

ニューラルネットワークによる
Kohn-Sham交換相関ポテンシャルの構成
**(Construction of the Kohn-Sham exchange-correlation
potential with the neural network)**

Ryosuke AKASHI (明石遼介)

Dept. Phys., Univ. of Tokyo, Japan

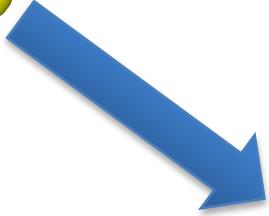
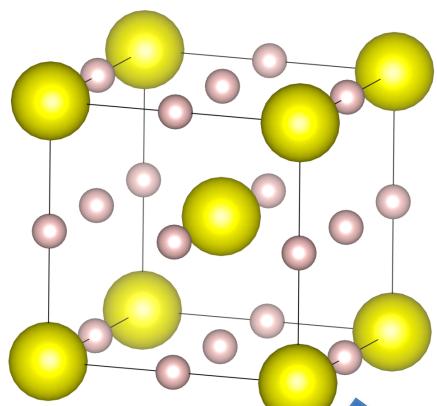
R. Nagai (M2, Sugino lab., ISSP Kashiwa)



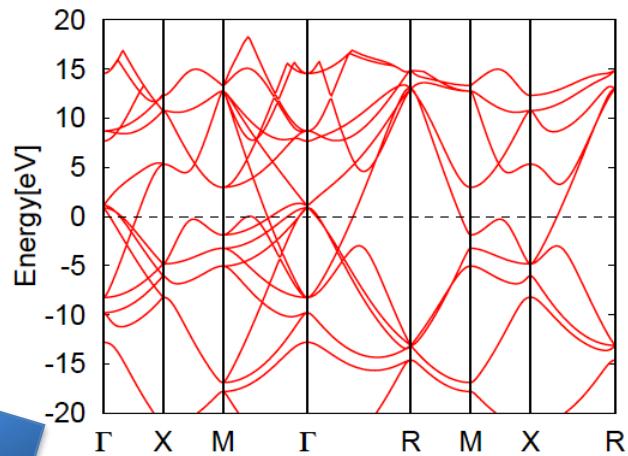
Email: akashi_at_cms.phys.s.u-tokyo.ac.jp
r-nag_at_issp.u-tokyo.ac.jp

Inside *ab initio* simulation codes

Crystal structure



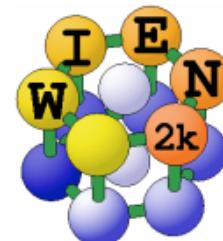
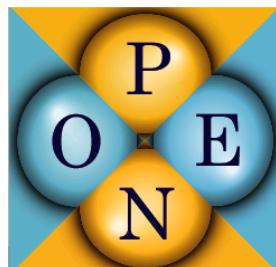
Electronic structure; etc.



Something complicated



QUANTUMESPRESSO



Kohn-Sham equation

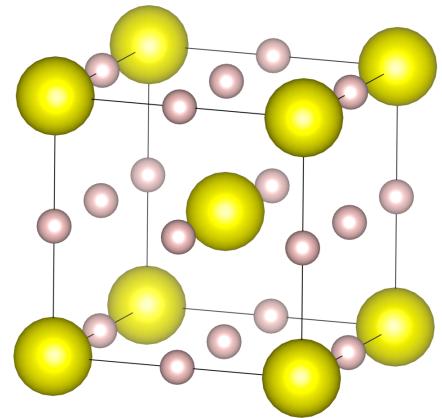
The Hamiltonian (non-relativistic, Classical ions)

$$\mathcal{H} = \int d\mathbf{r} \hat{\psi}_\sigma^\dagger(\mathbf{r}) \left[-\frac{\nabla^2}{2} + V_{\text{ion}}(\mathbf{r}) \right] \hat{\psi}_\sigma(\mathbf{r}) + \frac{1}{2} \int d\mathbf{r} d\mathbf{r}' \hat{\psi}_\sigma^\dagger(\mathbf{r}) \hat{\psi}_{\sigma'}^\dagger(\mathbf{r}') \frac{1}{|\mathbf{r} - \mathbf{r}'|} \hat{\psi}_{\sigma'}(\mathbf{r}') \hat{\psi}_\sigma(\mathbf{r})$$

Particle anticommutation relation

$$[\hat{\psi}_\sigma(\mathbf{r}), \hat{\psi}_{\sigma'}^\dagger(\mathbf{r}')] = \delta_{\sigma\sigma'} \delta(\mathbf{r} - \mathbf{r}')$$

The variety of the electronic properties originates from the ionic potential



Kohn-Sham equation

Many particle Schroedinger equation

$$\mathcal{H}|\Psi_{\text{GS}}\rangle = E_{\text{GS}}|\Psi_{\text{GS}}\rangle$$



Corresponding non-interacting system

W. Kohn and L. J. Sham, Phys. Rev. 140, A1133 (1965)



Kohn-Sham equation

$$\left[-\frac{\nabla^2}{2} + V_{\text{ion}}(\mathbf{r}) + \int d\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + V_{\text{xc}}([n]; \mathbf{r}) \right] \psi_i(\mathbf{r}) = \varepsilon_i \psi_i(\mathbf{r})$$

Solution of the KS equation reproduces the GS electron density $n(\mathbf{r})$
if appropriate functional $V_{\text{xc}}[n]$ is used.

Today's topic

Exchange correlation potential and energy

$$V_{\text{xc}}(\mathbf{r}) \equiv \frac{\delta E_{\text{xc}}}{\delta n(\mathbf{r})}$$

← Exchange-correlation energy

Definition

$$E[n] = T_s[n] + \int dr V_{\text{ion}}([n]; r) n(r) + E_H[n] + E_{\text{xc}}[n]$$

$$T_s[n] \equiv \min_{\Phi \in \text{Slater}} \langle \Phi[n] | \int dr \hat{\psi}^\dagger(r) \left(-\frac{\nabla^2}{2} \right) \hat{\psi}(r) | \Phi[n] \rangle$$

$$E_H[n] \equiv \frac{1}{2} \int dr dr' n(r) \frac{1}{|r - r'|} n(r')$$

Example: Perdew-Zunger LDA

W. Kohn and L. J. Sham, Phys. Rev. 140, A1133 (1965)

J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981) cited ~20000 times

$$\begin{aligned} E_{\text{xc}} &\equiv \int d\mathbf{r} n(\mathbf{r}) \epsilon_{\text{xc}}([n]; \mathbf{r}) \\ &\simeq \int d\mathbf{r} n(\mathbf{r}) \epsilon_{\text{xc}}(n(\mathbf{r})) \end{aligned}$$

Example: Perdew-Zunger LDA

W. Kohn and L. J. Sham, Phys. Rev. 140, A1133 (1965)

J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981) cited ~20000 times

$$\begin{aligned} E_{xc} &\equiv \int d\mathbf{r} n(\mathbf{r}) \epsilon_{xc}([n]; \mathbf{r}) \\ &\simeq \int d\mathbf{r} n(\mathbf{r}) \epsilon_{xc}(n(\mathbf{r})) \end{aligned}$$

$\epsilon_{xc}(n)$: Refer to the uniform electron gas ($[n]$ is represented by one parameter n)

Example: Perdew-Zunger LDA

W. Kohn and L. J. Sham, Phys. Rev. 140, A1133 (1965)

J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981) cited ~20000 times

$$\begin{aligned} E_{xc} &\equiv \int d\mathbf{r} n(\mathbf{r}) \epsilon_{xc}([n]; \mathbf{r}) \\ &\simeq \int d\mathbf{r} n(\mathbf{r}) \epsilon_{xc}(n(\mathbf{r})) \end{aligned}$$

$\epsilon_{xc}(n)$: Refer to the uniform electron gas ($[n]$ is represented by one parameter n)

1, Calculate total energy density by an accurate wave function method

$$\epsilon(n) = \epsilon_{kin}(n) + \epsilon_H(n) + \epsilon_x(n) + \epsilon_c(n)$$

Example: Perdew-Zunger LDA

W. Kohn and L. J. Sham, Phys. Rev. 140, A1133 (1965)

J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981) cited ~20000 times

$$\begin{aligned} E_{xc} &\equiv \int d\mathbf{r} n(\mathbf{r}) \epsilon_{xc}([n]; \mathbf{r}) \\ &\simeq \int d\mathbf{r} n(\mathbf{r}) \epsilon_{xc}(n(\mathbf{r})) \end{aligned}$$

$\epsilon_{xc}(n)$: Refer to the uniform electron gas ($[n]$ is represented by one parameter n)

1, Calculate total energy density by an accurate wave function method

$$\epsilon(n) = \epsilon_{kin}(n) + \epsilon_H(n) + \epsilon_x(n) + \epsilon_c(n)$$

$$\epsilon_{kin}(n) = \frac{(3\pi^2)^{5/3}}{10\pi^2 m} n^{2/3} \text{(Free electron gas)} \quad \epsilon_H(n) = 0 \text{ (Canceled with the el-ion pot.)}$$

$$\epsilon_x(n) = -\frac{3(3\pi^2)^{1/3}}{4\pi} n^{1/3} \quad \text{(Coulomb pot. of plane wave Slater state)}$$

Example: Perdew-Zunger LDA

W. Kohn and L. J. Sham, Phys. Rev. 140, A1133 (1965)

J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981) cited ~20000 times

$$\begin{aligned} E_{xc} &\equiv \int d\mathbf{r} n(\mathbf{r}) \epsilon_{xc}([n]; \mathbf{r}) \\ &\simeq \int d\mathbf{r} n(\mathbf{r}) \epsilon_{xc}(n(\mathbf{r})) \end{aligned}$$

$\epsilon_{xc}(n)$: Refer to the uniform electron gas ($[n]$ is represented by one parameter n)

1, Calculate total energy density by an accurate wave function method

$$\epsilon(n) = \epsilon_{kin}(n) + \epsilon_H(n) + \epsilon_x(n) + \epsilon_c(n)$$

$$\epsilon_{kin}(n) = \frac{(3\pi^2)^{5/3}}{10\pi^2 m} n^{2/3} \text{(Free electron gas)} \quad \epsilon_H(n) = 0 \text{ (Canceled with the el-ion pot.)}$$

$$\epsilon_x(n) = -\frac{3(3\pi^2)^{1/3}}{4\pi} n^{1/3} \quad \text{(Coulomb pot. of plane wave Slater state)}$$

2, Numerically fit the remaining part

$$\epsilon_c(n) = \epsilon(n) - \epsilon_{kin}(n) - \epsilon_H(n) - \epsilon_x(n)$$

with a heuristic function that satisfies the exact asymptotic behavior.

Example: Perdew-Zunger LDA

J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

2, Numerically fit the remaining part

$$\epsilon_c(n) = \epsilon(n) - \epsilon_{\text{kin}}(n) - \epsilon_H(n) - \epsilon_x(n)$$

with a heuristic function that satisfies the exact asymptotic behavior.

High density expansion

Mäcke (1950); Gell-Mann-Brueckner (1957);
Onsager-Mittag-Stephen(1966); DuBois (1959); Carr-Maradudin (1964);
Endo-Horiuchi-Takada-Yasuhara (1999).

$$\epsilon_c(n) \sim 0.0311 \ln r_s - 0.048 + C r_s \ln r_s + D r_s$$

Low density expansion Wigner (1938); Carr Jr. (1961)

$$\epsilon_c(n) \sim \frac{f_0}{r_s} + \frac{f_1}{r_s^{3/2}} + \frac{f_2}{r_s^2} + O(r_s^{-5/2})$$

Example: Perdew-Zunger LDA

J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

2, Numerically fit the remaining part

$$\epsilon_c(n) = \epsilon(n) - \epsilon_{\text{kin}}(n) - \epsilon_H(n) - \epsilon_x(n)$$

with a heuristic function that satisfies the exact asymptotic behavior.

High density expansion

Mäcke (1950); Gell-Mann-Brueckner (1957);
Onsager-Mittag-Stephen(1966); DuBois (1959); Carr-Maradudin (1964);
Endo-Horiuchi-Takada-Yasuhara (1999).

$$\epsilon_c(n) \sim 0.0311 \ln r_s - 0.048 + C r_s \ln r_s + D r_s$$

$$\frac{1}{n} = \frac{4\pi r_s^3}{3}$$

Low density expansion

Wigner (1938); Carr Jr. (1961)

$$\epsilon_c(n) \sim \frac{f_0}{r_s} + \frac{f_1}{r_s^{3/2}} + \frac{f_2}{r_s^2} + O(r_s^{-5/2})$$

Example: Perdew-Zunger LDA

J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

2, Numerically fit the remaining part

$$\epsilon_c(n) = \epsilon(n) - \epsilon_{\text{kin}}(n) - \epsilon_H(n) - \epsilon_x(n)$$

with a heuristic function that satisfies the exact asymptotic behavior.

High density expansion

Mäcke (1950); Gell-Mann-Brueckner (1957);
Onsager-Mittag-Stephen(1966); DuBois (1959); Carr-Maradudin (1964);
Endo-Horiuchi-Takada-Yasuhara (1999).

$$\epsilon_c(n) \sim 0.0311 \ln r_s - 0.048 + Cr_s \ln r_s + Dr_s$$

Interpolation formula

Ceperley, Phys. Rev. B **18**, 3126 (1978)

$$\epsilon_c(n) = \frac{\gamma}{1 + \beta_1 \sqrt{r_s} + \beta_2 r_s}$$

- Select the form so that $r_s \rightarrow \infty$ is reproduced
- Determine $(C, D, \gamma, \beta_1, \beta_2)$ so that
 - it connects to the $r_s \rightarrow 0$ form
 - it well fits the wave-function theory result

Low density expansion

Wigner (1938); Carr Jr. (1961)

$$\epsilon_c(n) \sim \frac{f_0}{r_s} + \frac{f_1}{r_s^{3/2}} + \frac{f_2}{r_s^2} + O(r_s^{-5/2})$$

Example: Perdew-Burke-Ernzerhof GGA

J. P. Perdew, K. Burke and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996) cited ~100000 times

$$\begin{aligned}\epsilon_{\text{xc}}([n]; \mathbf{r}) &\simeq \epsilon_{\text{xc}}(n(\mathbf{r}), \nabla n(\mathbf{r})) \\ &\equiv \epsilon_{\text{x}}(n(\mathbf{r}), s(\mathbf{r})) + \epsilon_{\text{c}}(n(\mathbf{r}), t(\mathbf{r}))\end{aligned}$$

$$\begin{aligned}s &= \frac{|\nabla n|}{2k_{\text{F}}n} & t &= \frac{|\nabla n|}{2k_{\text{s}}n} \\ k_{\text{F}} &= (3\pi^2 n)^{1/3} \\ k_{\text{s}} &= \sqrt{4k_{\text{F}}/\pi}\end{aligned}$$

Example: Perdew-Burke-Ernzerhof GGA

J. P. Perdew, K. Burke and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996) cited ~100000 times

$$\begin{aligned}\epsilon_{xc}([n]; \mathbf{r}) &\simeq \epsilon_{xc}(n(\mathbf{r}), \nabla n(\mathbf{r})) \\ &\equiv \epsilon_x(n(\mathbf{r}), s(\mathbf{r})) + \epsilon_c(n(\mathbf{r}), t(\mathbf{r}))\end{aligned}\quad s = \frac{|\nabla n|}{2k_F n} \quad t = \frac{|\nabla n|}{2k_s n}$$

Correlation

$$\epsilon_c \equiv \epsilon_c^{\text{UEG}}(n) + H(n, t)$$

Example: Perdew-Burke-Ernzerhof GGA

J. P. Perdew, K. Burke and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996) cited ~100000 times

$$\begin{aligned}\epsilon_{xc}([n]; \mathbf{r}) &\simeq \epsilon_{xc}(n(\mathbf{r}), \nabla n(\mathbf{r})) \\ &\equiv \epsilon_x(n(\mathbf{r}), s(\mathbf{r})) + \epsilon_c(n(\mathbf{r}), t(\mathbf{r}))\end{aligned}\quad s = \frac{|\nabla n|}{2k_F n} \quad t = \frac{|\nabla n|}{2k_s n}$$

Correlation

$$\epsilon_c \equiv \epsilon_c^{\text{UEG}}(n) + H(n, t)$$

$$H(n, t) \sim at^2 \quad (n \rightarrow \infty; t \sim 0)$$

$$\sim -\epsilon_c^{\text{UEG}} \quad (t \rightarrow \infty)$$

$$\sim blnt^2 \quad (n \rightarrow \infty; t \rightarrow \infty)$$

Example: Perdew-Burke-Ernzerhof GGA

J. P. Perdew, K. Burke and M. Ernzerhof, Phys. Rev. Lett. 77, 3865 (1996) cited ~100000 times

$$\begin{aligned}\epsilon_{xc}([n]; \mathbf{r}) &\simeq \epsilon_{xc}(n(\mathbf{r}), \nabla n(\mathbf{r})) \\ &\equiv \epsilon_x(n(\mathbf{r}), s(\mathbf{r})) + \epsilon_c(n(\mathbf{r}), t(\mathbf{r}))\end{aligned}\quad s = \frac{|\nabla n|}{2k_F n} \quad t = \frac{|\nabla n|}{2k_s n}$$

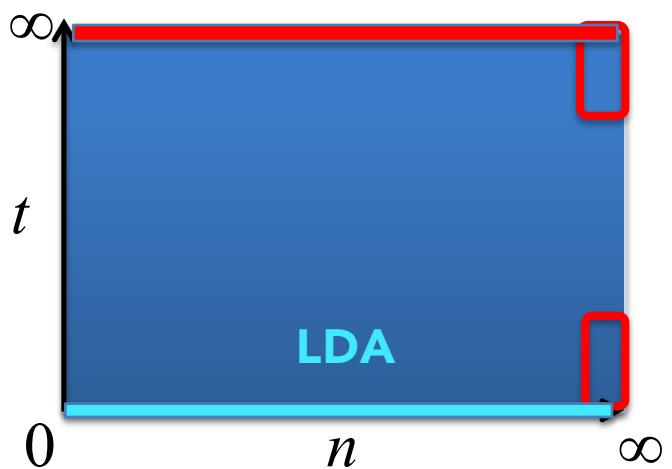
Correlation

$$\epsilon_c \equiv \epsilon_c^{\text{UEG}}(n) + H(n, t)$$

$$H(n, t) \sim at^2 \quad (n \rightarrow \infty; t \sim 0)$$

$$\sim -\epsilon_c^{\text{UEG}} \quad (t \rightarrow \infty)$$

$$\sim blnt^2 \quad (n \rightarrow \infty; t \rightarrow \infty)$$



Example: Perdew-Burke-Ernzerhof GGA

J. P. Perdew, K. Burke and M. Ernzerhof, Phys. Rev. Lett. 77, 3865 (1996) cited ~100000 times

$$\begin{aligned}\epsilon_{xc}([n]; \mathbf{r}) &\simeq \epsilon_{xc}(n(\mathbf{r}), \nabla n(\mathbf{r})) \\ &\equiv \epsilon_x(n(\mathbf{r}), s(\mathbf{r})) + \epsilon_c(n(\mathbf{r}), t(\mathbf{r}))\end{aligned}\quad s = \frac{|\nabla n|}{2k_F n} \quad t = \frac{|\nabla n|}{2k_s n}$$

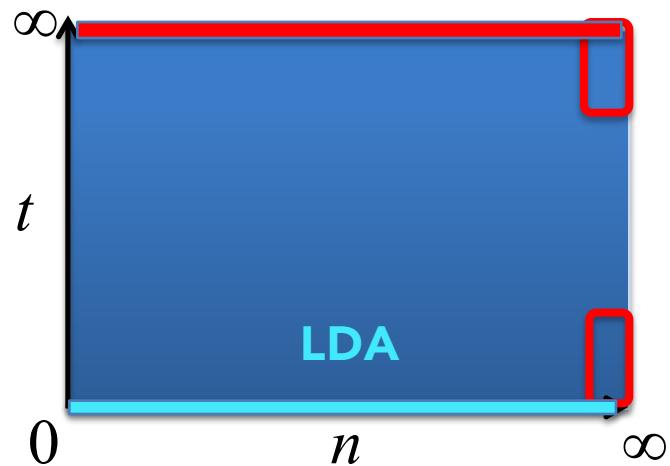
Correlation

$$\epsilon_c \equiv \epsilon_c^{\text{UEG}}(n) + H(n, t)$$

$$H(n, t) \sim at^2 \quad (n \rightarrow \infty; t \sim 0)$$

$$\sim -\epsilon_c^{\text{UEG}} \quad (t \rightarrow \infty)$$

$$\sim blnt^2 \quad (n \rightarrow \infty; t \rightarrow \infty)$$



A satisfactory interpolation function

$$H = bln \left\{ 1 + \frac{a}{b} t^2 \left[\frac{1 + At^2}{1 + At^2 + A^2 t^4} \right] \right\} \quad A = \frac{a}{b} \frac{1}{[\exp\{-\epsilon_c^{\text{UEG}}/b\} - 1]}$$

Example: Perdew-Burke-Ernzerhof GGA

J. P. Perdew, K. Burke and M. Ernzerhof, Phys. Rev. Lett. 77, 3865 (1996) cited ~100000 times

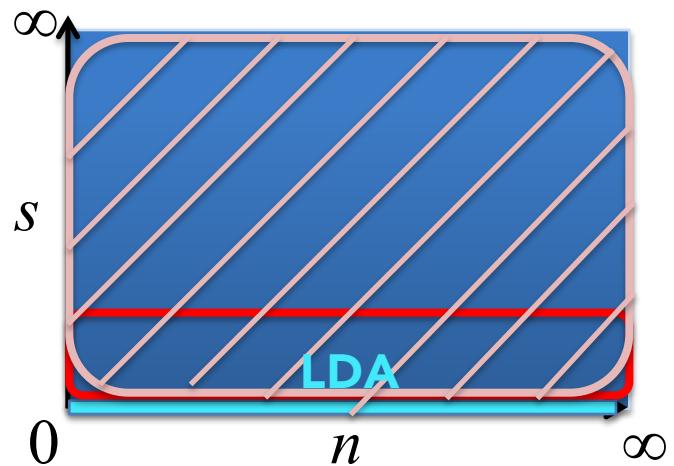
$$\begin{aligned}\epsilon_{\text{xc}}([n]; \mathbf{r}) &\simeq \epsilon_{\text{xc}}(n(\mathbf{r}), \nabla n(\mathbf{r})) \\ &\equiv \epsilon_{\text{x}}(n(\mathbf{r}), s(\mathbf{r})) + \epsilon_{\text{c}}(n(\mathbf{r}), t(\mathbf{r}))\end{aligned}\quad s = \frac{|\nabla n|}{2k_{\text{F}}n} \quad t = \frac{|\nabla n|}{2k_{\text{s}}n}$$

Exchange

$$\epsilon_{\text{x}} \equiv \epsilon_{\text{x}}^{\text{UEG}}(n) F_{\text{x}}(s)$$

$$F_{\text{x}}(s) \sim 1 + \mu s^2$$

$$F_{\text{x}}(s) \leq 1 + \kappa$$

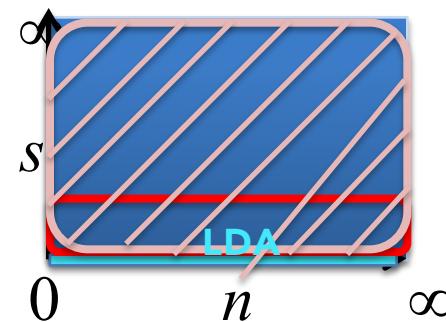
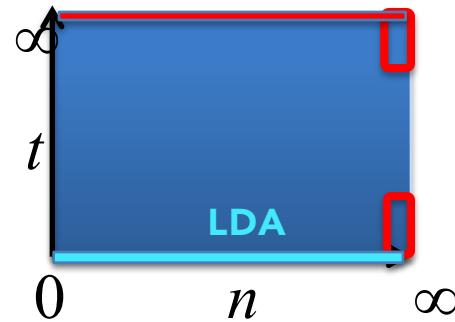


A satisfactory interpolation function

$$F_{\text{x}}(s) = 1 + \kappa - \frac{\kappa^2}{\kappa + \mu s^2}$$

Strategy for functionals

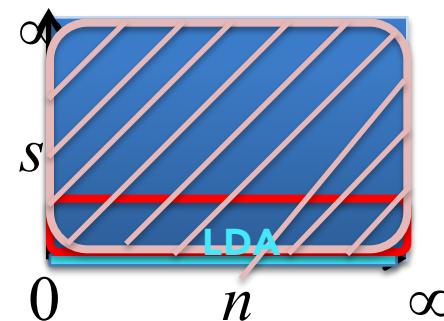
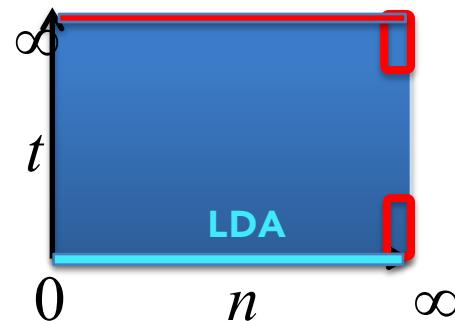
1, Derive the asymptotic formulae and/or exact constraints



2, Construct a smooth function that satisfies them, where
fitting to reference data (accurate calculations and/or experimental values)
is employed (cf. B3LYP, SCAN, . . .).

Strategy for functionals

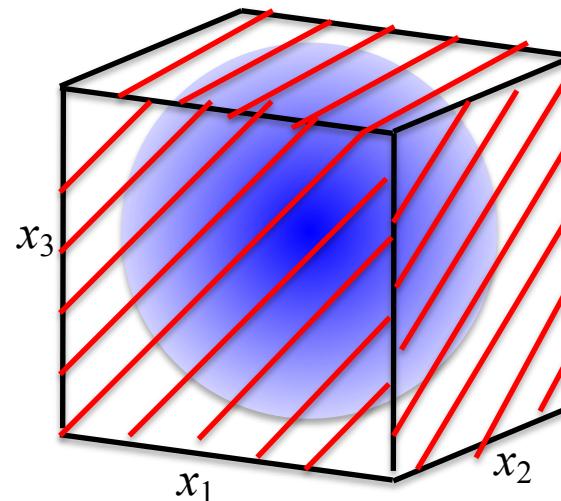
1, Derive the asymptotic formulae and/or exact constraints



2, Construct a smooth function that satisfies them, where
fitting to reference data (accurate calculations and/or experimental values)
is employed (cf. B3LYP, SCAN, . . .).

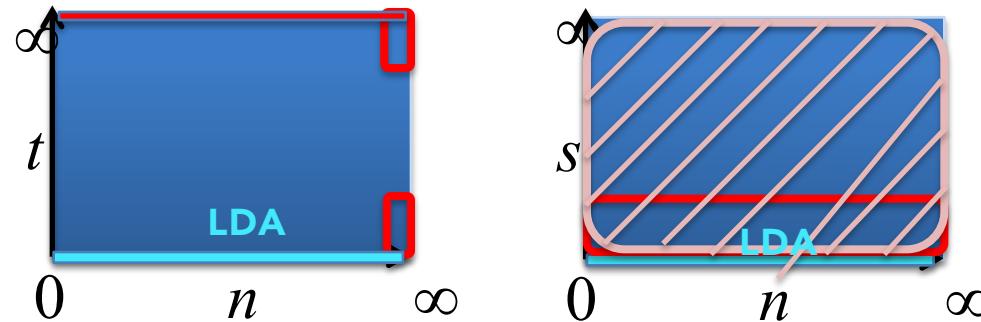
In more dimensions,

- (i) It becomes more difficult
to derive asymptotic formulae . . .
- (ii) Options in the intermediate regime
seems diverging . . .



Strategy for functionals

1, Derive the asymptotic formulae and/or exact constraints

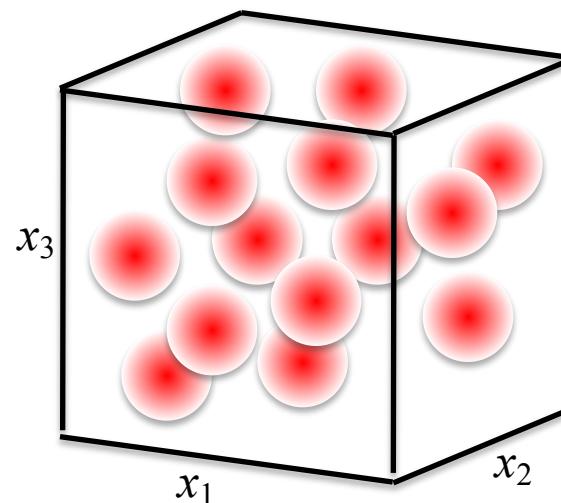


2, Construct a smooth function that satisfies them, where
fitting to reference data (accurate calculations and/or experimental values)
is employed (cf. B3LYP, SCAN, . . .).

In more dimensions,

- (i) **Machine learning could provide us alternative.**
- (ii) **Machine learning could provide us alternative.**

seems diverging . . .

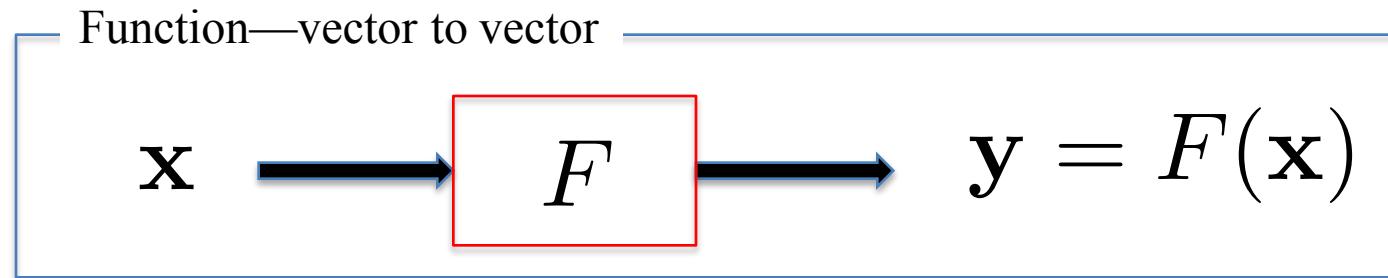


Possible use of the machine-learning methods

A (biased) view on the (supervised) machine learning

What is the machine learning (ML)?:

To formulate a “function” in a mathematical (programmable, calculable) language, for later efficient utilization.



When is (supervised) ML useful?:

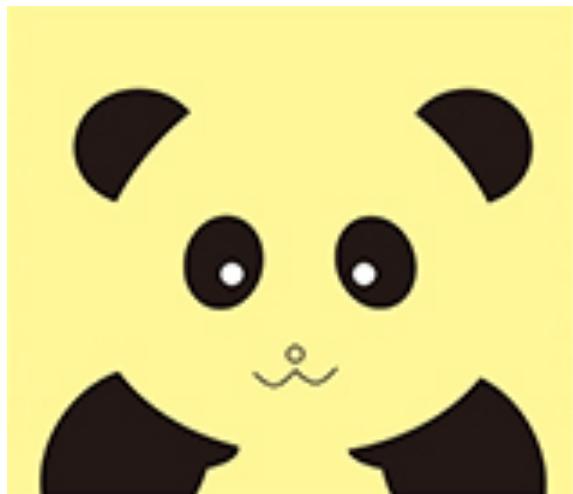
We can get a (massively) few pairs $(\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2), \dots$ but operation F is expensive.

- requires black box
- computationally complicated
- . . .

Possible use of the machine-learning methods

Example: Image recognition

x: picture (as pixel data)



Copyright: 錯視の科学館

F: human



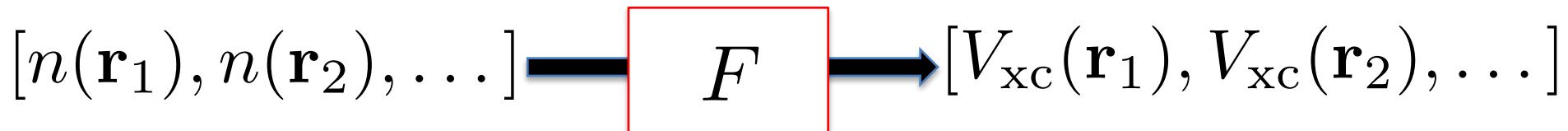
y: classification

Panda !

Possible use of the machine-learning methods

n to V_{xc}

$$E_{xc} \equiv E_{xc}[n] \quad V_{xc}(\mathbf{r}) \equiv V_{xc}([n]; \mathbf{r}) = \frac{\delta E_{xc}}{\delta n(\mathbf{r})}$$



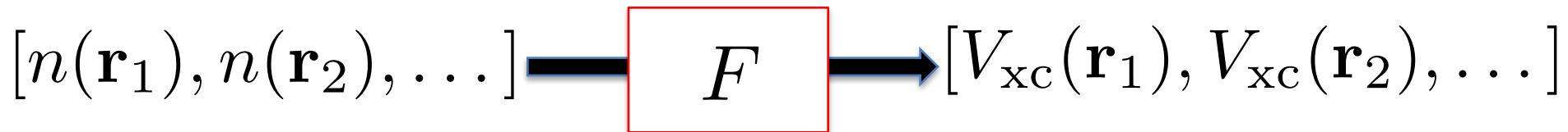
F :

- (i) Solve the Kohn-Sham equation with varying total potential V until it reproduces the given n
- (ii) $V_{xc} = V - V_{ion}$

Possible use of the machine-learning methods

n to V_{xc}

$$E_{xc} \equiv E_{xc}[n] \quad V_{xc}(\mathbf{r}) \equiv V_{xc}([n]; \mathbf{r}) = \frac{\delta E_{xc}}{\delta n(\mathbf{r})}$$



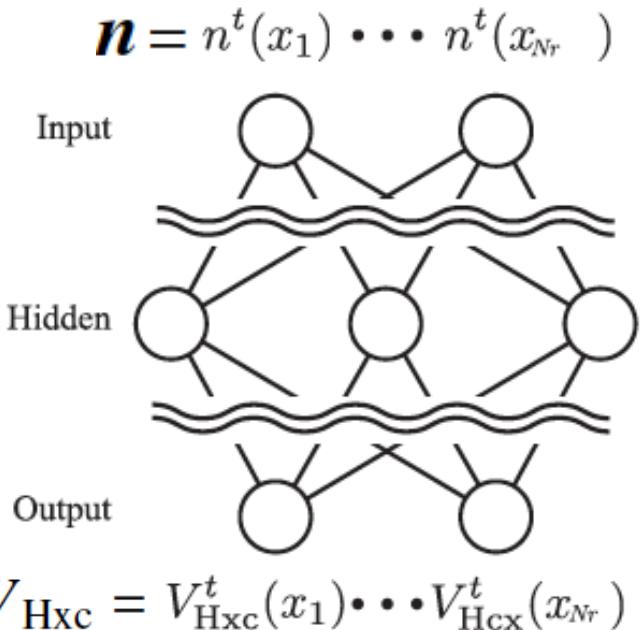
F :

- (i) Solve the Kohn-Sham equation with varying total potential V until it reproduces the given n
- (ii) $V_{xc} = V - V_{\text{ion}}$

ML implementation of the vector-to-vector mapping F ?

Neural network construction of Vxc

R. Nagai, RA, S. Sasaki, and S. Tsuneyuki, J. Chem. Phys. 148, 241737 (2018);
R. Nagai, RA, and O. Sugino, arXiv:1903.00238



Fully connected NN formula

$$V_{Hxc} = \cdots f[W^{(2)}f[W^{(1)}\mathbf{n} + \mathbf{b}^{(1)}] + \mathbf{b}^{(2)}] \cdots$$

$W^{(i)}$: parameter matrix

$\mathbf{b}^{(i)}$: bias vector

f : Nonlinear activation function
(e.g., step function)

Advantages:

- Computational cost \sim repeated matrix-vector multiplications
- Universal approximation theorem: Any continuous functions can be mimicked with arbitrary accuracy by increasing the number of nodes.

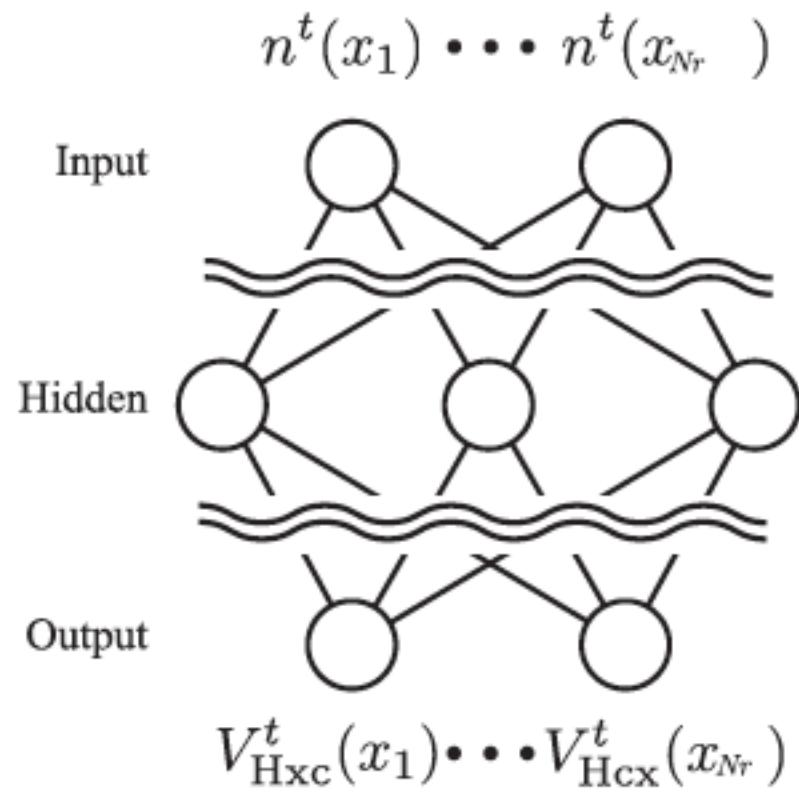
K. Hornik, Neural Networks 4, 251 (1991);

G. Cybenko, Math. Control. Signals Syst. 2, 303 (1989).

Neural network construction of Vxc

R. Nagai, RA, S. Sasaki, and S. Tsuneyuki, J. Chem. Phys. 148, 241737 (2018);
R. Nagai, RA, and O. Sugino, arXiv:1903.00238

R. Nagai
(M2, Sugino lab., ISSP Kashiwa)



Our motivation: How's the NN Vxc transferrable?

1, Fully nonlocal V_{xc} for a simple model

R. Nagai, RA, S. Sasaki, and S. Tsuneyuki, J. Chem. Phys. 148, 241737 (2018)

Two spinless fermions in one dimension

Schrodinger eq.

$$\left[\sum_{k=1,2} \left(-\frac{\nabla_k^2}{2} + V(x_k) \right) + \frac{1}{|x_1 - x_2|} \right] \Psi = E_{\text{tot}} \Psi$$

$$V(x) = -A \exp(-x^2/B^2).$$

Attractive Gaussian potential

Corresponding Kohn-Sham eq.

$$\left[-\frac{\nabla^2}{2} + V(x) + V_{\text{Hxc}}([n]; x) \right] \psi_k(x) = \varepsilon_k \psi_k(x)$$

$$n(x) = \sum_{k=1,2} |\psi_k(x)|^2$$

1, Fully nonlocal V_{xc} for a simple model

R. Nagai, RA, S. Sasaki, and S. Tsuneyuki, J. Chem. Phys. 148, 241737 (2018)

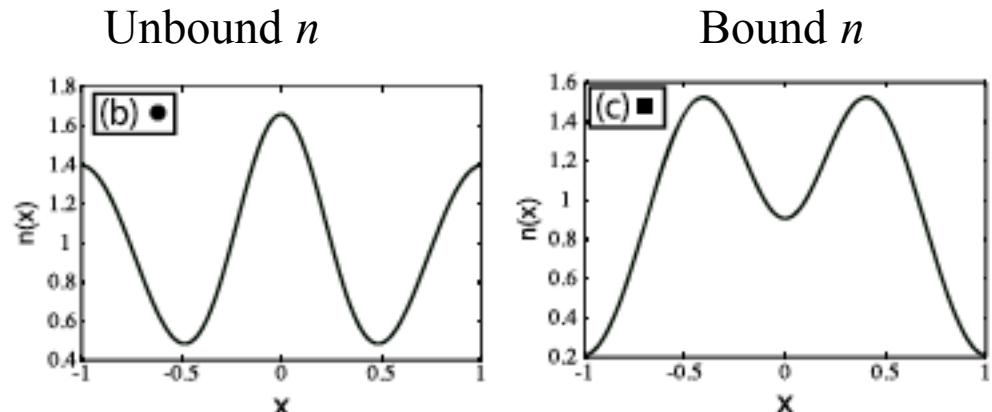
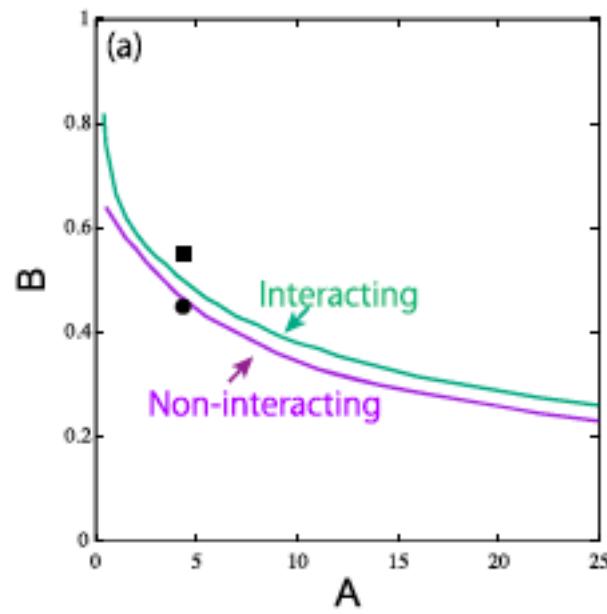
Two spinless fermions in one dimension

$$\left[\sum_{k=1,2} \left(-\frac{\nabla_k^2}{2} + V(x_k) \right) + \frac{1}{|x_1 - x_2|} \right] \Psi = E_{\text{tot}} \Psi$$

$$V(x) = -A \exp(-x^2/B^2).$$

Property:

- One particle GS is always bound.
- Lowest excited states can be either bound or unbound, depending on the potential.



1, Fully nonlocal V_{xc} for a simple model

R. Nagai, RA, S. Sasaki, and S. Tsuneyuki, J. Chem. Phys. 148, 241737 (2018)

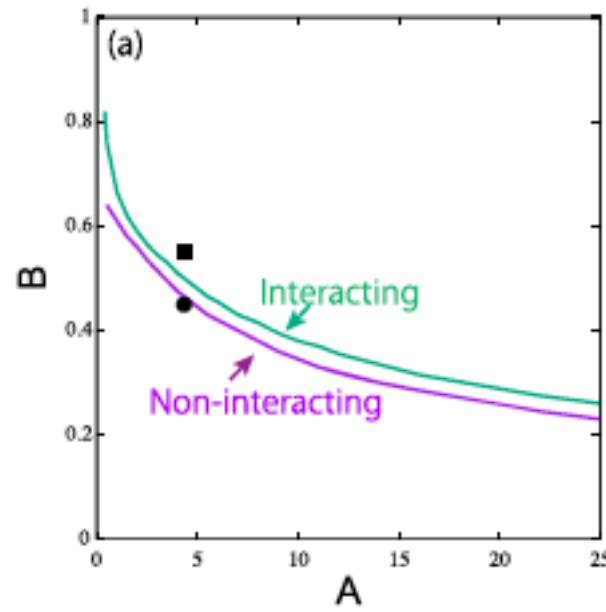
Two spinless fermions in one dimension

$$\left[\sum_{k=1,2} \left(-\frac{\nabla_k^2}{2} + V(x_k) \right) + \frac{1}{|x_1 - x_2|} \right] \Psi = E_{\text{tot}} \Psi$$

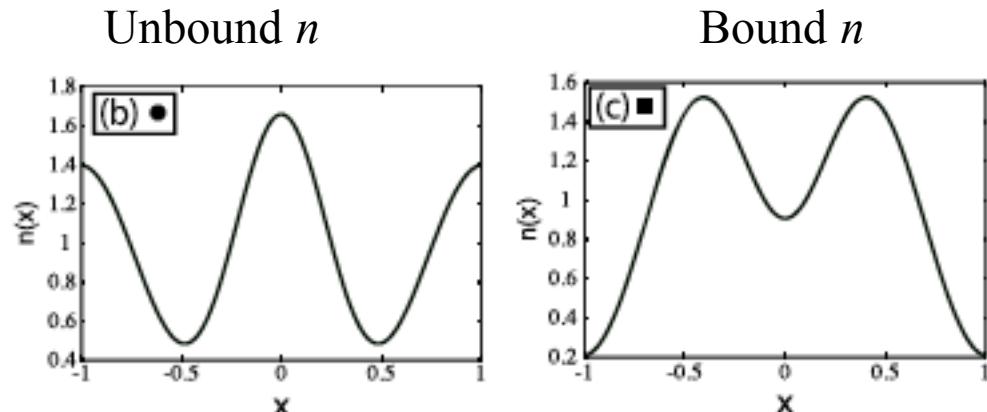
$$V(x) = -A \exp(-x^2/B^2).$$

Property:

- One particle GS is always bound.
- Lowest excited states can be either bound or unbound, depending on the potential.



If we train VHxc with unbound n , is it possible to apply it to bound region (and vice versa)?



1, Fully nonlocal V_{xc} for a simple model

R. Nagai, RA, S. Sasaki, and S. Tsuneyuki, J. Chem. Phys. 148, 241737 (2018)

Two spinless fermions in one dimension

Derivation of V_{Hxc} —collect the $(n \rightarrow V_{\text{Hxc}})$ pairs

- (1) Solve the Schrödinger eq. exactly for parameter (A, B) to get physical n and E_{tot}
- (2) Solve the Kohn-Sham eq. with varying V_{Hxc} so that the given n is reproduced

1, Fully nonlocal V_{xc} for a simple model

R. Nagai, RA, S. Sasaki, and S. Tsuneyuki, J. Chem. Phys. 148, 241737 (2018)

Two spinless fermions in one dimension

Derivation of V_{Hxc} —collect the $(n \rightarrow V_{\text{Hxc}})$ pairs

- (1) Solve the Schrödinger eq. exactly for parameter (A, B) to get physical n and E_{tot}
- (2) Solve the Kohn-Sham eq. with varying V_{Hxc} so that the given n is reproduced
- (2') Determine the constant part of V_{Hxc} so that E_{tot} is reproduced by the following

$$E_{\text{tot}} = \sum_{k=1,2} \varepsilon_k$$

Cf. Levy-Zahariev shifted formalism

M. Levy and F. Zahariev,
Phys. Rev. Lett. 113, 113002 (2014)

$$E_{\text{tot}} = \sum_{k=1,2} (\varepsilon_k + c[n]) + E_{\text{Hxc}} - \int dx \{V_{\text{Hxc}}(x) + c[n]\} n(x)$$

Set $c[n]$ so that this part =0

Advantages:

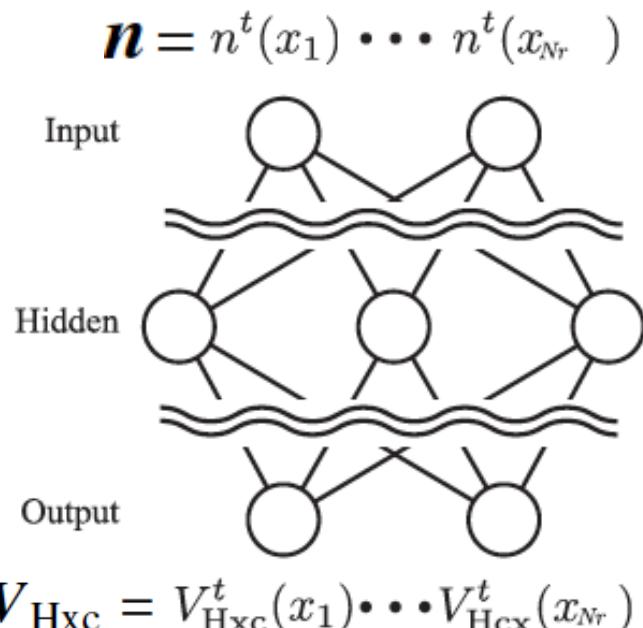
- (i) E_{tot} can be evaluated without explicit treatment of E_{Hxc} .
- (ii) Drastic n dependence of V_{Hxc} via step (2) is cured.

1, Fully nonlocal V_{xc} for a simple model

R. Nagai, RA, S. Sasaki, and S. Tsuneyuki, J. Chem. Phys. 148, 241737 (2018)

Two spinless fermions in one dimension

Derivation of V_{Hxc} —train the $n \rightarrow V_{\text{Hxc}}$ network



Fully connected NN formula

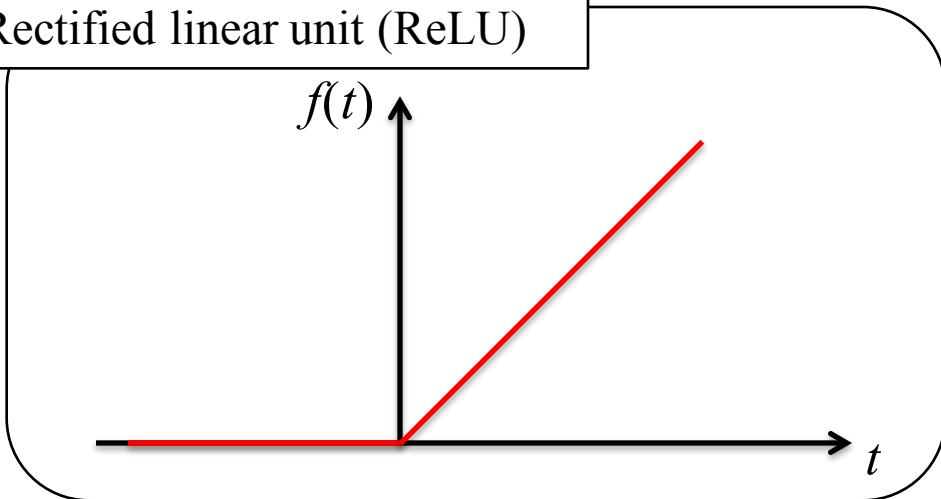
$$V_{\text{Hxc}} = \cdots f[W^{(2)}f[W^{(1)}\mathbf{n} + \mathbf{b}^{(1)}] + \mathbf{b}^{(2)}] \cdots$$

of real space grid : 100

of hidden layers : 2

of nodes per layer: 300

Nonlinear activation function:
Rectified linear unit (ReLU)

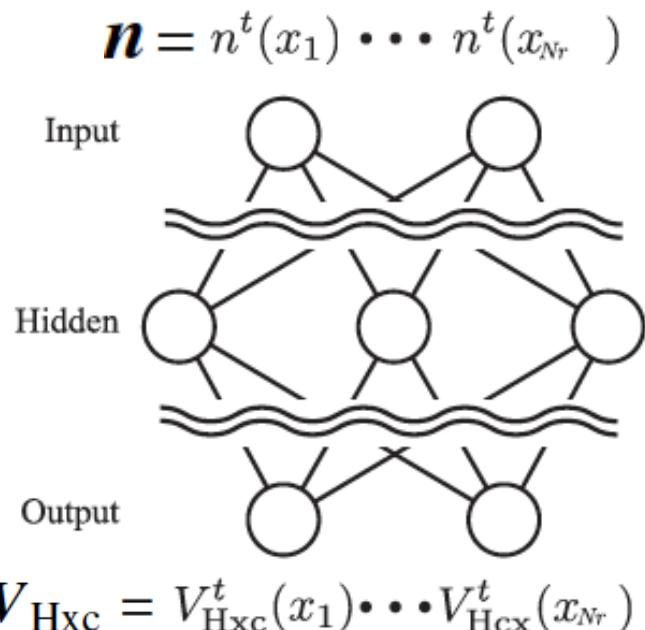


1, Fully nonlocal V_{xc} for a simple model

R. Nagai, RA, S. Sasaki, and S. Tsuneyuki, J. Chem. Phys. 148, 241737 (2018)

Two spinless fermions in one dimension

Derivation of V_{Hxc} —train the $n \rightarrow V_{\text{Hxc}}$ network



Fully connected NN formula

$$V_{\text{Hxc}} = \cdots f[W^{(2)}f[W^{(1)}\mathbf{n} + \mathbf{b}^{(1)}] + \mathbf{b}^{(2)}] \cdots$$

of real space grid : 100

of hidden layers : 2

of nodes per layer: 300

Error function to be minimized

$$\text{Err}(\mathbf{w}) = \frac{1}{N_{\text{train}} N_r} \sum_{i=1}^{N_{\text{train}}} \sum_{j=1}^{N_r} \left[V_{\text{Hxc}}^{\text{NN}}(n^{(i)}; x_j) - V_{\text{Hxc}}^{(i)}(x_j) \right]^2.$$

1, Fully nonlocal V_{xc} for a simple model

R. Nagai, RA, S. Sasaki, and S. Tsuneyuki, J. Chem. Phys. 148, 241737 (2018)

Two spinless fermions in one dimension

Result—transferrability test

- (i) Solve the Kohn-Sham equation for a given (A, B)
using the trained NN- V_{Hxc}
- (ii) Evaluate the errors from the exact diagonalization result

Test errors

$$\Delta n = \sqrt{\sum_{i=1}^{N_r} (n(x_i) - n^{\text{exact}}(x_i))^2 / N_r}, \quad \Delta E = |E_{\text{tot}} - E_{\text{tot}}^{\text{exact}}|$$

1, Fully nonlocal V_{xc} for a simple model

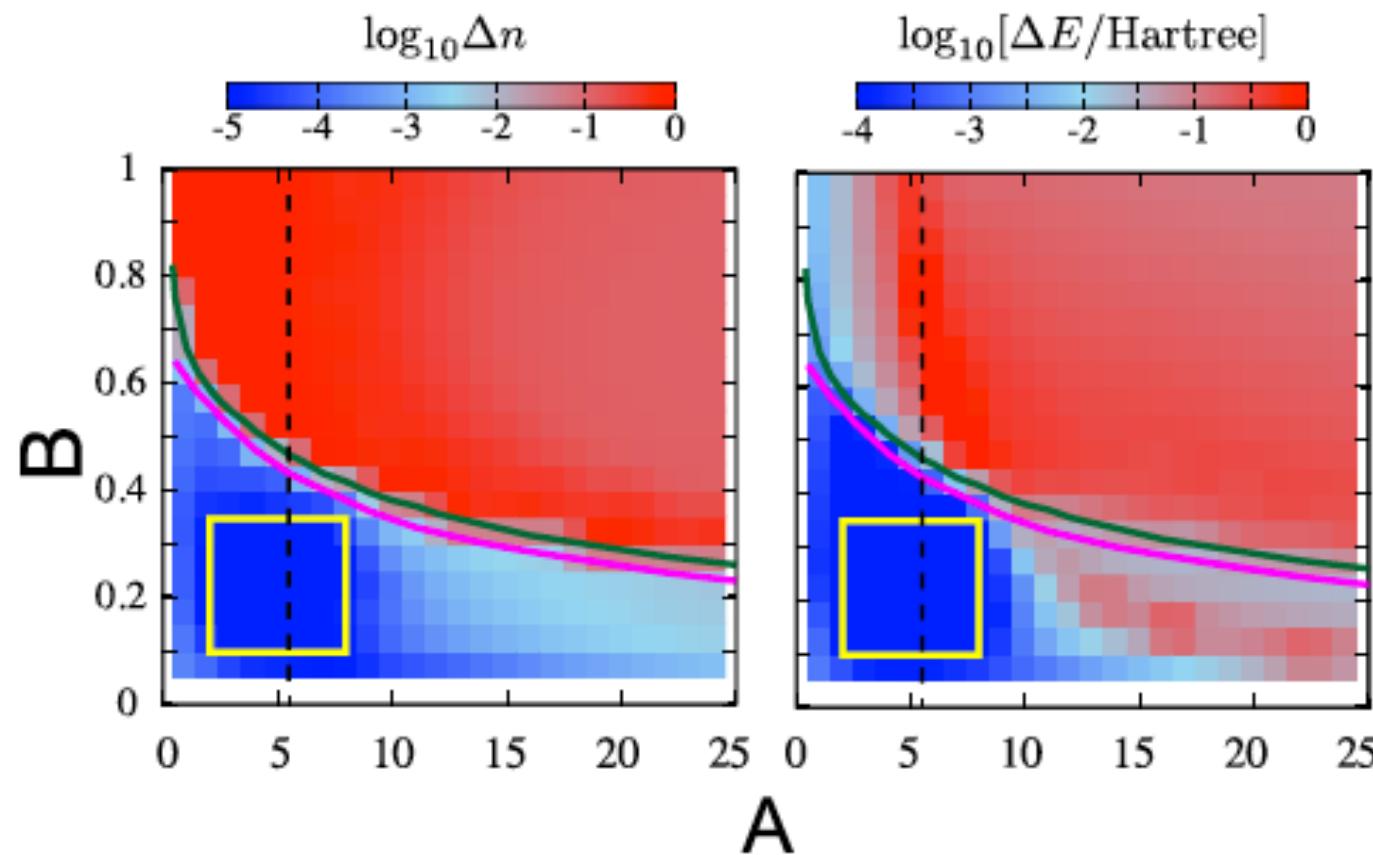
R. Nagai, RA, S. Sasaki, and S. Tsuneyuki, J. Chem. Phys. 148, 241737 (2018)

Two spinless fermions in one dimension

Result—transferrability test

Test errors

$$\Delta n = \sqrt{\sum_{i=1}^{N_r} (n(x_i) - n^{\text{exact}}(x_i))^2 / N_r}, \quad \Delta E = |E_{\text{tot}} - E_{\text{tot}}^{\text{exact}}|$$



1, Fully nonlocal V_{xc} for a simple model

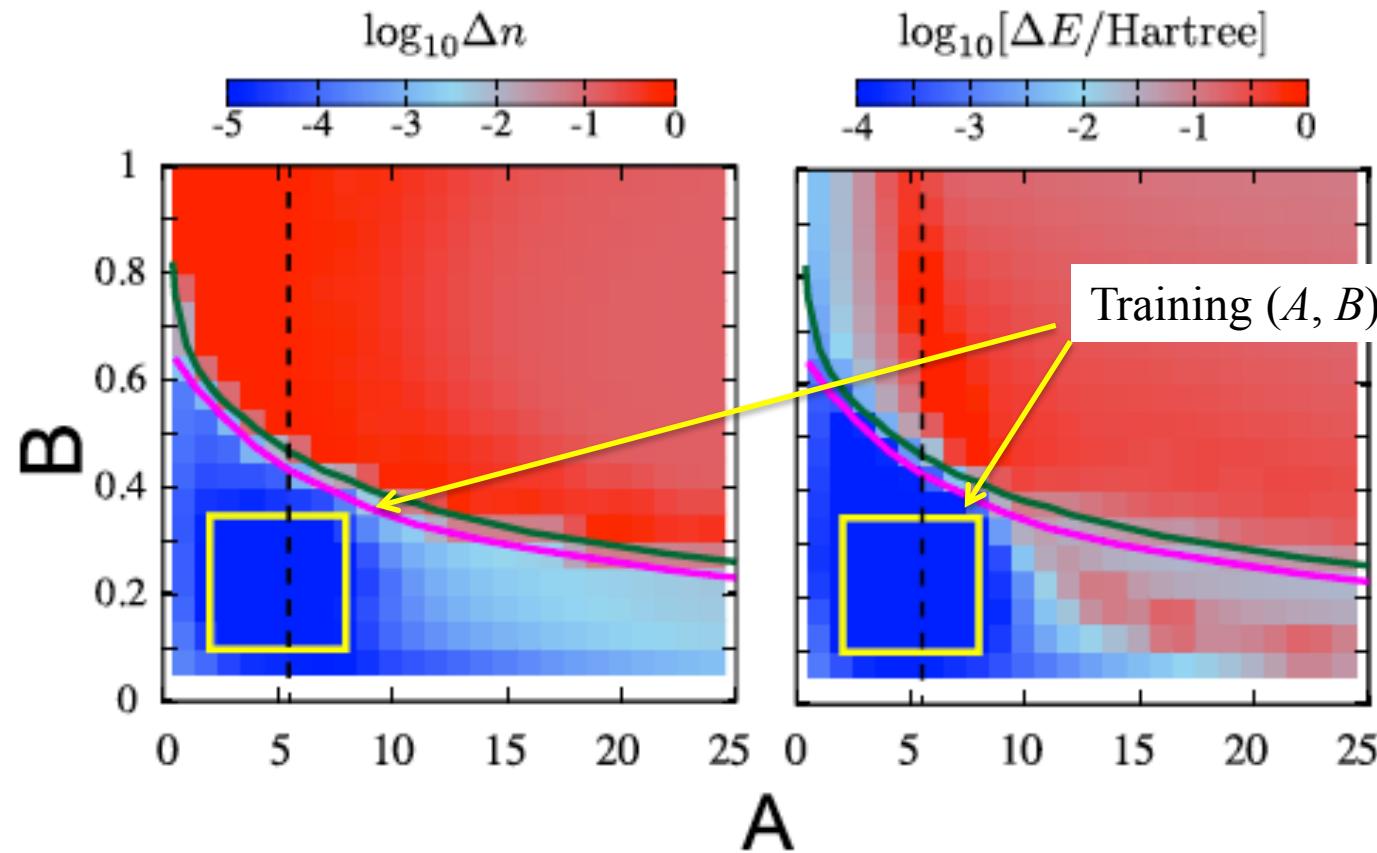
R. Nagai, RA, S. Sasaki, and S. Tsuneyuki, J. Chem. Phys. 148, 241737 (2018)

Two spinless fermions in one dimension

Result—transferrability test

Test errors

$$\Delta n = \sqrt{\sum_{i=1}^{N_r} (n(x_i) - n^{\text{exact}}(x_i))^2 / N_r}, \quad \Delta E = |E_{\text{tot}} - E_{\text{tot}}^{\text{exact}}|$$



1, Fully nonlocal V_{xc} for a simple model

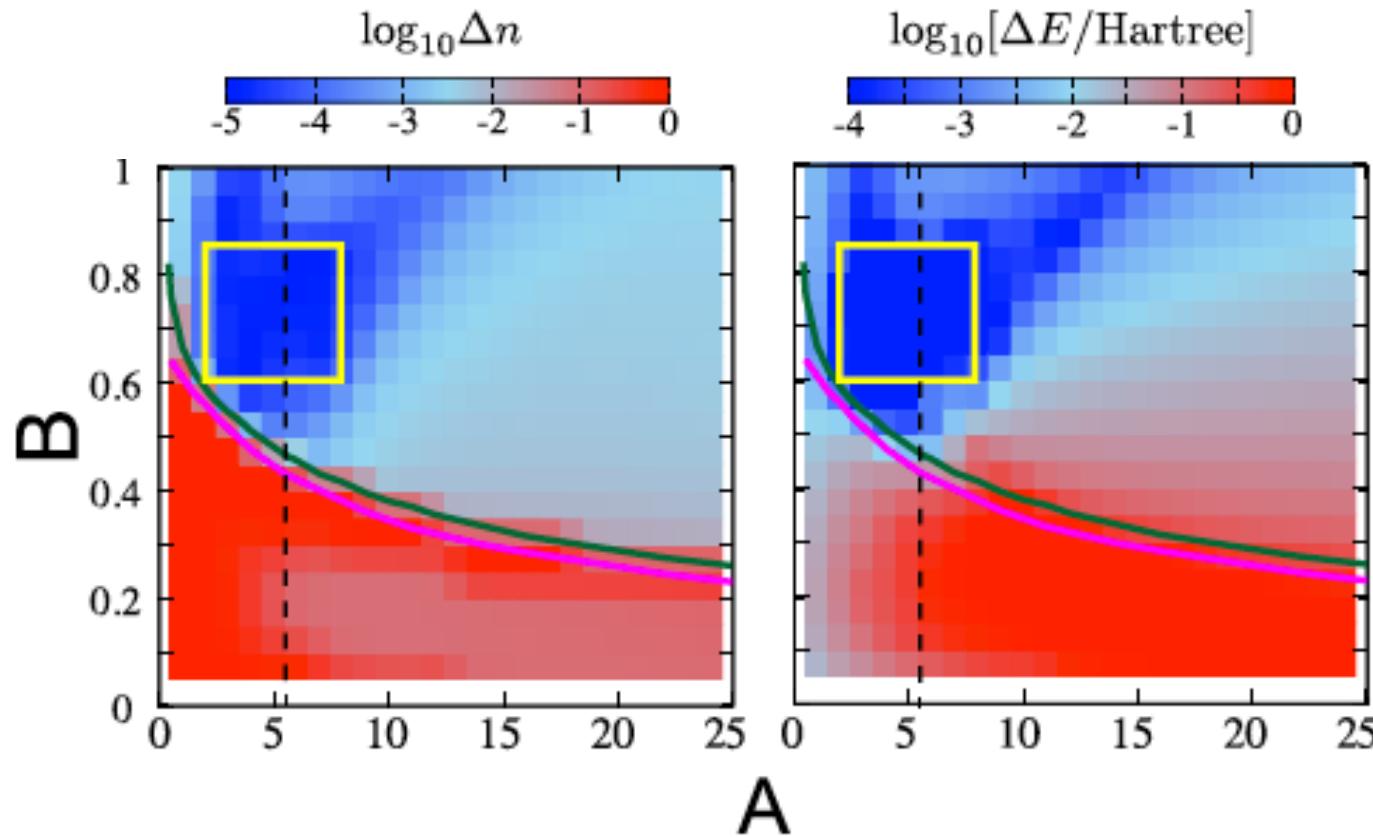
R. Nagai, RA, S. Sasaki, and S. Tsuneyuki, J. Chem. Phys. 148, 241737 (2018)

Two spinless fermions in one dimension

Result—transferrability test

Test errors

$$\Delta n = \sqrt{\sum_{i=1}^{N_r} (n(x_i) - n^{\text{exact}}(x_i))^2 / N_r}, \quad \Delta E = |E_{\text{tot}} - E_{\text{tot}}^{\text{exact}}|$$



1, Fully nonlocal V_{xc} for a simple model

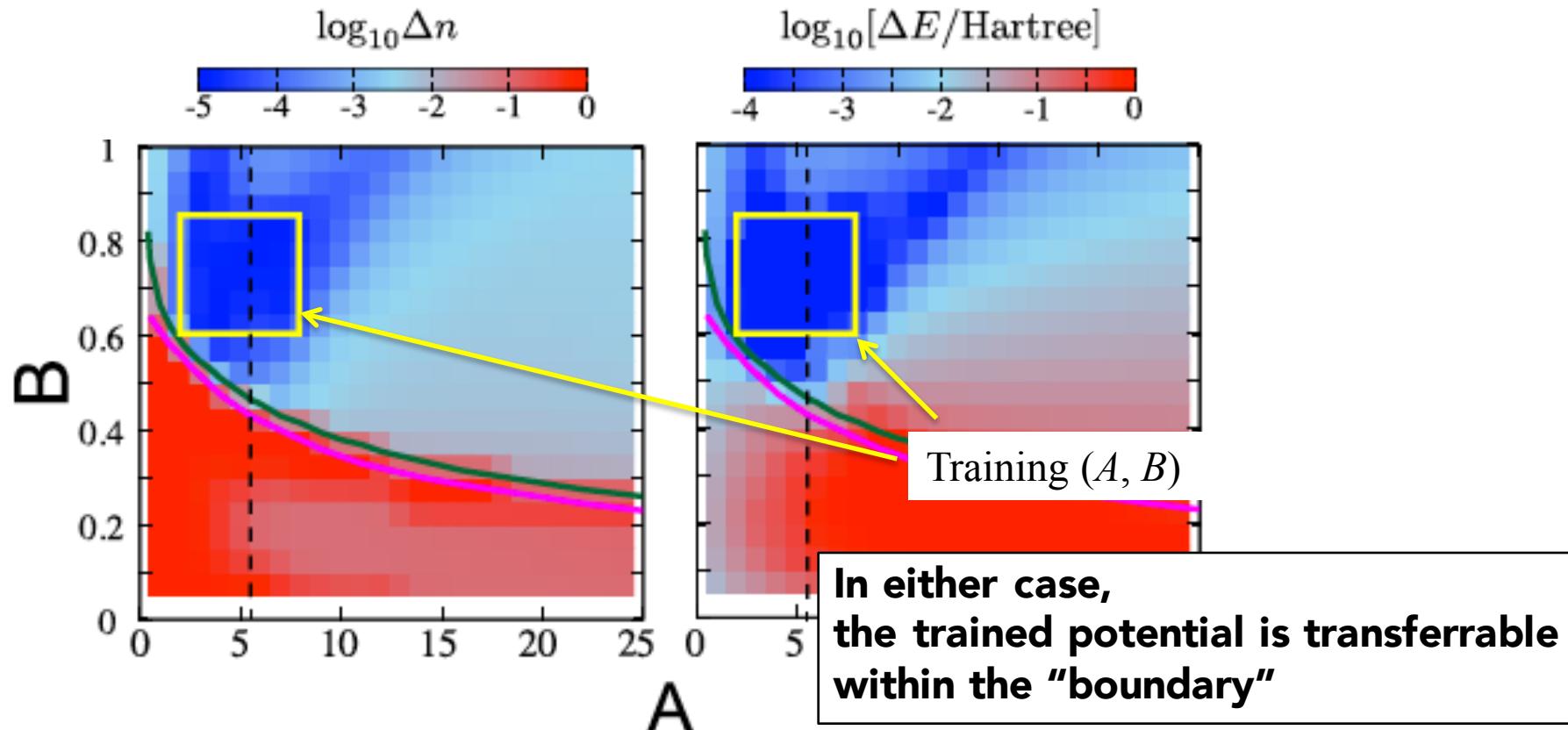
R. Nagai, RA, S. Sasaki, and S. Tsuneyuki, J. Chem. Phys. 148, 241737 (2018)

Two spinless fermions in one dimension

Result—transferrability test

Test errors

$$\Delta n = \sqrt{\sum_{i=1}^{N_r} (n(x_i) - n^{\text{exact}}(x_i))^2 / N_r}, \quad \Delta E = |E_{\text{tot}} - E_{\text{tot}}^{\text{exact}}|$$



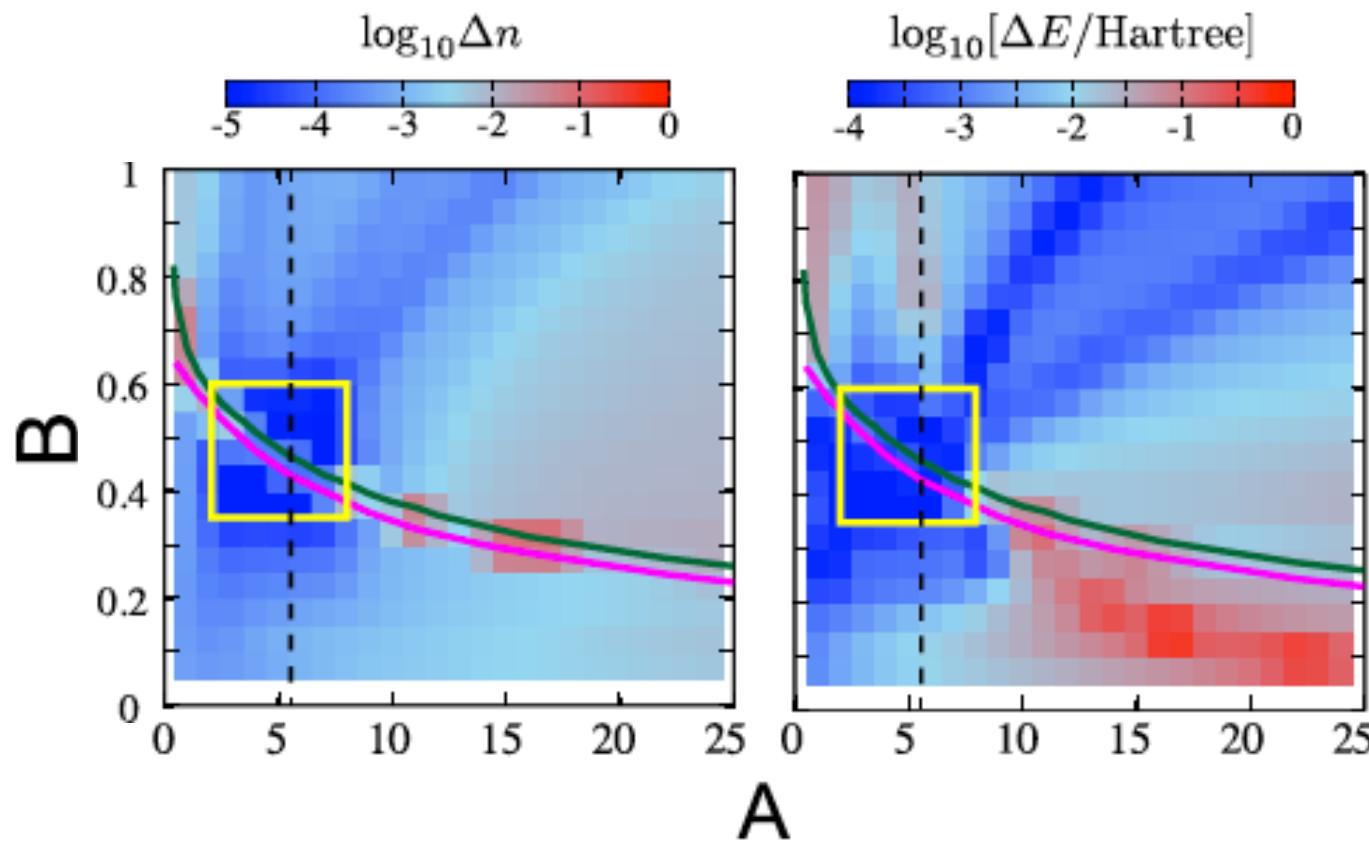
1, Fully nonlocal V_{xc} for a simple model

R. Nagai, RA, S. Sasaki, and S. Tsuneyuki, J. Chem. Phys. 148, 241737 (2018)

Two spinless fermions in one dimension

Result—transferrability test

Training across the “boundary”



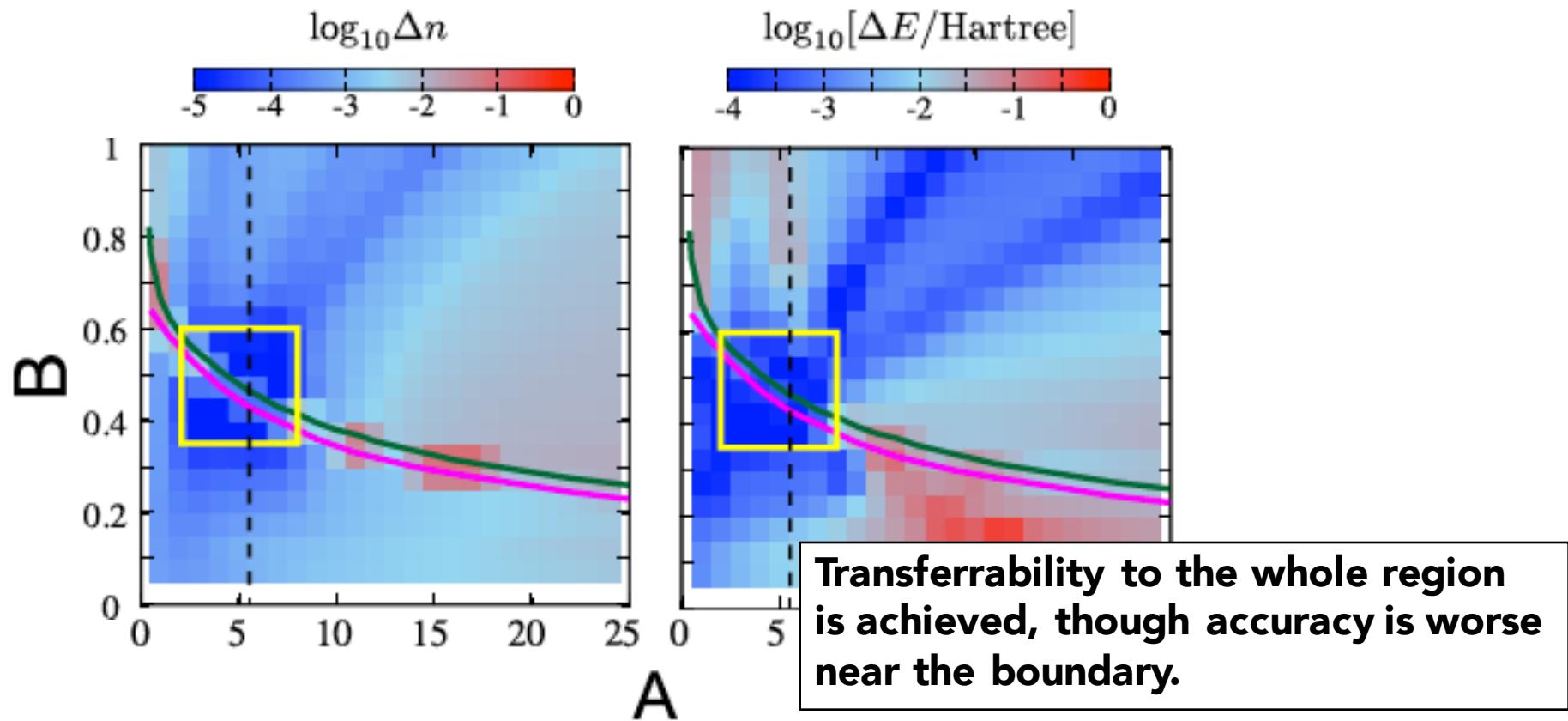
1, Fully nonlocal V_{xc} for a simple model

R. Nagai, RA, S. Sasaki, and S. Tsuneyuki, J. Chem. Phys. 148, 241737 (2018)

Two spinless fermions in one dimension

Result—transferrability test

Training across the “boundary”



1, Fully nonlocal V_{xc} for a simple model

R. Nagai, RA, S. Sasaki, and S. Tsuneyuki, J. Chem. Phys. 148, 241737 (2018)

Relative transferrability

“By-passing” the Schrodinger (Kohn-Sham) equation

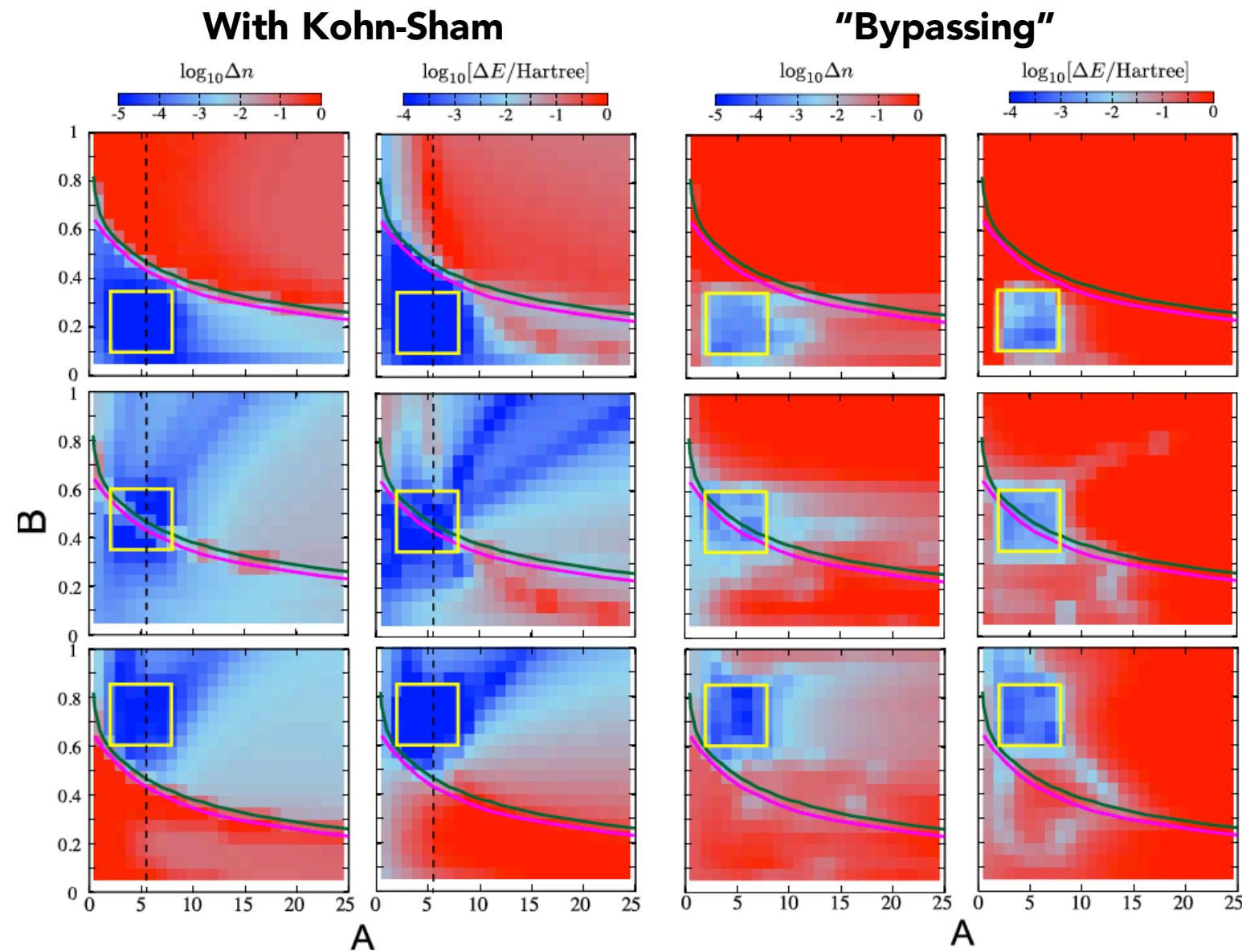
F. Brockherde, K. Burke et al., Nat. Commun. 8, 872 (2017)

$$[V_{\text{ion}}(\mathbf{r}_1), V_{\text{ion}}(\mathbf{r}_2), \dots] \xrightarrow{F} [n(\mathbf{r}_1), n(\mathbf{r}_2), \dots]$$

F : Solve the Schrodinger (or Kohn-Sham) equation

1, Fully nonlocal V_{xc} for a simple model

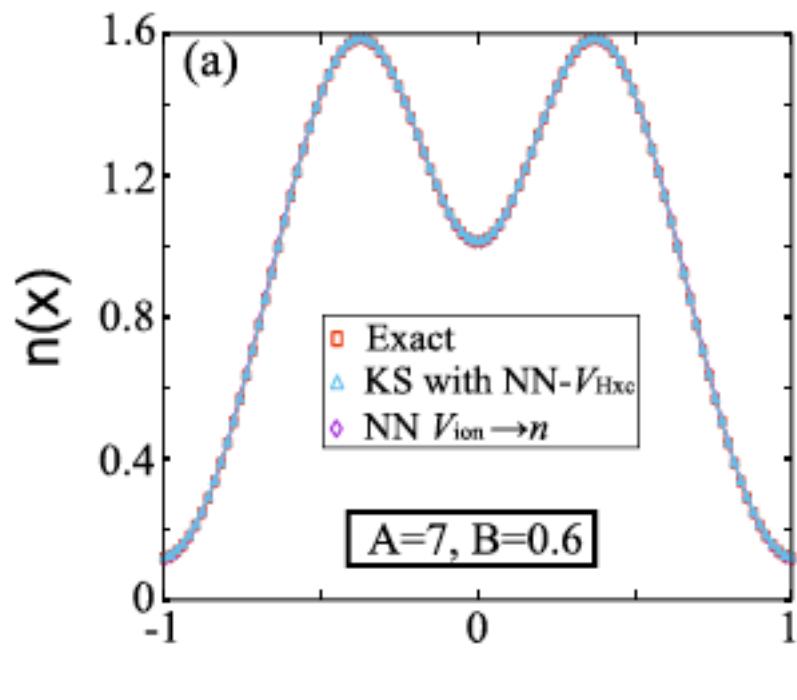
R. Nagai, RA, S. Sasaki, and S. Tsuneyuki, J. Chem. Phys. 148, 241737 (2018)



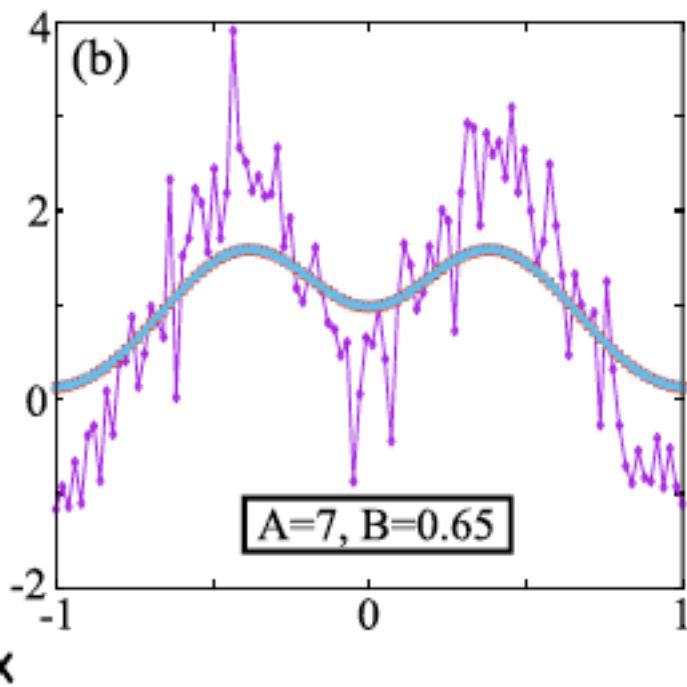
1, Fully nonlocal V_{xc} for a simple model

R. Nagai, RA, S. Sasaki, and S. Tsuneyuki, J. Chem. Phys. 148, 241737 (2018)

In the training range



Out of the training range



$$\left[-\frac{\nabla^2}{2} + V(x) + V_{\text{Hxc}}([n]; x) \right] \psi_k(x) = \varepsilon_k \psi_k(x)$$

Kinetic term mitigates the spurious oscillation, yielding better transferrability

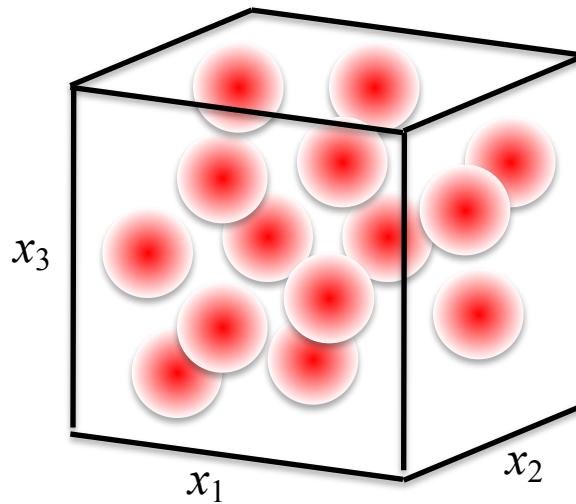
1, Fully nonlocal V_{xc} for a simple model

R. Nagai, RA, S. Sasaki, and S. Tsuneyuki, J. Chem. Phys. **148**, 241737 (2018)

Summary

The extrapolation error of the NN V_{Hxc} is not so crucial, as long as the localized/delocalized character of n does not vary drastically, suggesting the exact V_{Hxc} is locally mimicked numerically.

→ “Patchwork” numerical construction of V_{Hxc} is possible?



2, Semilocal V_{xc} for small molecules

R. Nagai, RA, and O. Sugino, arXiv:1903.00238

Let's see E_{xc} as the function.

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

Local descriptor vector

$$\mathbf{g}[n](\mathbf{r}) = \left(n(\mathbf{r}), \zeta(\mathbf{r}) \equiv \frac{n_\uparrow(\mathbf{r}) - n_\downarrow(\mathbf{r})}{n(\mathbf{r})} \right) \quad (\text{LSDA})$$

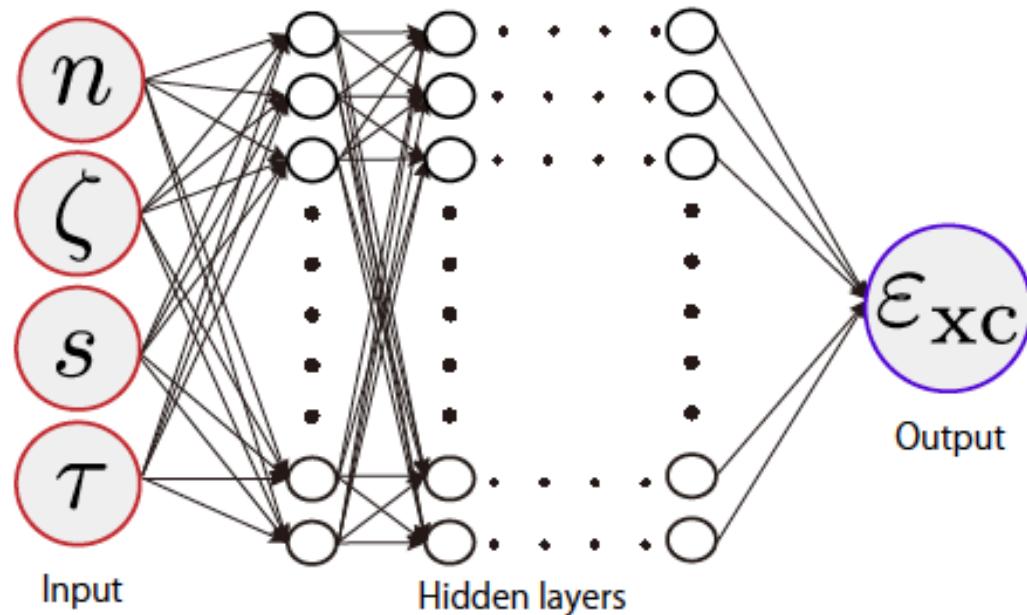
$$\left(n(\mathbf{r}), \zeta(\mathbf{r}), s(\mathbf{r}) \equiv \frac{|\nabla n(\mathbf{r})|}{n^{4/3}} \right) \quad (\text{GGA})$$

$$\left(n(\mathbf{r}), \zeta(\mathbf{r}), s(\mathbf{r}), \tau(\mathbf{r}) \equiv \frac{1}{2} \sum_i^{\text{occ}} |\nabla \phi_i(\mathbf{r})|^2 \right) \quad (\text{MetaGGA})$$

...

2, Semilocal \mathbf{V}_{xc} for small molecules

R. Nagai, RA, and O. Sugino, arXiv:1903.00238



4 hidden layers
W1: $N \times 100$ (N is the size of \mathbf{g})
W2: 100×100
W3: 100×100
W4: 100×1

$$V_{xc}(\mathbf{r}) = \frac{\delta E_{xc}}{\delta n(\mathbf{r})} = \varepsilon_{xc}(\mathbf{r}) + n(\mathbf{r}) \frac{\delta \mathbf{g}(\mathbf{r})}{\delta n(\mathbf{r})} \cdot \frac{\partial \varepsilon_{xc}(\mathbf{r})}{\partial \mathbf{g}(\mathbf{r})}$$

Demonstration

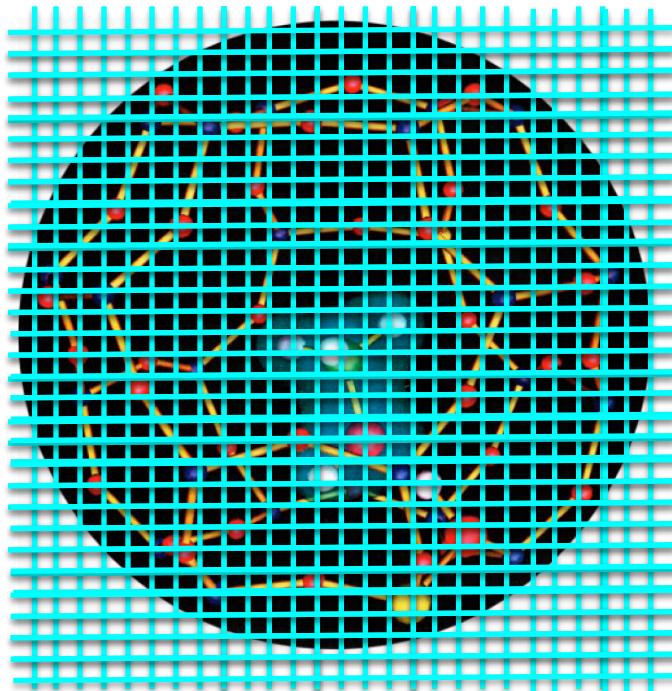
- (i) Train the NN with reference data from accurate wave function theory (CCSD)
- (ii) Apply the trained NN to unreference systems

2, Semilocal V_{xc} for small molecules

R. Nagai, RA, and O. Sugino, arXiv:1903.00238

Training

$g(r)$ data from CCSD



Even in one system, various $g(\mathbf{r})$ values are obtained.

Parameters of the NN can be determined with fitting of the *whole* $n(\mathbf{r})$ via the KS eq. using the *common* NN to all the \mathbf{r} points.

2, Semilocal V_{xc} for small molecules

R. Nagai, RA, and O. Sugino, arXiv:1903.00238

Training

Error function referring to *three* molecules

$$\Delta_{\text{err}} = a \times (\Delta^{\text{CCSD}} n_{\text{H}_2\text{O}} + \Delta^{\text{CCSD}} n_{\text{NH}_3} + \Delta^{\text{CCSD}} n_{\text{NO}})$$

n-error

$$+ (\Delta^{\text{G2}} AE_{\text{H}_2\text{O}} + \Delta^{\text{G2}} AE_{\text{NH}_3} + \Delta^{\text{G2}} AE_{\text{NO}})$$

atomization energy error

Our criterion to choose the molecules: respect the diversity of \mathbf{g}

- (i) low symmetry
- (ii) electrically polarized
- (iii) spin polarized (NO)

2, Semilocal V_{xc} for small molecules

R. Nagai, RA, and O. Sugino, arXiv:1903.00238

Training

Error function referring to *three* molecules

$$\Delta_{\text{err}} = a \times (\Delta^{\text{CCSD}} n_{\text{H}_2\text{O}} + \Delta^{\text{CCSD}} n_{\text{NH}_3} + \Delta^{\text{CCSD}} n_{\text{NO}})$$

n-error

$$+ (\Delta^{\text{G2}} AE_{\text{H}_2\text{O}} + \Delta^{\text{G2}} AE_{\text{NH}_3} + \Delta^{\text{G2}} AE_{\text{NO}})$$

atomization energy error

NN optimization cycle

Change $\{W_i\}, \{b_i\}$ slightly



DFT calculation using NN functional
and evaluate cost function Δ_{err}



Accept/Reject the $\{W_i\}, \{b_i\}$
(Metropolis-type update)

2, Semilocal V_{xc} for small molecules

R. Nagai, RA, and O. Sugino, arXiv:1903.00238

Systematic extension to nonlocality

Near-region approximation (NRA)

$$\mathbf{g}[n](\mathbf{r}) = (n(\mathbf{r}), \zeta(\mathbf{r}), s(\mathbf{r}), \tau(\mathbf{r}), R(\mathbf{r}))$$

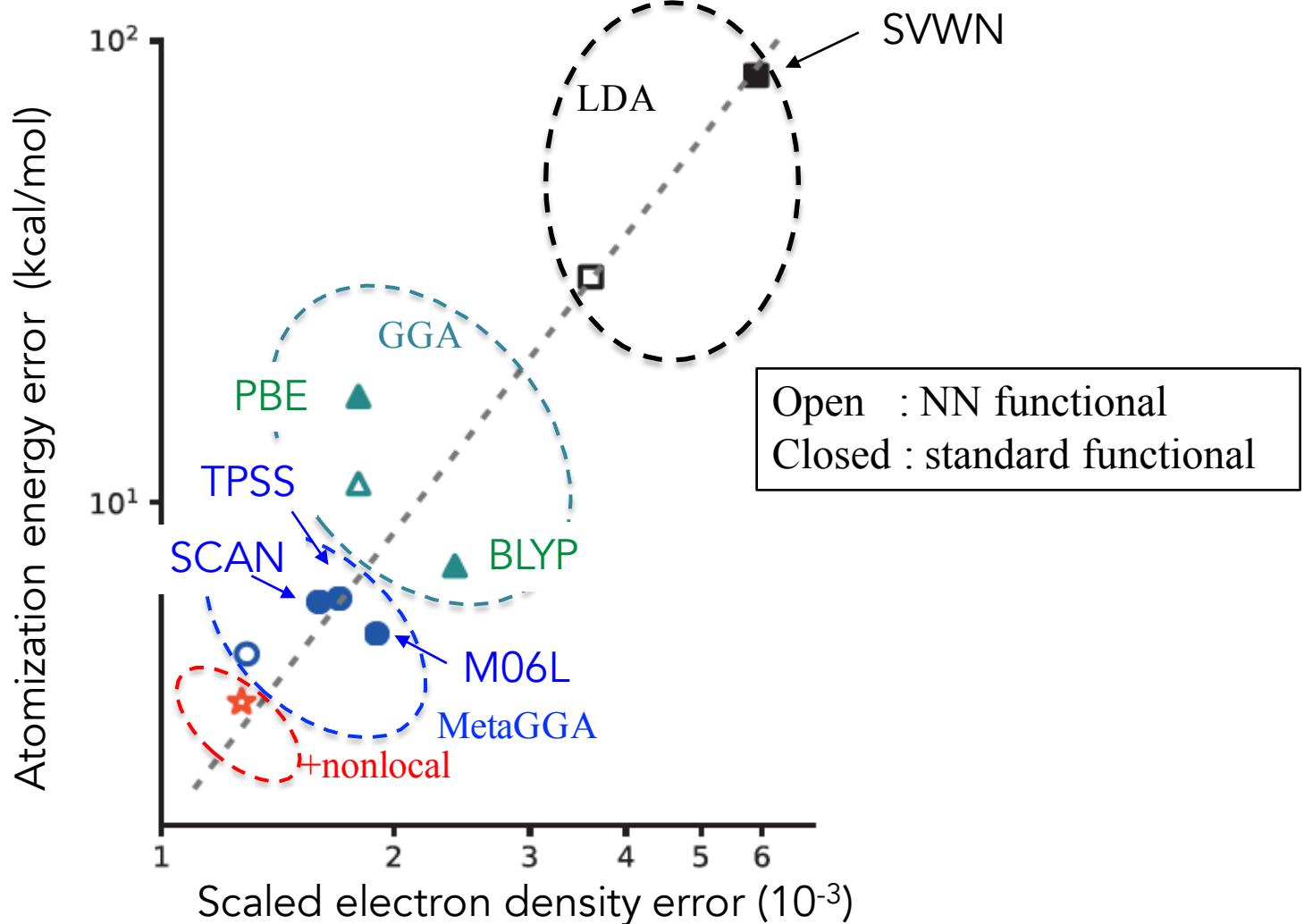
$$R(\mathbf{r}) = \int d\mathbf{r}' n(\mathbf{r}') \exp\left(-\frac{|\mathbf{r} - \mathbf{r}'|}{\sigma}\right) \quad (\sigma = 0.2)$$

Exploratory extension is easy with no reference asymptotic formulae.

2, Semilocal V_{xc} for small molecules

R. Nagai, RA, and O. Sugino, arXiv:1903.00238

Result: averaged error over 147 molecules

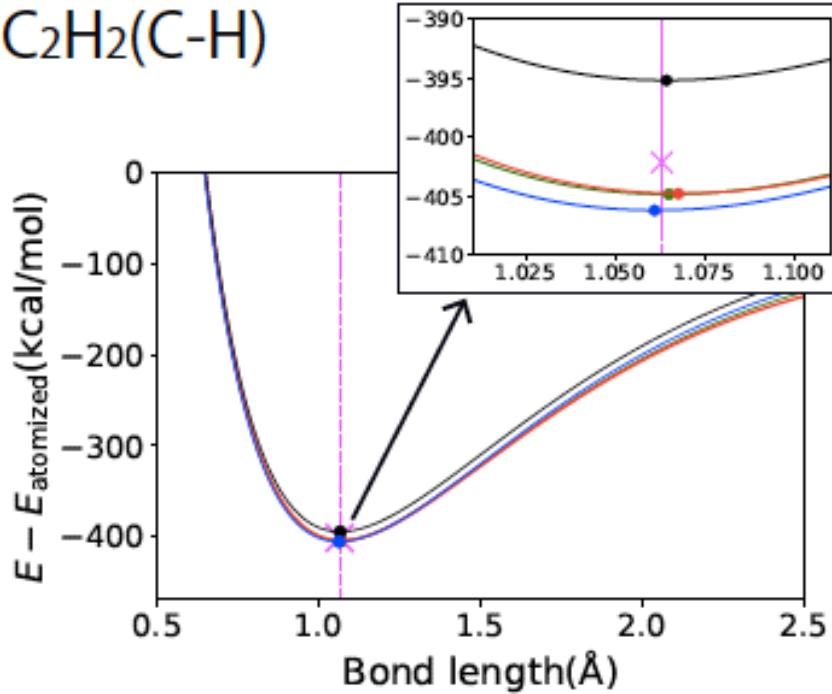


2, Semilocal V_{xc} for small molecules

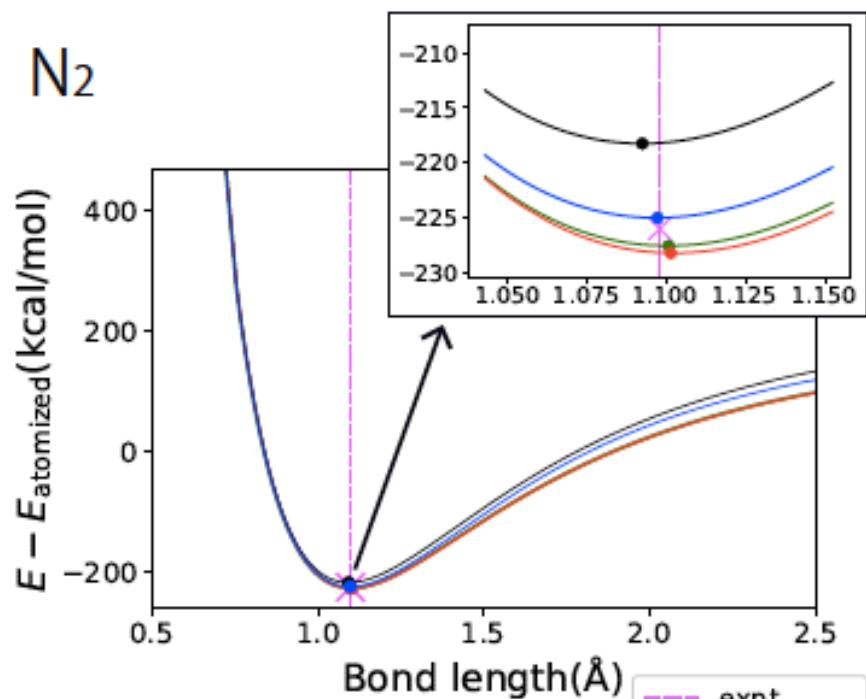
R. Nagai, RA, and O. Sugino, arXiv:1903.00238

Result: Dissociation curves for unreference molecules

$\text{C}_2\text{H}_2(\text{C}-\text{H})$



N_2



- expt
- TPSS
- SCAN
- M06-L
- NN-mGGA

2, Semilocal V_{xc} for small molecules

R. Nagai, RA, and O. Sugino, arXiv:1903.00238

Summary

Ask the semilocal relation $g \rightarrow \epsilon_{xc}$ to real molecules, every a few of which provide much amount of information of the ideal ϵ_{xc} .

Code availability

<https://github.com/ml-electron-project/NNfunctional>

| ml-electron-project Update README.md | | |
|--------------------------------------|----------------------|---------------------------------|
| | | Latest commit 3f4a979 on 21 Jul |
| File | Action | Time |
| GGA.py | Add files via upload | 7 months ago |
| LICENSE | Update LICENSE | 3 months ago |
| LSDA.py | Add files via upload | 7 months ago |
| NNGGA | Add files via upload | 7 months ago |
| NNLSDA | Add files via upload | 7 months ago |
| NNNRA | Add files via upload | 7 months ago |
| NNmGGA | Add files via upload | 7 months ago |
| NRA.py | Add files via upload | 7 months ago |
| README.md | Update README.md | 3 months ago |
| metaGGA.py | Add files via upload | 7 months ago |
| numint.py | Add files via upload | 3 months ago |
| numint_1.5.py | Add files via upload | 3 months ago |

2, Semilocal V_{xc} for small molecules

R. Nagai, RA, and O. Sugino, arXiv:1903.00238

Code availability

<https://github.com/ml-electron-project/NNfunctional>

| Branch: master | | New pull request | Create new file | Upload files | Find file | Clone or download |
|--------------------------------------|----------------------|------------------|-----------------|--------------|-----------|---------------------------------|
| ml-electron-project Update README.md | | | | | | Latest commit 3f4a979 on 21 Jul |
| GGA.py | Add files via upload | | | | | 7 months ago |
| LICENSE | Update LICENSE | | | | | 3 months ago |
| LSDA.py | Add files via upload | | | | | 7 months ago |
| NNGGA | Add files via upload | | | | | 7 months ago |
| NNLSDA | Add files via upload | | | | | 7 months ago |
| NNNRA | Add files via upload | | | | | 7 months ago |
| NNmGGA | Add files via upload | | | | | 7 months ago |
| NRA.py | Add files via upload | | | | | 7 months ago |
| README.md | Update README.md | | | | | 3 months ago |
| metaGGA.py | Add files via upload | | | | | 7 months ago |
| numint.py | Add files via upload | | | | | 3 months ago |
| numint_1.5.py | Add files via upload | | | | | 3 months ago |

0, Install PySCF (<https://sunqm.github.io/pyscf/install.html>) and
PyTorch (<https://pytorch.org/>)

- 1, Put the NN parameter file (NNLSDA, NNGGA, NNmGGA, NNNRA)
- 2, *\$ python GGA.py*
- 3, Enjoy!

Final remark

@ a panel discussion in 物性研究所スパコン共同利用・CCMS合同研究会 (2016), ISSP

RA



出自はさておき解が見つかれば確度検証は容易な
タイプの問題と機械学習は相性がいいと思う。
一例は結晶構造探索だが他には何かないか？

ML seems useful for a type of problems where once
a solution is heuristically found its validation is easy.
An example is crystal structure search; any others?

Final remark

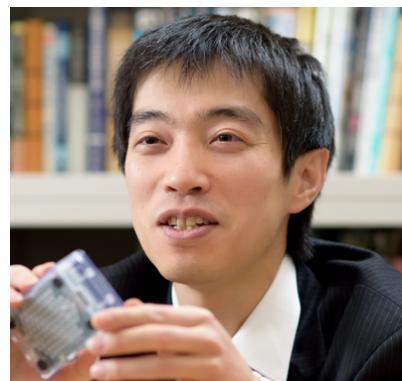
@ a panel discussion in 物性研究所スパコン共同利用・CCMS合同研究会 (2016), ISSP

RA



出自はさておき解が見つかれば確度検証は容易な
タイプの問題と機械学習は相性がいいと思う。
一例は結晶構造探索だが他には何かないか?
ML seems useful for a type of problems where once
a solution is heuristically found its validation is easy.
An example is crystal structure search; any others?

Prof. Yoshihide Yoshimoto
(UTokyo)



交換相關ポテンシャル、擬ポテンシャル、...
などはよいのではないか
Exchange-correlation potential, pseudopotential, etc.?

Final remark

@ a panel discussion in 物性研究所スパコン共同利用・CCMS合同研究会 (2016), ISSP

RA



Prof. Yoshihide Yoshimoto
(UTokyo)

出自はさておき解が見つかれば確度検証は容易な
タイプの問題と機械学習は相性がいいと思う。
一例は結晶構造探索だが他には何かないか?
ML seems useful for a type of problems where once
a solution is heuristically found its validation is easy.
An example is crystal structure search; any others?

交換相關ポテンシャル、擬ポテンシャル、...
などはよいのではないか

Exchange-correlation potential, pseudopotential, etc.?



Any others, guys?

Refs.

- R. Nagai, RA, S. Sasaki, and S. Tsuneyuki, J. Chem. Phys. **148**, 241737 (2018)
R. Nagai, RA, and O. Sugino, arXiv:1903.00238