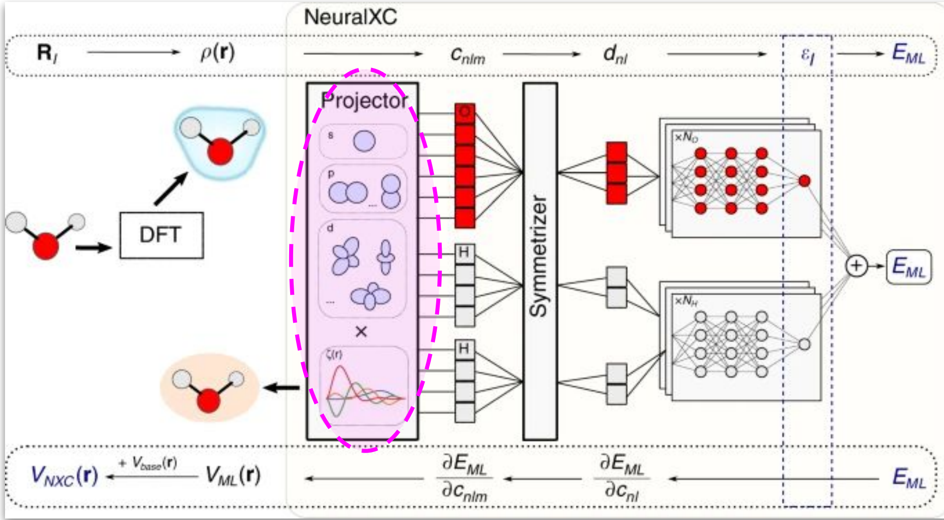
Non-gradient or gradient based?

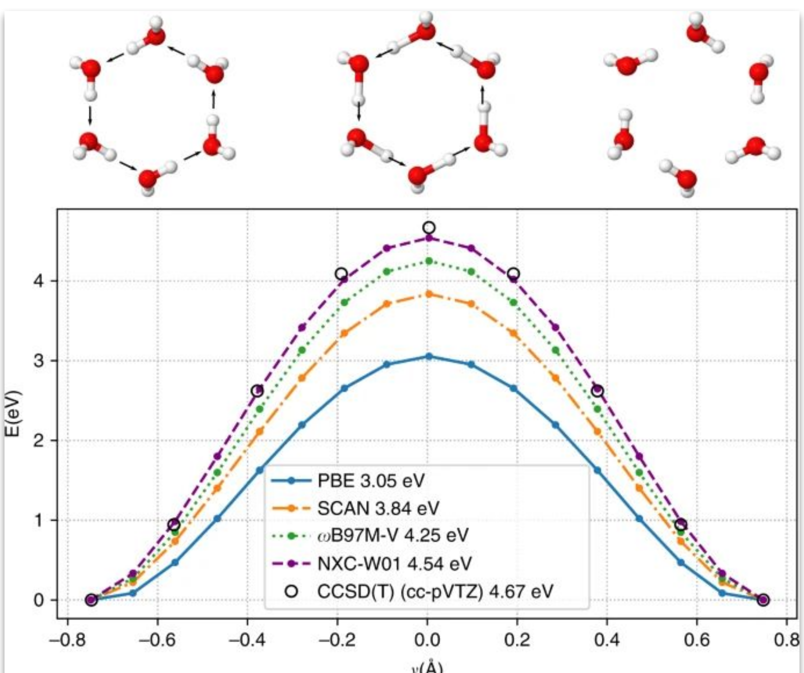
Can we afford random search?

From 10.2166/h2oj.2020.128

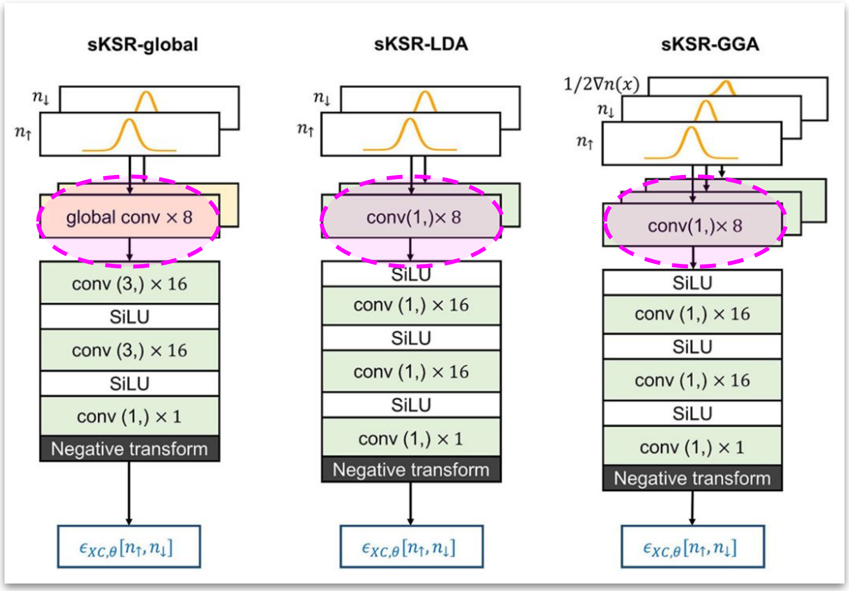

1D toy model
Learning Kinetic Energy
$T[n(\mathbf{r})]$ (Snyder & Burke 2012)
Organics (COH)
Learning HK map
$n[v(\boldsymbol{r})]$ (Li & Burke 2017)
Energetics 3rd period of the periodic table
Learning the E_{xc} Functional
(Nagai, 2018)
Projectors as n(r) Descriptors
(Dick &Fernandez-Serra, 2020)
Same as before + validation in weakly bounded system
Energetics of Organics (COH)
KS equations as regularizer
(Li & Burke, 2021)
Sparse Data on Training
(Kasim & Vinko, 2021)
Same as before +properties not seen on training
DM21 - Fractional Charge/Spin
(DeepMind, 2021)
Same as before +properties not seen on training
Same as before + a lattice parameter of many solids
1D H and He systems
Convolutional NN as E_{xc}
(Li & Burke, 2022)
Physical Constraints
(Nagai, 2022)
Transition Metal Complexes
Adiabatic Energy Differences
(My work!)

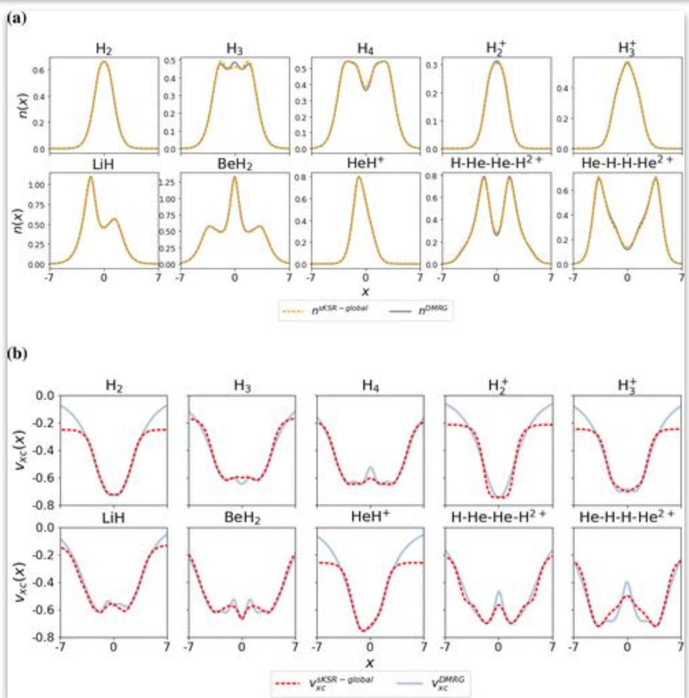
Projectors as n(r) Descriptors
(Dick & Fernandez-Serra, 2020)



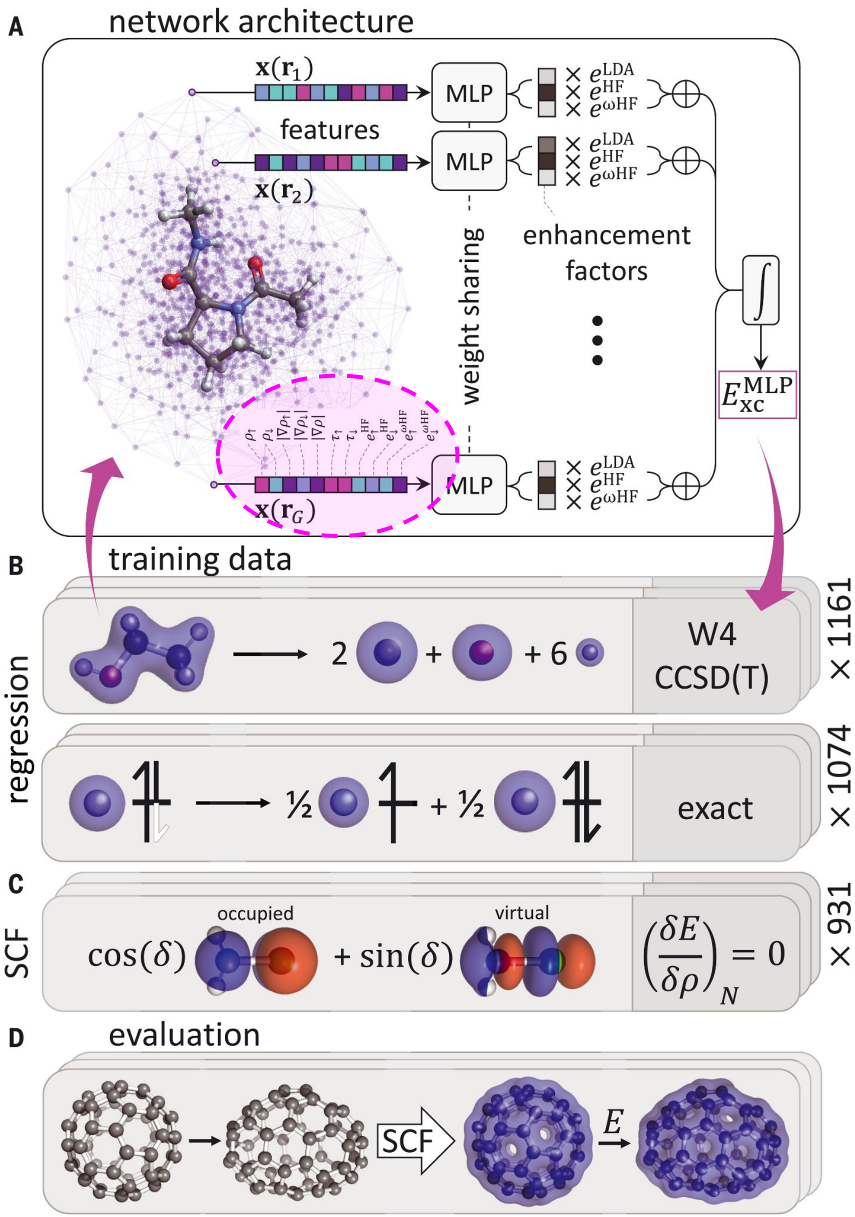


Convolutional NN as EXC
(Li & Burke, 2022)

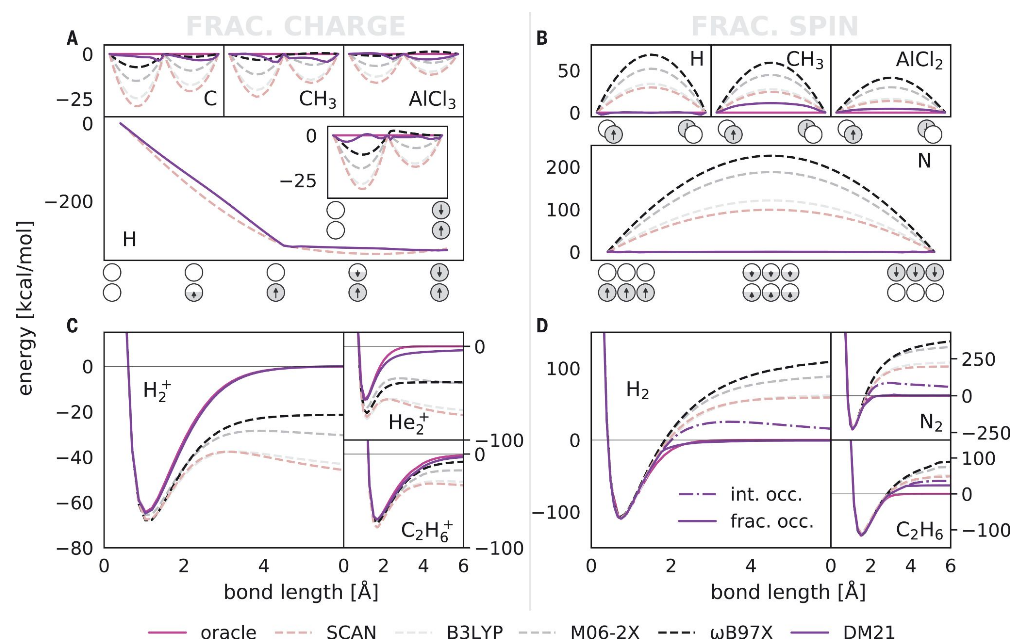




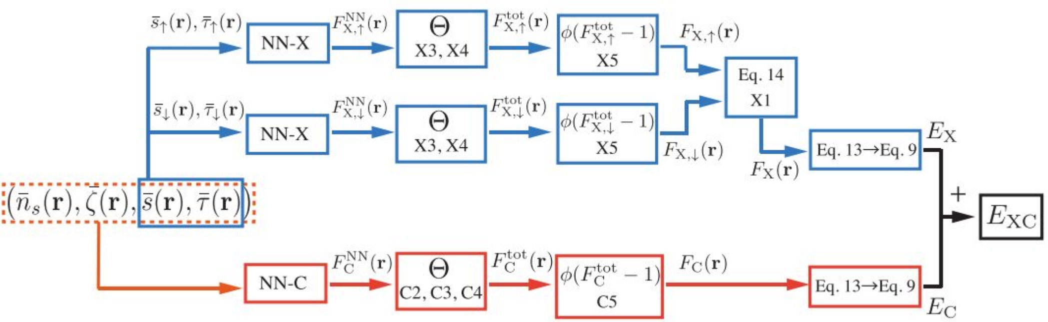
DM21 - DFT Functional from DeepMind, 2021



“The resulting functional, DM21 (DeepMind 21), correctly describes typical examples of **artificial charge delocalization** and **strong correlation** and performs better than traditional functionals on thorough **benchmarks for main-group atoms and molecules.**”



Physical Constraints (Nagai, 2022)



	Physical constraint	$\mathbf{x_0}$
X1	Correct uniform coordinate density-scaling behavior	
X2	Exact spin scaling relation	
X3	Uniform electron gas limit	$(0, 0)$
X4	F_{X} vanishes as $s^{-1/2}$ at $s \to \infty$	$(1, \bar{\tau})$
X5	Negativity of ε_{X}	
C1	Uniform electron gas limit	$(\bar{n}_s, \bar{\zeta}, 0, 0)$
C2	Uniform density scaling to low-density limit	$(0, \bar{\zeta}, \bar{s}, \bar{\tau})$
C3	Weak dependence on ζ in low-density region	
C4	Uniform density scaling to high-density limit	$(1, \bar{\zeta}, \bar{s}, \bar{\tau})$
C5	Nonpositivity of ε_{C}	

TABLE II. Benchmark results for atomization energies of 144 molecules. (MAE, mean absolute error; ME, mean error; SD, standard deviation of signed error.)

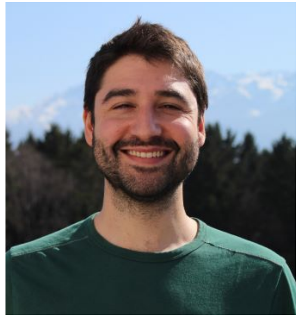
XC	MAE (kcal/mol)	ME (kcal/mol)	SD (kcal/mol)
PBE	17.3	16.2	13.1
SCAN	6.2	−4.5	5.7
NN-based	4.8	1.8	6.3
pcNN-based	3.6	0.3	4.5

TABLE III. Benchmark results for lattice constants of 48 solids. Parentheses in the “NN-based” row indicate that the numerical calculations for six solids did not converge; thus, only the converged calculations were used for the statistics.

XC	MAE (mÅ)	ME (mÅ)	SD (mÅ)
PBE	38.1	33.9	44.2
SCAN	22.3	−7.5	28.5
NN-based	(22.9)	(0.8)	(32.0)
pcNN-based	19.1	-2.5	26.5

Artificial Neural Network Meta-GGA based Density Functional for **Adiabatic Energy Differences** in **Transition Metal Complexes**

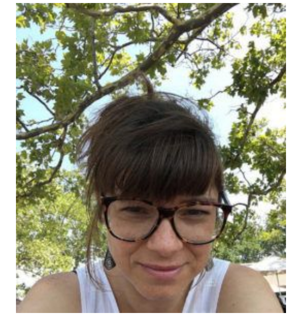
Others involved in this work:



Lorenzo Mariano (Postdoc, Trinity College)

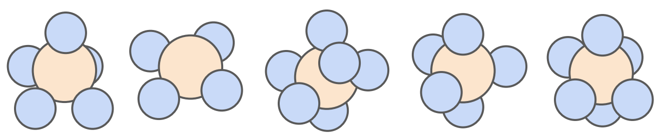
Emilie Devijver (CNRS, LIG)

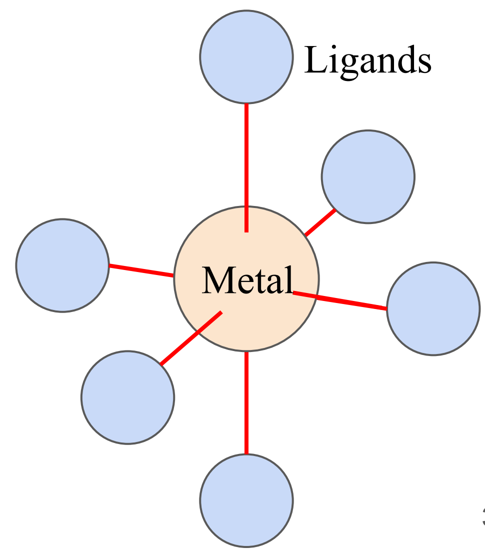
Noel Jakse (G-INP, SIMaP)



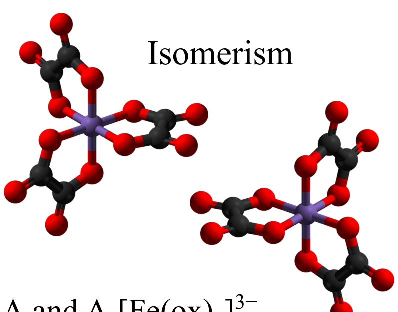
Roberta Poloni (CNRS, SIMaP)

Available at **[arXiv:2304.07899]**

Many possible symmetries:  …

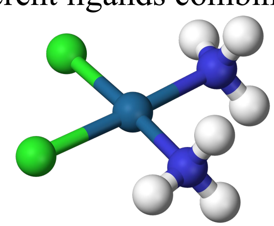


Isomerism

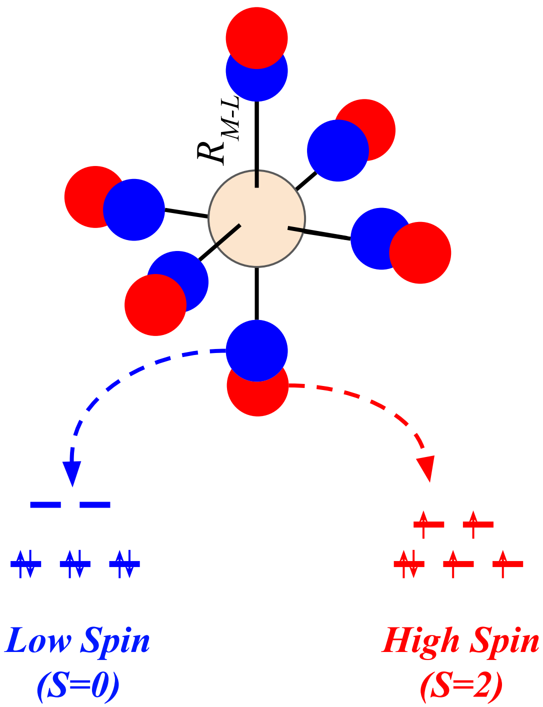


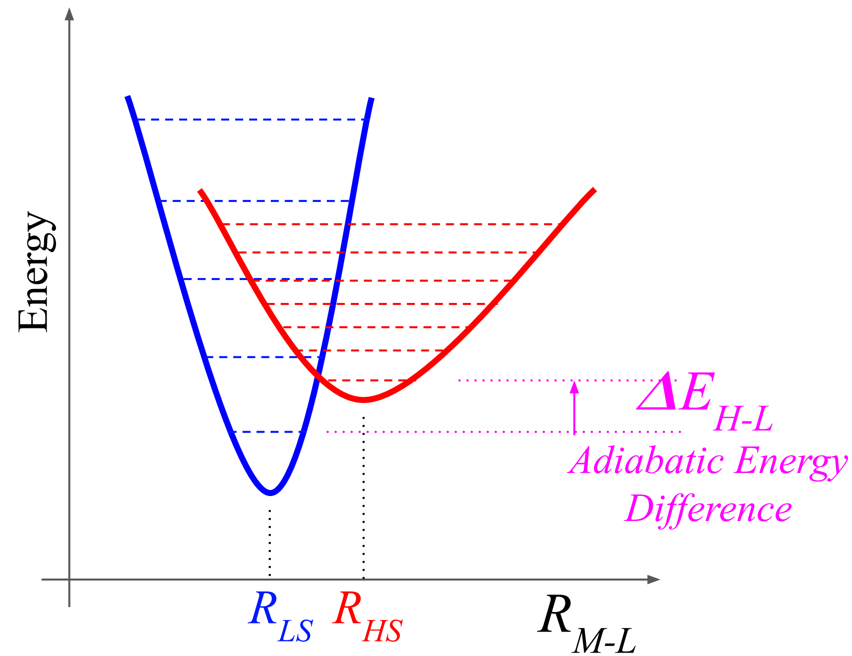
Λ and Δ-$[Fe(ox)_3]^{3-}$

Different ligands combination



$PtCl_2(NH_3)_2$



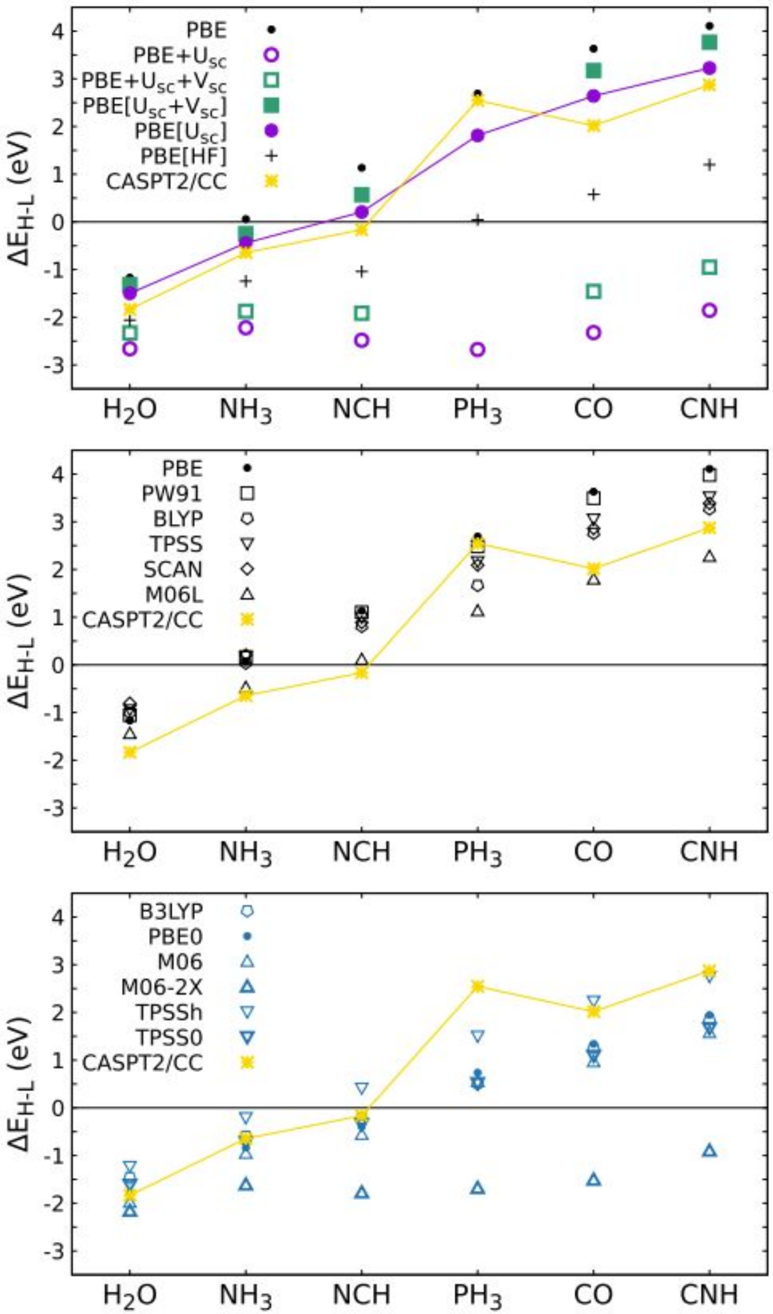


Spin crossover (SCO) is a phenomenon that occurs in some metal complexes wherein the spin state of the complex changes due to an external stimulus

Adiabatic energy difference: energy difference between two spin states computed at the corresponding geometry

$$\Delta E_{H\text{-}L}=E_{HS}\text{-}E_{LS}$$

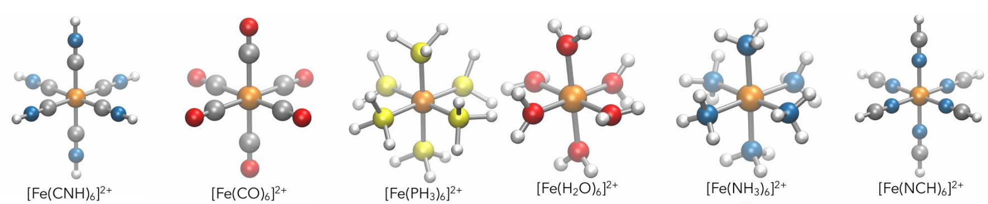
Challenge for any ab initio methods ⟶**Lack of error cancellation**



CASPT2/CC is a multiconfigurational method that we take as reference for $\Delta E_{H\text{-}L}$ (expensive)

Different density functionals yield very different values of $\Delta E_{H\text{-}L}$.

Can a DFT functional be build (using machine learning) to predict accurate values of $\Delta E_{H\text{-}L}$?



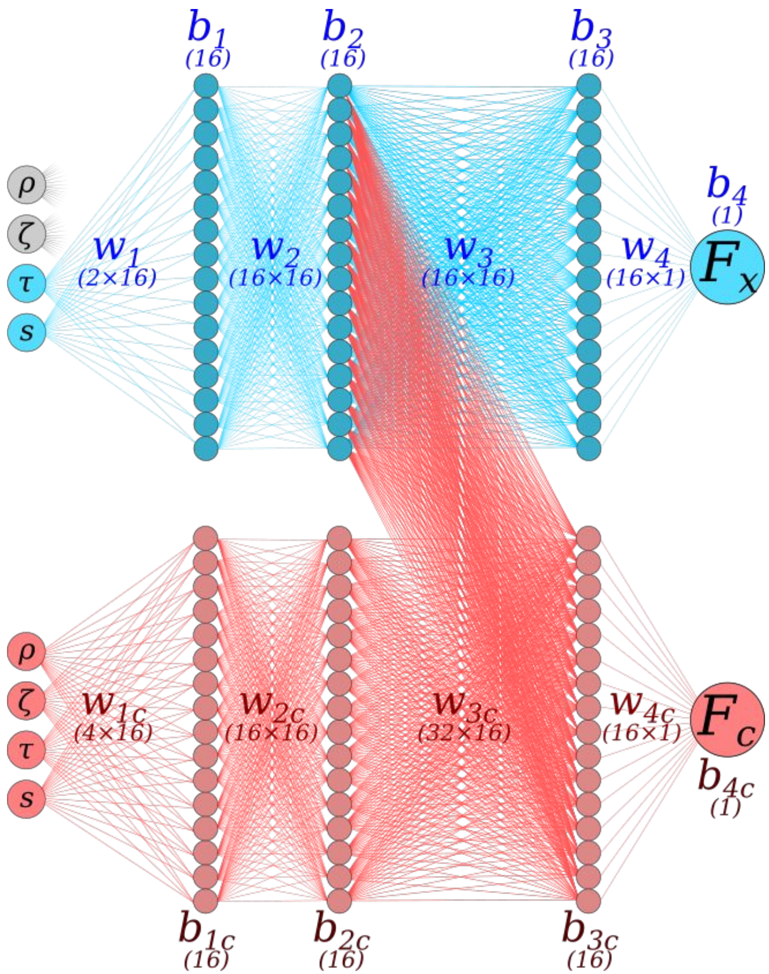
L. A. Mariano et al. *Journal of Chemical Theory and Computation* 17.5 (2021): 2807-2816.

Our Artificial Neural Network Functional

Architecture

$$E_{XC}[\rho] = \int \rho(\mathbf{r})\varepsilon_{XC}(\mathbf{g}[\rho(\mathbf{r})])d\mathbf{r}$$

$$\varepsilon_{XC} = F_X\varepsilon_X^{\text{R2SCAN}} + F_C\varepsilon_C^{\text{R2SCAN}}$$



1503 parameters

Loss Function

$$\Delta_{\text{err}} = c_1 \sum_i |AE_i^{\text{ref}} - AE_i^{\text{DFT}}|$$

Atomization Energies (AE)

$$+ c_2 \sum_j \frac{1}{N_e} \int |\rho_j^{\text{ref}} - \rho_j^{\text{DFT}}| dV$$

Electronic Densities (ρ)

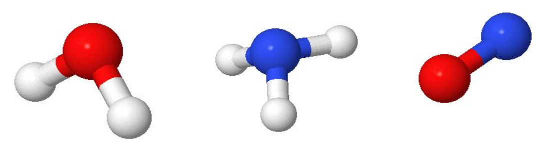
$$+ c_3 \sum_k |\Delta E_{HL,k}^{\text{ref}} - \Delta E_{HL,k}^{\text{DFT}}|$$

Adiabatic Energy Differences (ΔE_{HL})

Training set

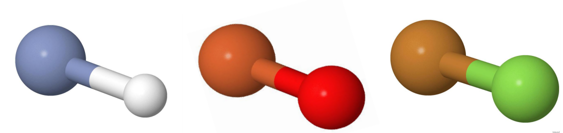
Different combinations of :

H_2O NH_3 NO



$\rho^{CCSD(T)}$ AE^{ref}

CrH FeO CuF

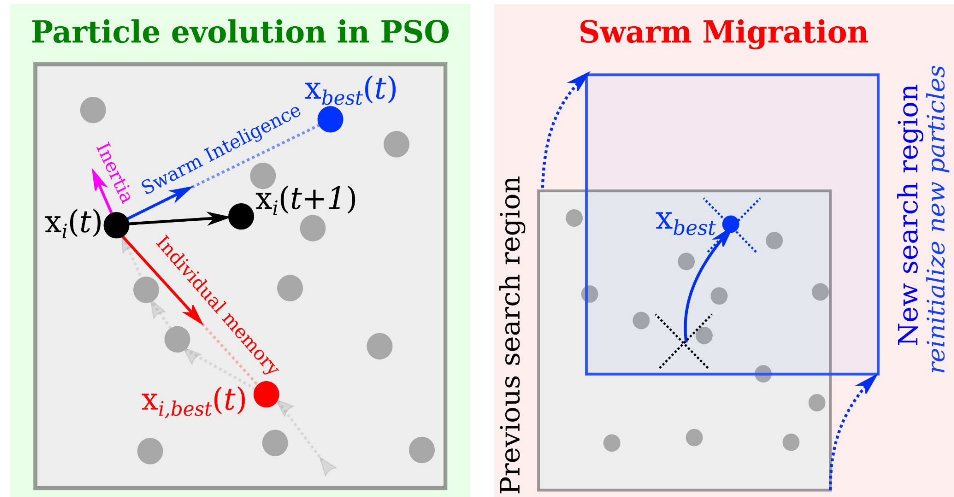


$\rho^{CCSD(T)}$ $\Delta E_{HL}^{Exp.}$

Training Algorithm: Modified PSO via Migrations

Idea:

In practice:

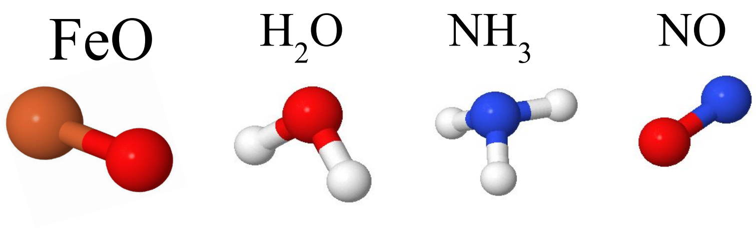


Particle Swarm Optimization (PSO)

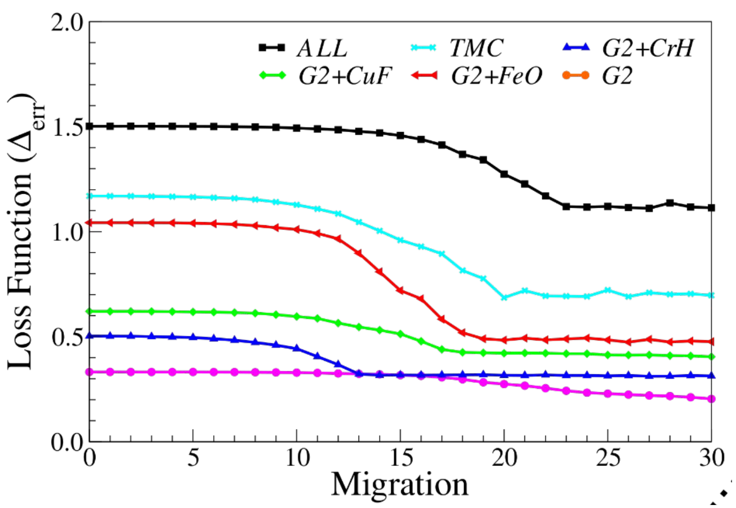
- No need of evaluating the derivatives of the loss function with respect to the parameters of the functional
- Improves over Monte Carlo optimization exploring the phenomena of swarm intelligence

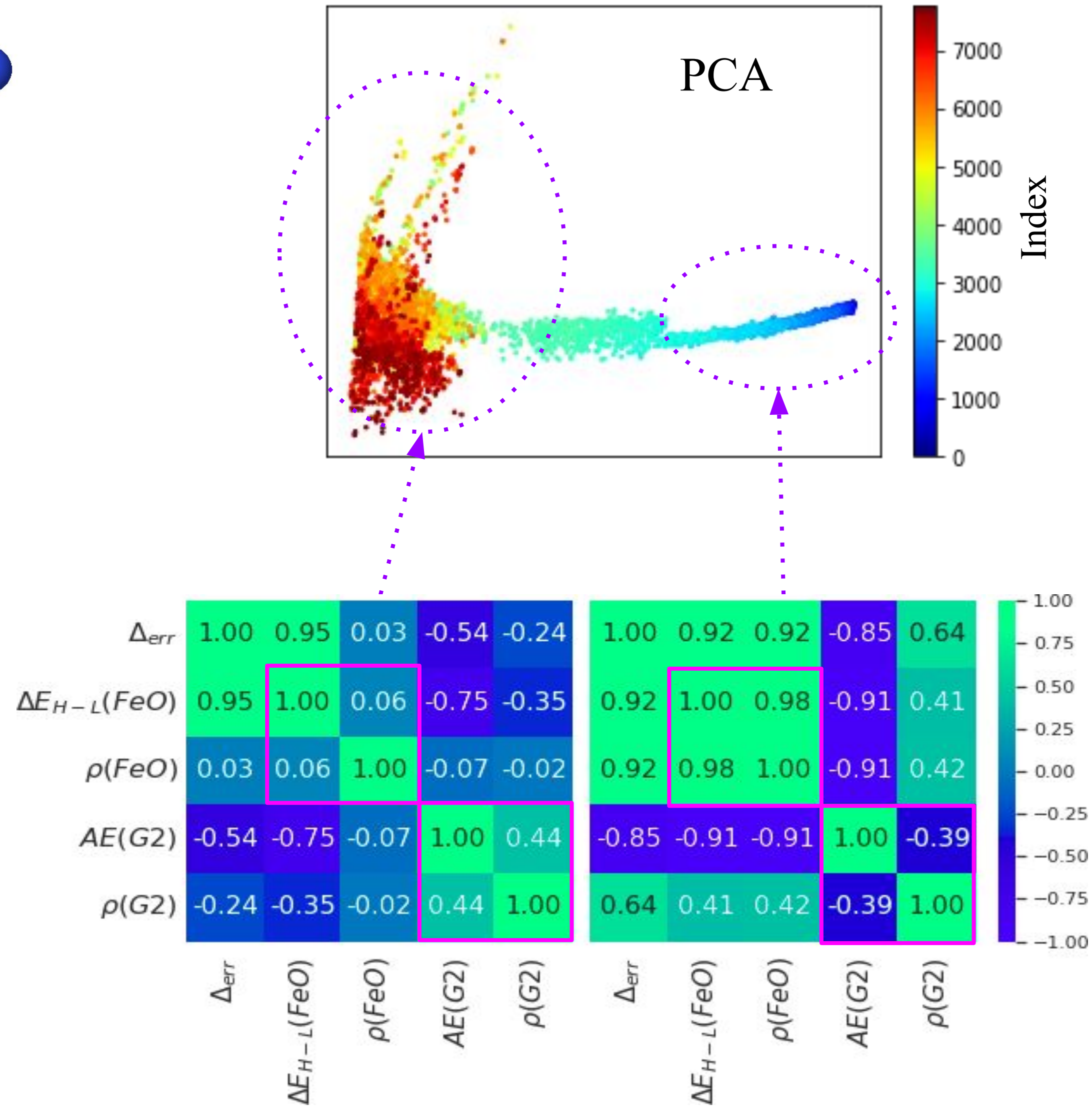
Migrations

- Changing periodically the search box to follow the best solution
- Improves on the stability not generating searches too far from the best solution found in the present epoch



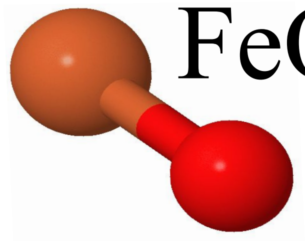
The optimization process is dominated by energetic properties and limited by the precision on $\Delta E_{H\text{-}L}$.





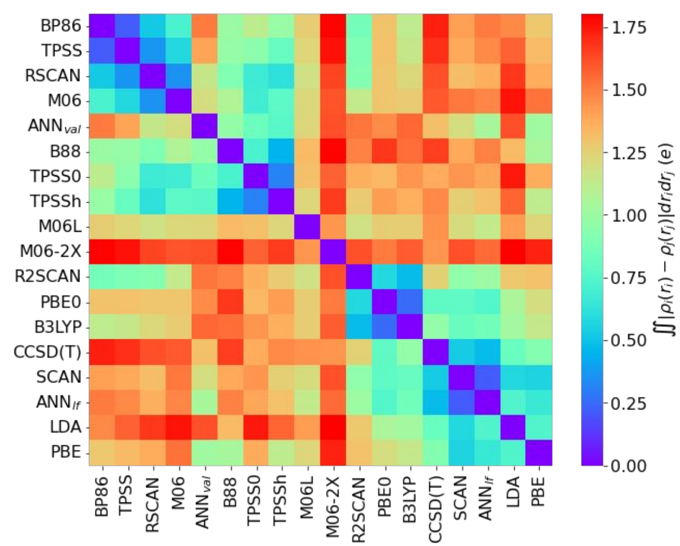
Spearman Correlations

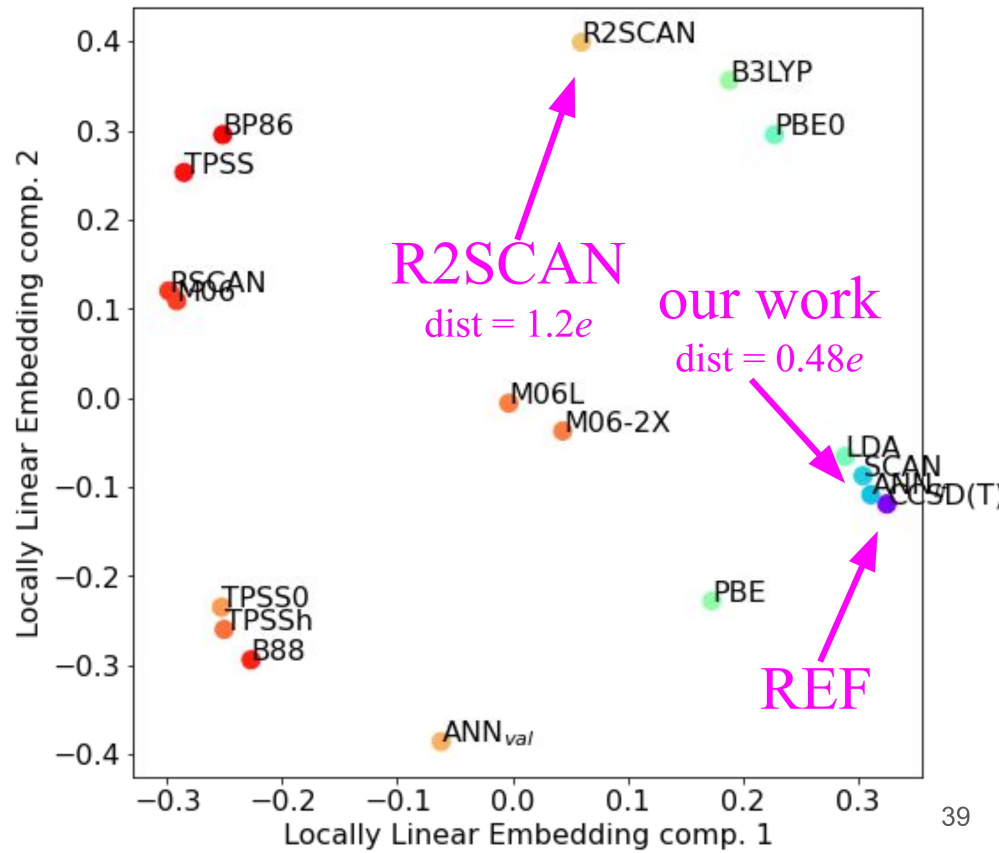
The trained functional predicts densities closer to the CCSD(T) reference than to the original R2SCAN.



FeO

$$\mathrm{dist}_{ij} = \iint |\rho_i(\mathbf{r}) - \rho_j(\mathbf{r})| \mathrm{d}\mathbf{r}$$





Validation Set: Iron +2 Complexes

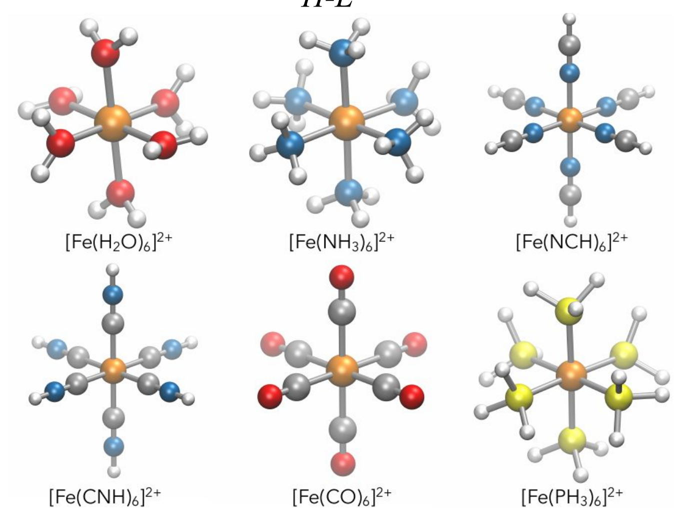
	ΔE_{H-L} (eV)									
	PBE	**SCAN**	**NAS**[a]	**B3LYP**	**DM21**	**R2SCAN**	**PBE[U]**[b]	ANN_{lf}^{ALL}	ANN_{val}^{ALL}	**REF.**[c]
$[Fe(H_2O)_6]^{2+}$	-1.17	-0.81	-0.74	-1.44	-0.63	-1.36	-1.50	-2.51	-1.83	-1.83
$[Fe(NH_3)_6]^{2+}$	-0.06	0.21	0.21	-0.58	0.26	-0.29	-0.44	-1.27	-0.71	-0.64
$[Fe(NCH)_6]^{2+}$	1.14	0.89	0.90	-0.21	0.64	0.41	0.21	-0.39	0.08	-0.16
$[Fe(PH_3)_6]^{2+}$	2.69	2.09	2.21	0.62	2.24	1.91	1.81	2.05	1.94	2.54
$[Fe(CO)_6]^{2+}$	3.63	2.86	2.93	1.25	2.35	2.58	2.64	2.42	2.50	2.02
$[Fe(NCH)_6]^{2+}$	4.11	3.38	3.42	1.86	2.99	3.06	3.23	2.93	2.99	2.87
MAE[c]	*0.92*	*0.79*	*0.80*	*0.70*	*0.61*	*0.46*	*0.44*	*0.41*	*0.25*	

[a] NAS is the functional obtained by R. Nagai, R. Akashi, and O. Sugino.[29]

[b] PBE[U] represents the use of Hubbard U_{sc}-corrected density in the PBE functional.[28]

[c] Reference values come from CASPT2/CC.[28]

Ref. for $\Delta E_{H\text{-}L}$: CASPT2/CC



29 - Nagai, R., Akashi, R., and Sugino, O. “Machine-learning-based exchange correlation functional with physical asymptotic constraints.” *Physical Review Research*, 4.1 (2022): 013106.

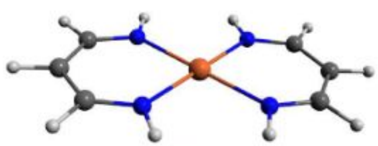
28 - Mariano, Lorenzo A., Bess Vlaisavljevich, and Roberta Poloni. "Improved Spin-State Energy Differences of Fe (II) molecular and crystalline complexes via the Hubbard U-corrected Density." *Journal of Chemical Theory and Computation* 17.5 (2021): 2807-2816.

Other Metal Complexes

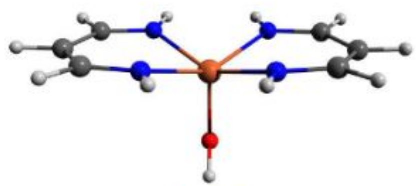
		ΔE_{H-L} (eV)							
	2S+1	ANN_{lf}^{ALL}	ANN_{val}^{ALL}	R2SCAN	TPSS	PBE	PBE[U]	DM21	REF.[b]
FeL_2	$1 \rightarrow 5$	-1.439	-0.944	-0.611	0.227	-	-1.096	-1.968	-1.487
	$3 \rightarrow 5$	0.193	0.513	0.746	1.259	1.245	0.756	0.814	0.213
MnL_2	$2 \rightarrow 6$	-2.078	-1.427	-0.957	-0.055	-0.111	-0.961	-	-1.782
	$4 \rightarrow 6$	-0.061	0.207	0.414	0.835	0.919	0.491	-0.739	-0.455
FeL_2SH	$2 \rightarrow 6$	-0.464	0.081	0.470	-0.416	1.087	-	-	0.399
	$4 \rightarrow 6$	-0.488	-0.130	0.123	0.588	0.521	0.245	-	-0.017
$[Co(NCH)_6]^{2+}$	$2 \rightarrow 4$	-0.372	-0.170	-0.027	0.31	0.476	0.165	0.057	-0.581
NiCp(acac)	$1 \rightarrow 3$	-0.204	-0.177	-0.157	0.046	0.163	0.095	0.414	0.117
$MnCp_2$	$2 \rightarrow 6$	-0.220	-0.354	0.184	0.649	0.981	0.432	-	0.304
MAE[b]		*0.349*	*0.406*	*0.474*	*0.945*	*0.885*[a]	*0.482*[a]	*0.460*[a]	

[a] Average computed over available values.

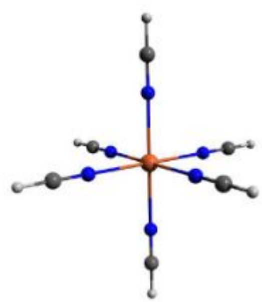
[b] Reference values come from CASPT2/CC.[68]



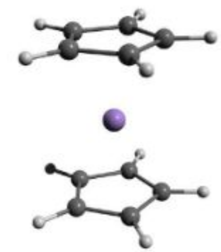
$M^{II}L_2$



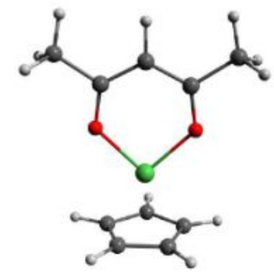
$Fe^{III}L_2XH$
X = O, S



$[M^{II}(NCH)_6]^{2+}$
M = Fe, Co



$Mn^{II}Cp_2$



$Ni^{II}Cp(acac)$

Molecules from: *"Spin state energetics in first-row transition metal complexes: contribution of (3s3p) correlation and its description by second-order perturbation theory."* Kristine Pierloot, Quan Manh Phung, and Alex Domingo *Journal of chemical theory and computation* 13.2 (2017): 537-553.

Preparing for the Practical Section

"Here be Dragons"

Quantum Chemistry Software for DFT

LIBXC

https://tddft.org/programs/libxc/

- Abinit - a software suite to calculate the optical, mechanical, vibrational, and other observable properties of materials
- ACE-Molecule - a quantum chemistry package based on a real-space numerical grid
- ADF - a density functional theory program for molecules and condensed matter
- APE - a computer package designed to generate and test norm-conserving pseudopotentials within density functional theory
- AtomPAW - a program for generating projector augmented wave functions
- BAGEL - a parallel electronic-structure program
- BigDFT - a fast, precise, and flexible density functional theory code for ab-initio atomistic simulation
- CP2K - a program to perform atomistic and molecular simulations of solid state, liquid, molecular, and biological systems
- DFT-FE - a massively parallel real-space code for first principles based materials modelling using Kohn-Sham density functional theory
- DP - a linear response time-dependent density functional theory code with a plane wave basis set
- Chronus Quantum - a computational chemistry software package focused on explicitly time-dependent and post-SCF methods
- Elk - an all-electron full-potential linearised augmented-plane wave code
- entos - a software package for Gaussian-basis ab initio molecular dynamics calculations on molecular and condensed-phase chemical reactions and other processes
- ERKALE - a DFT/HF molecular electronic structure code based on Gaussian-type orbitals
- exciting - a full-potential all-electron density-functional-theory package implementing the families of linearized augmented planewave methods
- FHI-AIMS - an efficient, accurate, all-electron, full-potential electronic structure code package for computational molecular and materials science
- GAMESS (US) - a general ab initio quantum chemistry package
- GPAW - a density-functional theory Python code based on the projector-augmented wave method
- HelFEM - Finite element methods for electronic structure calculations on small systems
- Horton - Python development platform for electronic structure methods
- INQ - a modern GPU-accelerated computational framework for (time-dependent) density functional theory
- JDFTx - plane-wave code designed for joint density functional theory
- MADNESS - a multiwave adaptive numerical grid program for electroni
- MOLGW - many-body perturbation theory for atoms, molecules, and clusters
- Molpro - a comprehensive system of ab initio programs for advanced molecular electronic structure calculations
- MRCC - a suite of ab initio and density functional quantum chemistry programs for high-accuracy electronic structure calculations
- NWChem - an open source, high-performance computational chemistry program
- Octopus - a program aimed at the ab initio virtual experimentation on a hopefully ever-increasing range of system types
- OpenMolcas - a quantum chemistry software package specializing in multiconfigurational approaches
- ORCA - ab initio quantum chemistry program that contains modern electronic structure methods
- PROFESS - orbital-free density functional theory implementation to simulate condensed matter and molecules
- Psi4 - an open-source suite of ab initio quantum chemistry programs designed for efficient, high-accuracy simulations of molecular properties
- PySCF - Python-based Simulations of Chemistry Framework
- QuantumATK - code including pseudopotential-based density functional theory methods with LCAO and plane-wave basis sets in one framework
- Quantum Espresso - an integrated suite of open source computer codes for electronic-structure calculations and materials modeling at the nanoscale
- Turbomole - a program package for electronic structure calculations
- Vasp - the Vienna Ab initio Simulation Package: atomic scale materials modelling from first principles
- WIEN2k - program for electronic structure calculations of solids using density functional theory based on the full-potential (linearized) augmented plane-wave + local orbitals method
- Yambo - a program that implements many-body perturbation theory methods such as GW and BSE and time-dependent density functional theory

pyscf

WE GAMESS

ORCA

CP2K

FHI-aims
The ab initio materials simulation package

Psi4

BigDFT

siesta

VASP

QUANTUMESPRESSO

Some tools and libraries for Neural Networks

PyTorch

TensorFlow

ITensor

Lasagne

Microsoft Cognitive Toolkit

opennn neural networks

neurolab

theano

Caffe

Keras

DL4J

What we will see on the Practical Section:

1. **Basics of PySCF**
   - Running basic calculations using DFT to compute energies of molecular systems
   - Additional informations on setting tricky calculations for charged/magnetic systems and controlling output options.
2. **Running PySCF DFT with a customized functional**
   - Exemple where we code PW86 functional by hand and mix it to PBE, highlighting the details on how it is done.
3. **Crafting a basic multilayer perceptron with TensorFlow**
   - Sample code of a very elementar implementation of a feed forward neural network using Tensorflow
4. **Using a multilayer perceptron as functional**
   - Writing a external functional in PySCF that uses a Neural Network to correct PBE.
5. **Using a pre-trained ANN functional**
   - We will use PySCF to run some computations with a Neural Network correction PBE trained only using H_2O atomization energy.

The main goal of this tutorial is to allow you to comprehend how to implement and use a neural network functional.

This tutorial uses the very simple case of an machine learning functional that correct PBE and try to predict energies closer to the B3LYP ones for molecules only composed of O and H atoms.

Even with its simplicity, this exemple show all the tools and elements necessary to extend to more complex cases.

What you will need:

I strongly recommend to use the Google Colab!

You will need internet and maybe a google account…

pyscf https://pyscf.org/install.html

tensorflow https://www.tensorflow.org/

If you want to do the things in the harder way:

https://pyscf.org/install.html#pyscf-docker-image

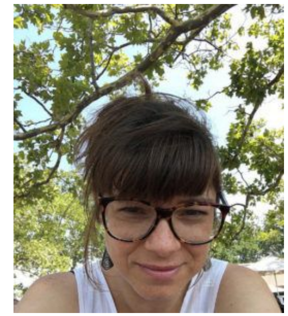
Acknowledgements

The team

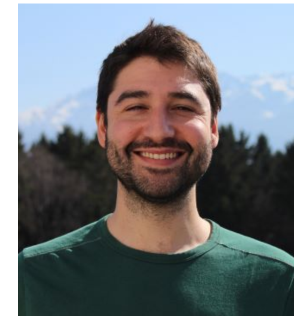
Ashna Jose
(PhD, SIMaP/LIG)

Emilie Devijver
(CNRS, LIG)

Noel Jakse
(G-INP, SIMaP)



Roberta Poloni
(CNRS, SIMaP)



Lorenzo Mariano
(Postdoc, Trinity College)

Fundings:

- ANR - Agence Nationale pour la Recherche
- MIAI - Multidisciplinary Institute of Artificial intelligence (Grenoble)

Thanks for your attention!

JP

Can you write a Sonnet about Density Functional Theory?

Of course! Here's a sonnet about Density Functional Theory:

In search of secrets hidden deep within,
The atom's heart we long to understand,
With Density Functional Theory we begin,
Our quest to solve the mystery at hand.

A world of electrons swirling round and round,
A dance of charges in an intricate array,
To map their movements we must be profound,
And grasp the laws that govern their display.

The density, the key to all we seek,
A function of the electron cloud's state,
With every particle, a bond unique,
A fingerprint that guides our search, innate.

And so we delve into this wondrous art,
To unlock the secrets of the atom's heart.

https://github.com/jpalastus
joao-paulo.almeida-de-mendonca@grenoble-inp.fr
@alastus