Fermi Flow: Ab initio study of fermions at finite temperature

Lei Wang (王磊)
Institute of Physics, CAS
https://wangleiphy.github.io

2105.08644

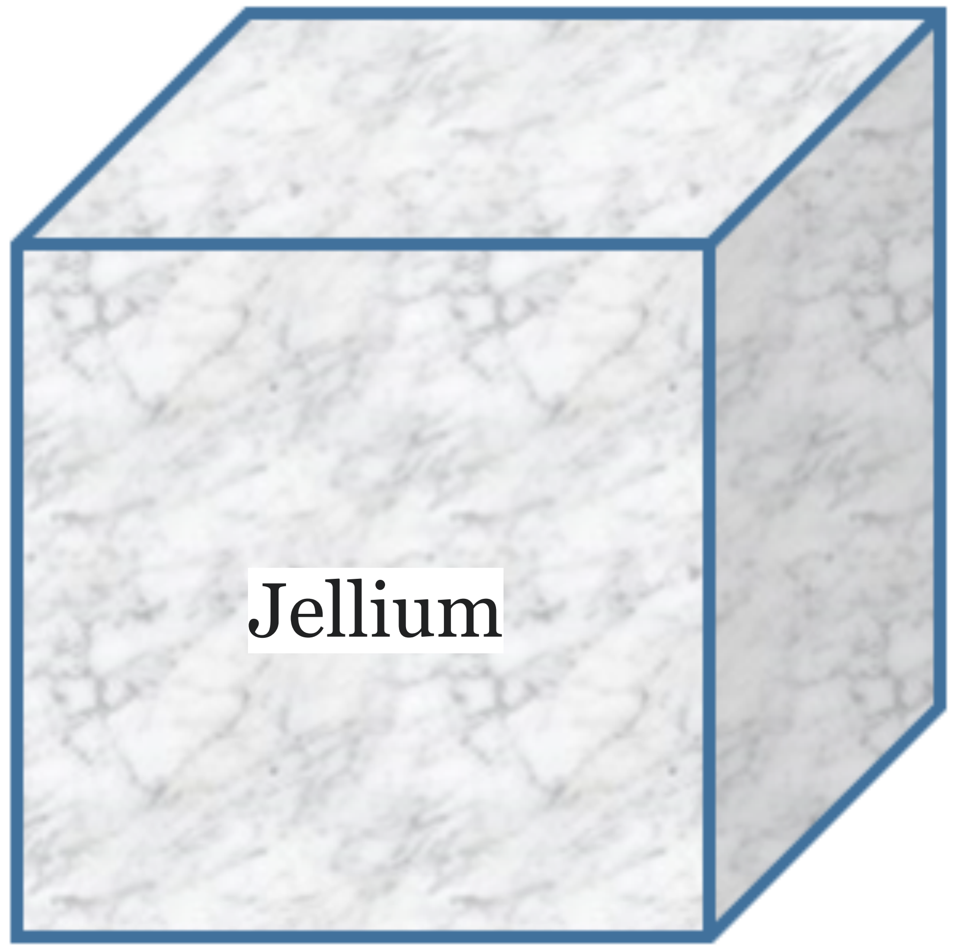
buwantaiji/FermiFlow

& work in progress

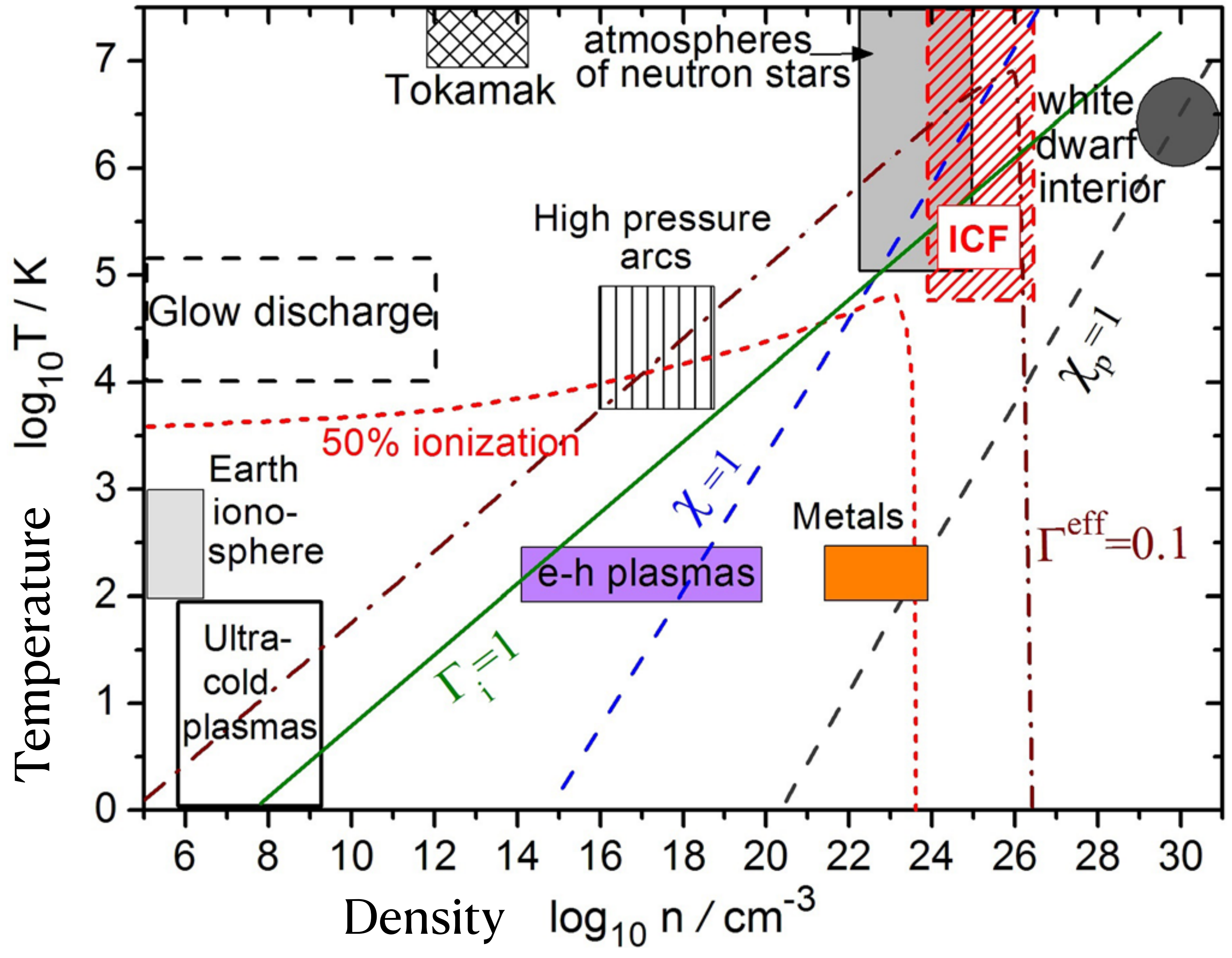
Uniform electron gas

The simplest model of interacting electrons ("the zeroth element")

$$H = -\sum_{i=1}^{N} \frac{\hbar^2 \nabla_i^2}{2m} + \sum_{i<j} \frac{e^2}{|\boldsymbol{x}_i - \boldsymbol{x}_j|}$$



Relevant to condensed matter, warm dense matter, thermal density functionals ...



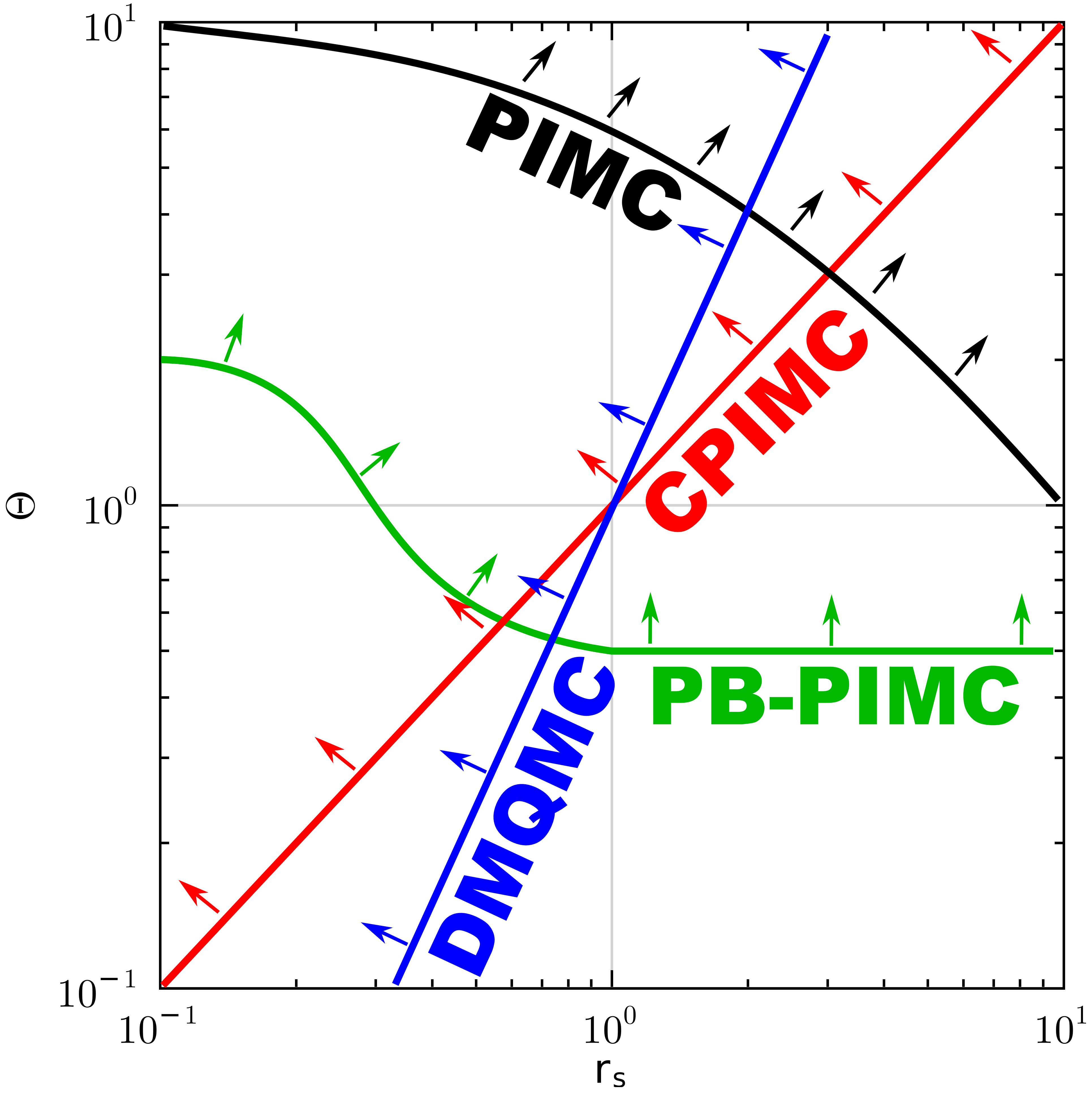
Bonitz et al, Phys. Plasmas '20

Limitation of current approaches

- There is no reliable data at low temperature and intermediate interaction, despite of decades of research

- The workhorse (Path Integral Monte Carlo methods) suffer from the notorious "sign problem"

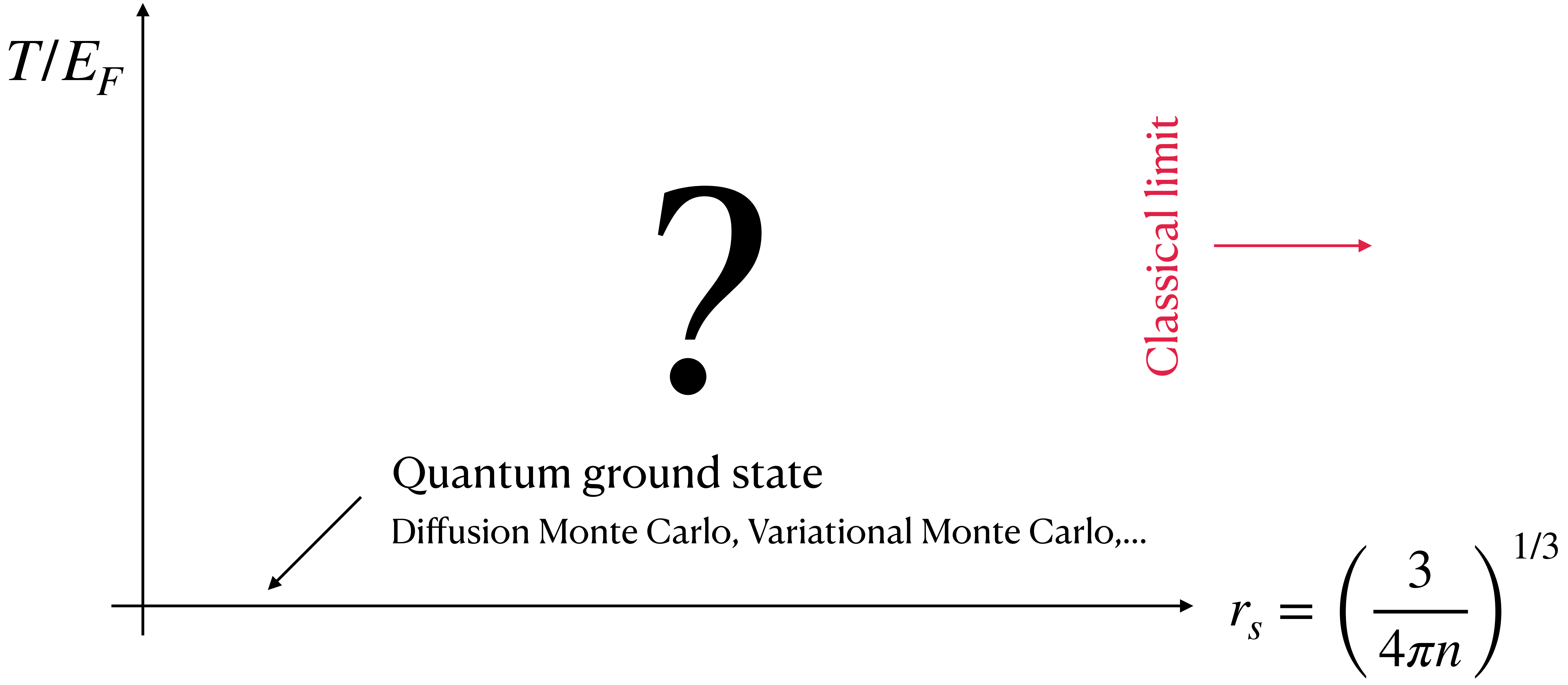
- Opportunity for deep learning



Dornheim et al, Phys. Plasmas '17

Warm up: Classical Coulomb gas

$$H = \sum_{i<j} \frac{e^2}{|\boldsymbol{x}_i - \boldsymbol{x}_j|} \qquad Z = \int d\boldsymbol{x}^{3N} e^{-\beta H(\boldsymbol{x})}$$



Warm up: Classical Coulomb gas

$$H = \sum_{i<j} \frac{e^2}{|\boldsymbol{x}_i - \boldsymbol{x}_j|} \qquad Z = \int d\boldsymbol{x}^{3N} e^{-\beta H(\boldsymbol{x})}$$

Variational free-energy $$\mathscr{L} = \mathop{\mathbb{E}}_{\boldsymbol{x}\sim p(\boldsymbol{x})} \left[\ln p(\boldsymbol{x}) + \beta H(\boldsymbol{x})\right] \geq -\ln Z$$

variational probability distribution

Turn the sampling problem to an optimization problem.

Not necessarily easy. But may better leverage deep learning engine.

…

Warm up: Classical Coulomb gas

$$H = \sum_{i<j} \frac{e^2}{|\boldsymbol{x}_i - \boldsymbol{x}_j|} \qquad Z = \int d\boldsymbol{x}^{3N} e^{-\beta H(\boldsymbol{x})}$$

Variational free-energy $$\mathscr{L} = \mathop{\mathbb{E}}_{\boldsymbol{x}\sim p(\boldsymbol{x})} \left[\ln p(\boldsymbol{x}) + \beta H(\boldsymbol{x})\right] \geq -\ln Z$$

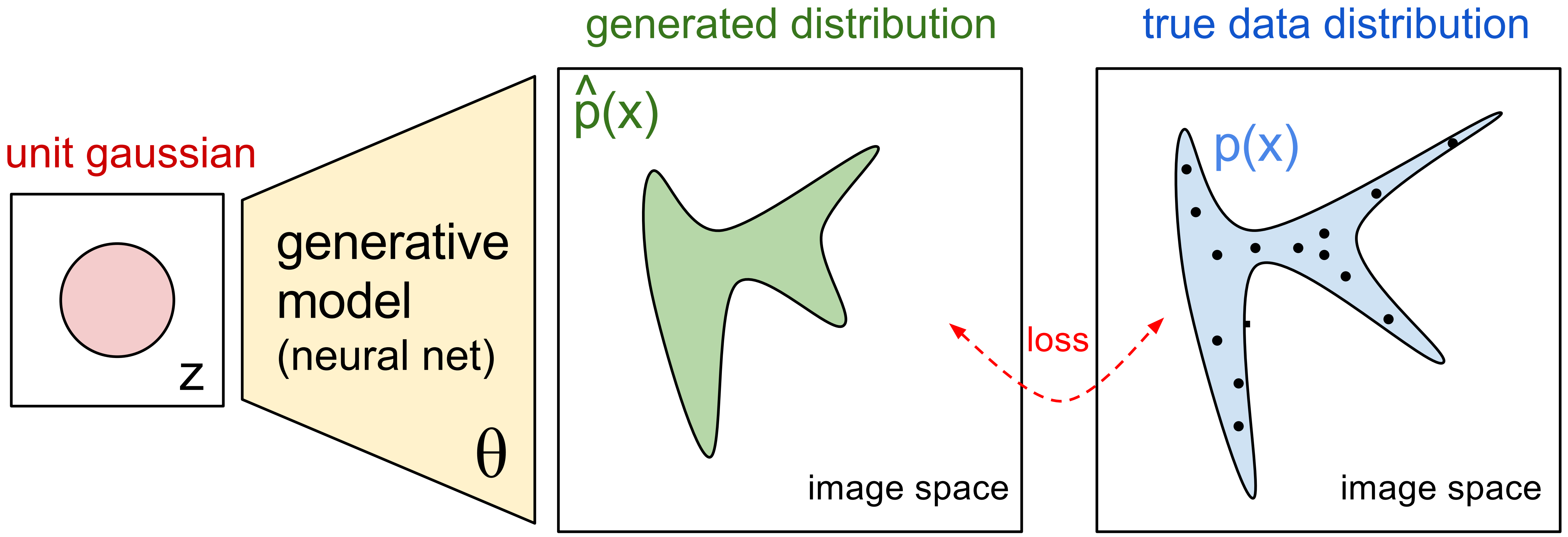
variational probability distribution

Turn the sampling problem to an optimization problem.
Not necessarily easy. But may better leverage deep learning engine.

"mean-field" approaches
Factorized probability
Pairwise interaction

Deep generative models

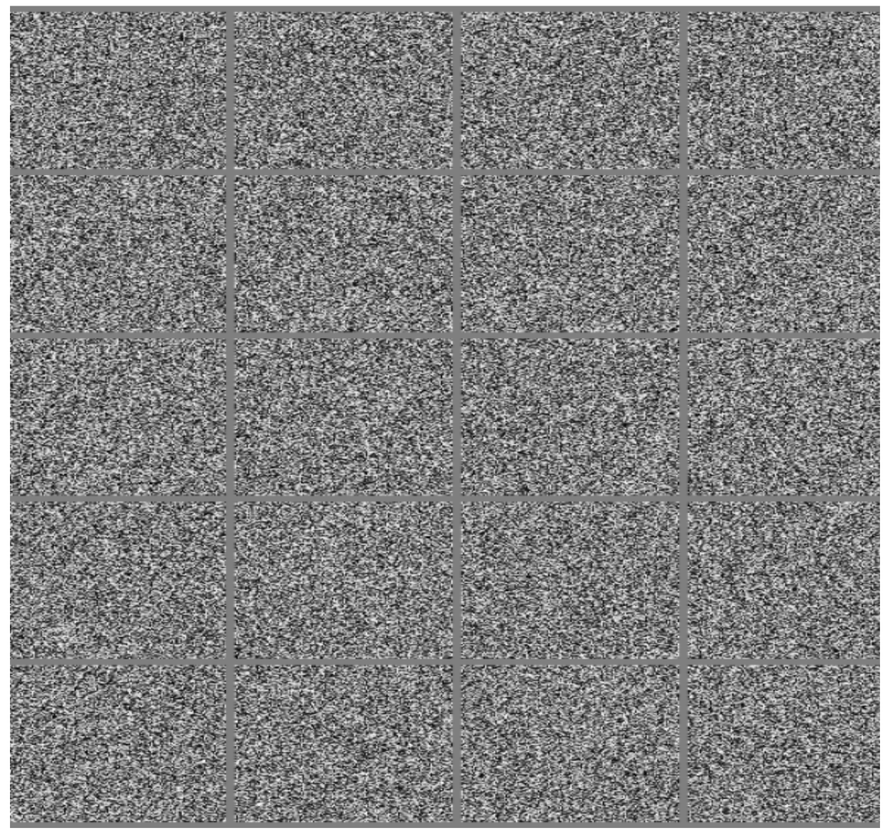
normalizing flow Li, LW, PRL '18
autoregressive model, Wu, LW, Zhang, PRL '19

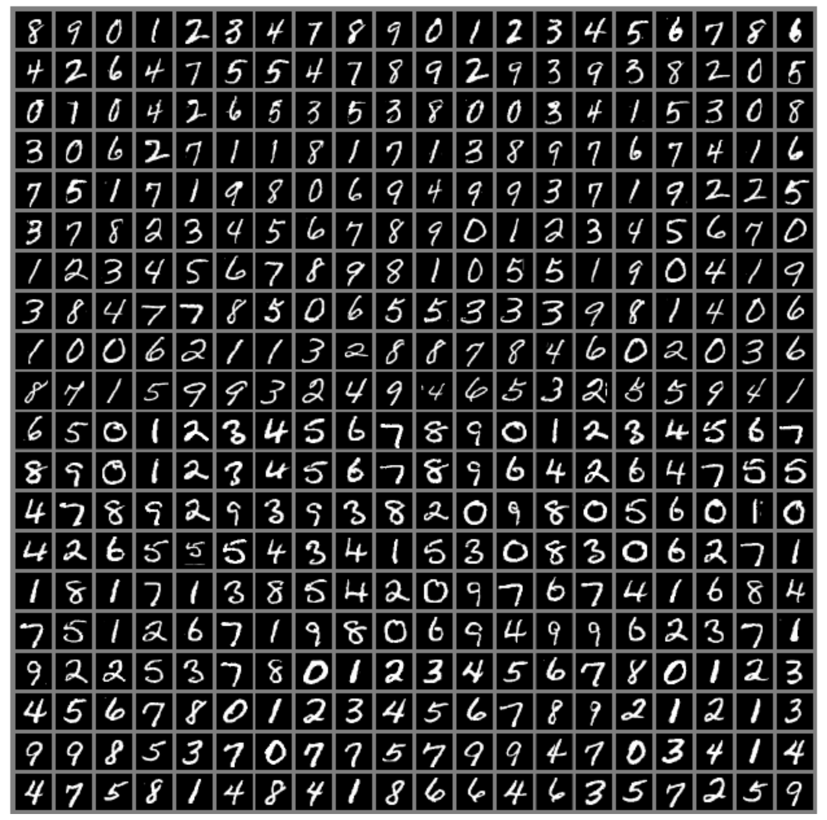


Probabilistic Generative Modeling

$$p(\boldsymbol{x})$$

How to express, learn, and sample from a high-dimensional probability distribution ?



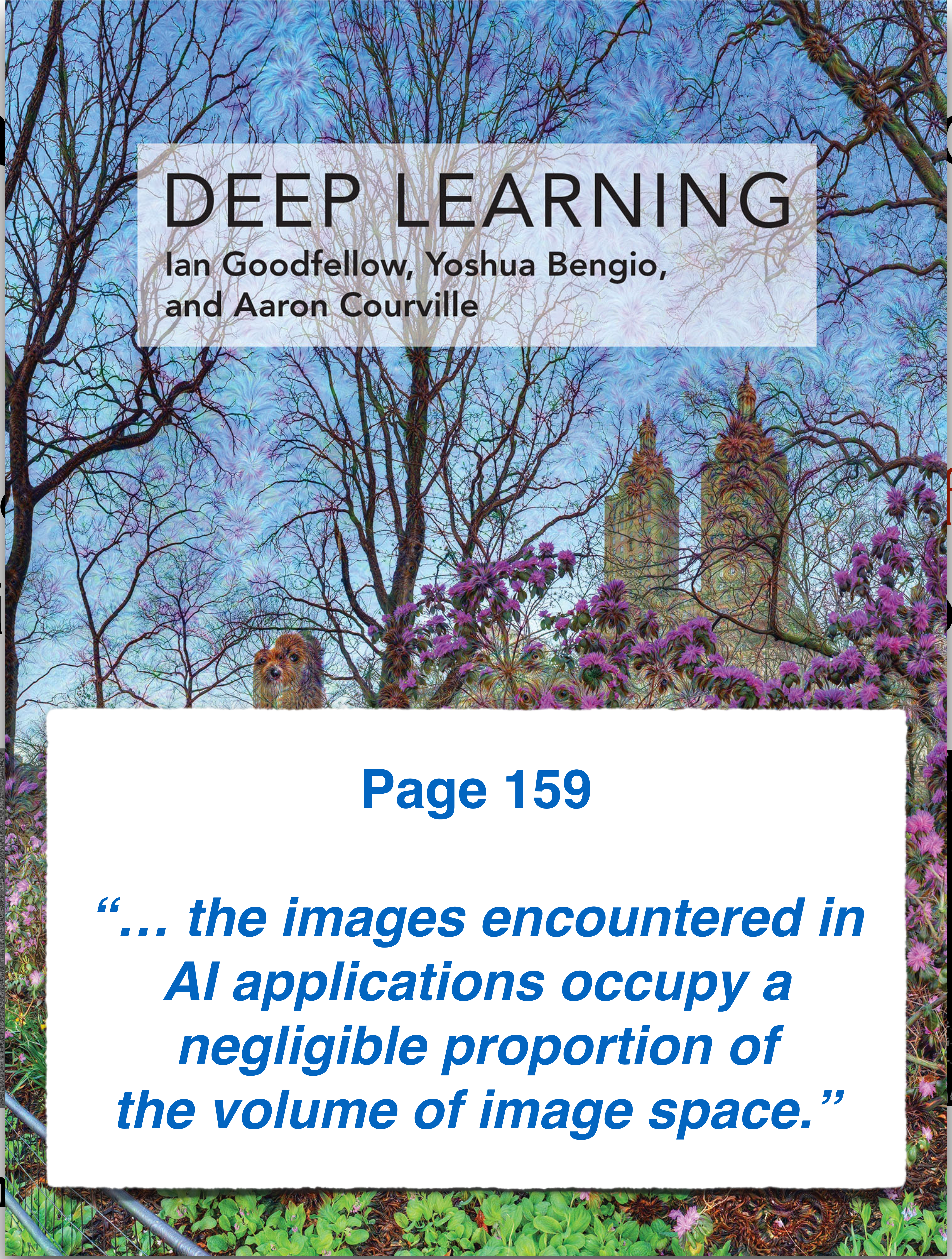


“random” images

“natural” images

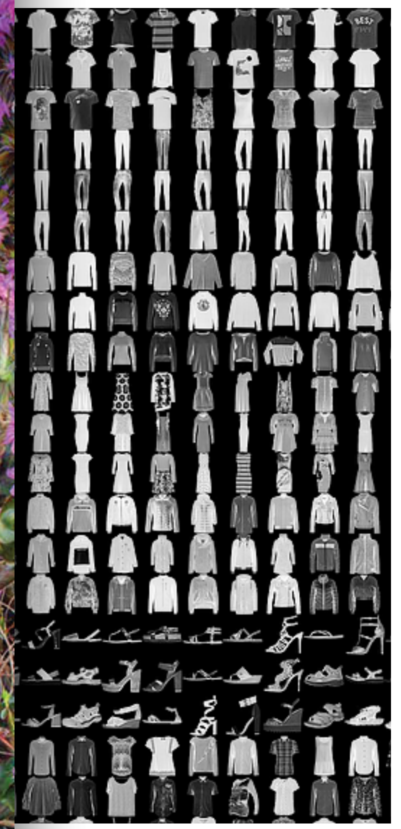
Proba odeling

How t from a
high-di ution ?



Page 159

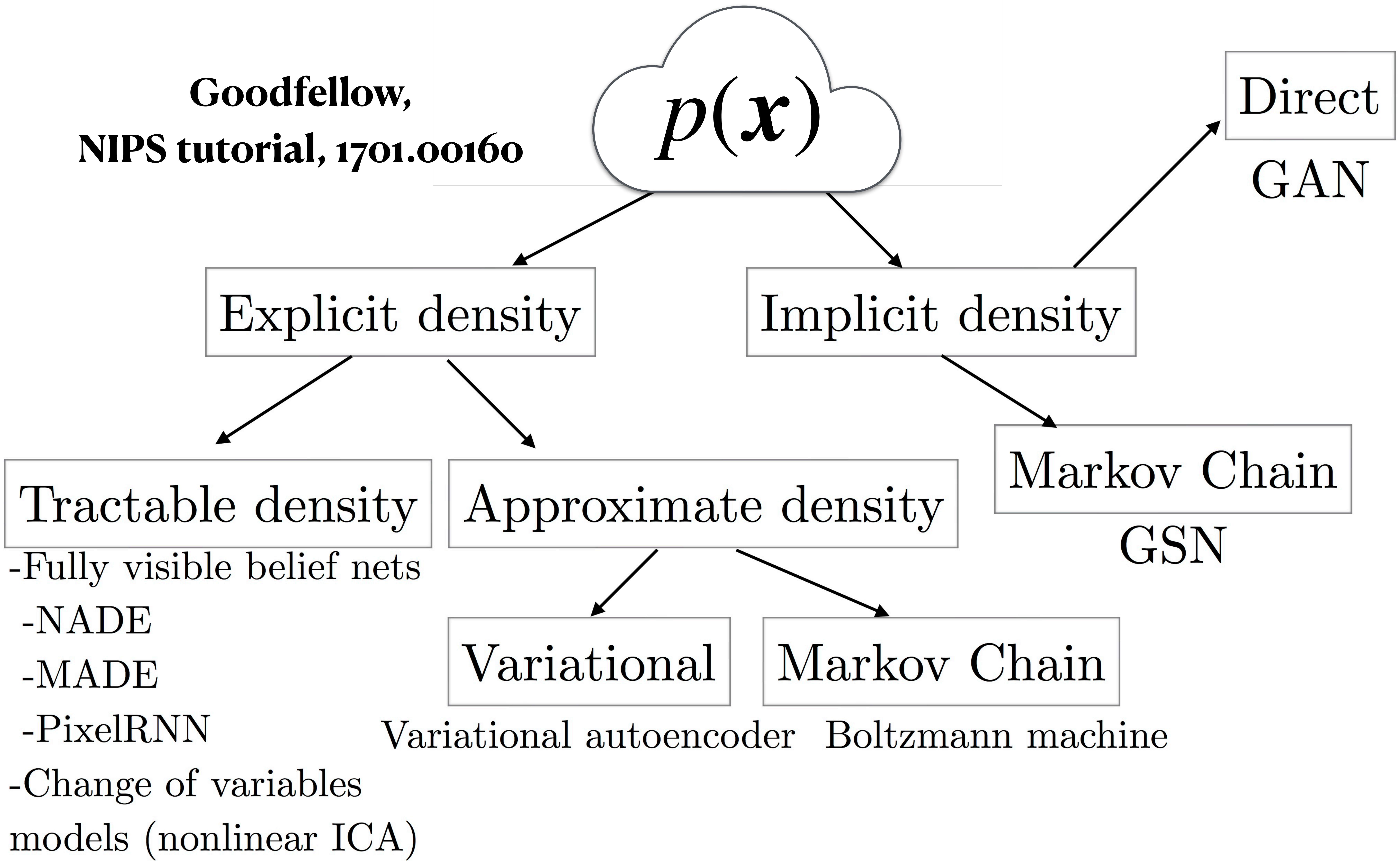
"… the images encountered in AI applications occupy a negligible proportion of the volume of image space."



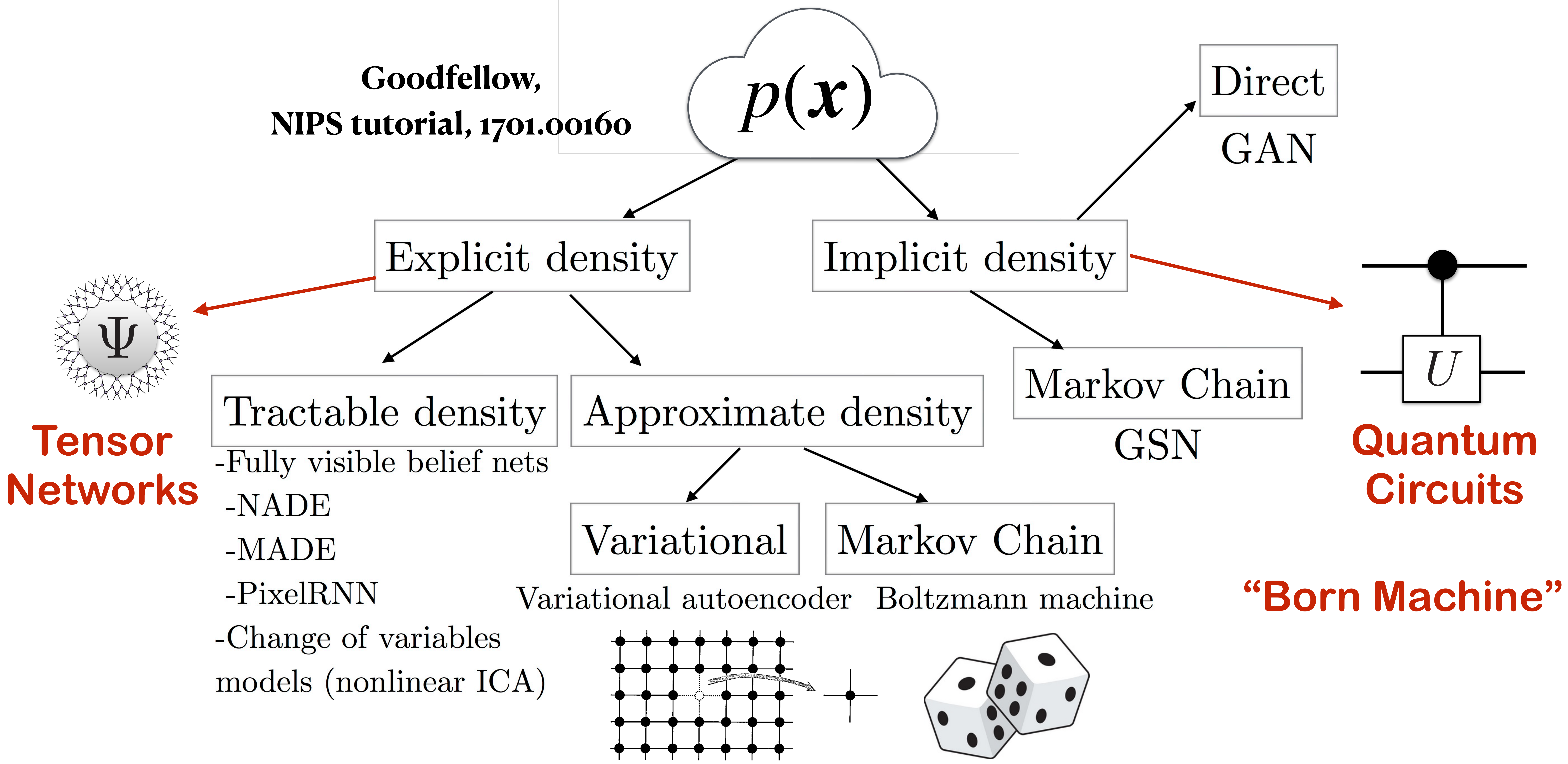
"random

Generative models and their physics genes

**Goodfellow,
NIPS tutorial, 1701.00160**



Generative models and their physics genes

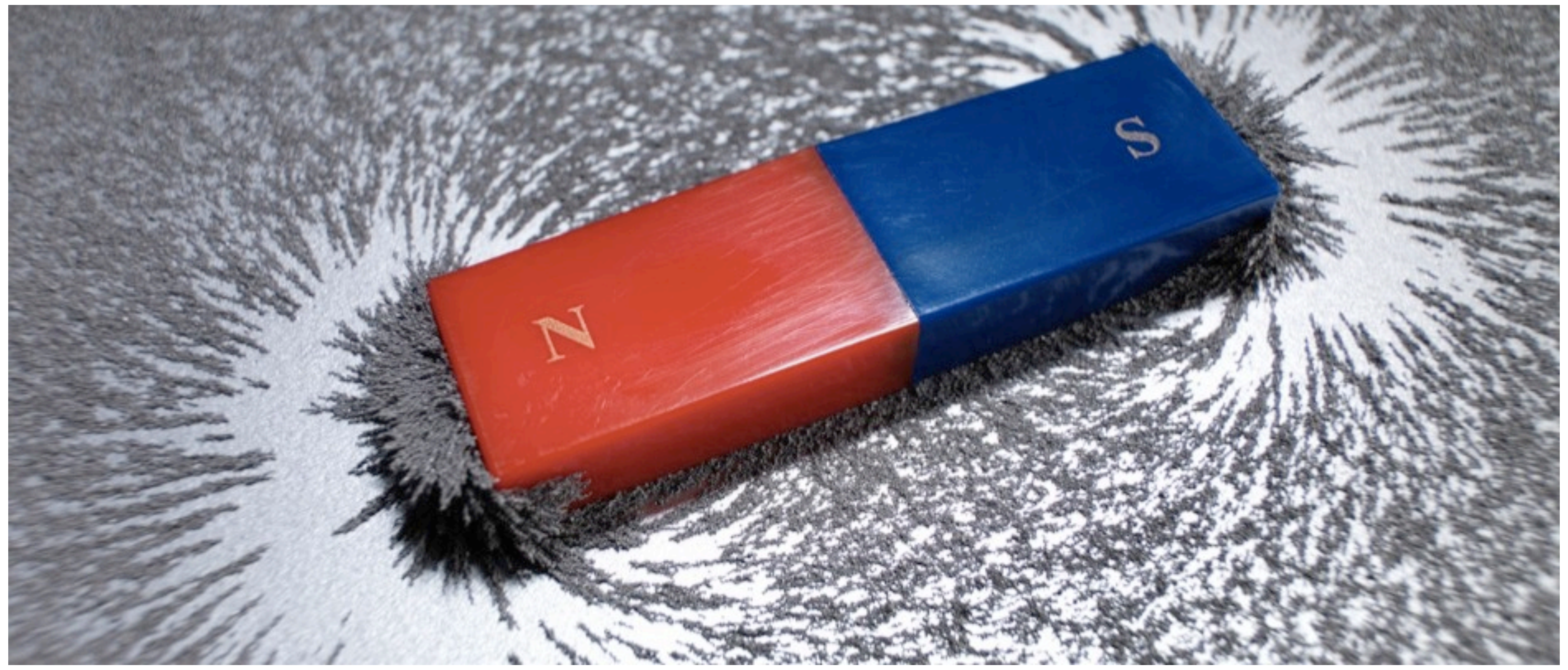


Generative modeling

Known: samples

Unknown: generating distribution

Statistical Physics



Known: energy function

Unknown: samples, partition function

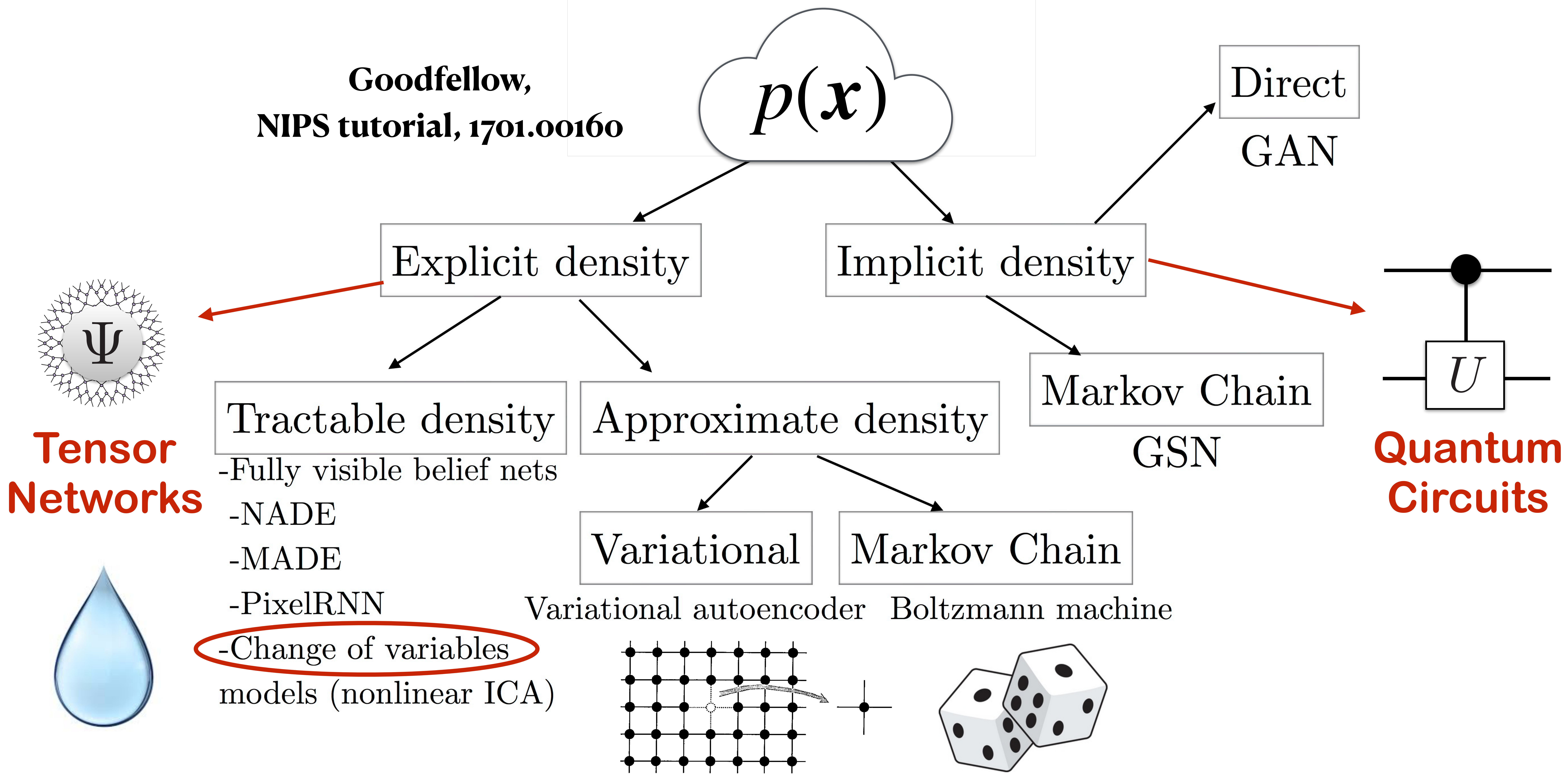
Modern generative models for physics

Physics of and for generative modeling

Generative models and their physics genes

Goodfellow, NIPS tutorial, 1701.00160

Generative models and their physics genes



Generative modeling with normalizing flows

Wavenet 1609.03499 1711.10433

https://deepmind.com/research/case-studies/wavenet

Glow 1807.03039

https://blog.openai.com/glow/

Generative modeling with normalizing flows

Wavenet 1609.03499 1711.10433

Glow 1807.03039

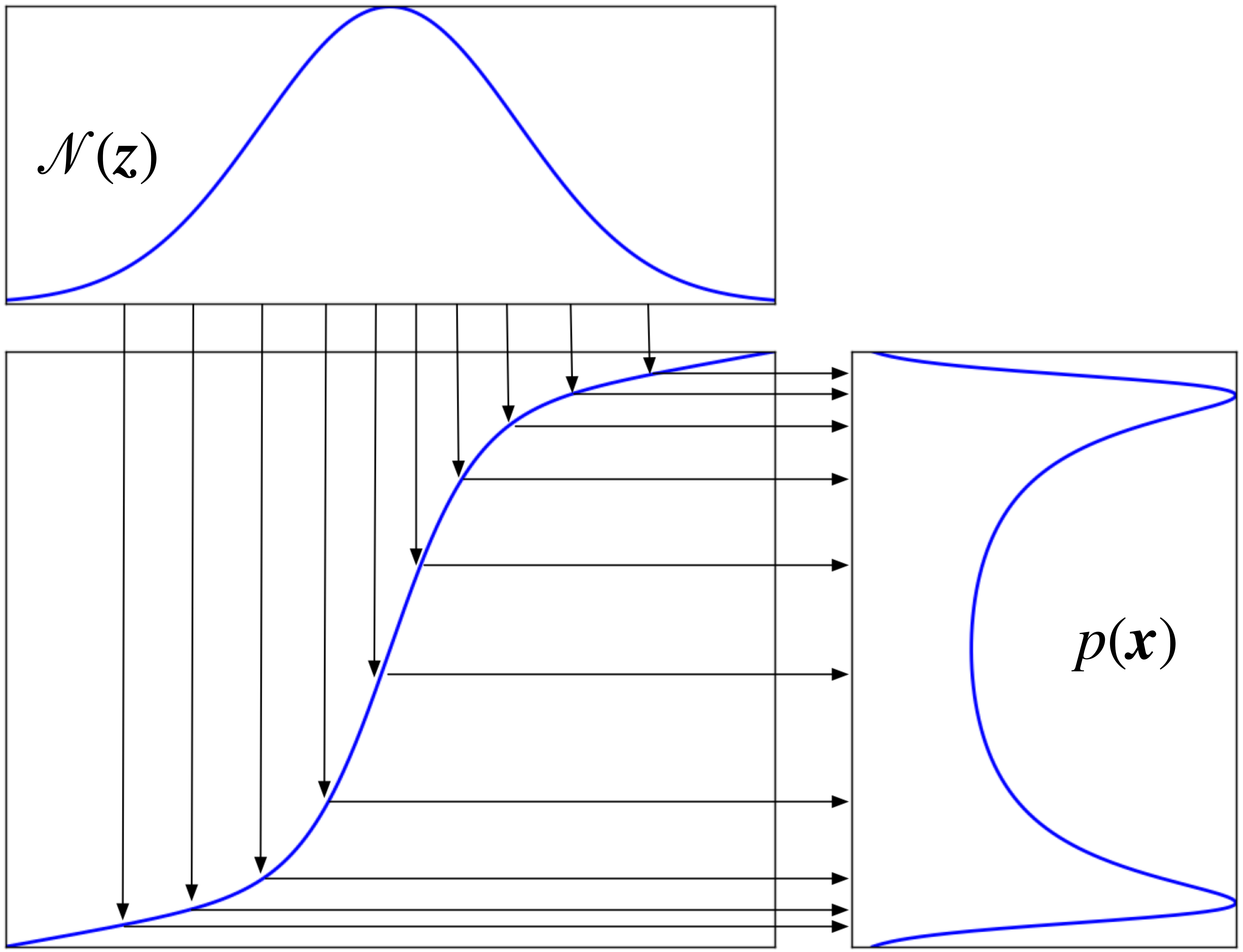
https://deepmind.com/research/case-studies/wavenet

https://blog.openai.com/glow/

Normalizing flow in a nutshell

latent space

"neural net" with 1 neuron



physical space

Normalizing Flows

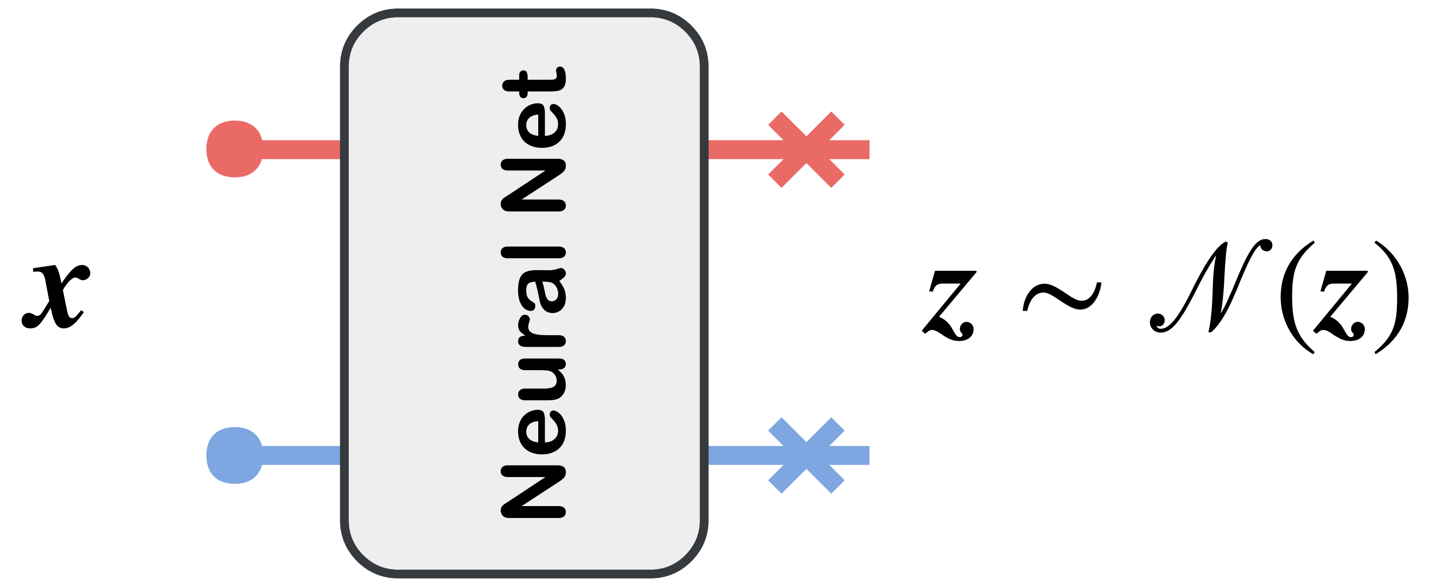
Change of variables $x \leftrightarrow z$ with deep neural nets

$$p(\boldsymbol{x}) = \mathcal{N}(z) \left| \det \left(\frac{\partial z}{\partial \boldsymbol{x}} \right) \right|$$

Review article

1912.02762

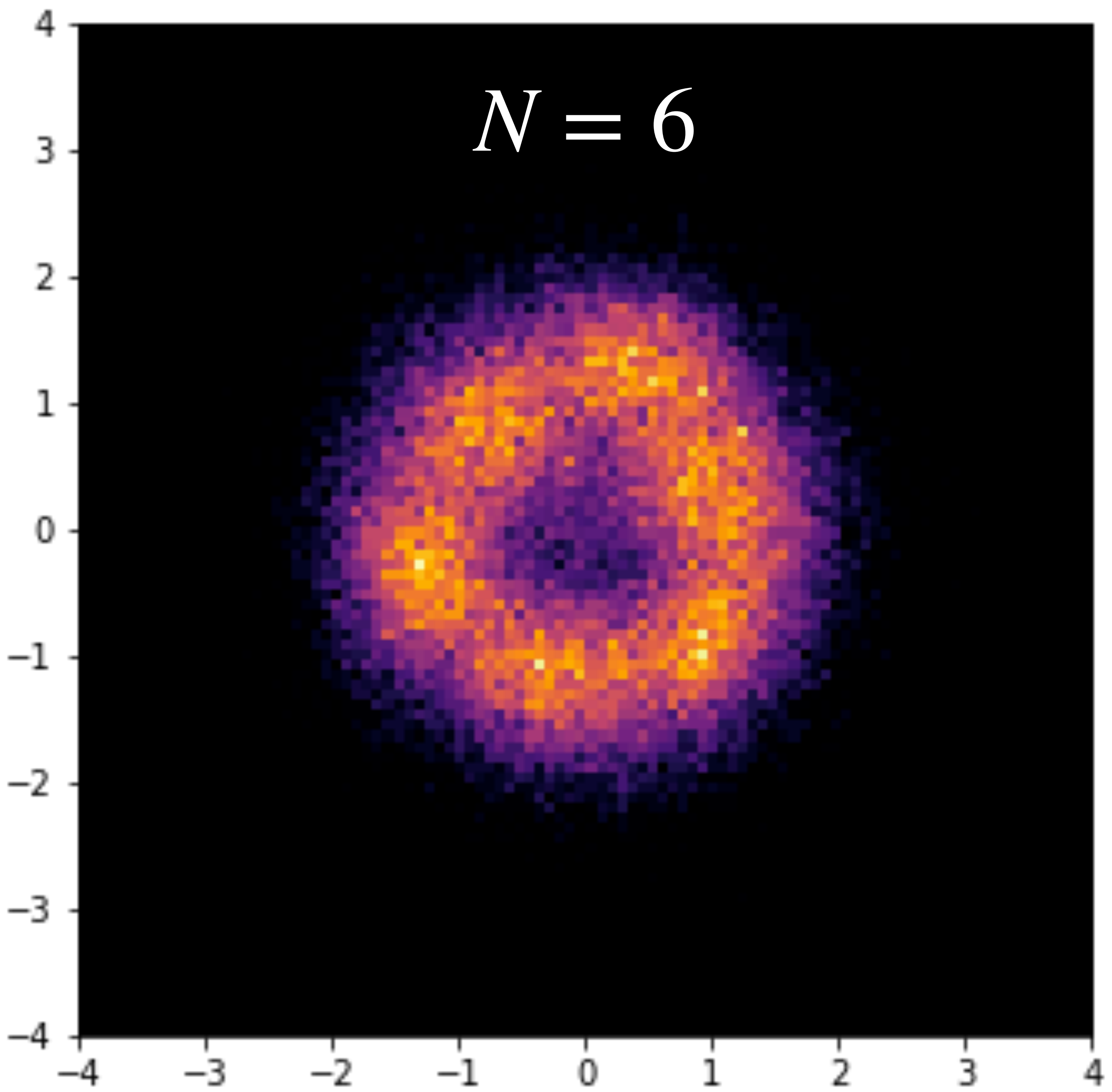
composable, differentiable, and invertible mapping between manifolds

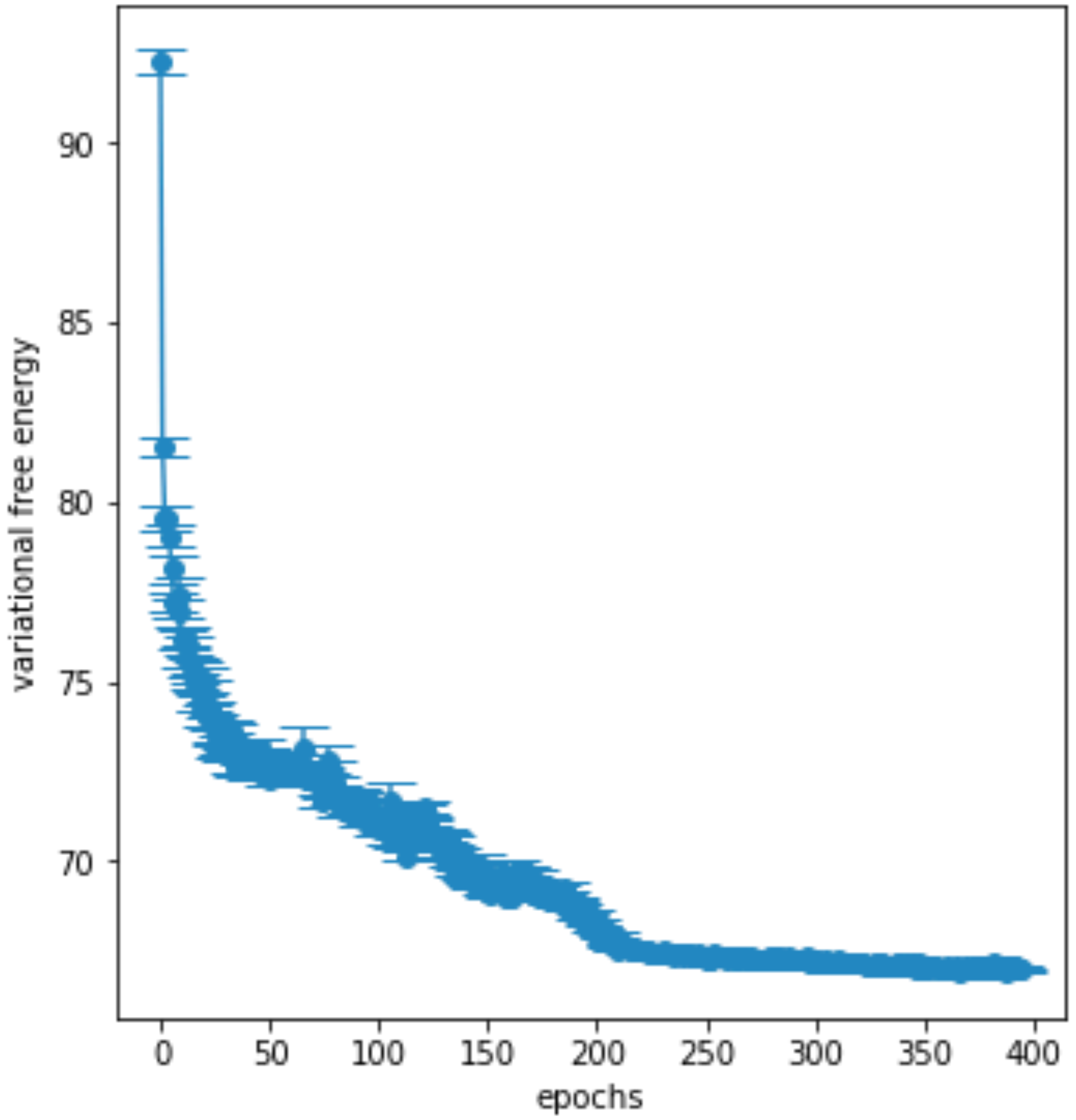


Learn probability transformations with normalizing flows

Classical Coulomb gas in a parabolic trap

$$H = \sum_{i<j} \frac{1}{|\boldsymbol{x}_i - \boldsymbol{x}_j|} + \sum_{i}^{N} \frac{\boldsymbol{x}_i^2}{2}$$

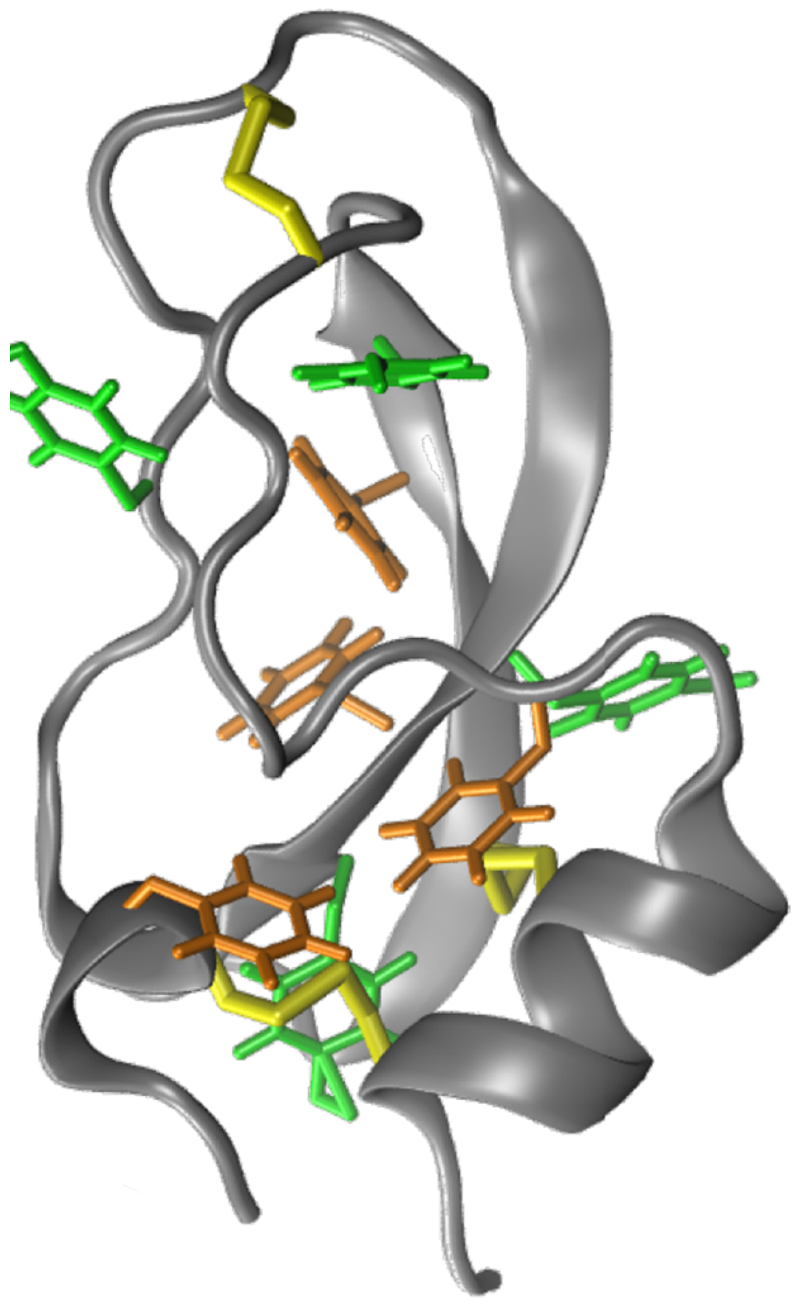




https://colab.research.google.com/drive/13wvsGtV4eTN4v6sXLf2Z52FCyCd6wQZP?usp=sharing

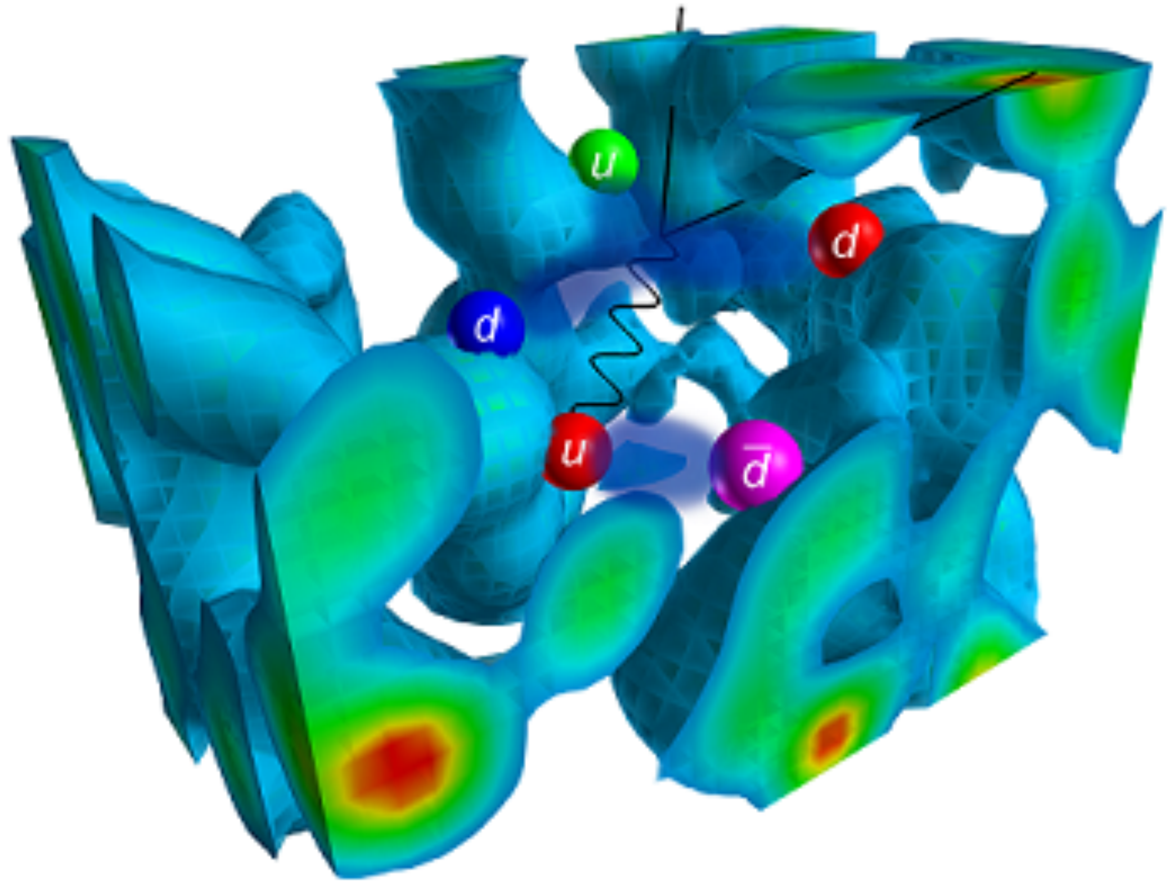
Other scientific applications of flow

Molecular simulation



Noe et al, Science '19

Lattice field theory



Kanwar et al, PRL '20

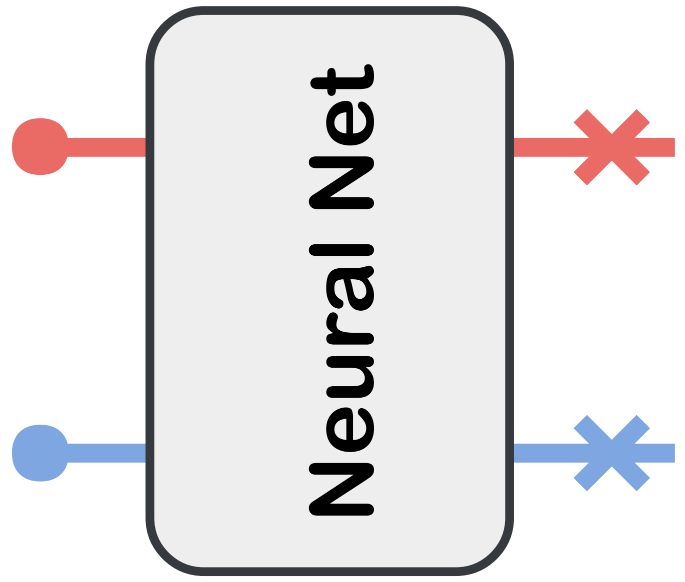
Two training approaches

Density estimation

"learn from data"

$$\mathscr{L} = -\mathbb{E}_{\boldsymbol{x}\sim\text{dataset}}\left[\ln p(\boldsymbol{x})\right]$$

$\boldsymbol{x} \sim \text{dataset}$



z

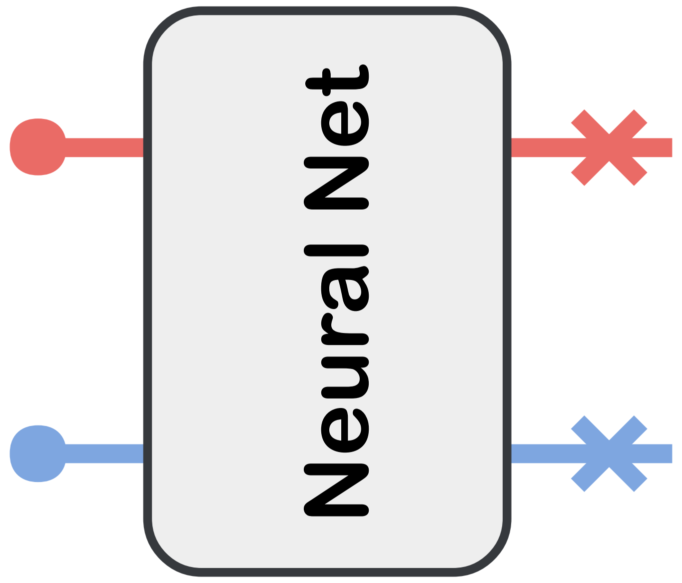
Samples from the given dataset

Variational calculation

"learn from Hamiltonian"

$$\mathscr{L} = \int d\boldsymbol{x}\, p(\boldsymbol{x})\left[\ln p(\boldsymbol{x}) + \beta H(\boldsymbol{x})\right]$$

$\boldsymbol{x}$



$z \sim \mathcal{N}(\mathbf{0}, \Sigma)$

Generate samples from the model

Two training approaches

Density estimation

"learn from data"

$$\mathscr{L} = -\mathbb{E}_{\boldsymbol{x}\sim\text{dataset}}\left[\ln p(\boldsymbol{x})\right]$$

$$\mathbb{KL}(\pi\,||\,p) = \sum_{\boldsymbol{x}} \pi \ln \pi - \underbrace{\sum_{\boldsymbol{x}} \pi \ln p}_{\mathscr{L}}$$

Variational calculation

"learn from Hamiltonian"

$$\mathscr{L} = \int d\boldsymbol{x}\, p(\boldsymbol{x})\left[\ln p(\boldsymbol{x}) + \beta H(\boldsymbol{x})\right]$$

$$\mathscr{L} + \ln Z = \mathbb{KL}\left(p\,||\,\frac{e^{-\beta H}}{Z}\right) \geq 0$$

Now, march into the quantum world

Classical world

Probability distribution p

Kullback-Leibler divergence

$$\mathbb{KL}\left(p\,||\,q\right)$$

Variational free-energy

$$\mathscr{L} = \int d\boldsymbol{x}\, p(\boldsymbol{x})\left[\ln p(\boldsymbol{x}) + \beta H(\boldsymbol{x})\right]$$

Quantum world

Density matrix ρ

Quantum relative entropy

$$S\left(\rho\,||\,\sigma\right)$$

Variational free-energy

$$\mathscr{L} = \mathrm{Tr}(\rho \ln \rho) + \beta \mathrm{Tr}(H\rho)$$

Density matrix

Classical probability

$$0 < \mu_n < 1$$

$$\sum_n \mu_n = 1$$

Quantum state

$$\Psi(\boldsymbol{x}) = \langle \boldsymbol{x} | \Psi \rangle$$

$$\int d\boldsymbol{x} \, |\Psi(\boldsymbol{x})|^2 = 1$$

$$\rho = \sum_n \mu_n |\Psi_n\rangle\langle\Psi_n|$$

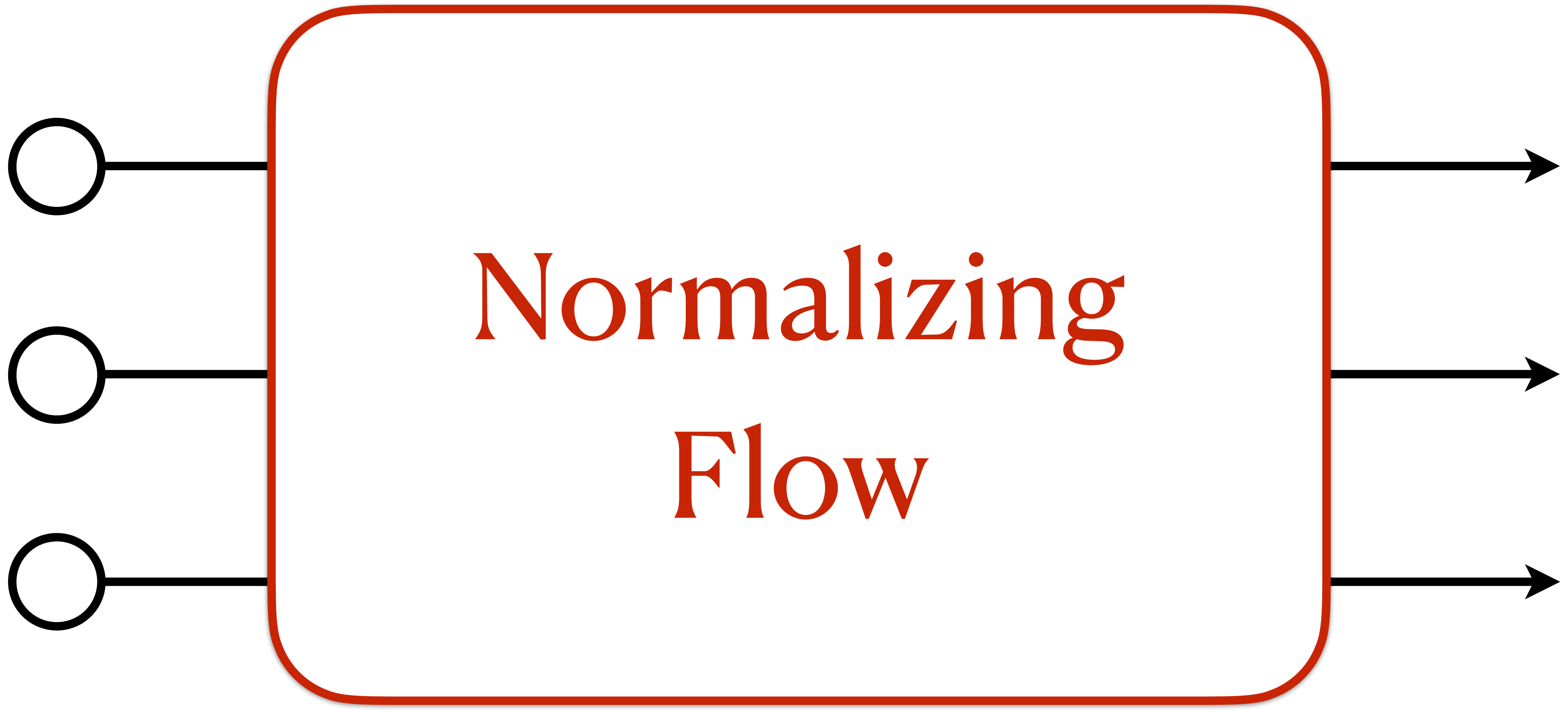
How to represent variational density matrix so it is physical & optimizable ?

$\mathrm{Tr}\rho = 1$ $\quad \rho \succ 0$ $\quad \rho^\dagger = \rho$ $\quad \langle \boldsymbol{x} | \rho | \boldsymbol{x}' \rangle = (-1)^{\mathscr{P}} \langle \mathscr{P}\boldsymbol{x} | \rho | \boldsymbol{x}' \rangle$

Idea: to parametrize a density, think about the transformation

Gaussian variables

$\mathcal{N}(z)$



Correlated variables

$p(\boldsymbol{x})$

$$\mathscr{L} = \int d\boldsymbol{x}\, p(\boldsymbol{x}) \left[\ln p(\boldsymbol{x}) + \beta H(\boldsymbol{x}) \right]$$

Imposing constraints to the transformation for physical densities

We need unitary transformations

ρ_0

$\rho = U\rho_0 U^\dagger$

$\mathrm{Tr}\rho = 1 \quad \rho \succ 0 \quad \rho^\dagger = \rho$ ✓

$$\mathscr{L} = \mathrm{Tr}(\rho \ln \rho) + \beta \mathrm{Tr}(H\rho)$$

How to parametrize and learn unitary transformations ?

PHYSICAL REVIEW VOLUME 85, NUMBER 4 FEBRUARY 15, 1952

Point Transformations in Quantum Mechanics

BRYCE SELIGMAN DEWITT*
Ecole d'Eté de Physique Théorique de l'Université de Grenoble, Les Houches, Haute Savoie, France

(Received September 14, 1951)

An isomorphism is shown to exist between the group of point transformations in classical mechanics and a certain subgroup of the group of all unitary transformations in quantum mechanics. This isomorphism is

The unitary representations of the point-transformation group may be obtained by determining the infinitesimal generators of the group. An infinitesimal point transformation may be expressed in the form

$$x'^i = x^i + \epsilon\Lambda^i(x), \tag{3.7}$$

$$p_i' = p_i - \tfrac{1}{2}\epsilon[(\partial/\partial x^i)\Lambda^j(x), p_j]_+, \tag{3.8}$$

unitary

generator

$S = \tfrac{1}{2}[\Lambda^i(x), p_i]_+.$

Unitary representation of coordinate transformation

Canonical transformations

Classical world

Symplectic

$$(\dot{\boldsymbol{x}}, \dot{\boldsymbol{p}}) = \nabla_{(\boldsymbol{x},\boldsymbol{p})} G(\boldsymbol{x}, \boldsymbol{p}) \begin{pmatrix} & -\mathbb{1} \\ \mathbb{1} & \end{pmatrix}$$

point transformation

$$G(\boldsymbol{x}, \boldsymbol{p}) = \boldsymbol{v}(\boldsymbol{x}) \cdot \boldsymbol{p}$$

Neural canonical transformations, Li, Dong, Zhang, LW, PRX '20

Quantum world

Unitary

$$(\dot{\boldsymbol{x}}, \dot{\boldsymbol{p}}) = i[(\boldsymbol{x}, \boldsymbol{p}), G(\boldsymbol{x}, \boldsymbol{p})]$$

"quantized" point transformation

$$G(\boldsymbol{x}, \boldsymbol{p}) = \frac{1}{2}\{\boldsymbol{v}(\boldsymbol{x}), \boldsymbol{p}\}$$

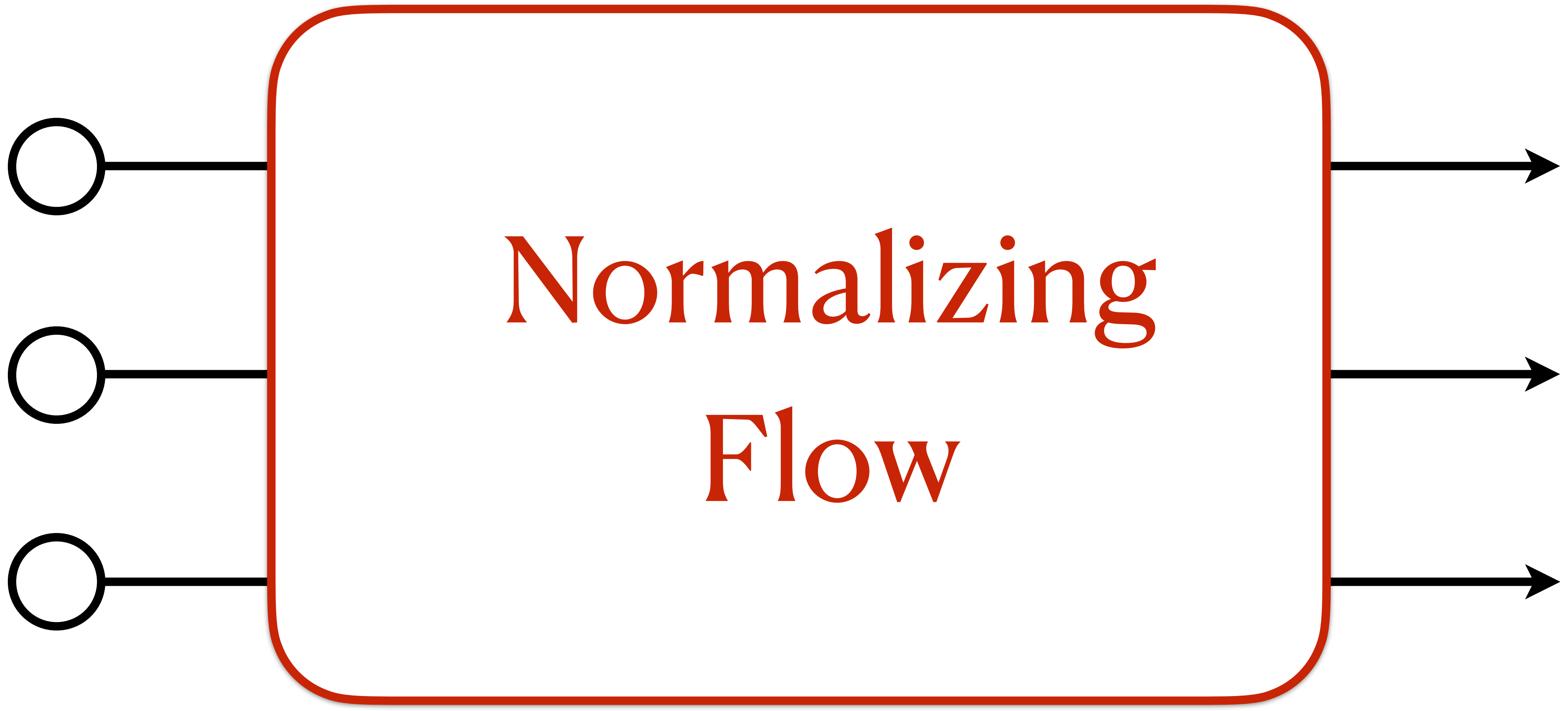
Quantum generalization to fermions, Xie, Zhang, LW, 2105.08644

see also Cranmer et al, 1904.05903

Neural canonical transformation

free electrons
coordinates

ρ_0



transformed
electron coordinates

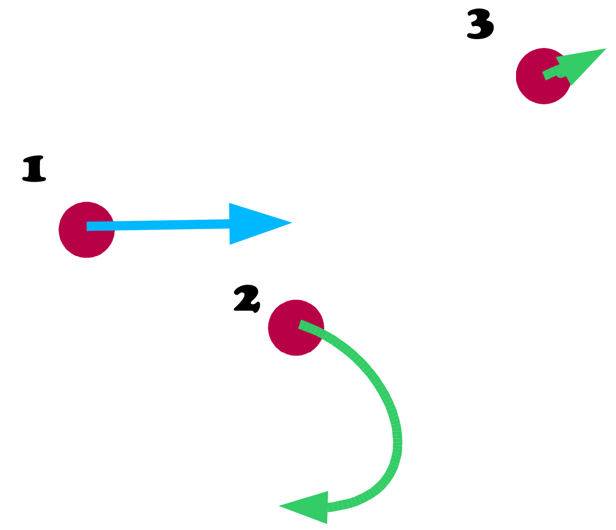
$\rho = U\rho_0 U^\dagger$

$$\mathscr{L} = \mathrm{Tr}(\rho \ln \rho) + \beta \mathrm{Tr}(H\rho)$$

Moreover, the flow should be permutation-equavariant
to preserve fermionic statistics $\langle \boldsymbol{x} | \rho | \boldsymbol{x}' \rangle = (-1)^{\mathscr{P}} \langle \mathscr{P}\boldsymbol{x} | \rho | \boldsymbol{x}' \rangle$

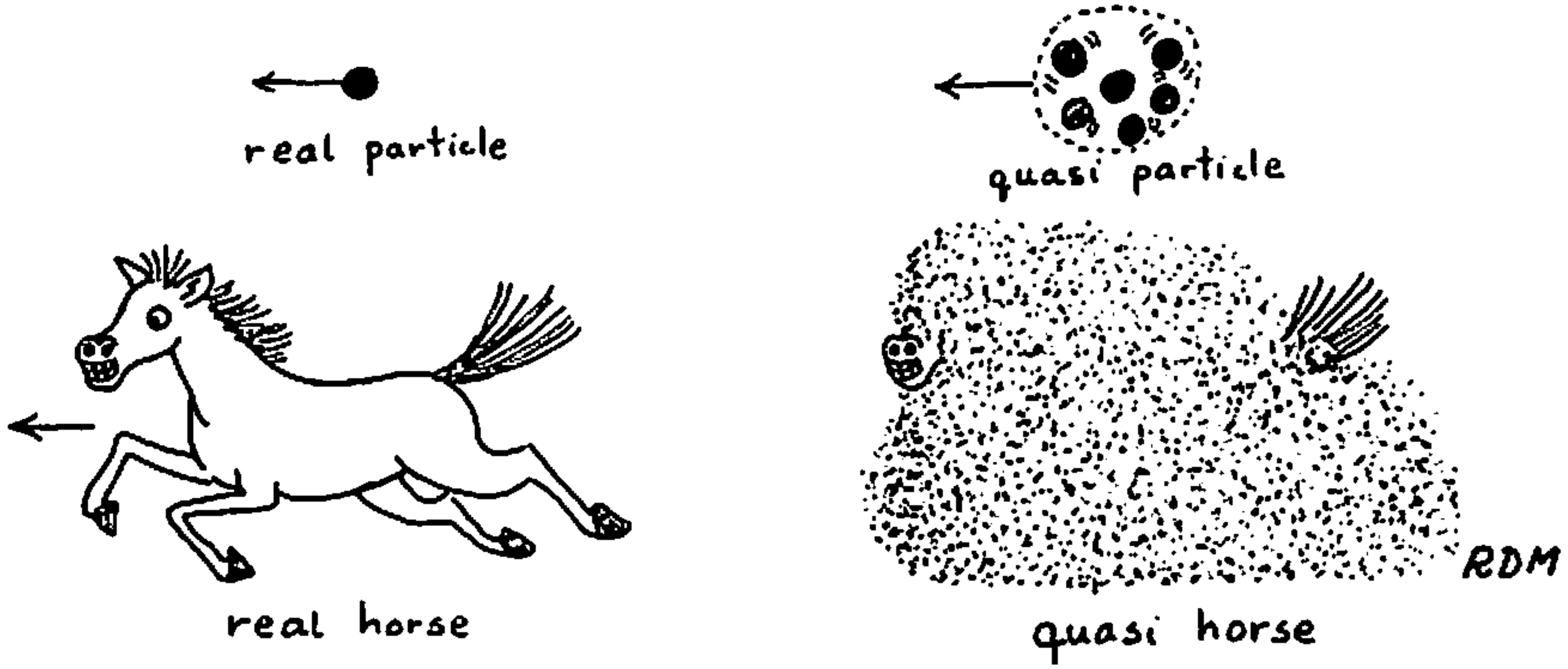
Backflow transformation

Collective coordinates



$$x_i' = x_i + \sum_{j \neq i} \eta(|x_i - x_j|)(x_j - x_i)$$

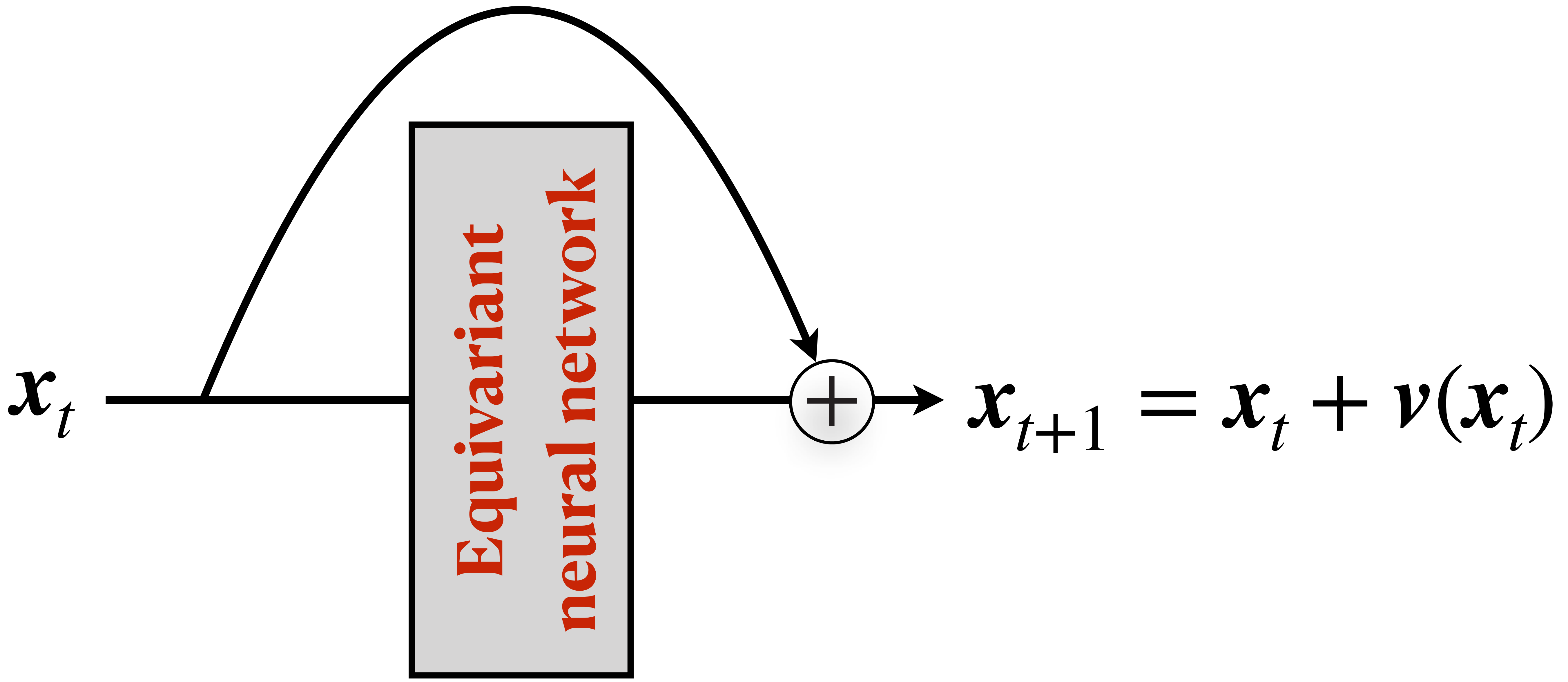
Wigner & Seitz 1934, Feynman 1954, …



$\Psi(\boldsymbol{x})$: independent particles $\quad\Psi(\boldsymbol{x}')$: interacting particles

Nowadays, view it as a residual network, or, discretization of a flow

Neural backflow transformation



$$\boldsymbol{v}(\mathscr{P}\boldsymbol{x}) = \mathscr{P}\boldsymbol{v}(\boldsymbol{x})$$

deep set, transformer, ...

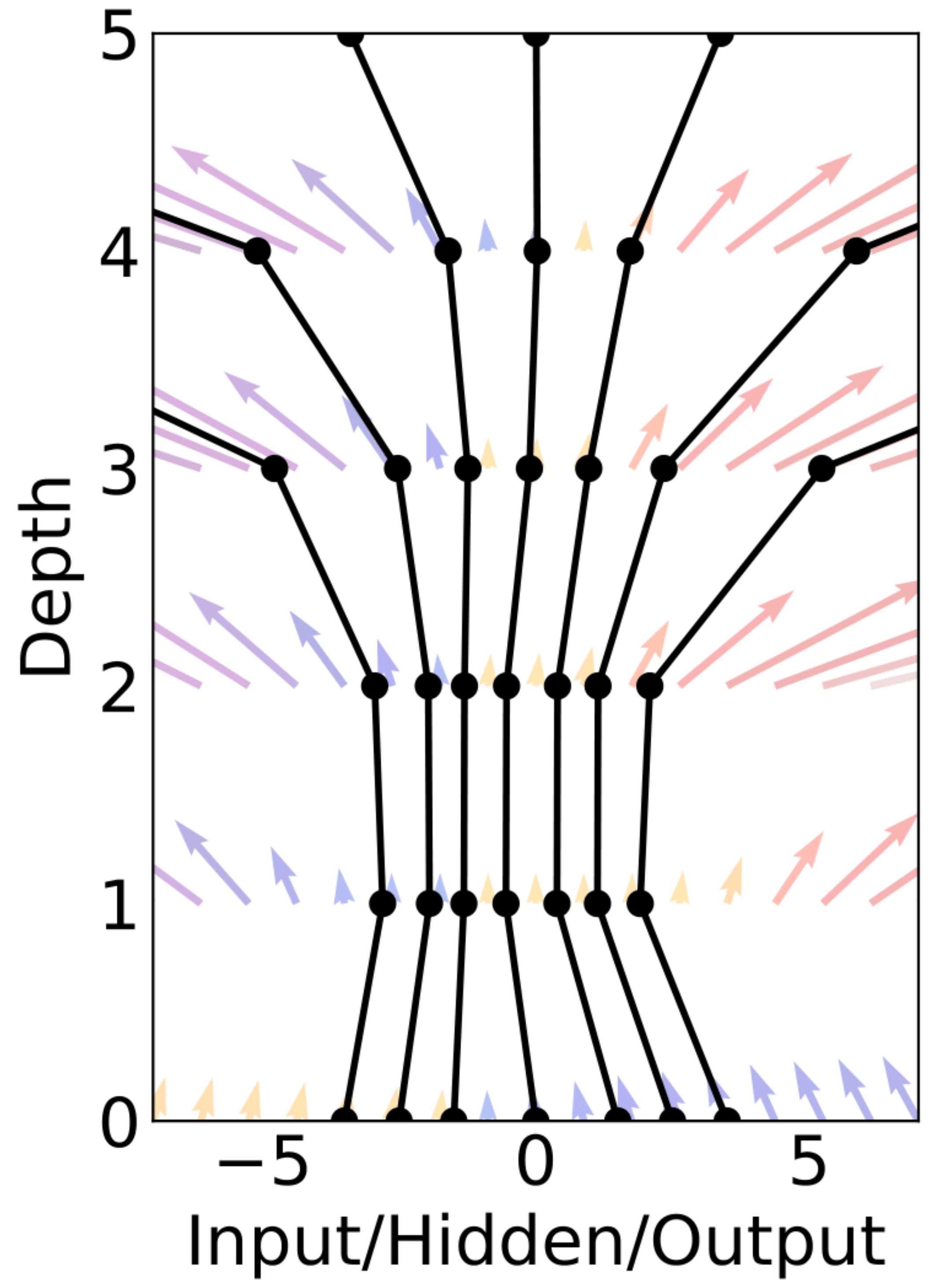
Pfau et al, 1909.02487
Hermann et al, 1909.08423

Köhler et al 1910.00753
Wirnsberger et al, 2002.04913

Li et al 2008.02676
Biloš et al 2010.03242

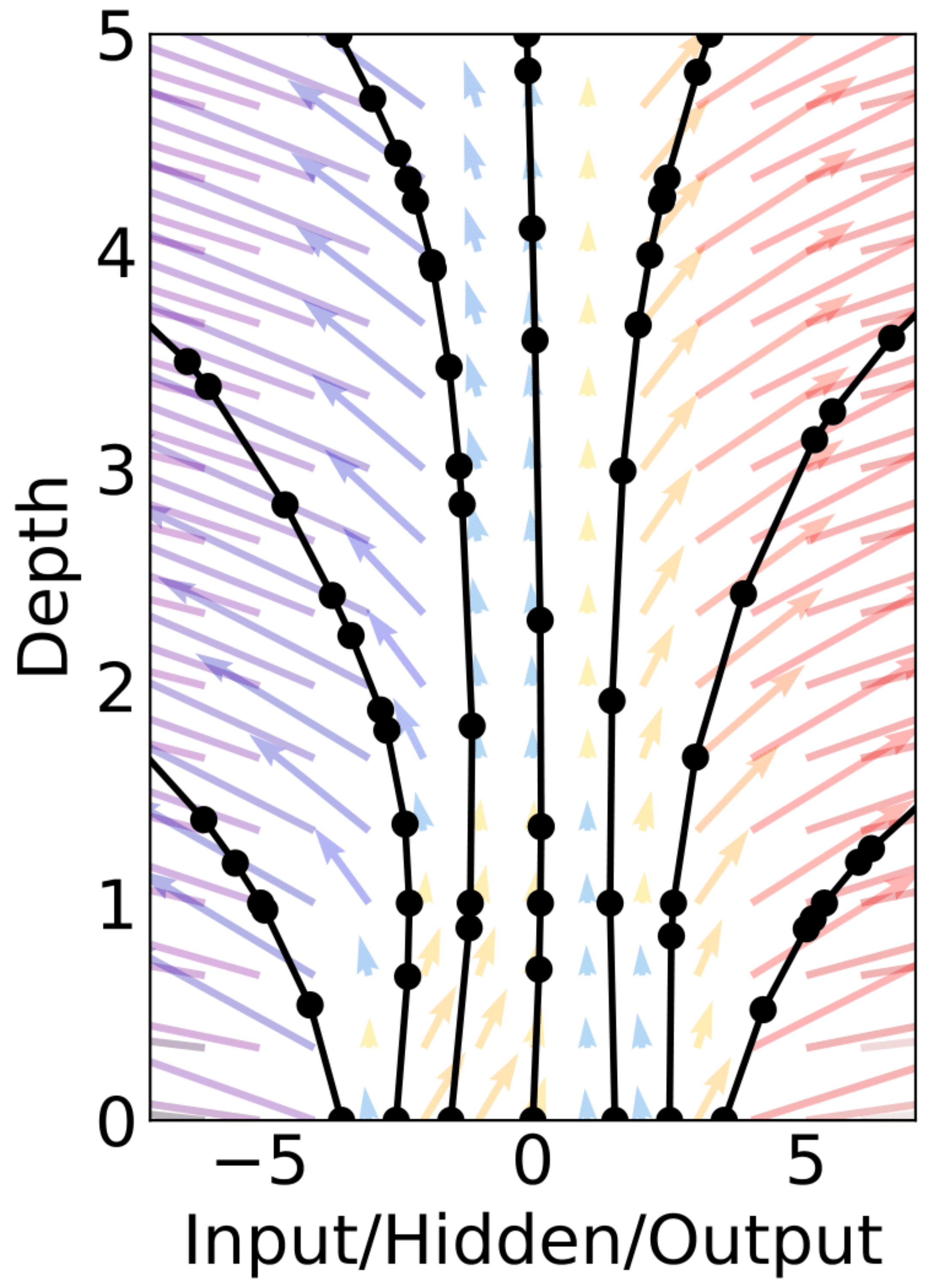
Continuous neural backflow transformation

Residual network



$$x_{t+1} = x_t + f(x_t)$$

Neural ODE



$$dx/dt = f(x)$$

Chen et al, 1806.07366

Harbor el al 1705.03341
Lu et al 1710.10121,
E Commun. Math. Stat 17'...

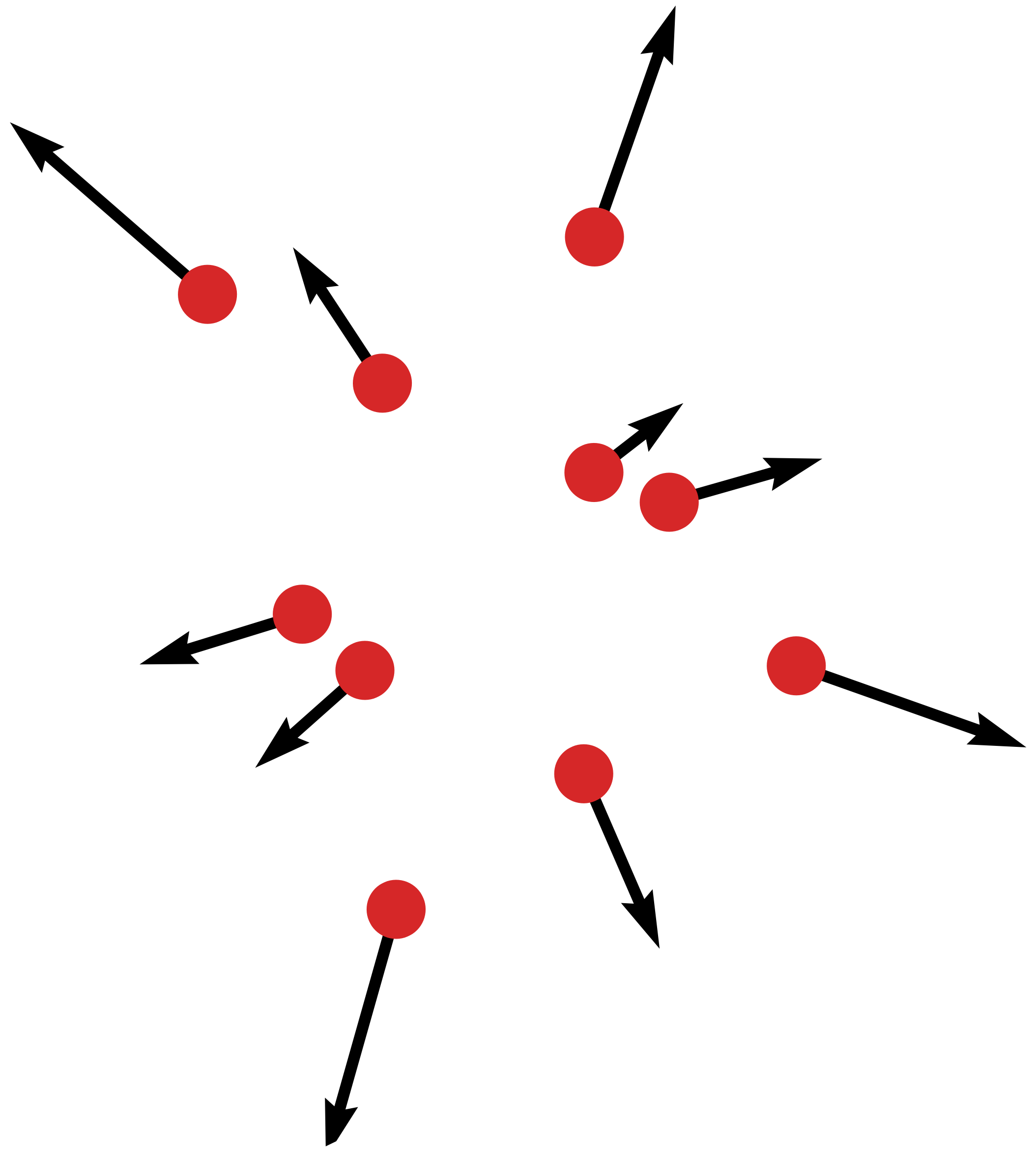
Continuous unitary transformation as a flow

$$i\frac{\partial}{\partial t}|\Psi_n\rangle = G|\Psi_n\rangle$$

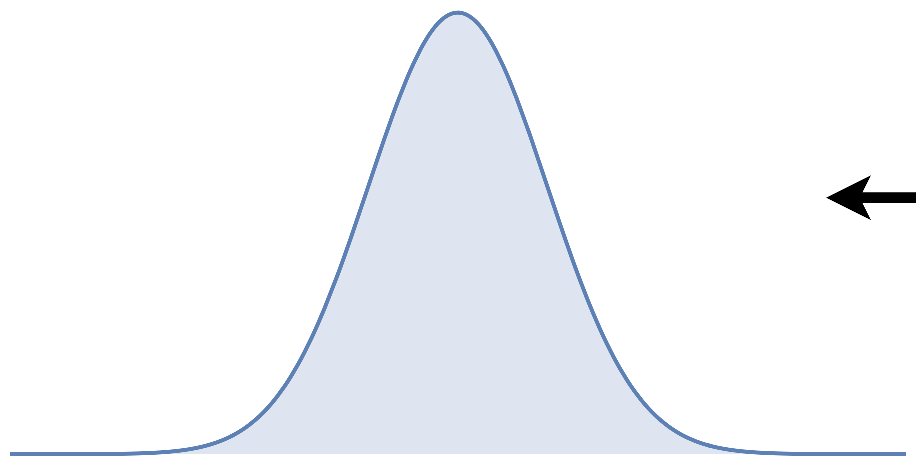
$$\frac{-i}{2}\{\boldsymbol{v}, \nabla\}$$ Dewitt 1952

$$\frac{\partial \left|\Psi_n(\boldsymbol{x},t)\right|^2}{\partial t} + \nabla\cdot\left(|\Psi_n(\boldsymbol{x},t)|^2\boldsymbol{v}\right) = 0$$

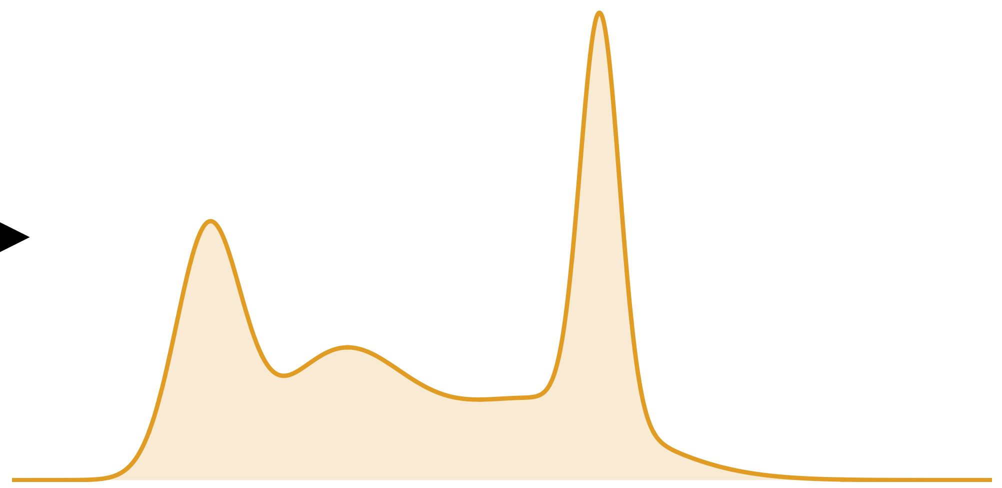
$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{v}$$



$$\frac{d\ln|\Psi_n(\boldsymbol{x},t)|^2}{dt} = -\nabla\cdot\boldsymbol{v}$$

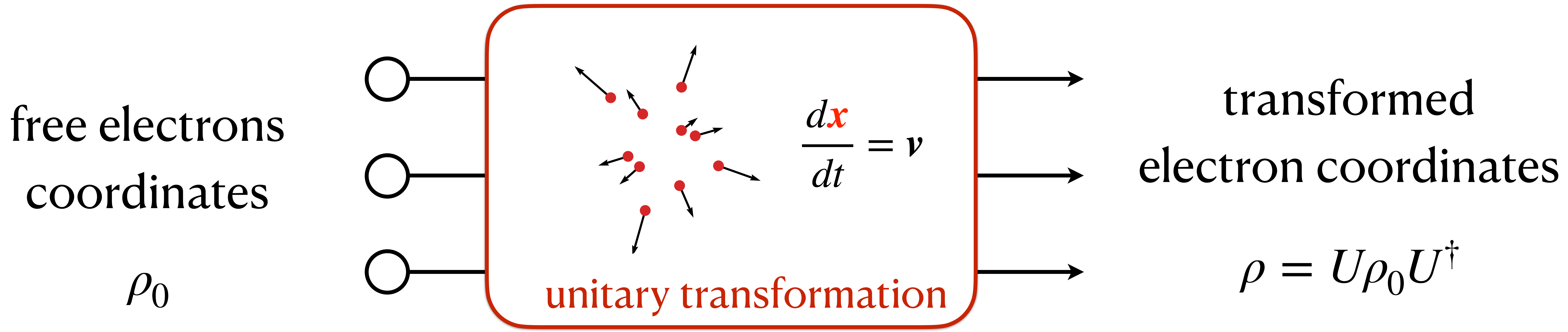


Simple density



Complex density

FermiFlow: Equivariant flow of fermions



ML technique: Perm-equavariant normalizing flow/Neural ODE/Invertible ResNet

Physical picture: Variational approximation of adiabatic preparation of thermal equilibrium

Mathematics: Optimal control a PDE with particle method

Details: Objective function

$$\mathcal{L} = \mathop{\mathbb{E}}_{n\sim\mu_n}\left[\ln\mu_n + \beta \mathop{\mathbb{E}}_{\boldsymbol{x}\sim p_n(\boldsymbol{x})}\left[E_n^{\mathrm{loc}}(\boldsymbol{x})\right]\right]$$

Boltzmann distribution

Born rule

Classical distribution
parametrized by μ_n for free electrons

$$0 < \mu_n < 1 \qquad \sum_n \mu_n = 1$$

Discrete probabilistic model
(Softmax, Autoregressive model,...)

Quantum distribution
parametrized by the equivariant drift $\boldsymbol{v}$

$$p_n(\boldsymbol{x}) = |\Psi_n(\boldsymbol{x})|^2$$

Continuous probabilistic model
(Permutation-equivariant continuous flow)

"Local energy"

$$E_n^{\mathrm{loc}}(\boldsymbol{x}) = -\frac{1}{4}\nabla^2 \ln p_n(\boldsymbol{x}) - \frac{1}{8}\left[\nabla \ln p_n(\boldsymbol{x})\right]^2 + \sum_{i<j}\frac{1}{|\boldsymbol{x}_i - \boldsymbol{x}_j|}$$

Details: Gradient estimators

Nested REINFORCE

Classical distribution

$$\nabla_{\boldsymbol{\phi}} \mathscr{L} = \mathop{\mathbb{E}}_{n \sim \mu_n} \left[\left(\ln \mu_n + \beta \mathop{\mathbb{E}}_{\boldsymbol{x} \sim p_n(\boldsymbol{x})} \left[E_n^{\mathrm{loc}}(\boldsymbol{x}) \right] \right) \nabla_{\boldsymbol{\phi}} \ln \mu_n \right]$$

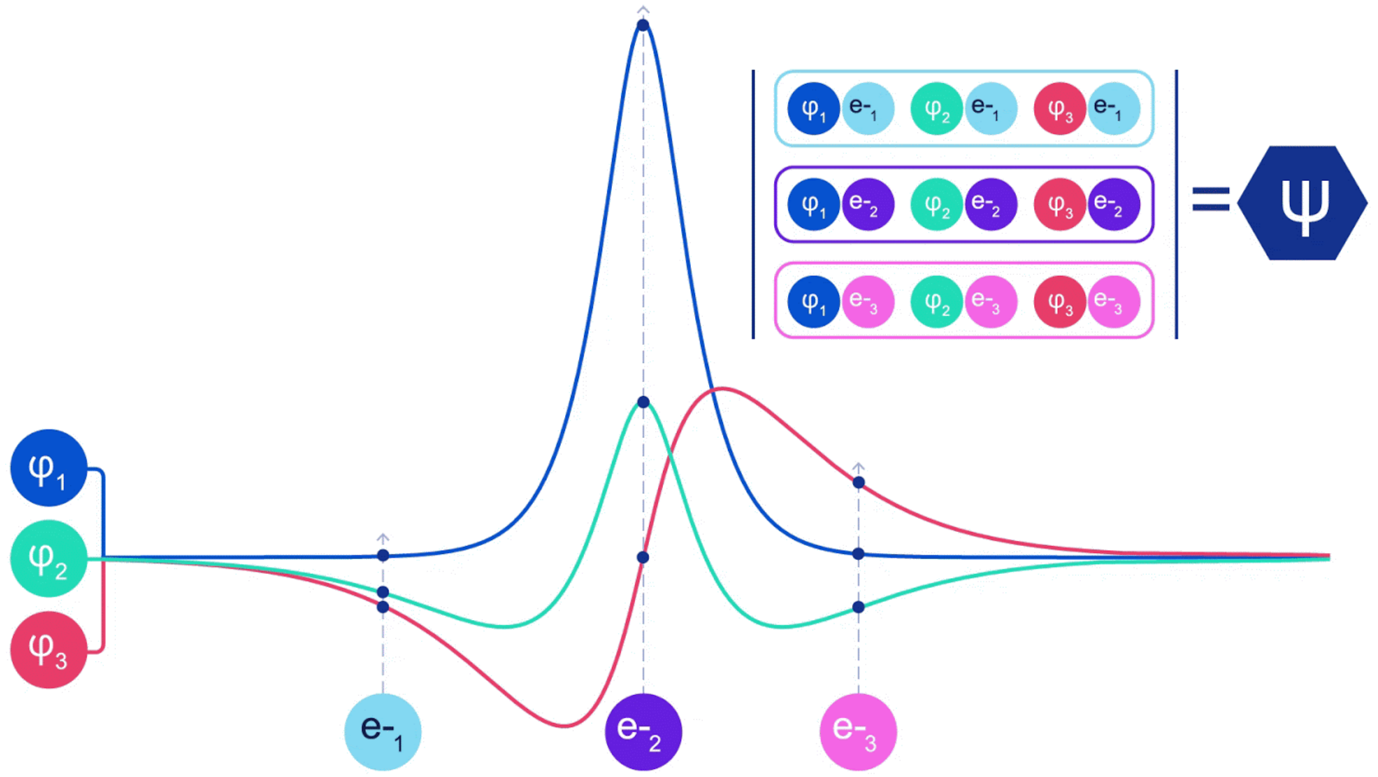
Quantum distribution

$$\nabla_{\boldsymbol{\theta}} \mathscr{L} = \beta \mathop{\mathbb{E}}_{n \sim \mu_n} \mathop{\mathbb{E}}_{\boldsymbol{x} \sim p_n(\boldsymbol{x})} \left[E_n^{\mathrm{loc}}(\boldsymbol{x}) \nabla_{\boldsymbol{\theta}} \ln p_n(\boldsymbol{x}) \right]$$

Variance reduction in both estimators by subtracting baseline

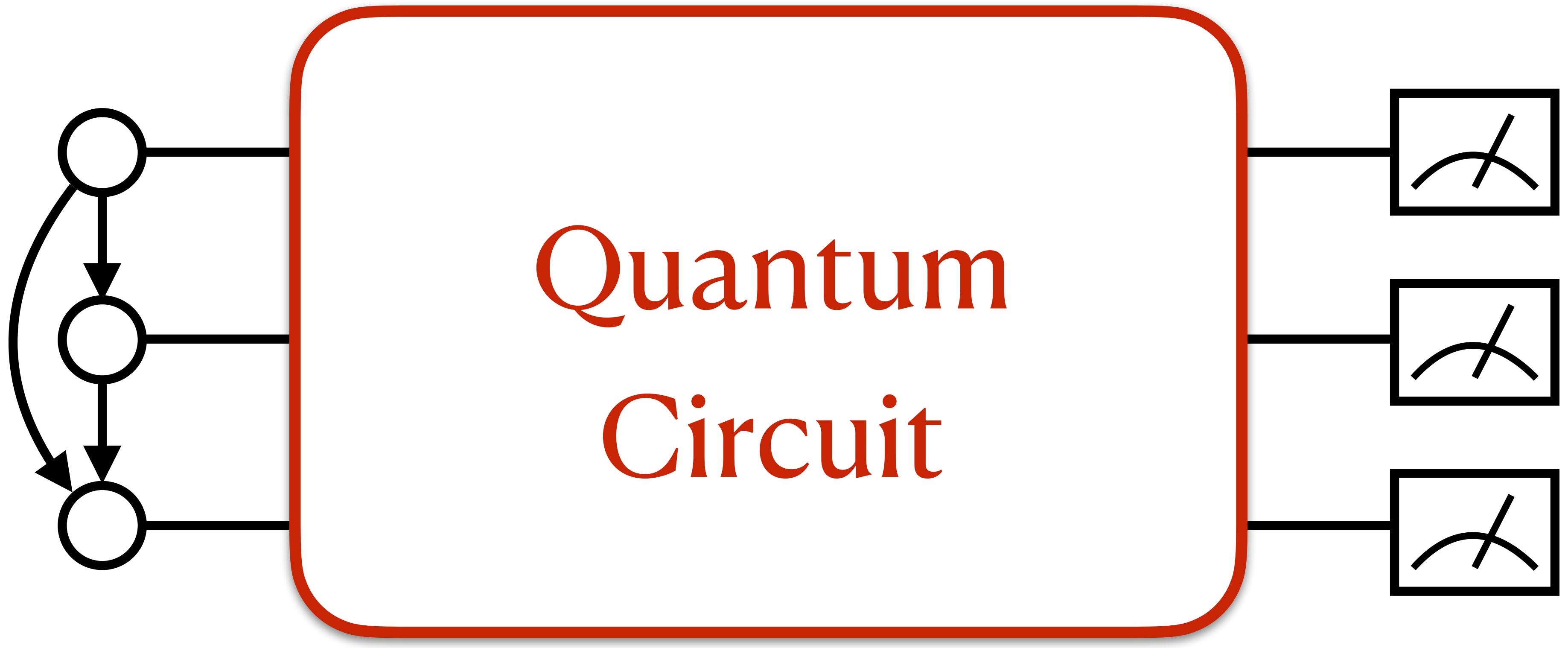
Related works

Neural nets for ground state



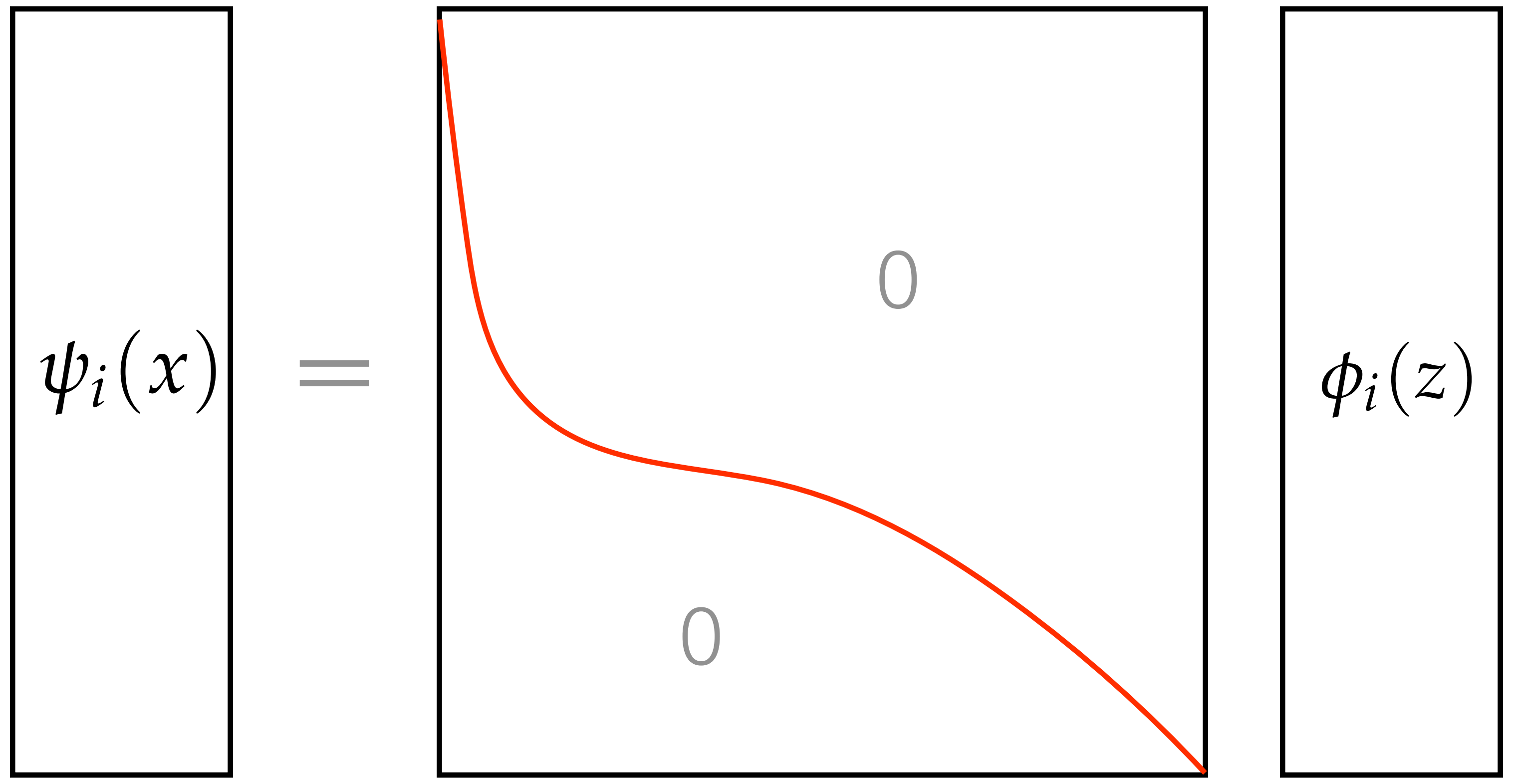
FermiNet, Pfau et al 1909.02487
PauliNet, neural backflow, Iterative backflow…

Quantum algorithms



β-VQE, Liu, Mao, Zhang, LW 1912.11381
Martyn et al 1812.01015, Verdon et al 1910.02071

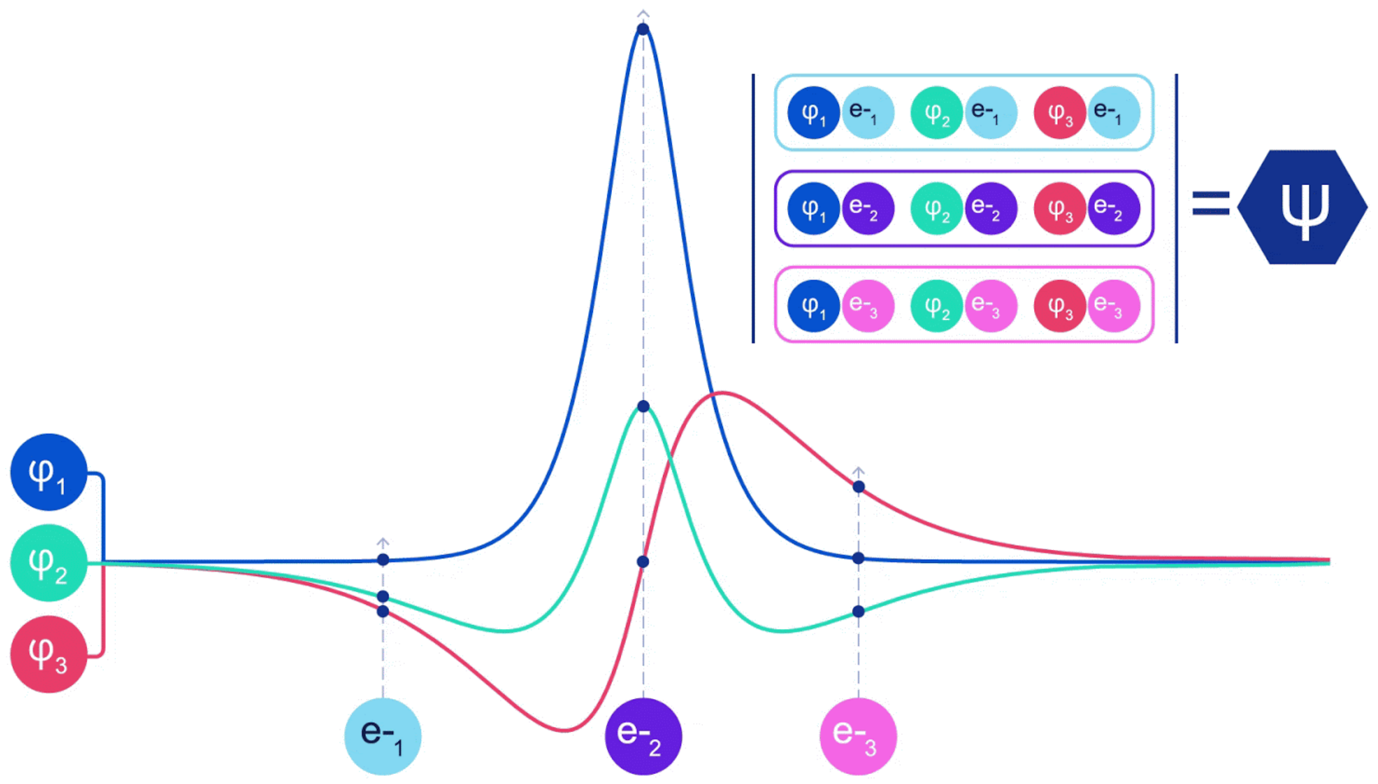
Quantum flows



Cranmer et al, 1904.05903

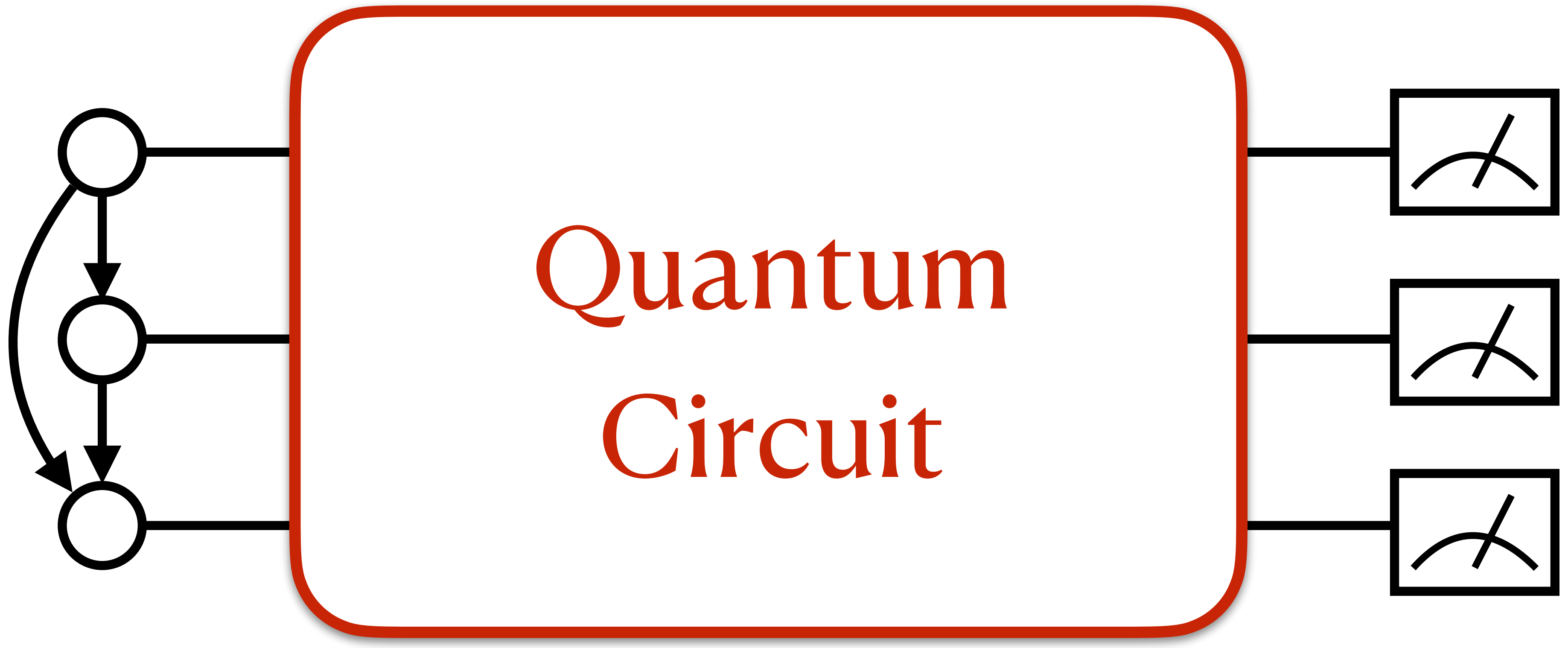
Related works

Neural nets for ground state



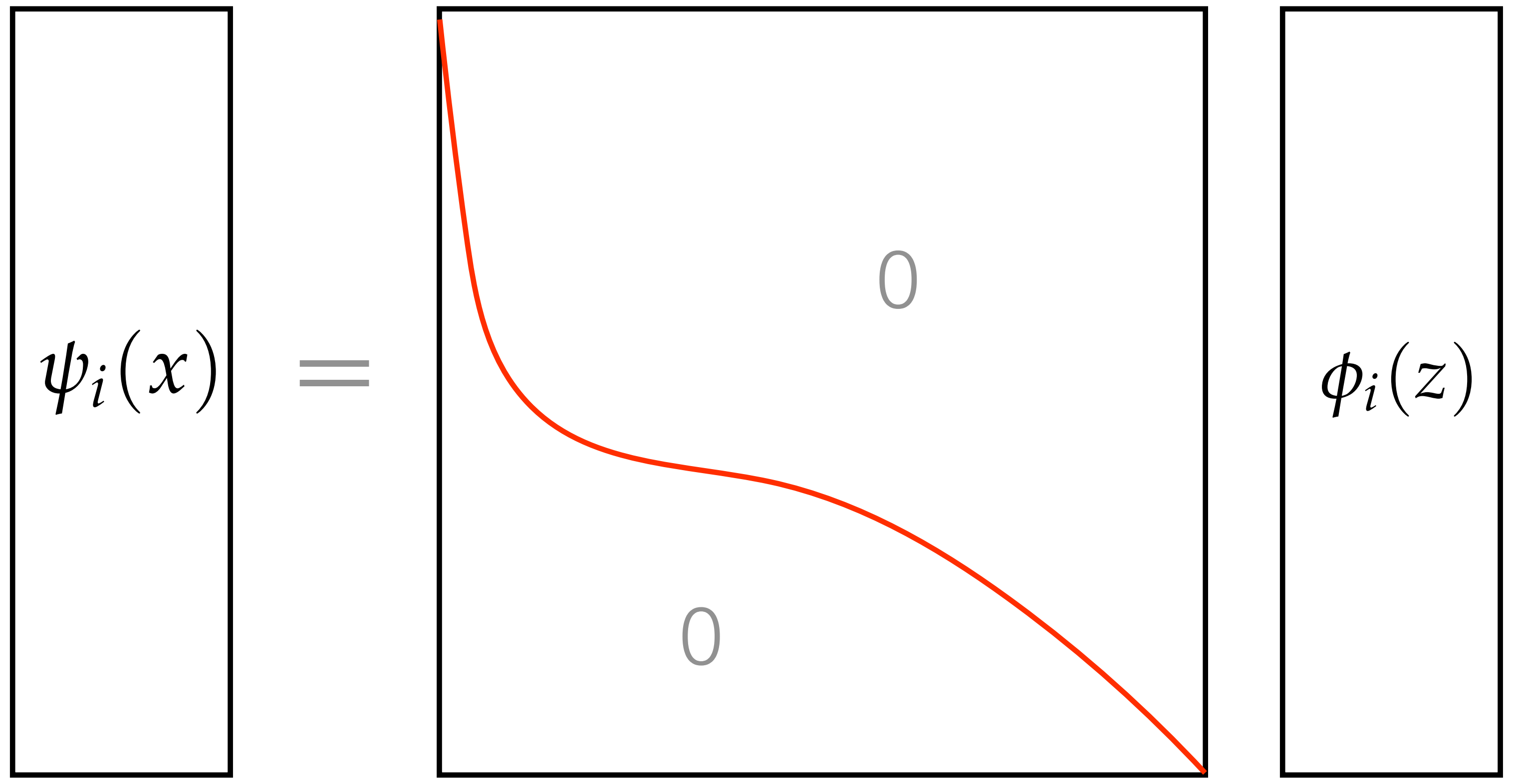
FermiNet, Pfau et al 1909.02487
PauliNet, neural backflow, Iterative backflow...

Quantum algorithms



β-VQE, Liu, Mao, Zhang, LW 1912.11381
Martyn et al 1812.01015, Verdon et al 1910.02071

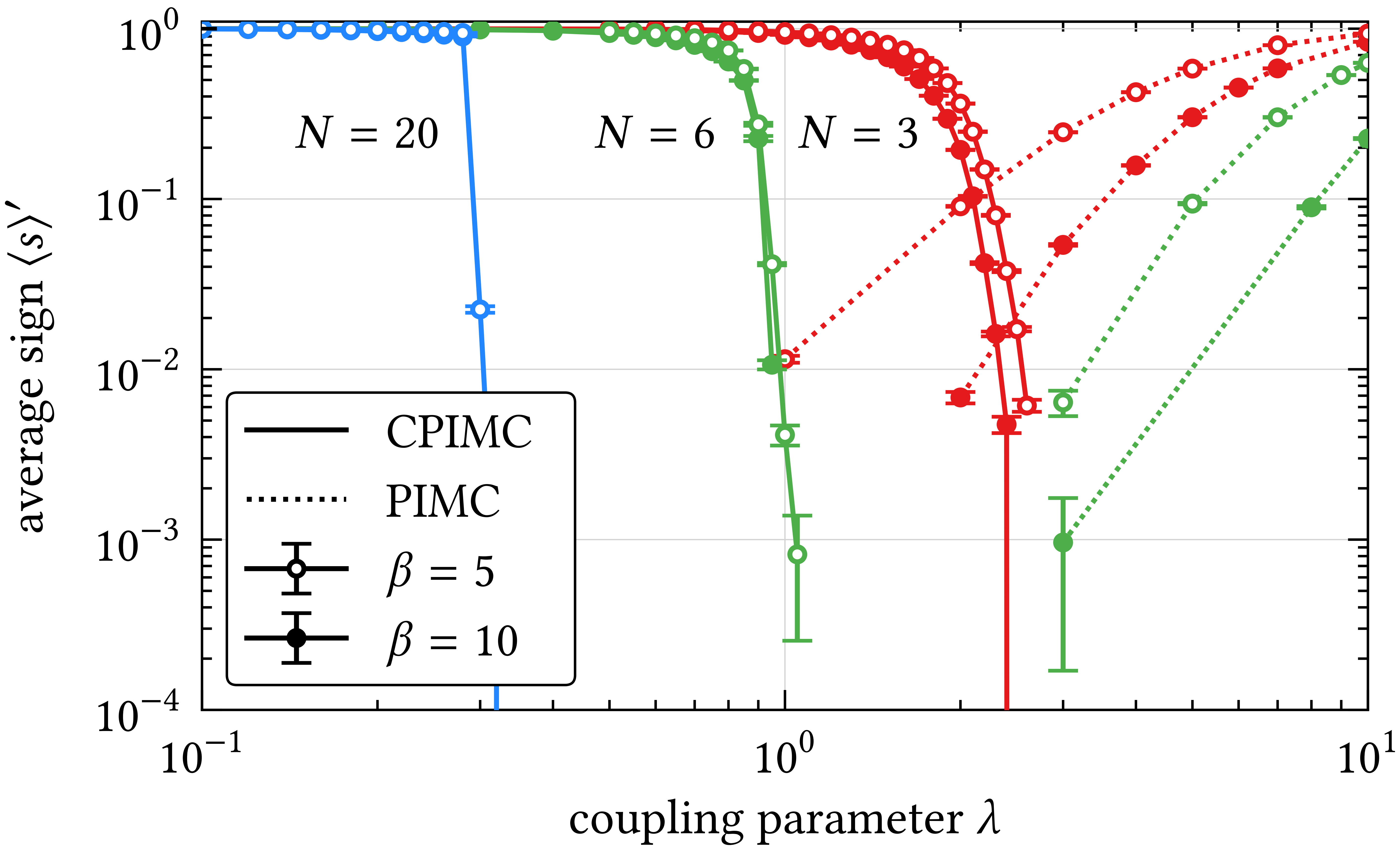
Quantum flows



Cranmer et al, 1904.05903

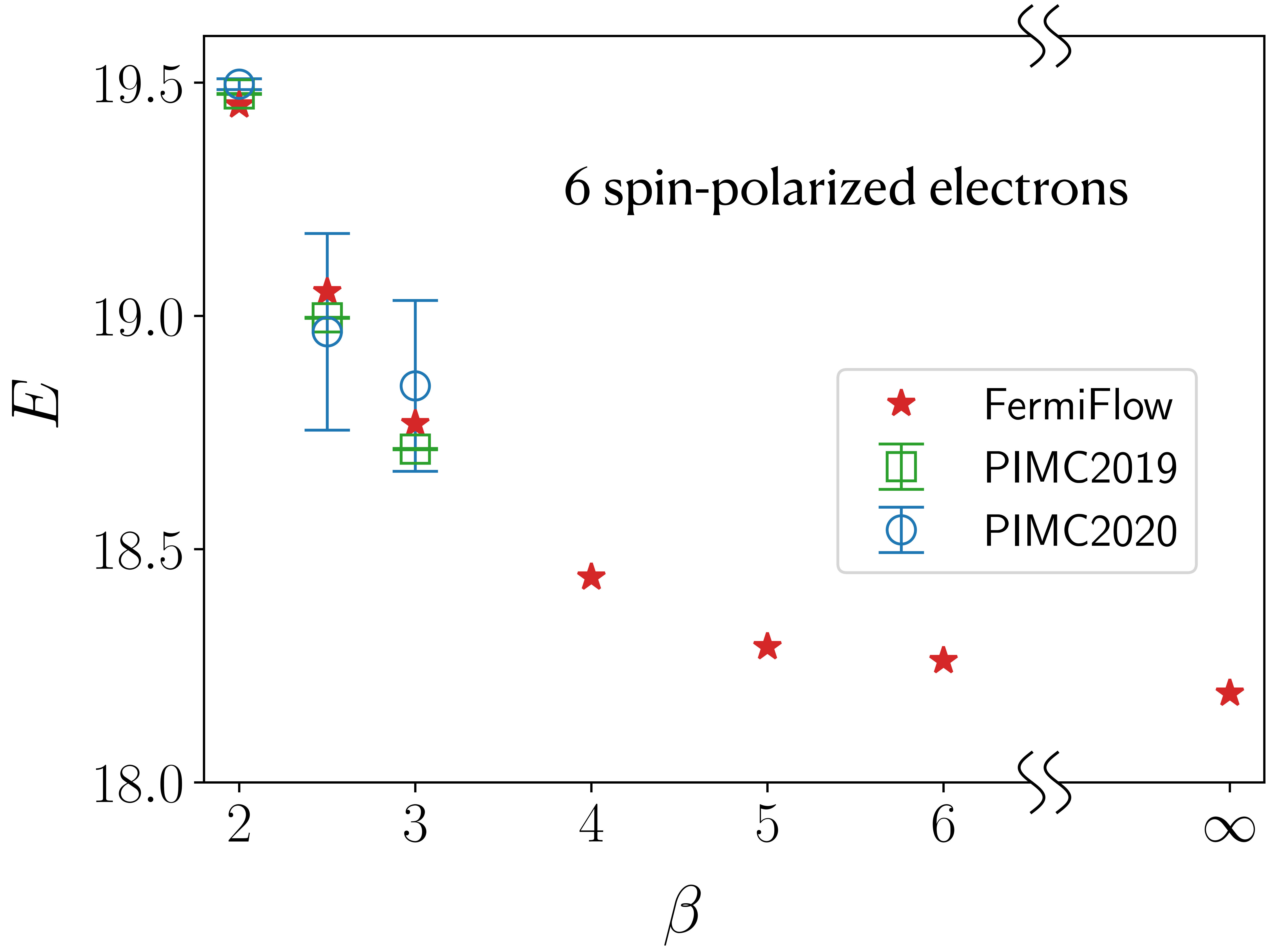
Demo: electrons in a 2D quantum dot

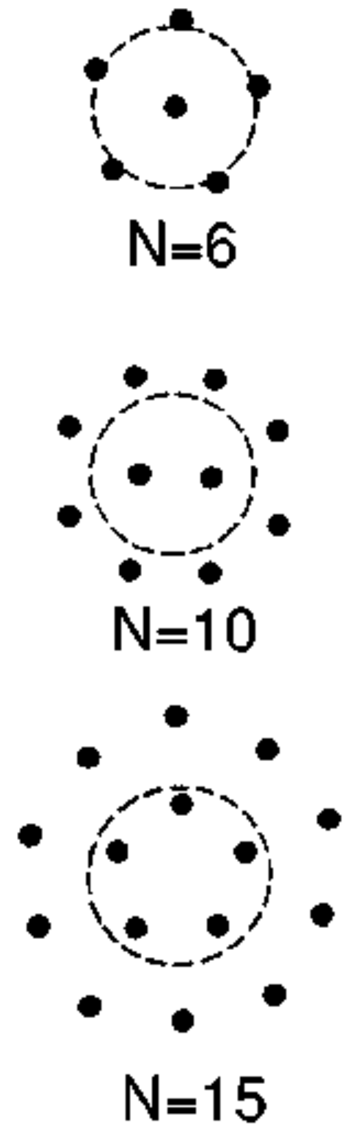
$$H = \sum_i \left(-\frac{\nabla_i^2}{2} + \frac{\boldsymbol{x}_i^2}{2} \right) + \sum_{i<j} \frac{\lambda}{|\boldsymbol{x}_i - \boldsymbol{x}_j|}$$



PhD Thesis of
Tim Schoof, '16

Demo: electrons in a 2D quantum dot



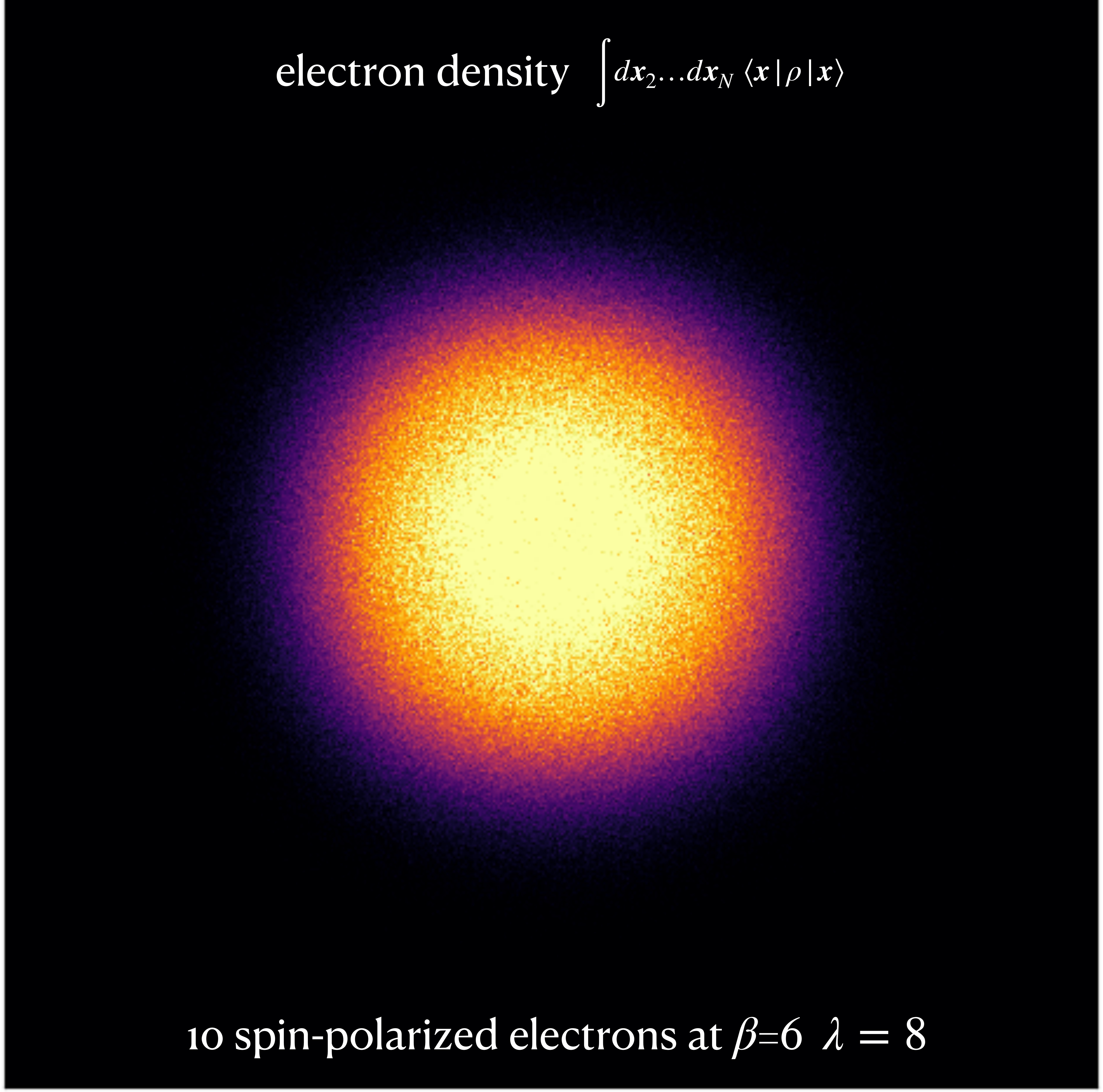


Reimann et al, RMP '02

free electrons

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{v}$$

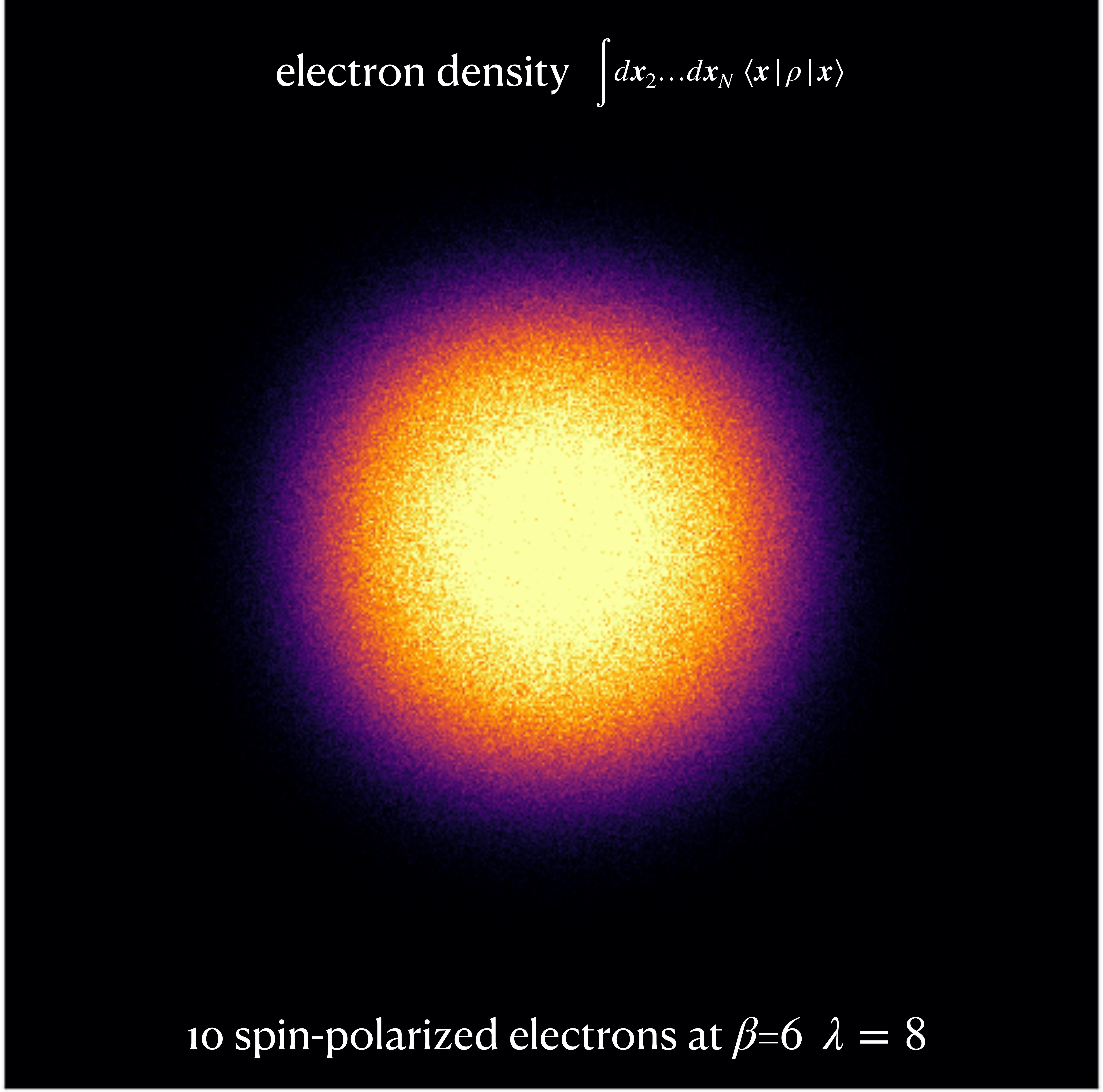
interacting electrons



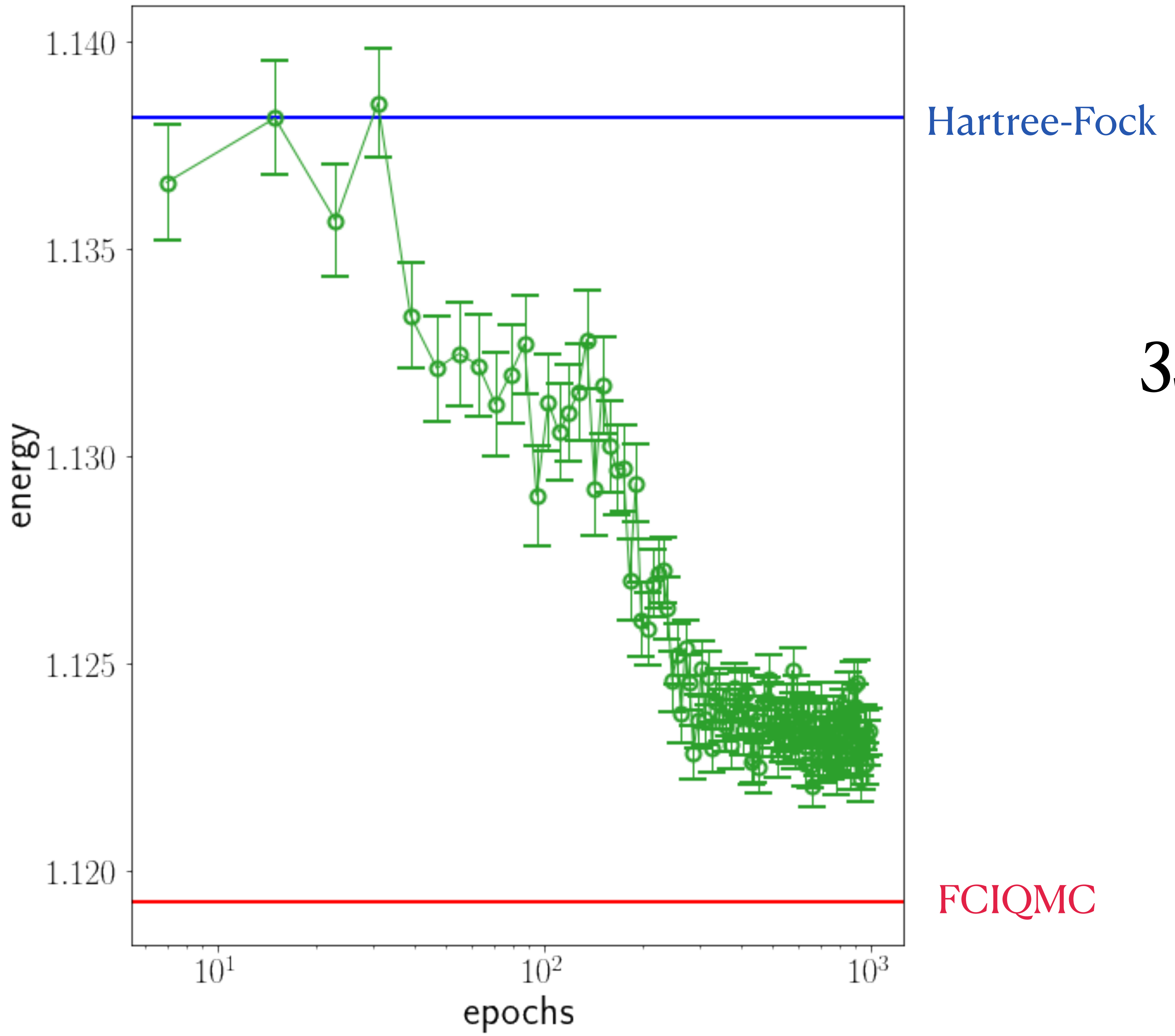
free electrons

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{v}$$

interacting electrons



Back to the uniform electron gas



33 spin polarized electrons @rs=1.0

Reach 0.004 Hartree/electron

ground state accuracy

30 hours training on 8 GPUs

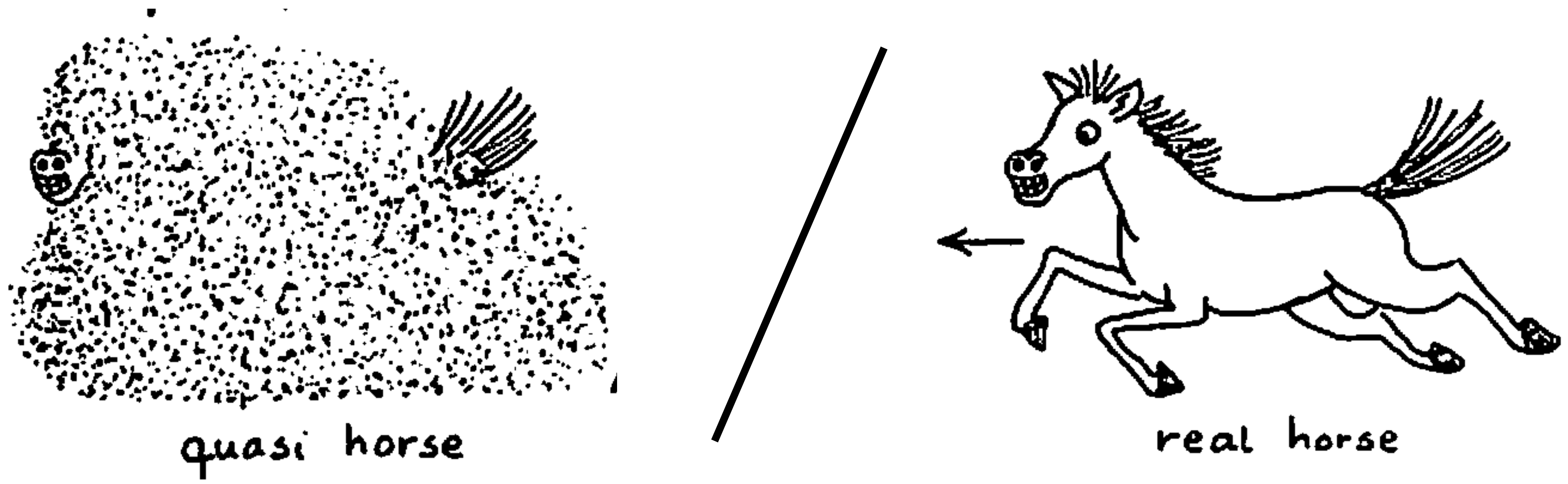
Finite temperature
benchmark data

Brown et al, PRL **110**, 146405 (2013)
Schoof et al, PRL **115**, 130402 (2015)
Malone et al, PRL **117**, 115701 (2016)

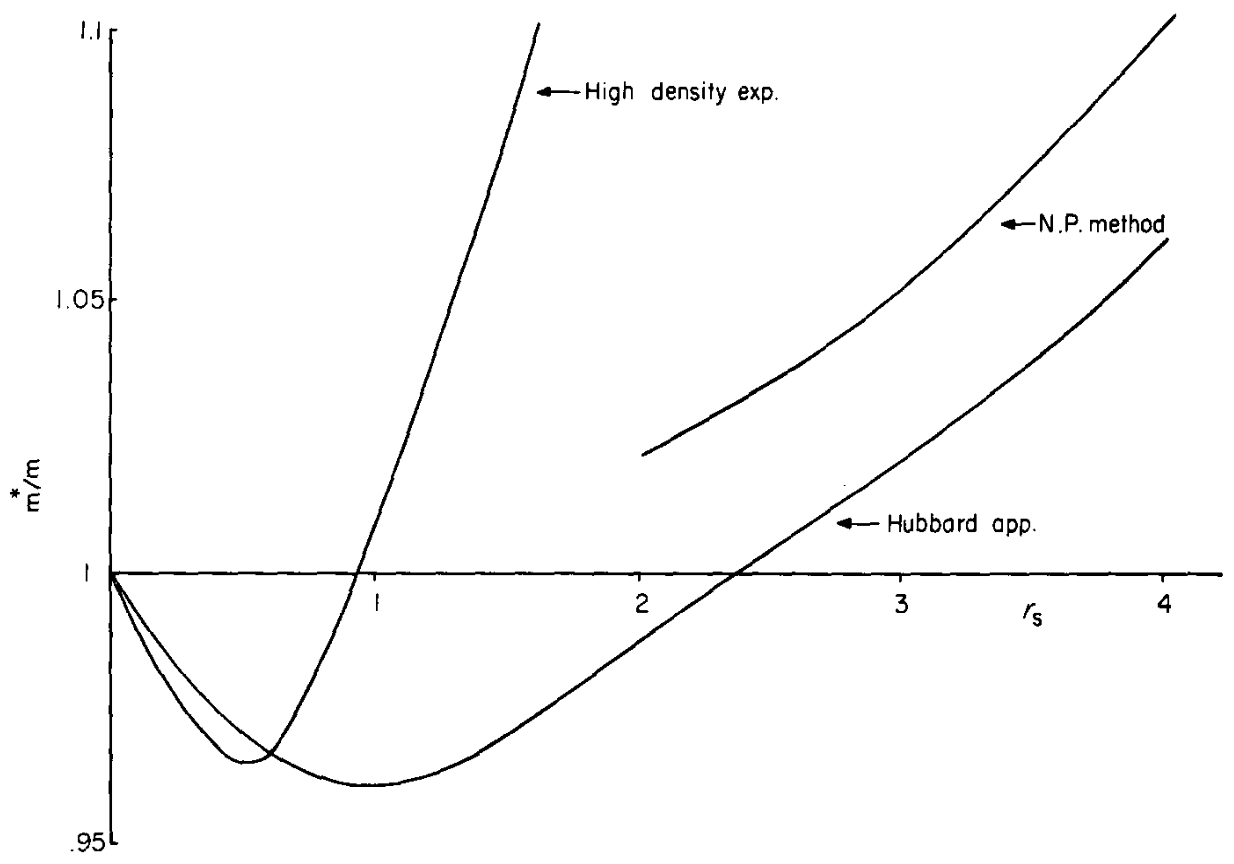
Application: effective mass of quasi-particles

Landau's Fermi liquid theory, 1956

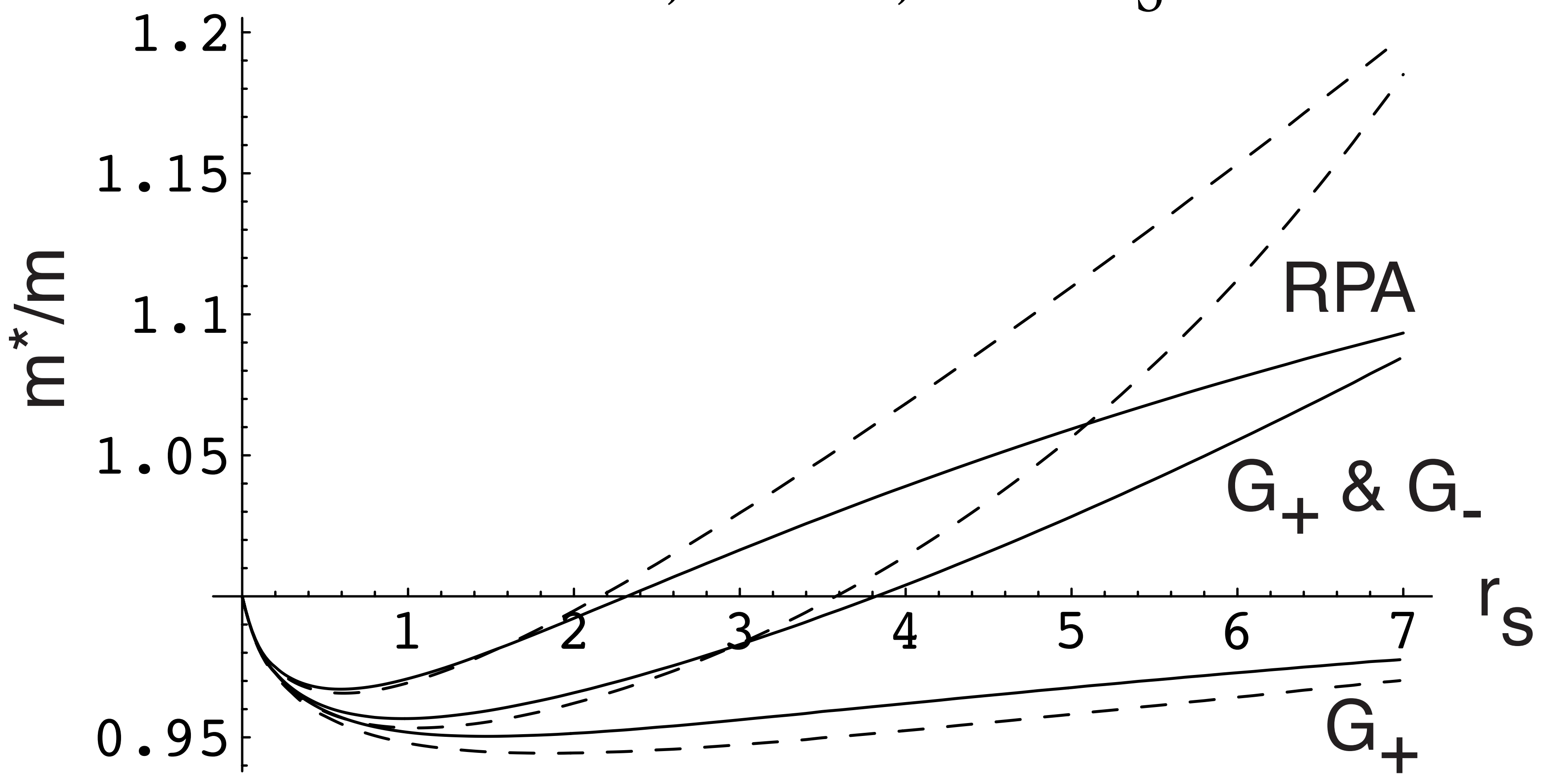
$$\frac{m^*}{m} =$$



T.M.Rice Ann. Phys. 1965

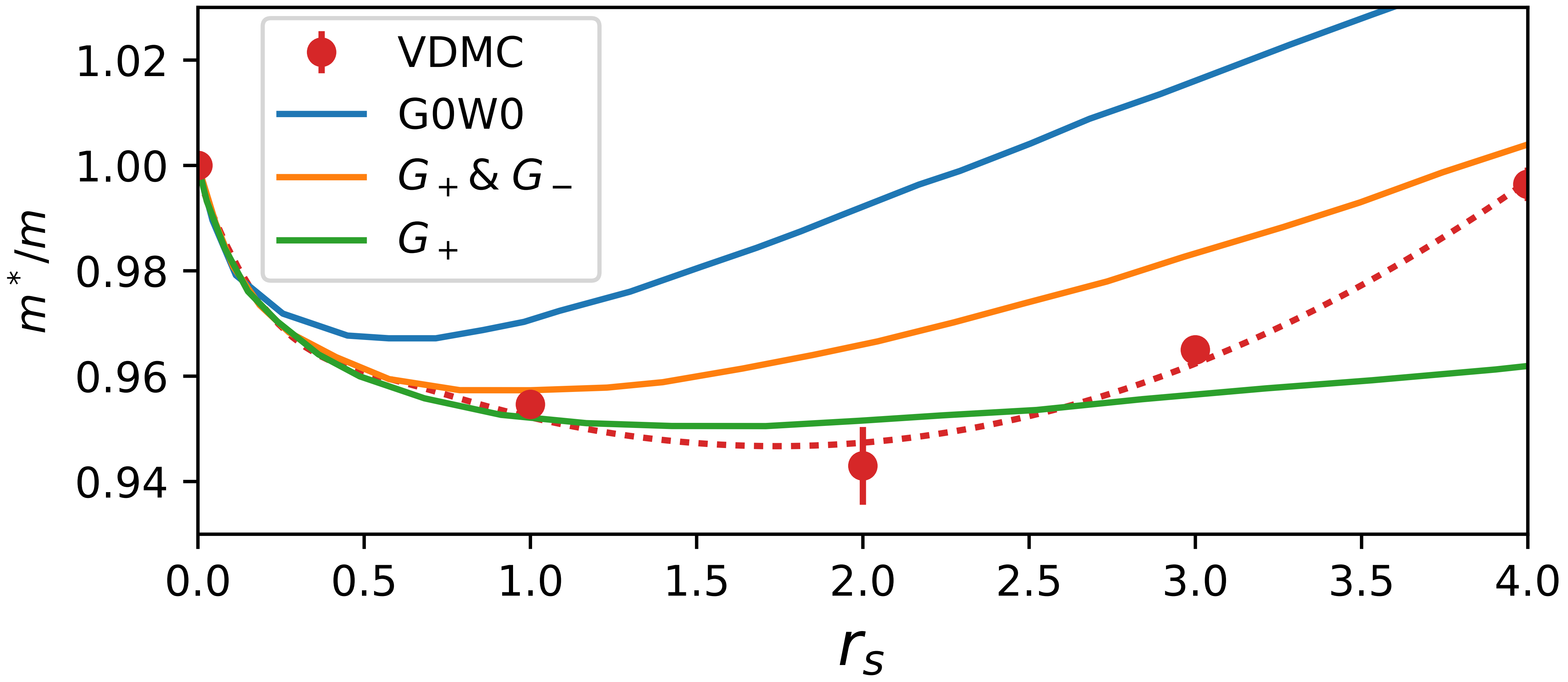


Simion, Guiliani, PRB 2005



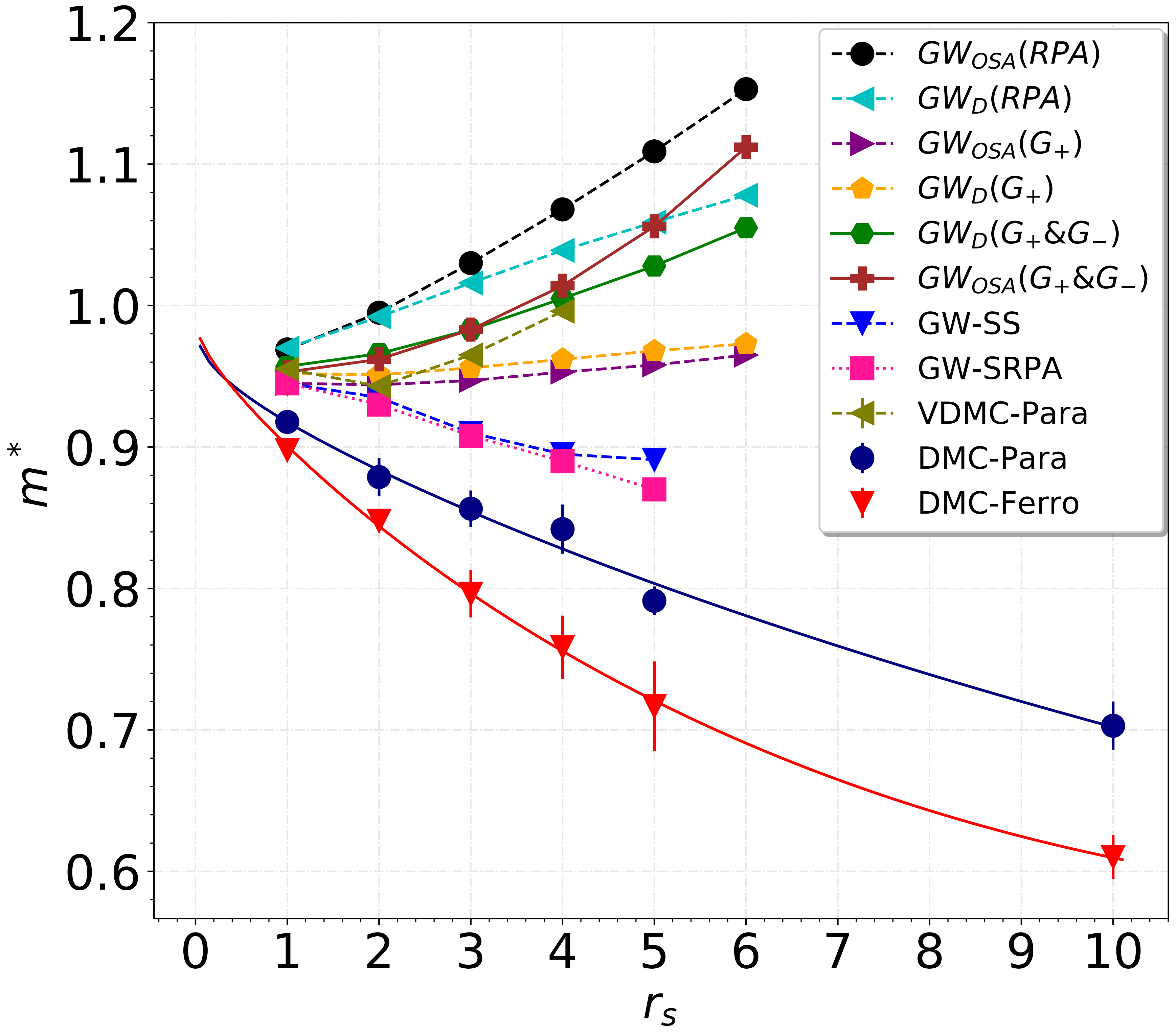
Inconclusive results from perturbative diagrammatic approaches

$$m^*/m = \left(1 - \frac{\partial \Sigma_\sigma(\mathbf{k},\omega)}{\partial \omega}\bigg|_{k=k_F}\right)\left(1 + \frac{m}{k}\frac{\partial \Sigma_\sigma(\mathbf{k},\omega)}{\partial k}\bigg|_{k=k_F}\right)^{-1}$$



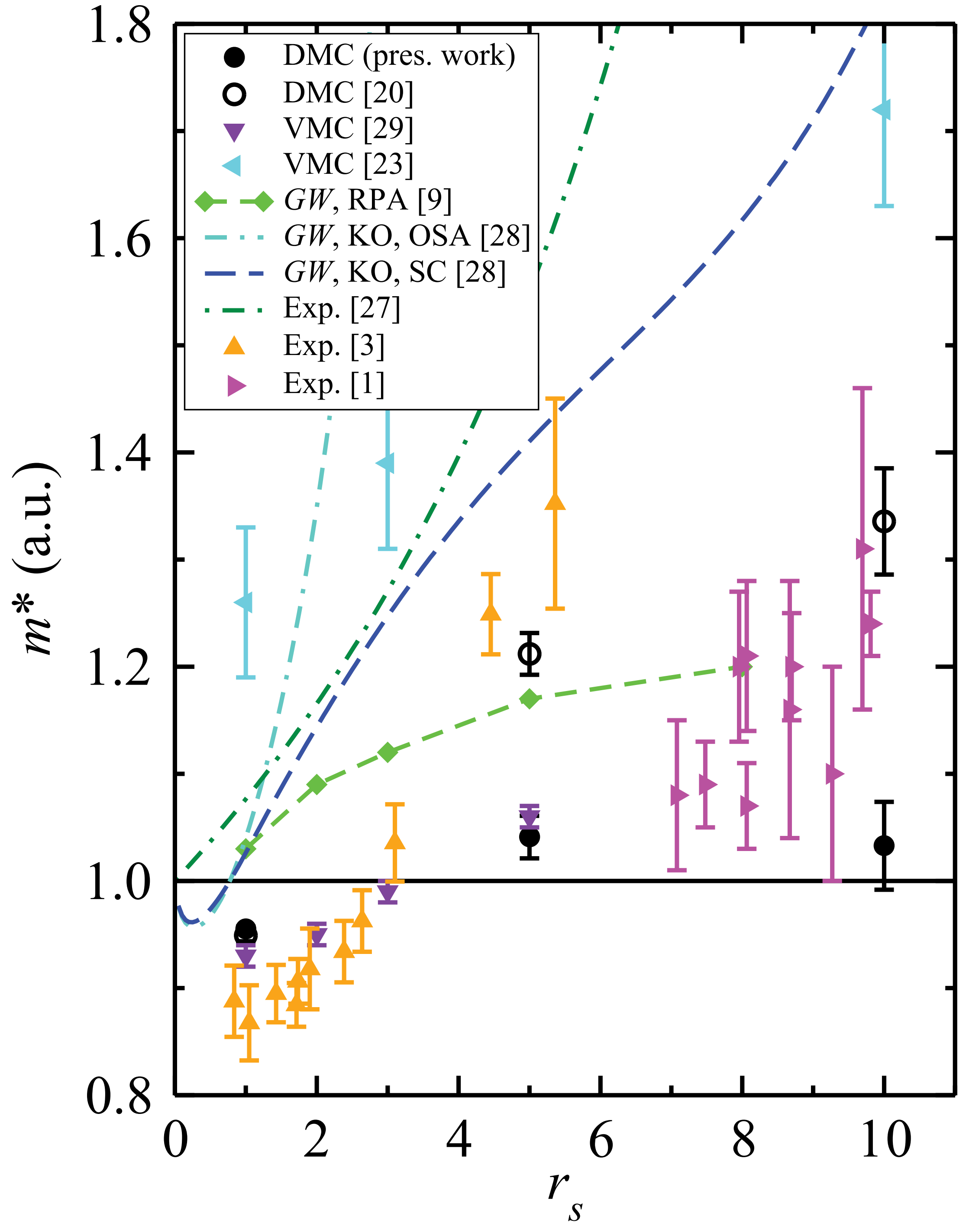
Haule, Chen, 2012.03146

$$m^* = k_F/(d\varepsilon/dk)_{k_F}$$



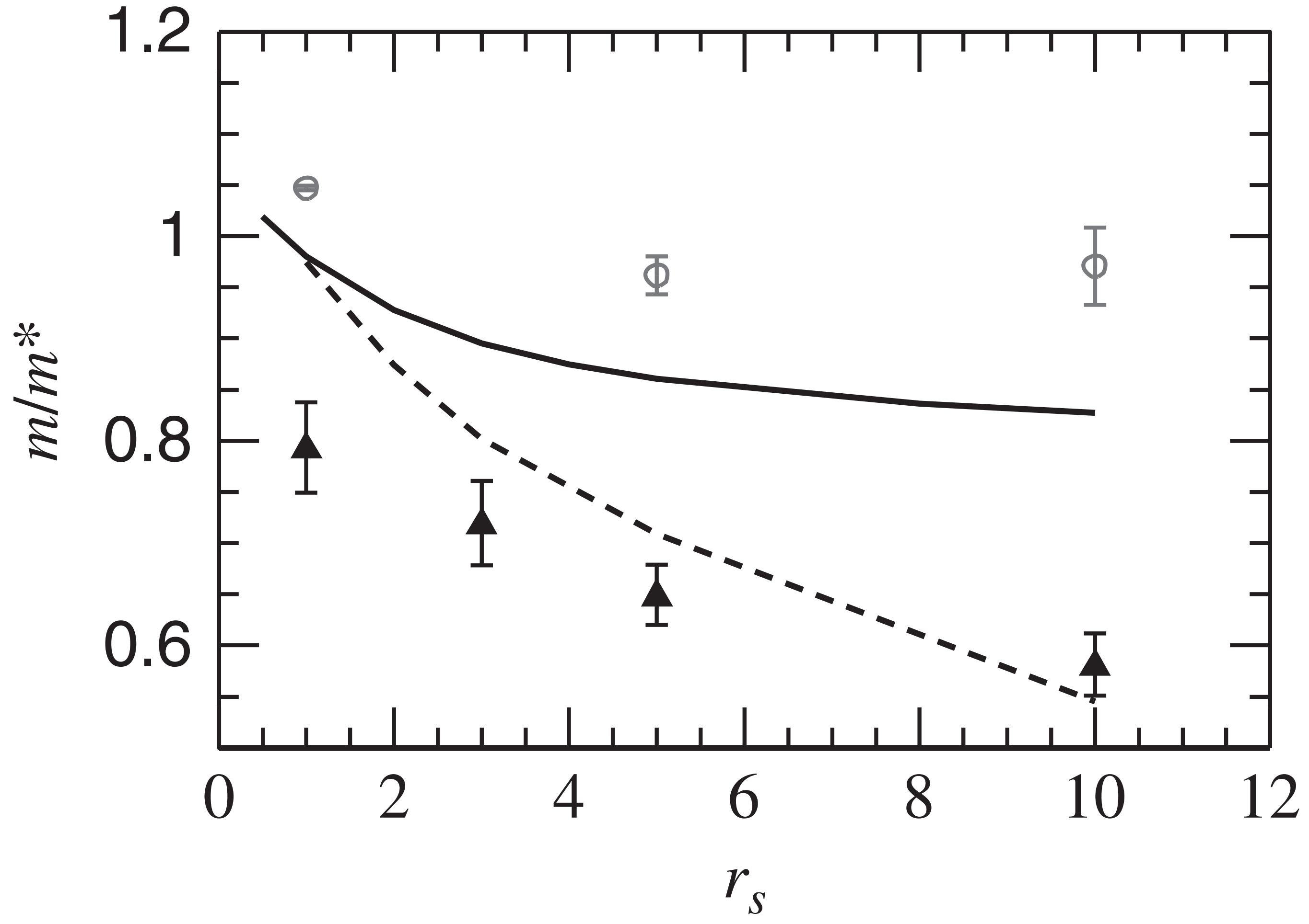
Azadi, Drummond, Foulkes, 2105.09139

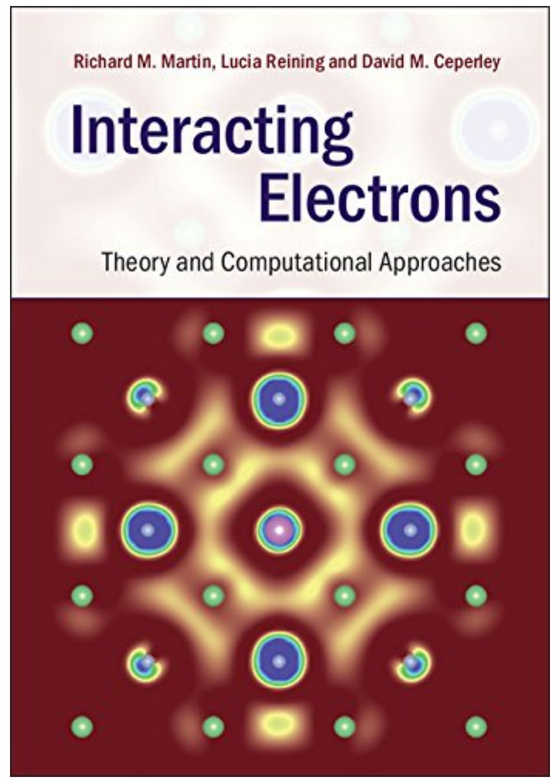
Non-perturbative methods, still inconclusive 🤦



Drummond, Needs, PRB '13

Two quite different QMC results for the 2D HEG are shown in Fig. 23.3 and compared with screened RPA and local field method results. The two different QMC calculations were done in a similar way, but the effective mass differs because of the way it is calculated from the QMC energies.





Martin, Reining, Ceperley, *Interacting Electrons* '16

More conflicting results for 2d electron gas

Effective mass from thermodynamics

Eich, Holzmann, Vignale, PRB '17

$$s_0 = \frac{mk_F}{3\hbar^2 n} k_B^2 T \qquad s = \frac{m^* k_F}{3\hbar^2 n} k_B^2 T$$

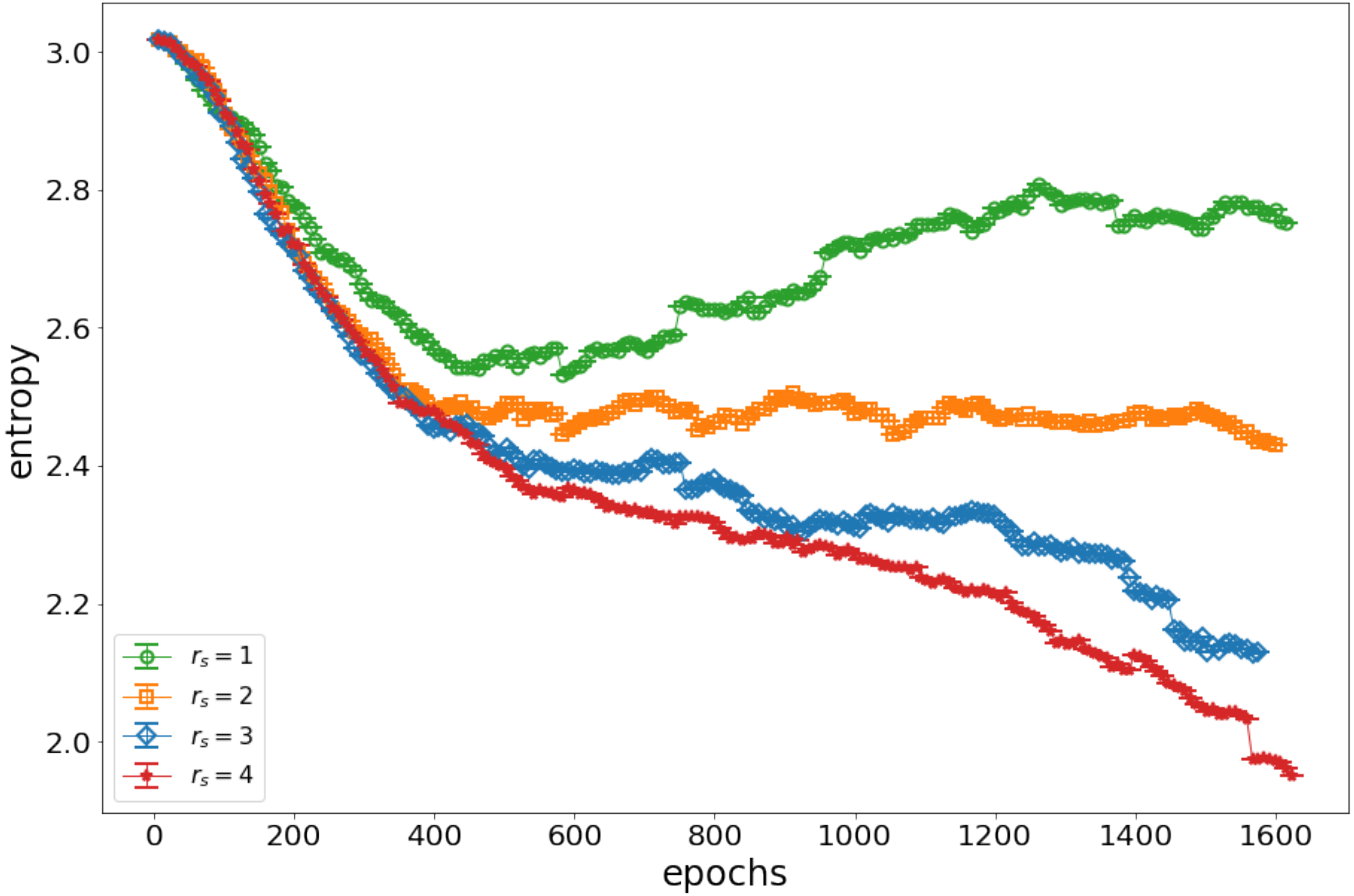
$$\frac{m^*}{m} = \frac{s}{s_0}$$

Interacting/noninteracting entropy ratio

Previous calculations can not reach the low temperature region

Moreover, entropy is not directly accessible to PIMC

14 electrons @ T/Ef=0.08



Outlooks

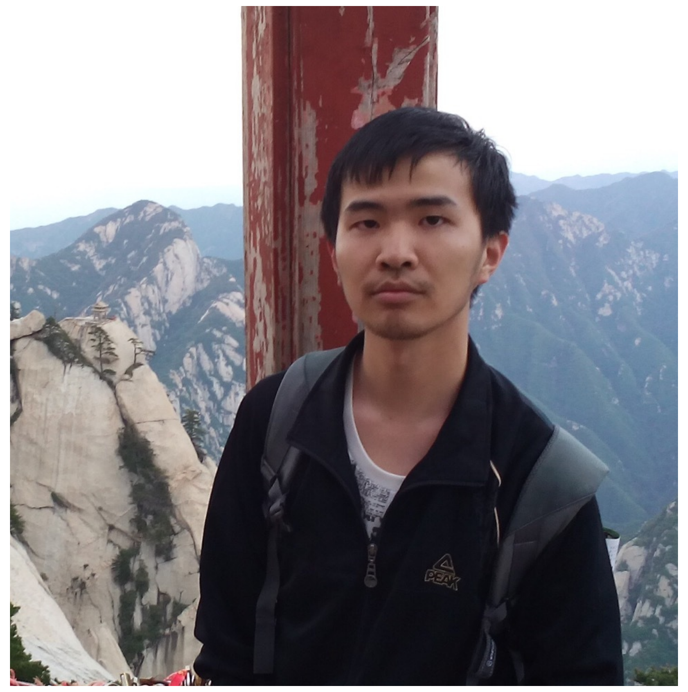
Ultracold fermi gases

Dense hydrogen

Warm dense matter

Thermal density functionals

Thank you!



Hao Xie

Linfeng Zhang

2105.08644

buwantaiji/FermiFlow