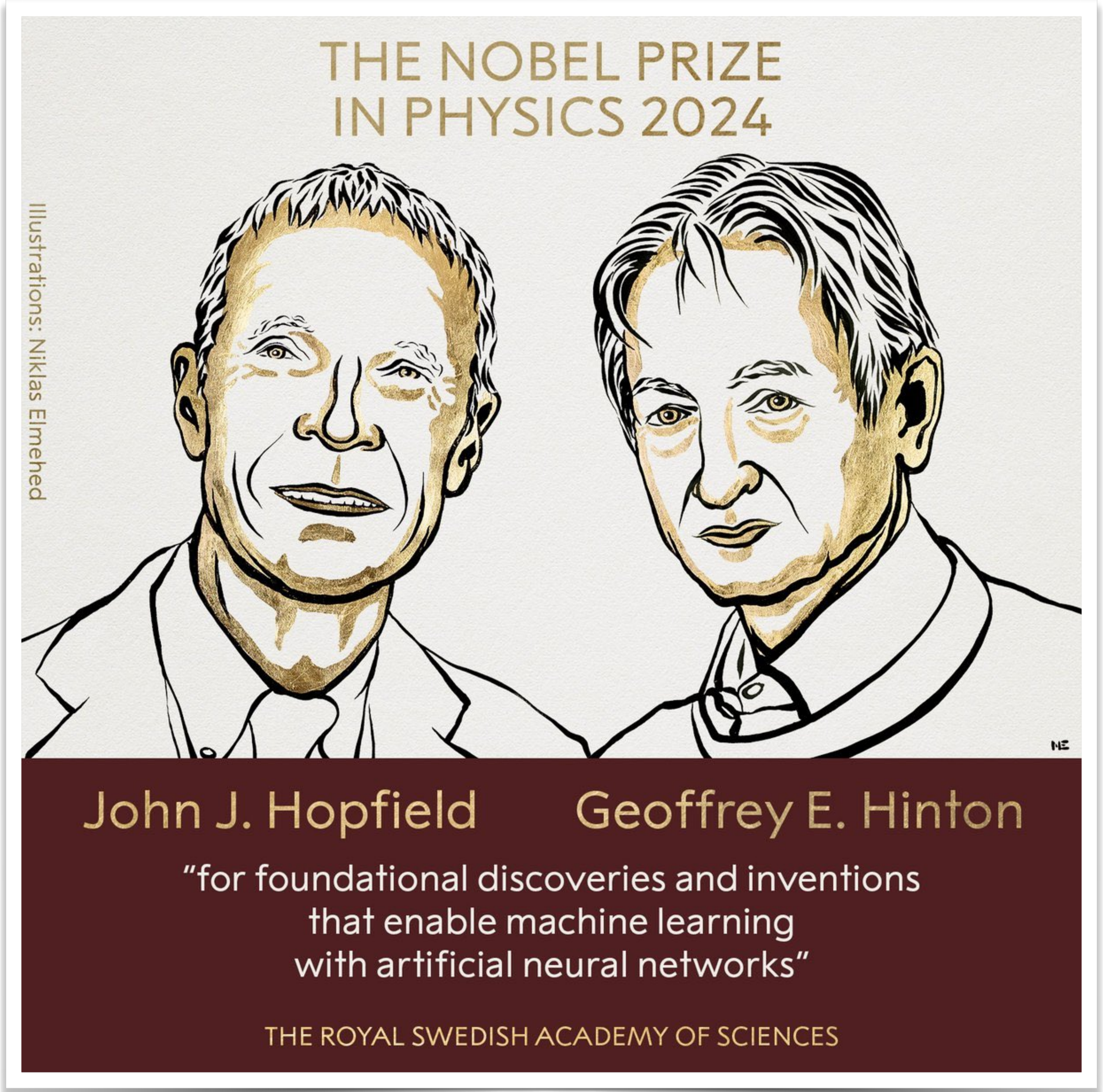
Learning Hadron Interactions with Deep Neural Networks

Lingxiao WANG(王 凌霄)
RIKEN-**iTHEMS**

Oct. 30, 2024
“Hadrons and Hadron Interactions in QCD 2024”
Nishinomiya-Yukawa memorial workshop at YITP

arXiv:2411.xxxxx; ;arXiv:2410.03082; Prog.Part.Nucl.Phys. 104084(2023).

Nobel Prize in Physics 2024

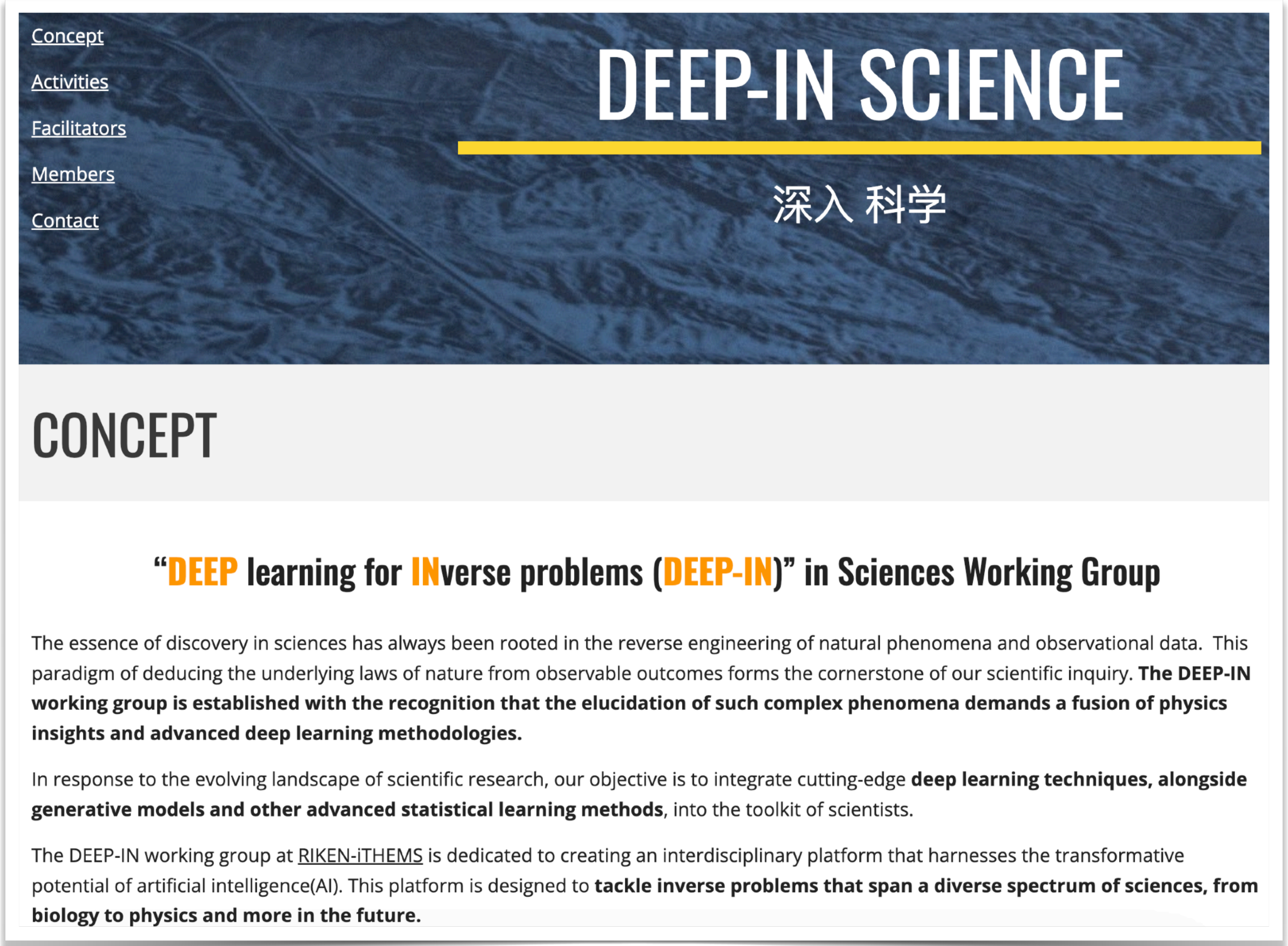


For physicists,

it is the best of times,

it is the worst of times.

DEEP-IN Working Group



[Concept](#)
[Activities](#)
[Facilitators](#)
[Members](#)
[Contact](#)

DEEP-IN SCIENCE

深入 科学

CONCEPT

“DEEP learning for INverse problems (DEEP-IN)” in Sciences Working Group

The essence of discovery in sciences has always been rooted in the reverse engineering of natural phenomena and observational data. This paradigm of deducing the underlying laws of nature from observable outcomes forms the cornerstone of our scientific inquiry. **The DEEP-IN working group is established with the recognition that the elucidation of such complex phenomena demands a fusion of physics insights and advanced deep learning methodologies.**

In response to the evolving landscape of scientific research, our objective is to integrate cutting-edge **deep learning techniques, alongside generative models and other advanced statistical learning methods**, into the toolkit of scientists.

The DEEP-IN working group at RIKEN-iTHEMS is dedicated to creating an interdisciplinary platform that harnesses the transformative potential of artificial intelligence(AI). This platform is designed to **tackle inverse problems that span a diverse spectrum of sciences, from biology to physics and more in the future.**

iTHEMS 理化学研究所 数理創造プログラム RIKEN Interdisciplinary Theoretical and Mathematical Sciences Program

About iTHEMS

HOME / About iTHEMS / Working Groups / DEEP-IN Working Group

DEEP-IN Working Group

“DEEP learning for INverse problems (DEEP-IN) in Sciences” working group (April 1st, 2024 -)

Lattice Computations

Gert Aarts, Swansea U.
Takumi Doi, iTHEMS
Andreas Ipp, TU Wien
Tetsuo Hatsuda, iTHEMS
Yan Lyu, iTHEMS

Now mostly physicists **->** **Future** more diverse scientists

BioPhysics: **Catherine Beauchemin,** iTHEMS
Condensed Matter Physics: **Steffen Backes,** iTHEMS
QCD Physics: **Kenji Fukushima**, UTokyo
Nuclear Physics: **Haozhao Liang**, UTokyo
Quantum Computing: **Enrico Rinaldi,** Quantinuum K.K./iTHEMS

Heavy-Ion Collisions

Long-Gang Pang, CCNU
Shuzhe Shi, THU
Kai Zhou, CUHK-ShenZhen

Astrophysics

Márcio Ferreira, Coimbra U.
Yuki Fujimoto, INT->iTHEMS
Akira Harada, NIT-Ibaraki
Zhenyu Zhu, TDLI->RIT

Machine Learning

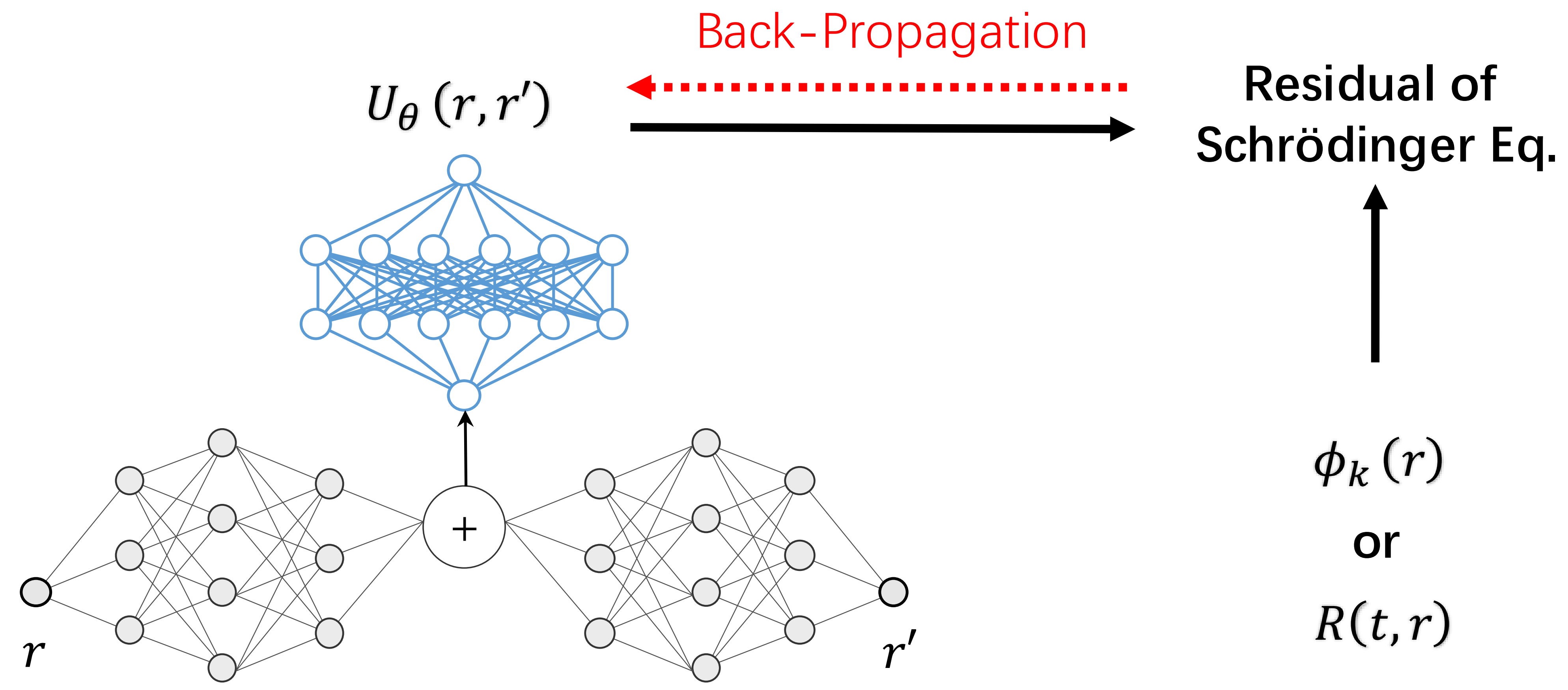
Akinori Tanaka, AIP/iTHEMS
Lingxiao Wang, iTHEMS

https://sites.google.com/view/deep-in-wg/homepage

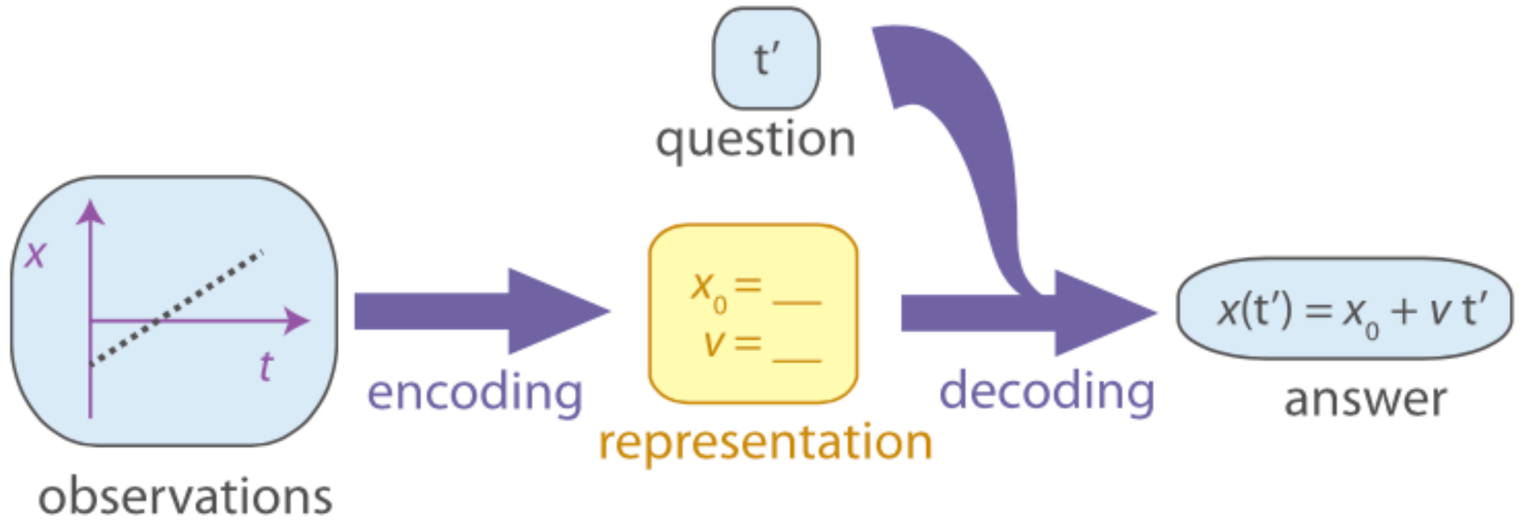
Lingxiao Wang (RIKEN iTHEMS) *Contact at lingxiao.wang@riken.jp

Outline

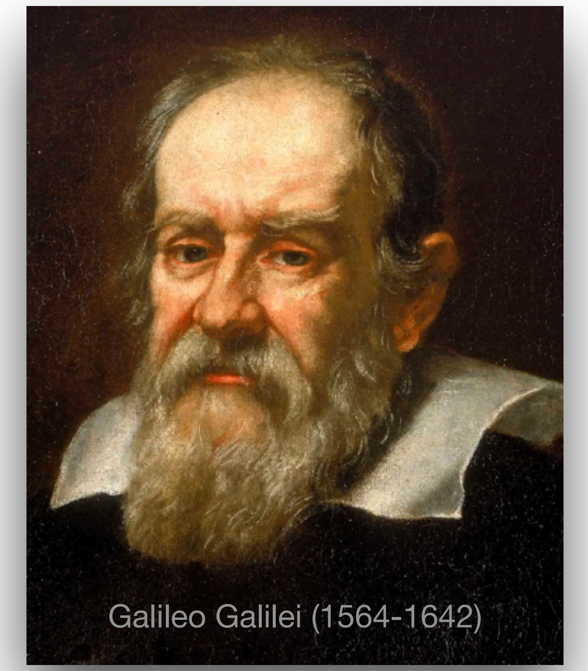
- **Machine Learning for Physics**
- **Hadron Interactions**
- **Inverse Femtoscopy**
- **HAL QCD meets DNNs**
- **Outlooks**



Machine Learning and Physics



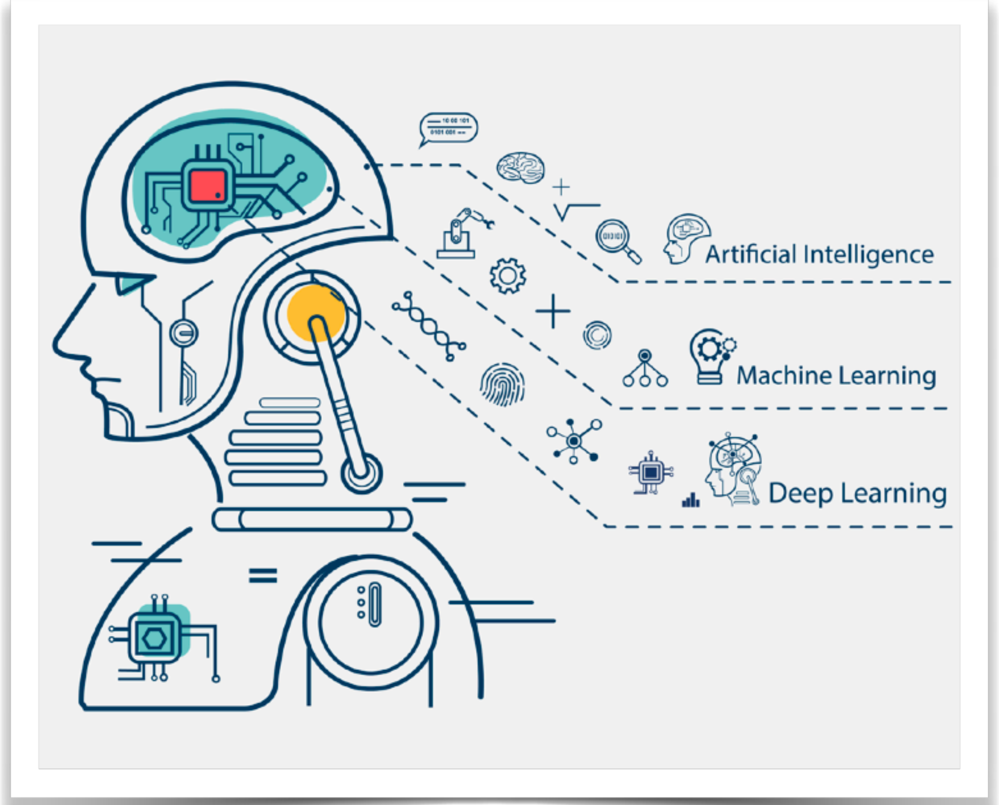
Phys.Rev. Lett. **124**, 010508 (2020)



Galileo Galilei (1564-1642)

An inverse problem in science is the process of inferring from a set of observations the causal factors that produced them.

Data, X



Machine, $\{\theta\}$

Prediction $p(X|\theta)$

Estimation

$$\hat{\theta} = \textbf{arg max}_{\theta}\{p(X|\theta)\}$$

Machine Learning and Inference

Maximum Likelihood Estimation(MLE)

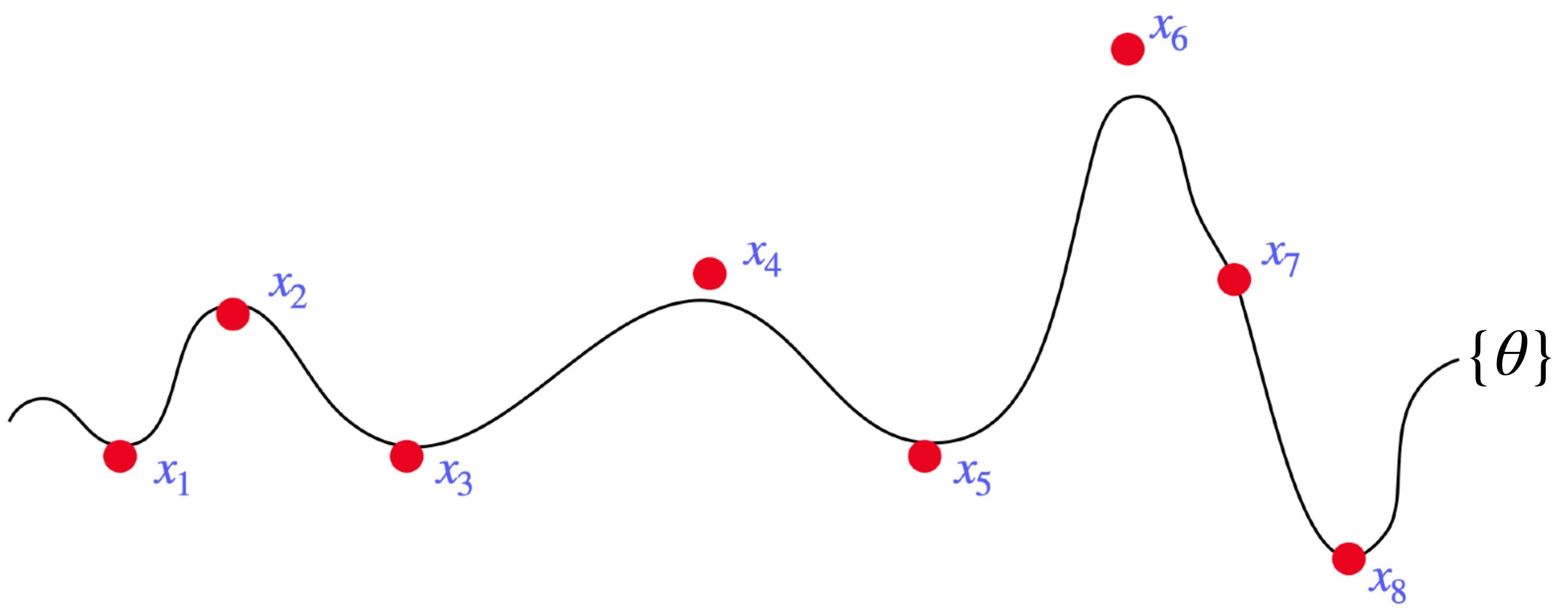
$$\max_{\theta} \prod_{i=1}^{N} p(\mathbf{x}_i \mid \theta)$$

Bayesian Inference

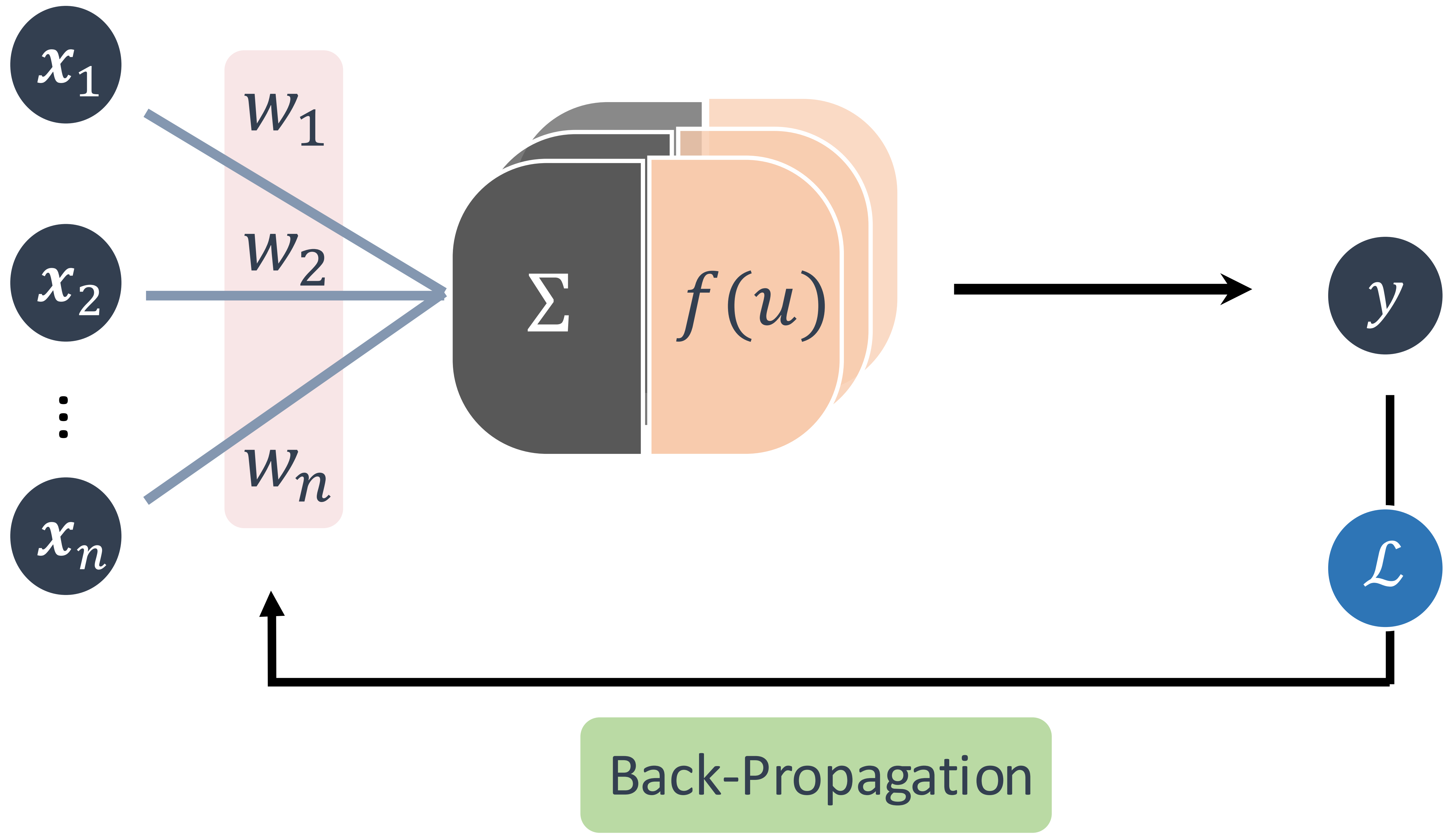
Maximum A Posterior(MAP)

$$p(\theta \mid X) = \frac{p(X \mid \theta)\pi(\theta)}{p(X)}$$

Posterior $p(\theta \mid X)$**, Prior** $\pi(\theta)$ **, Evidence** $p(X)$



Deep Model as Machine



Deep (neural network) Model

- **Inputs,** $\{x\} = x_1, x_2, \ldots, x_n$
- **Weights,** $\{w\} = w_1, w_2, \ldots, w_n$
- **Outputs,** y
- **Summation,** $\Sigma(\cdot)$
- **Non-Linear Activation Functions,** $f(u)$
- **Single Layer** $y = f(\Sigma_{i=1}^{n} x_i w_i)$

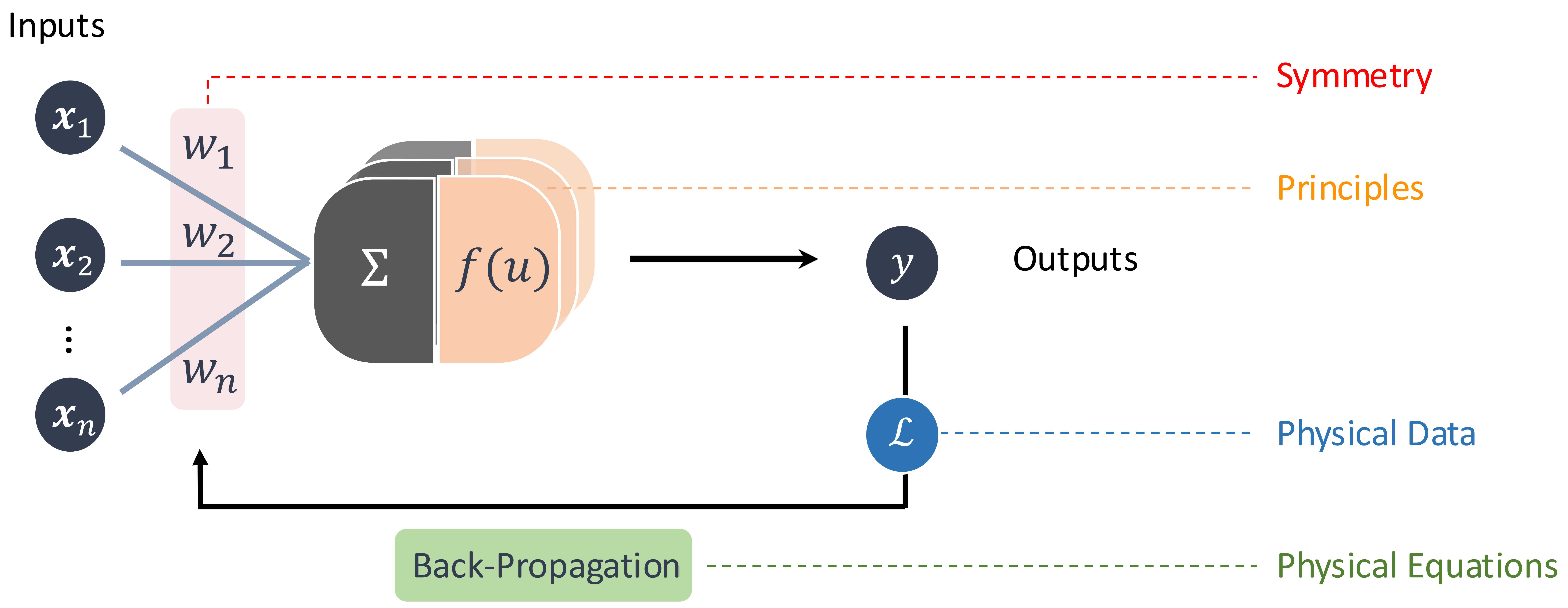
Objective

- **Loss Function,** $\mathscr{L}(y, \hat{y})$
- **Data,** $\hat{y}$

Optimization Algorithm

- **Back-Propagation,** $\frac{\partial \mathscr{L}}{\partial \omega}$
- **Stochastical Methods: SGD, Adam, …**

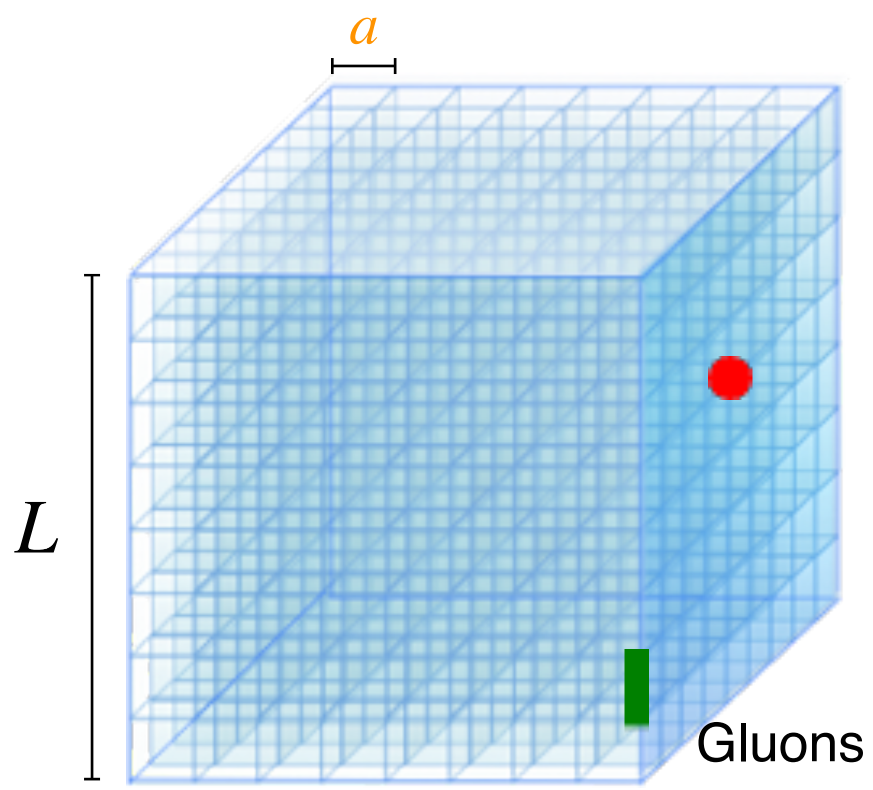
Physics-Driven Deep Learning



Hadron Interactions

Hadron Interactions

$$\mathscr{L} = -\frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a + \bar{q}\gamma^\mu(i\partial_\mu - gt^aA^a_\mu)q - m\bar{q}q$$



4-d Euclidean Lattice

Quarks on sites $q(x)$

Gluons on links $U_\mu(x, x+\mu) = \exp[-iaA_\mu]$

Huge integration variables

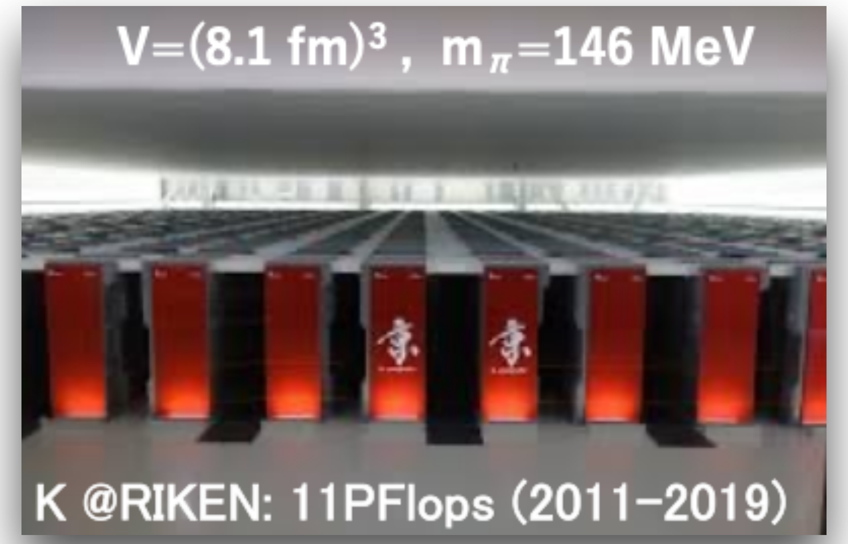
$\sim 10^{9-10}$ for 96^4 lattice, ~ 50 GB/config

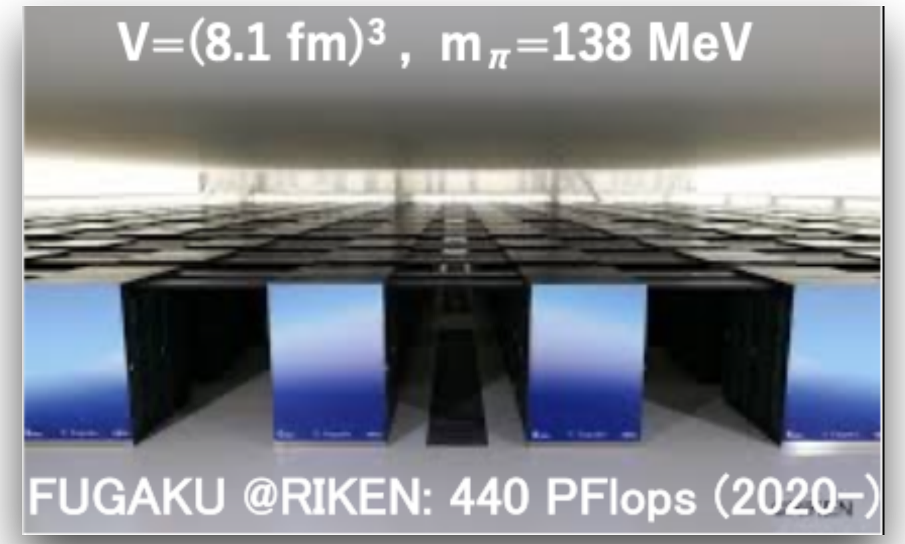
Importance Sampling

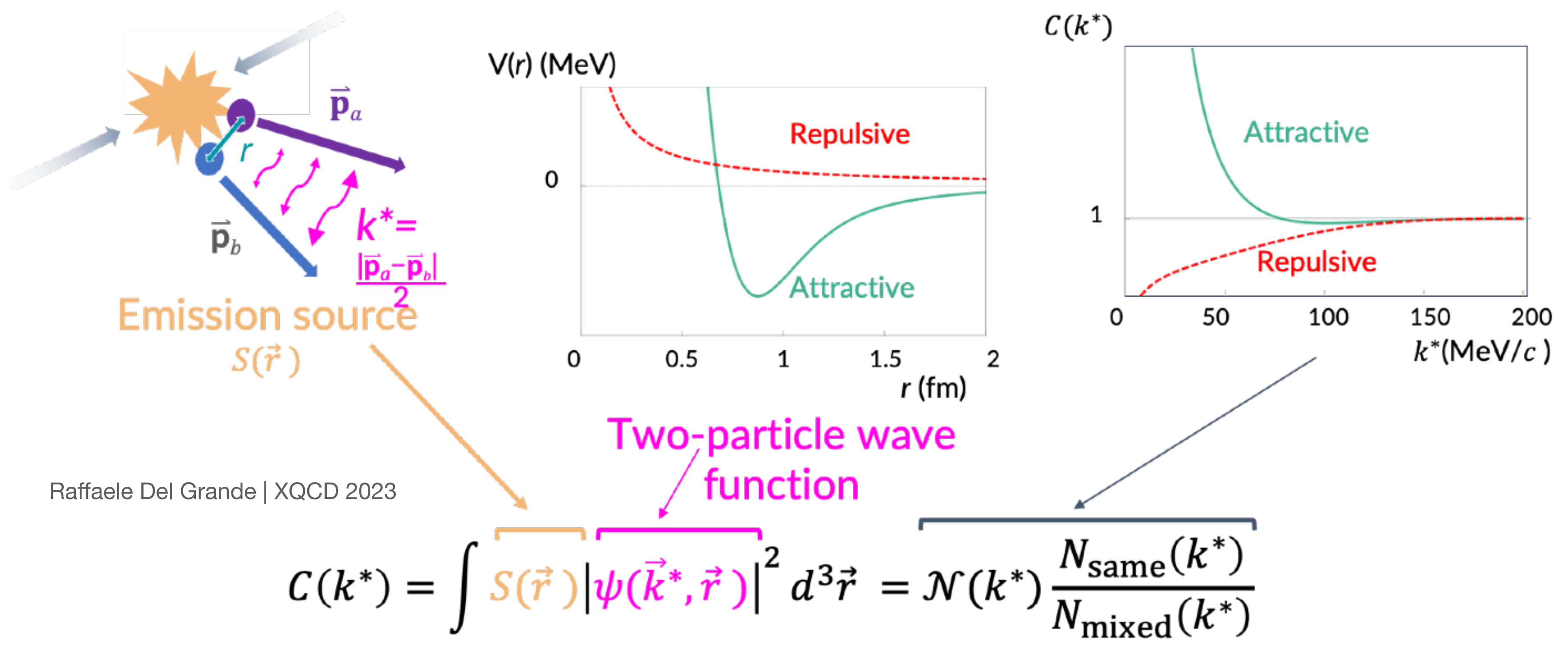
Hybrid MC = MD + Metropolis

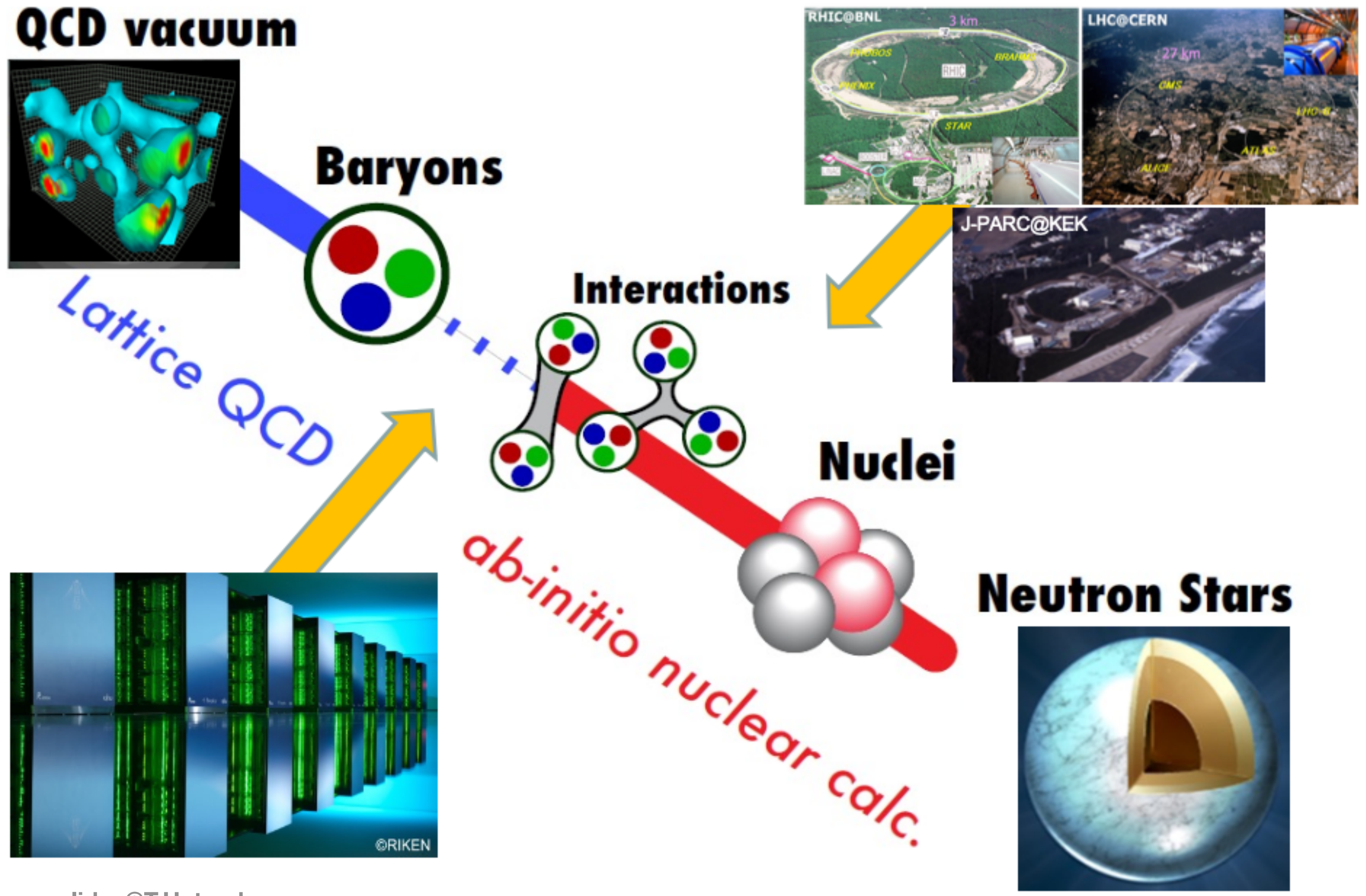
Continuum & Thermodynamic Limits

$a \to 0, L \to \infty$

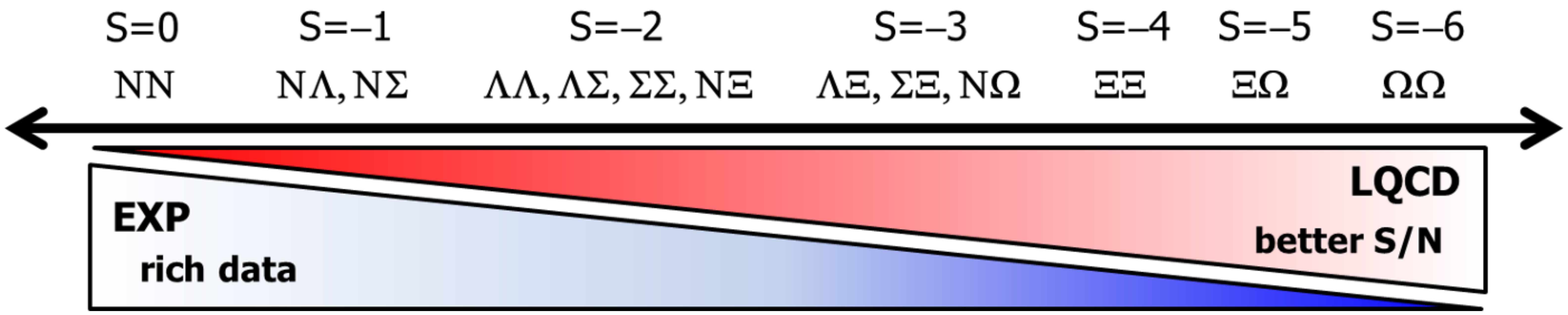




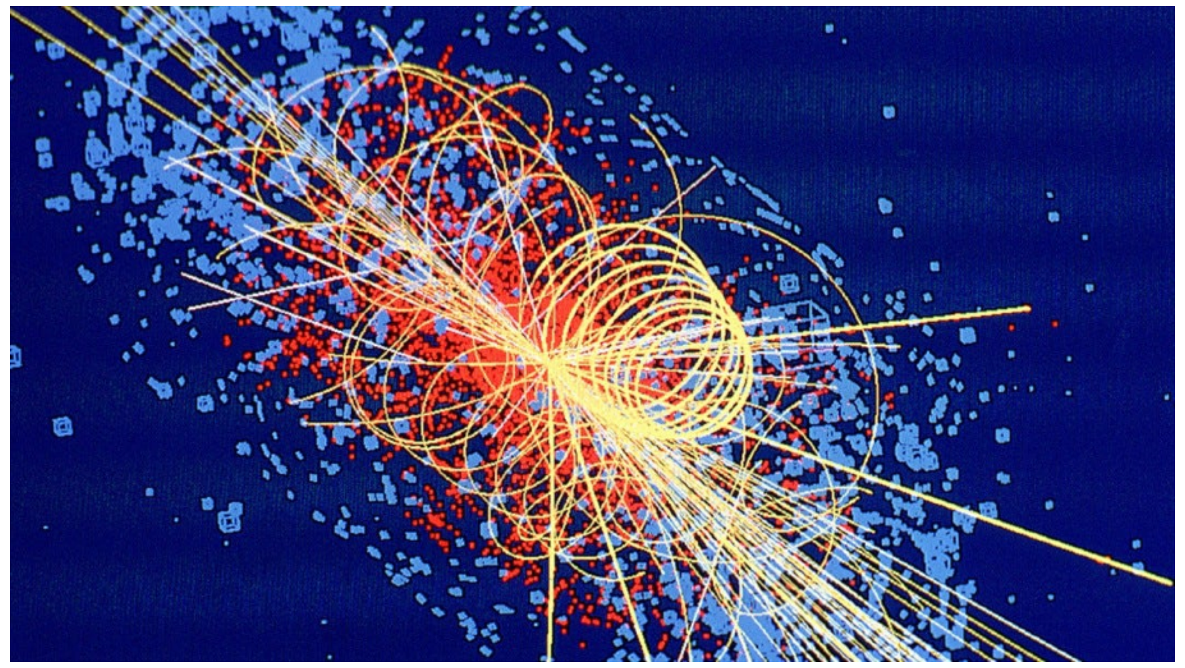


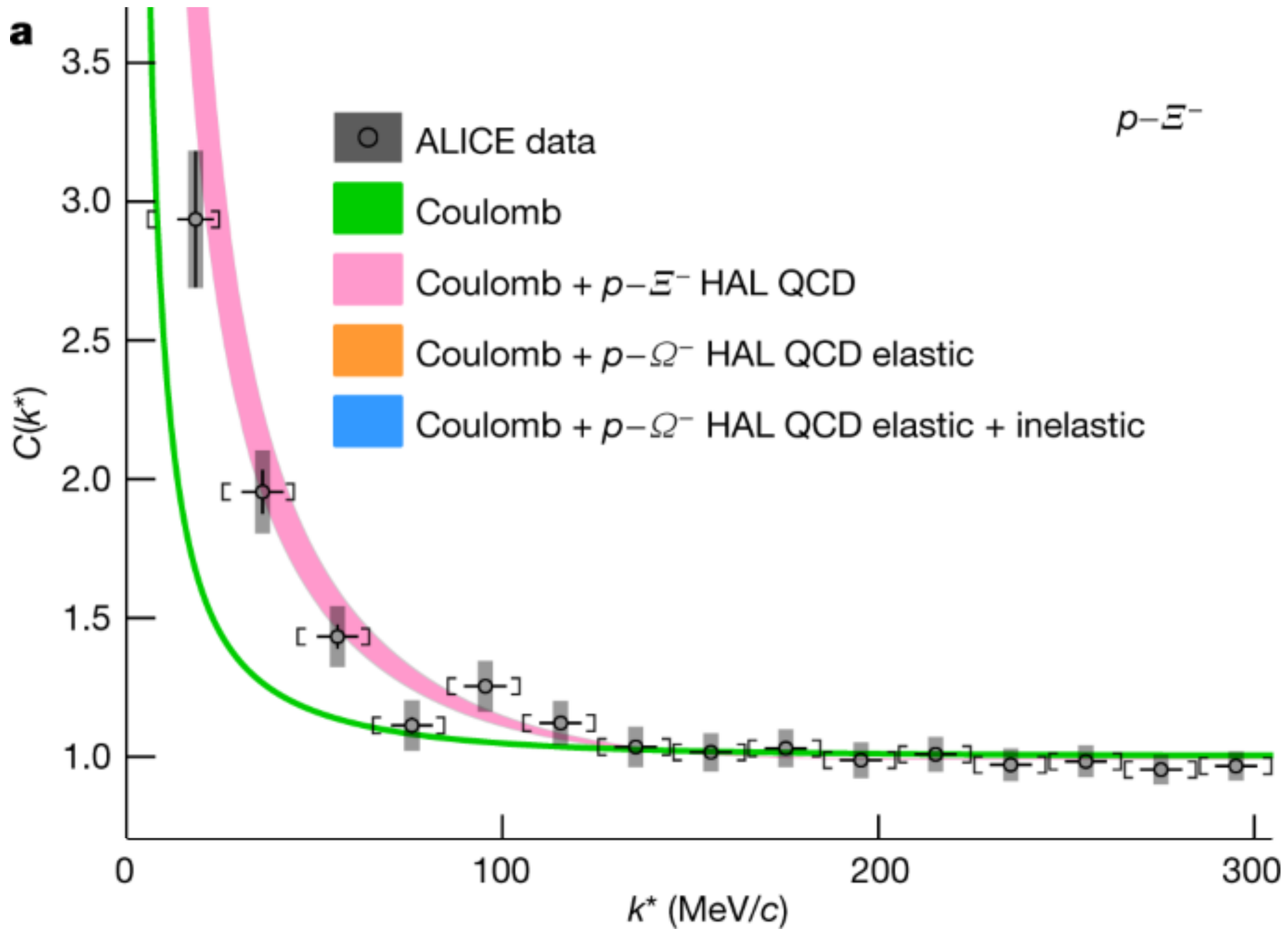


slides@T.Hatsuda



LHC@CERN, RHIC@BNL
J-PARC@KEK, FAIR@GSI, HIAF



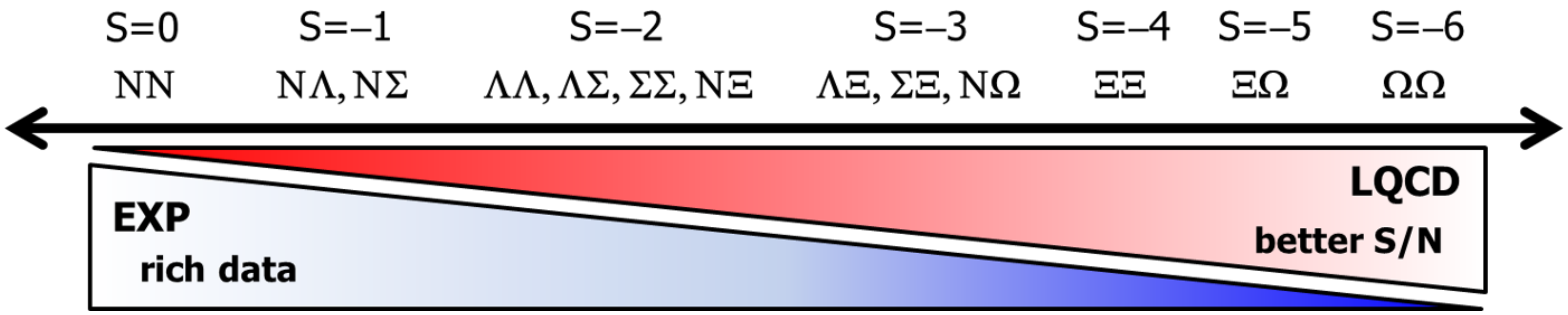


ALICE Collaboration, Nature 588, 232 (2020)

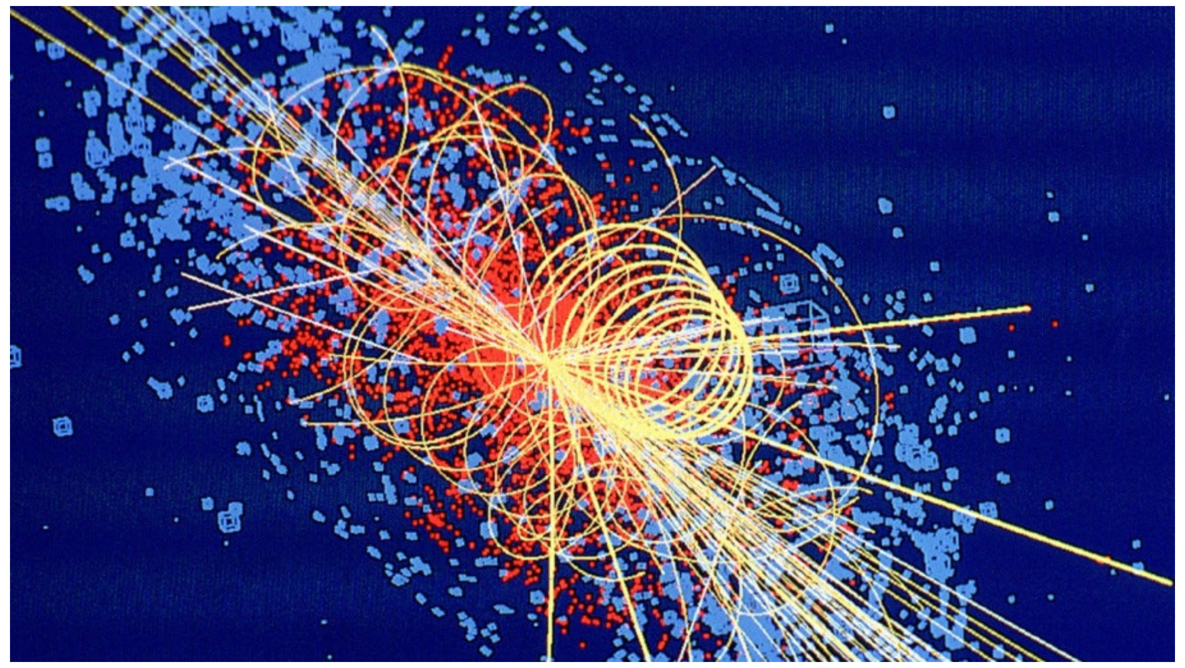
Hadrons to **A**tomic nuclei from **L**attice QCD
(**HAL** QCD Collaboration)

Hadrons to Atomic nuclei HAL from Lattice QCD

Hadron Interactions



LHC@CERN, RHIC@BNL
J-PARC@KEK, FAIR@GSI, HIAF



Femtoscopy
HAL QCD method

+

Deep Learning

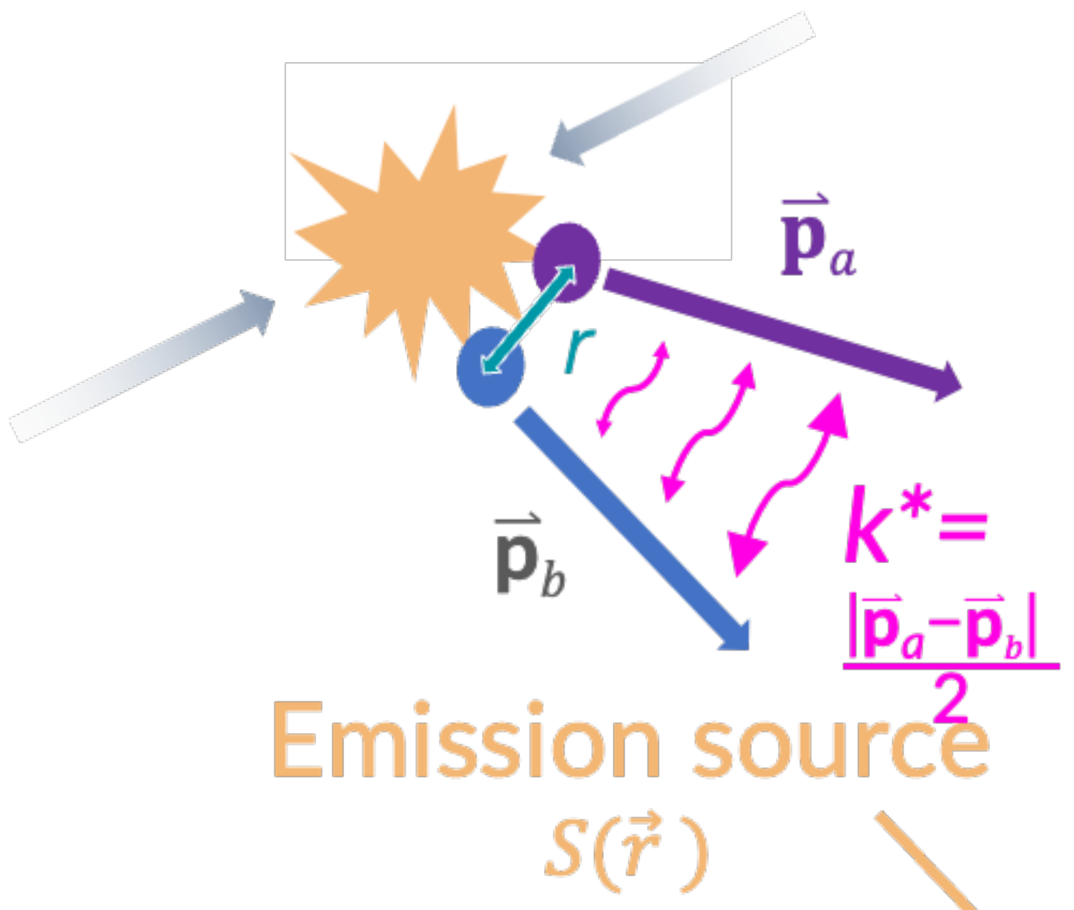
↓

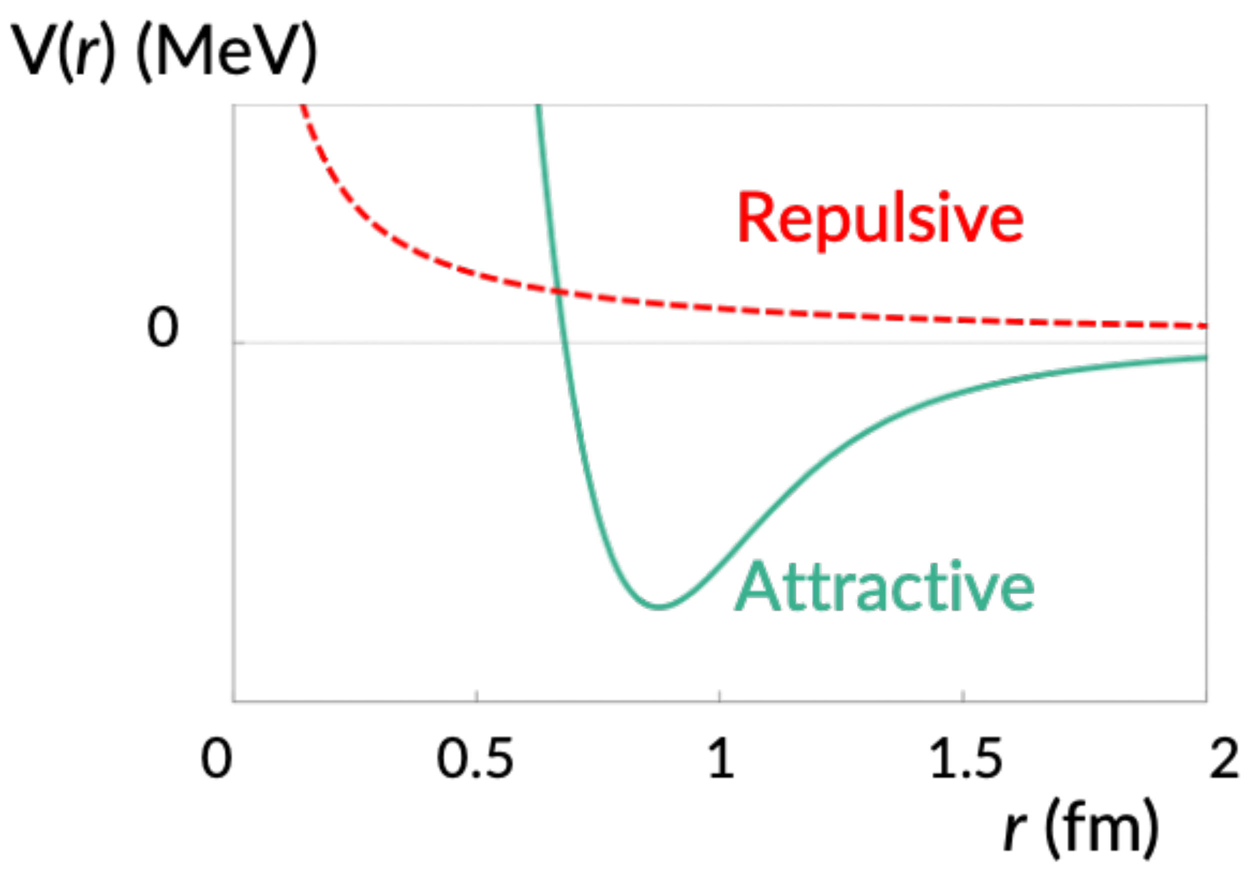
?

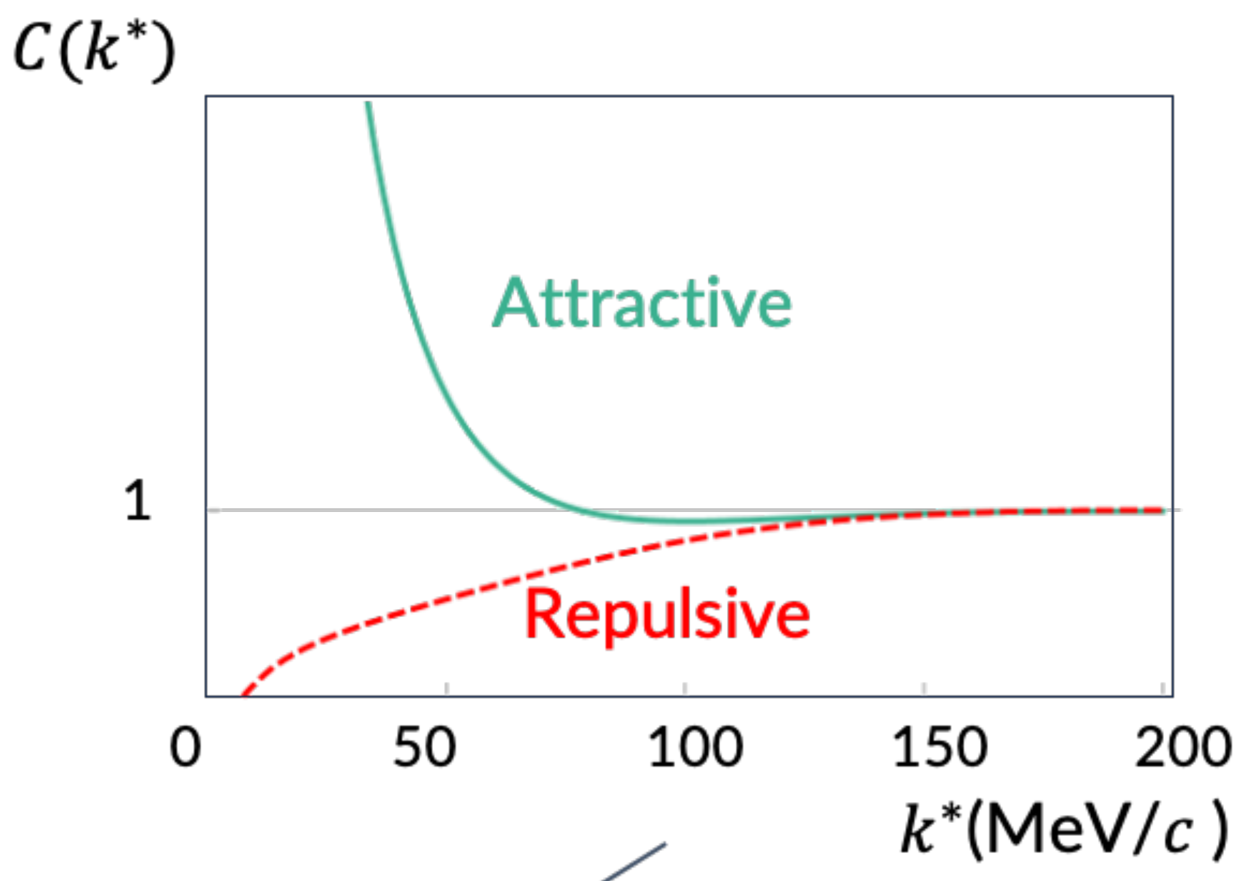
Hadrons to Atomic nuclei from Lattice QCD
(HAL QCD Collaboration)

Femtoscopy

Femtoscopy







Correlations

↑

Scattering Length
Effective Range

Two-particle wave function

$$C(k^*) = \int S(\vec{r})\,|\psi(\vec{k}^*,\vec{r})|^2\, d^3\vec{r} = \mathcal{N}(k^*)\frac{N_{\mathrm{same}}(k^*)}{N_{\mathrm{mixed}}(k^*)}$$

Raffaele Del Grande | XQCD 2023

Lednicky-Lyuboshits(LL) analytic model
(Asymptotic wave-function+
Effective range correlation+
Gaussian source)

Lednicky, Lyuboshits, Sov.J.Nucl.Phys. 35 (1982) 770

Femtoscopy

Asymptotic wave-function

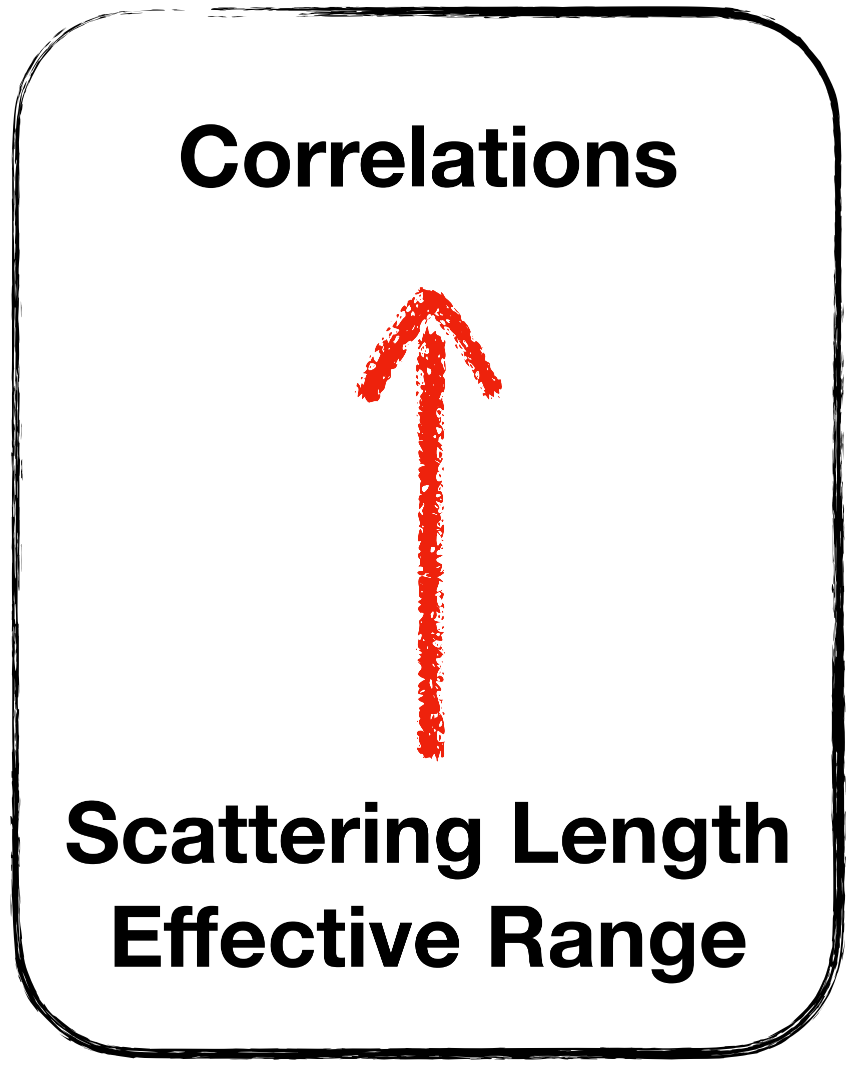
$$\psi_0(r) \to \psi_{\text{asy}}(r) = \frac{e^{-i\delta}}{qr}\sin(qr+\delta) = \mathcal{S}^{-1}\left[\frac{\sin qr}{qr} + f(q)\frac{e^{iqr}}{r}\right]$$

$$C_{\text{LL}}(q) = 1 + \int dr S_{12}(r)\left(|\psi_{\text{asy}}(r)|^2 - |j_0(qr)|^2\right)$$
$$= 1 + \frac{|f(q)|^2}{2R^2}F_3\left(\frac{r_{\text{eff}}}{R}\right) + \frac{2\text{Re}f(q)}{\sqrt{\pi}R}F_1(2x) - \frac{\text{Im}f(q)}{R}F_2(2x)$$

$x = qR$, R is Gaussian Size, F_1, F_2, F_3 are known functions

Scattering amplitude at low energies

$$q\cot\delta = -\frac{1}{a_0} + \frac{1}{2}r_{\text{eff}}q^2 + O(q^4) \to f(q) = (q\cot\delta - iq)^{-1}$$



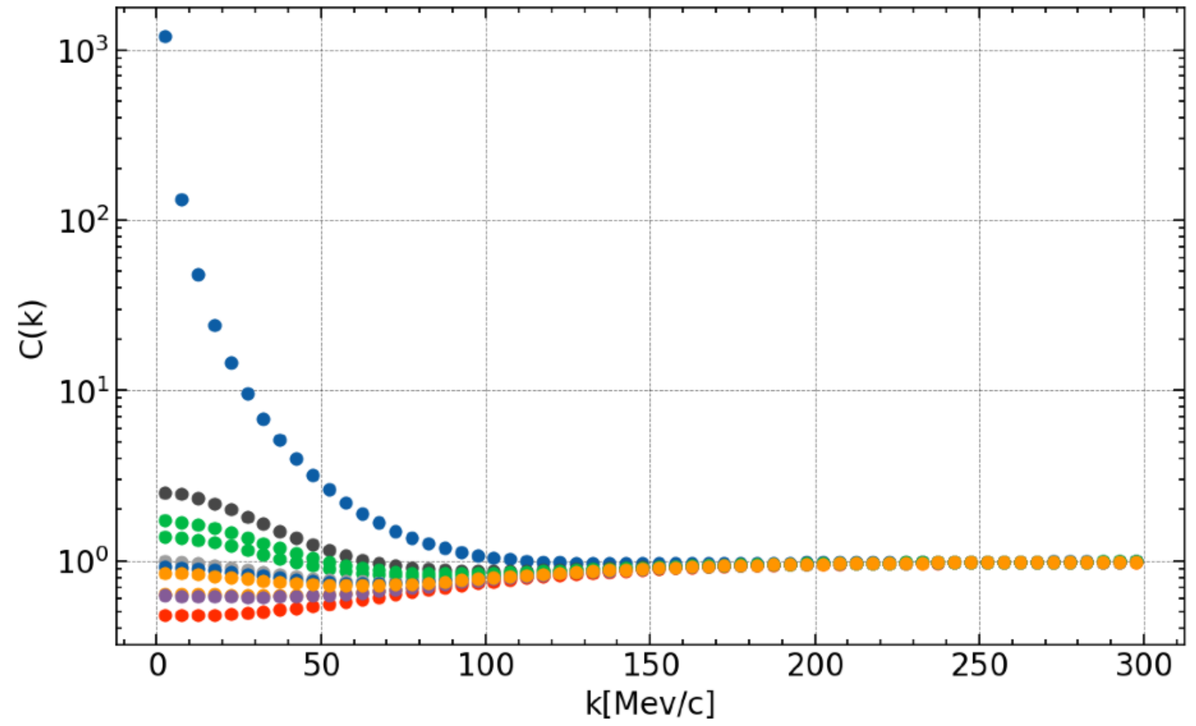
Lednicky-Lyuboshits(LL) analytic model
(Asymptotic wave-function+
Effective range correlation+
Gaussian source)

Lednicky, Lyuboshits, Sov.J.Nucl.Phys. 35 (1982) 770

Inverse Femtoscopy

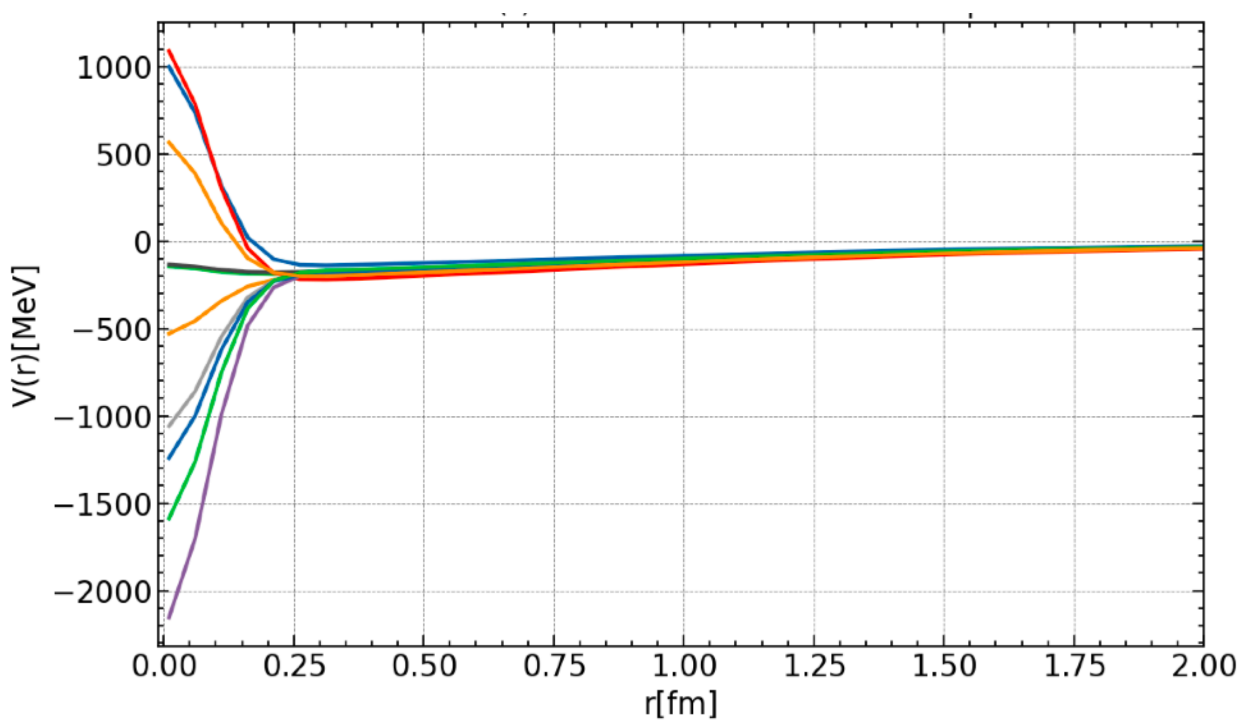
in Preparation **with Liang Zhang, Jiaxing Zhao, etc.**

$$C(k^*) = \int S(\vec{r}) \left| \psi(\vec{k}^*, \vec{r}\,) \right|^2 d^3\vec{r} \; = \mathcal{N}(k^*) \frac{N_{\text{same}}(k^*)}{N_{\text{mixed}}(k^*)}$$



Does this inverse mapping exist?

$$V(r)$$



Correlations

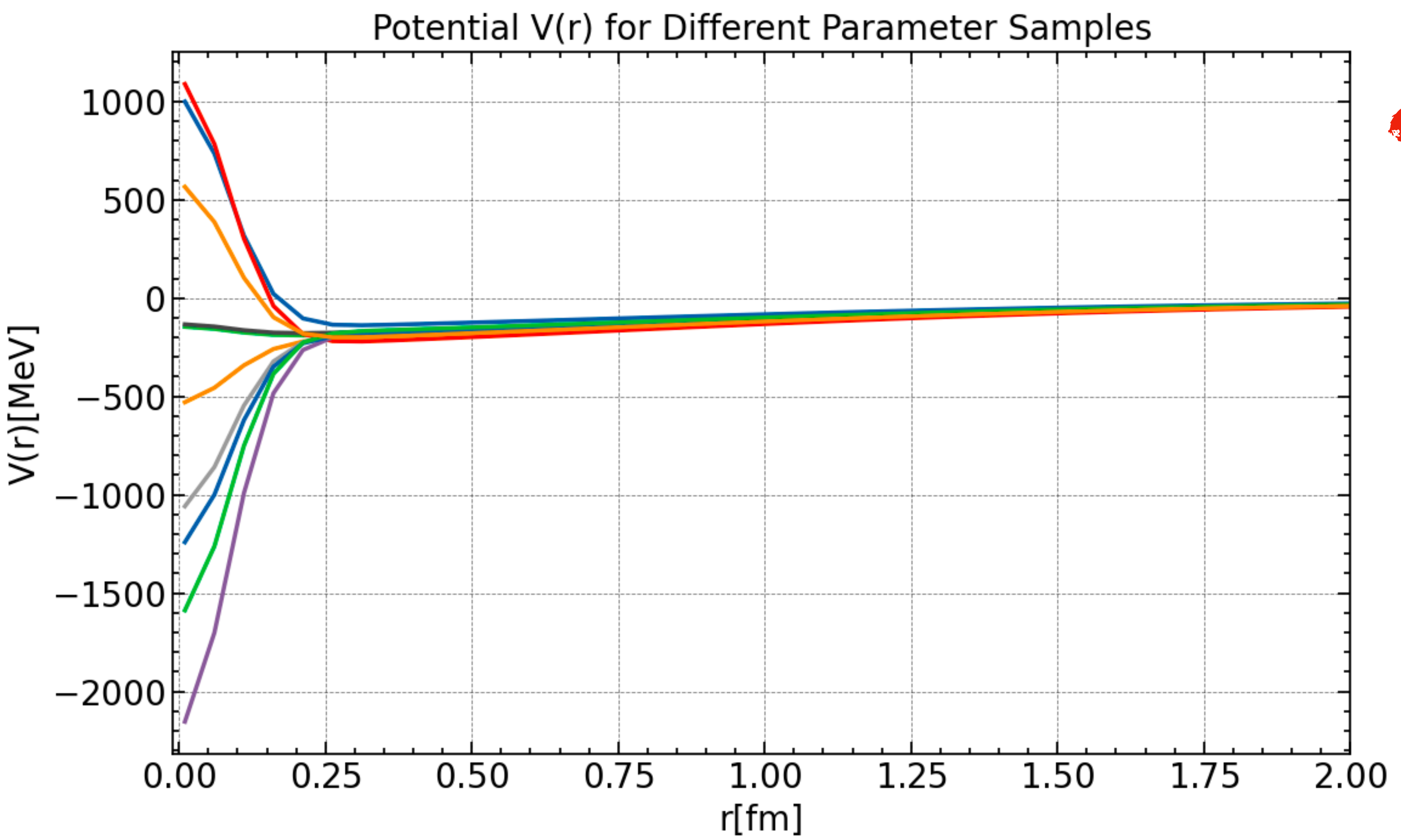
↓ **?**

Potentials

Inverse Femtoscopy

Potential Functions

$$V(r) = b_1 e^{-b_2 r^2} + b_3(1 - e^{-b_4 r^2})(\frac{e^{(-m_\pi r)}}{r})^{n_\pi}$$



Source Function

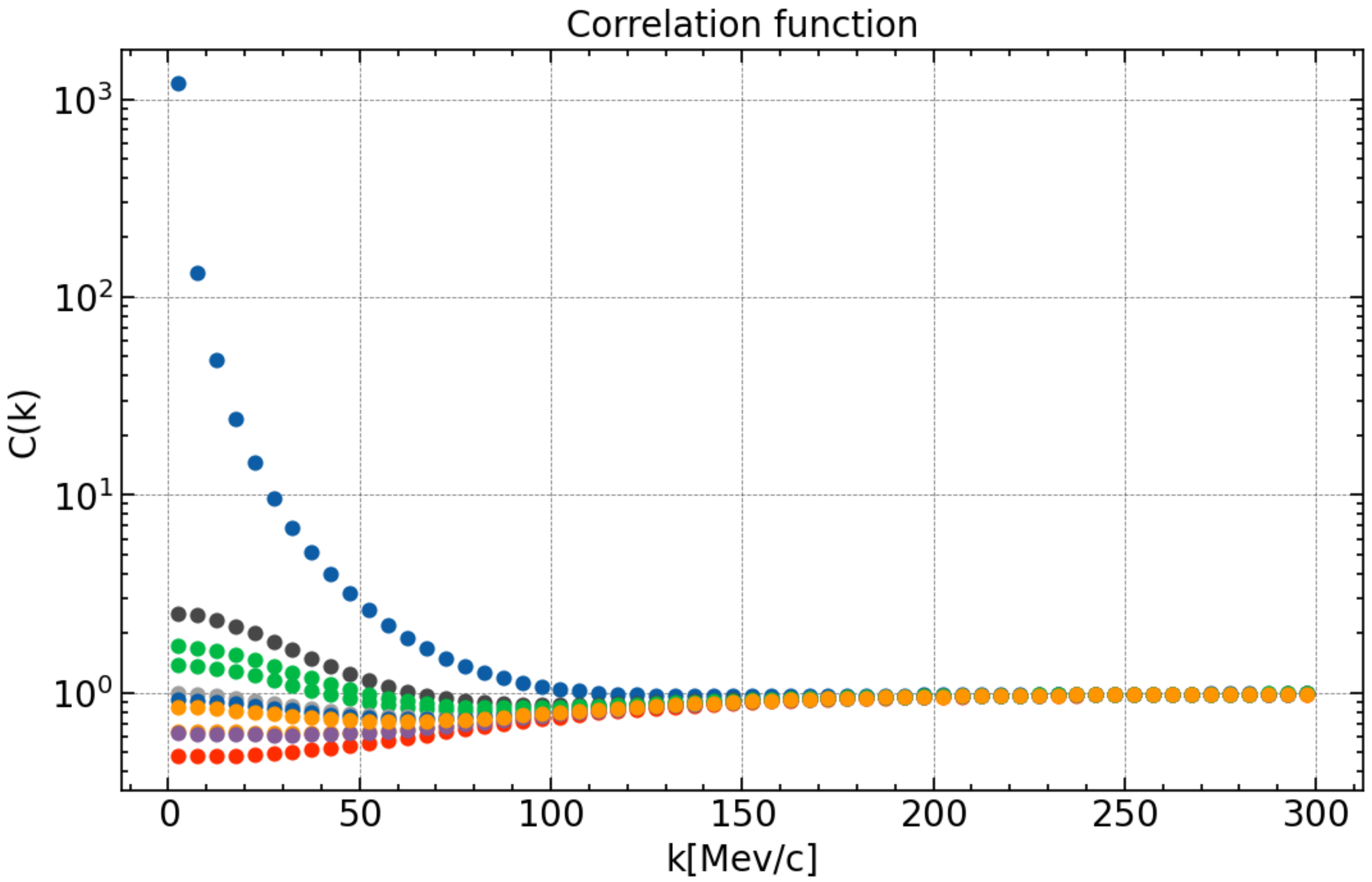
$$S(r) = (4\pi r_0^2)^{-3/2} e^{-\frac{r^2}{4r_0^2}} \qquad r_0 = 1.3 \text{ fm}$$

in Preparation **with Liang Zhang, Jiaxing Zhao, etc.**

Deep Neural Network (DNN)

Schrödinger eq.

CATS Framework: D. Mihaylov et al., Eur. Phys. J. C78 (2018) 394

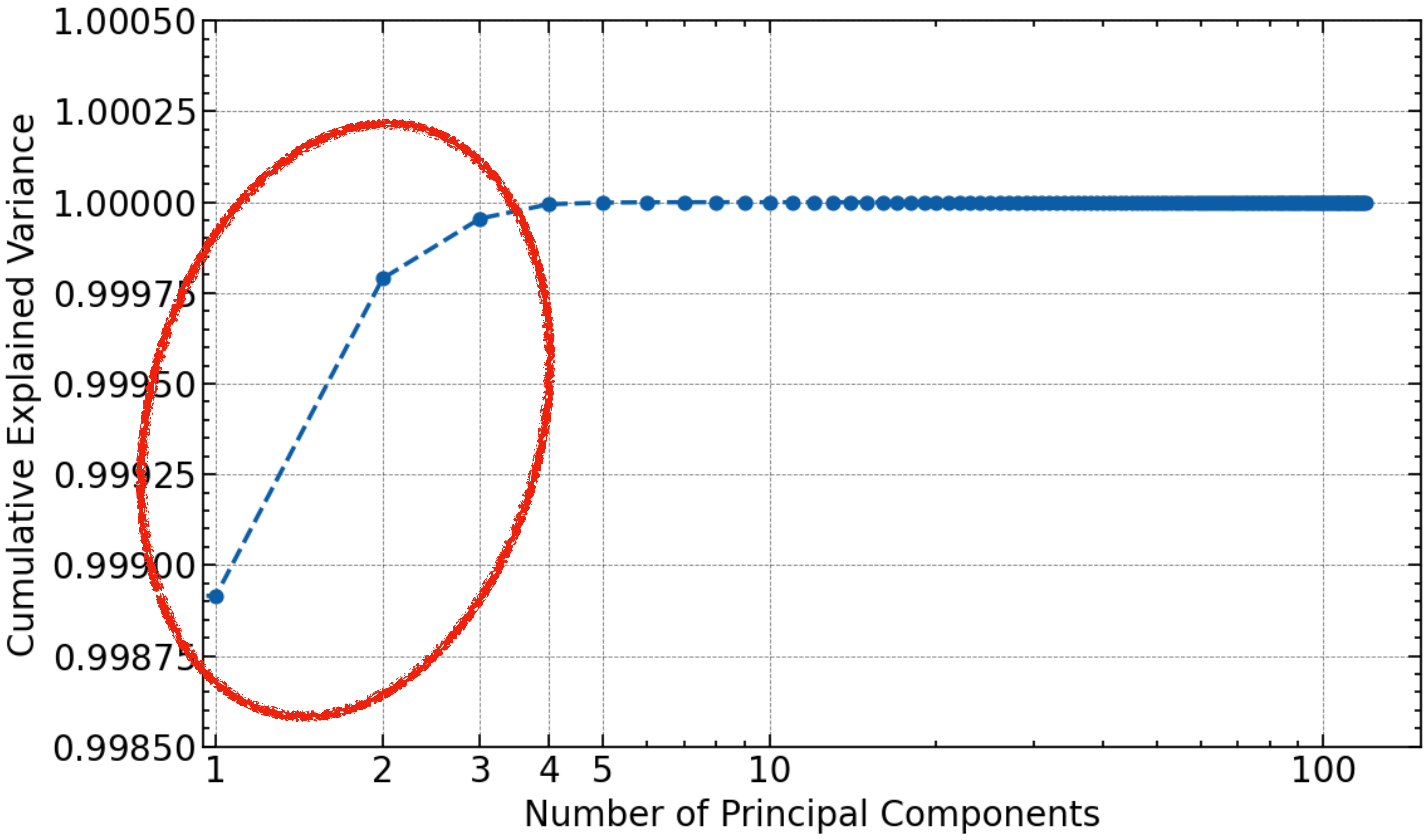


60 points(k), N_C correlations

Inverse Femtoscopy

in Preparation **with Liang Zhang, Jiaxing Zhao, etc.**

Principal Component Analysis(PCA)



$N_c = 25000$

r_0: $[0.52, 4.16]$ **fm**

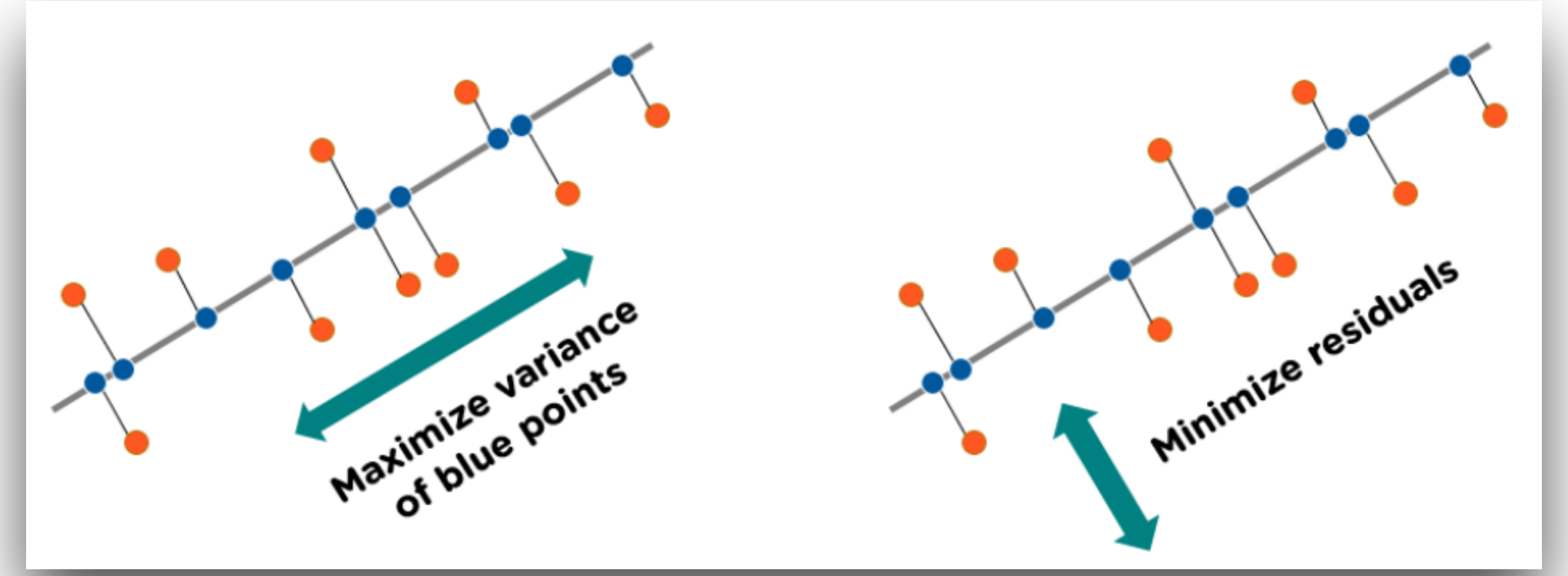
b_1: $[-2145.5, 1532.5]$ **MeV**

b_2: $[0.739, 147.761]$ **fm**$^{-2}$

b_3: $[-1064, 532]$ **MeV** $\cdot$ **fm**^{2}

b_4: $[0.078, 154.422]$ **fm**$^{-2}$

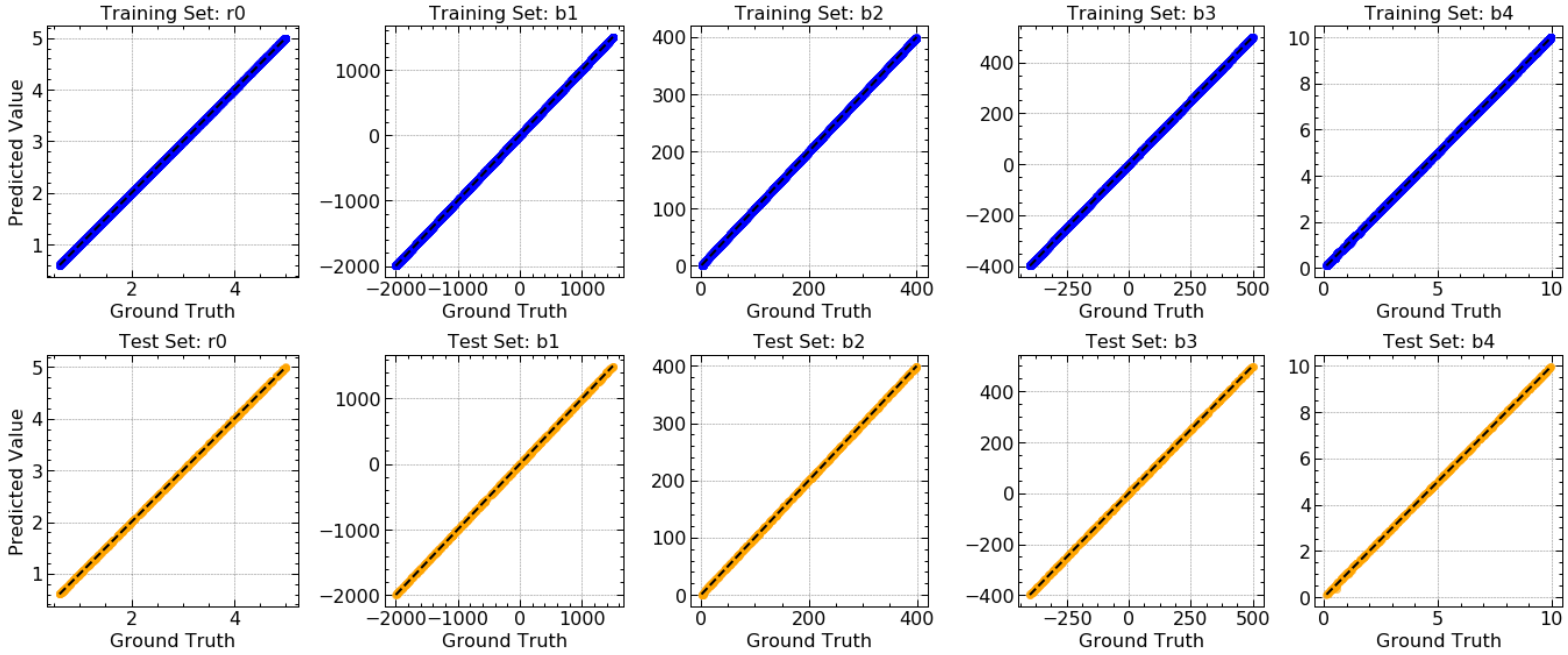
n_π: 2



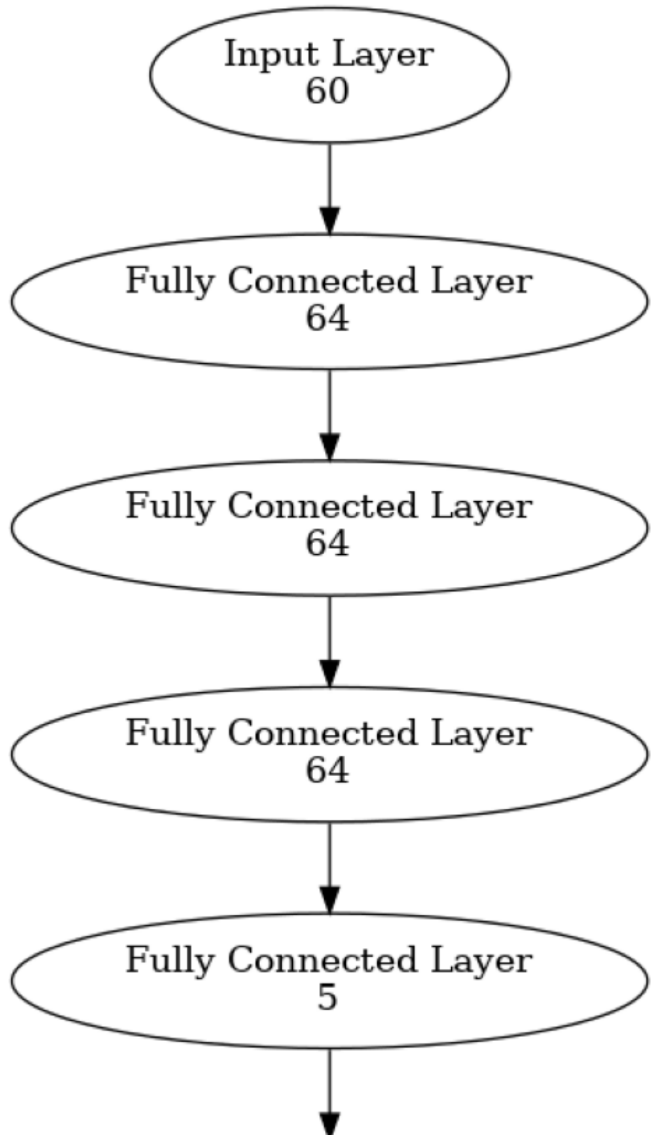
Inverse Femtoscopy

$N_c = 6400$ $r_0 = 1.3$ **fm**, $b_1 = -306.5, b_2 = 200, b_3 = -266, b_4 = 0.78, n_\pi = 2$

in Preparation **with Liang Zhang, Jiaxing Zhao, etc.**



60 points

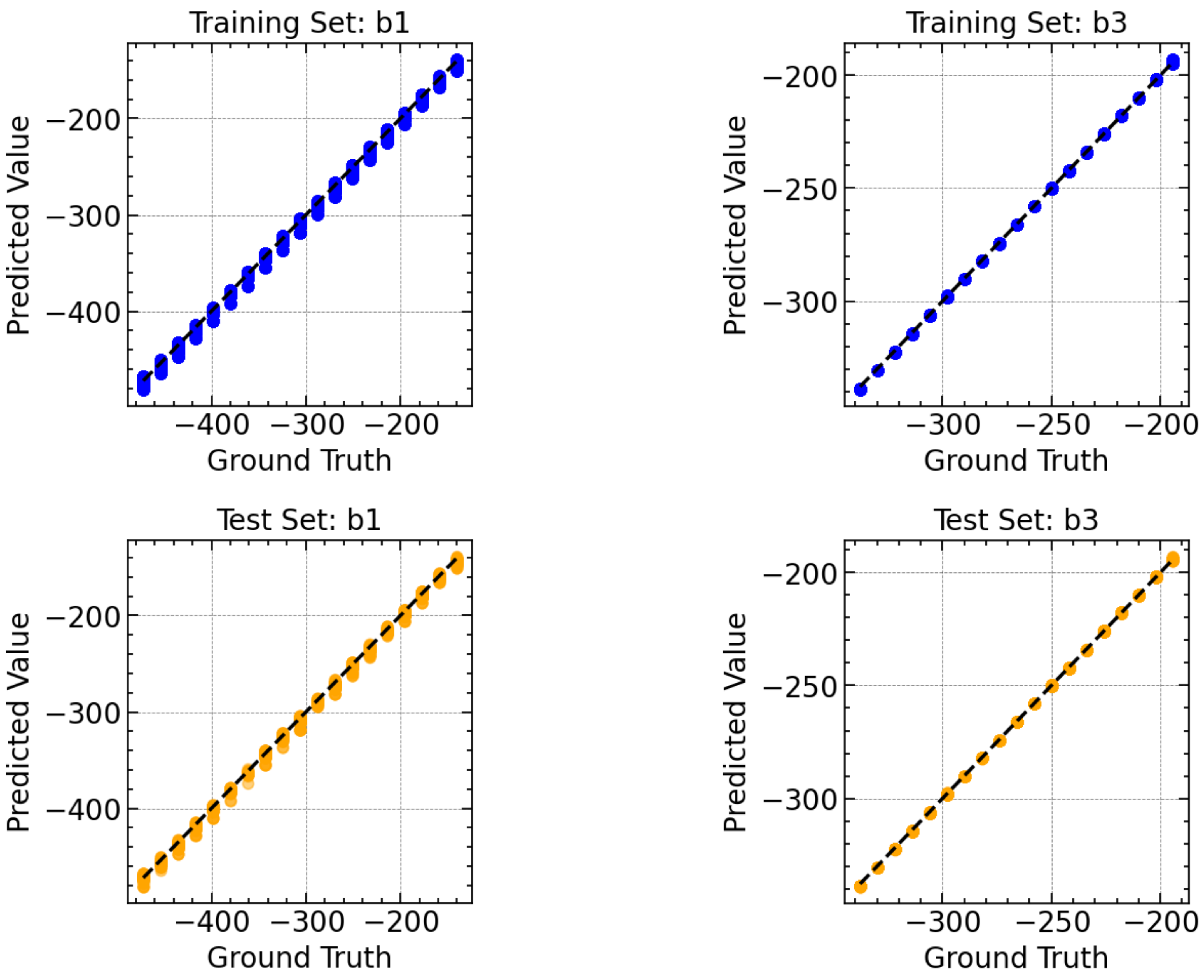


1 parameter

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}$$

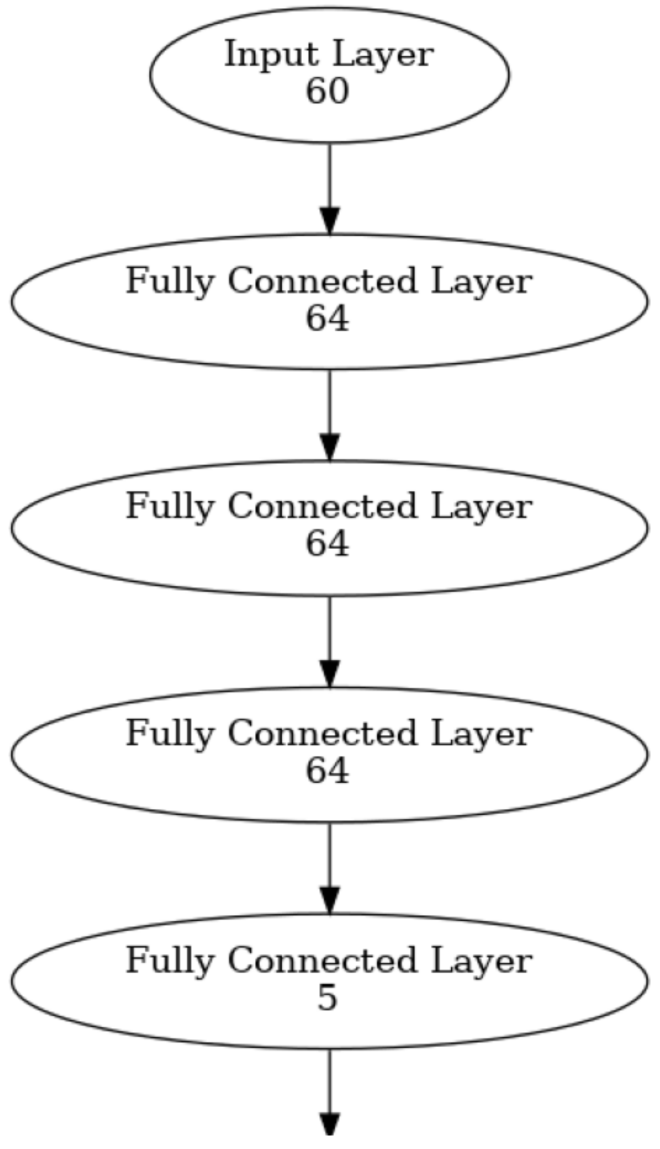
R-squared: 0.99, 0.99, 0.99, 0.99, 0.99

Inverse Femtoscopy



R-squared	b1	b3
Training	0.99	0.99
Testing	0.99	0.99

in Preparation **with Liang Zhang, Jiaxing Zhao, etc.**



60 points

b_1, b_3

2 parameters

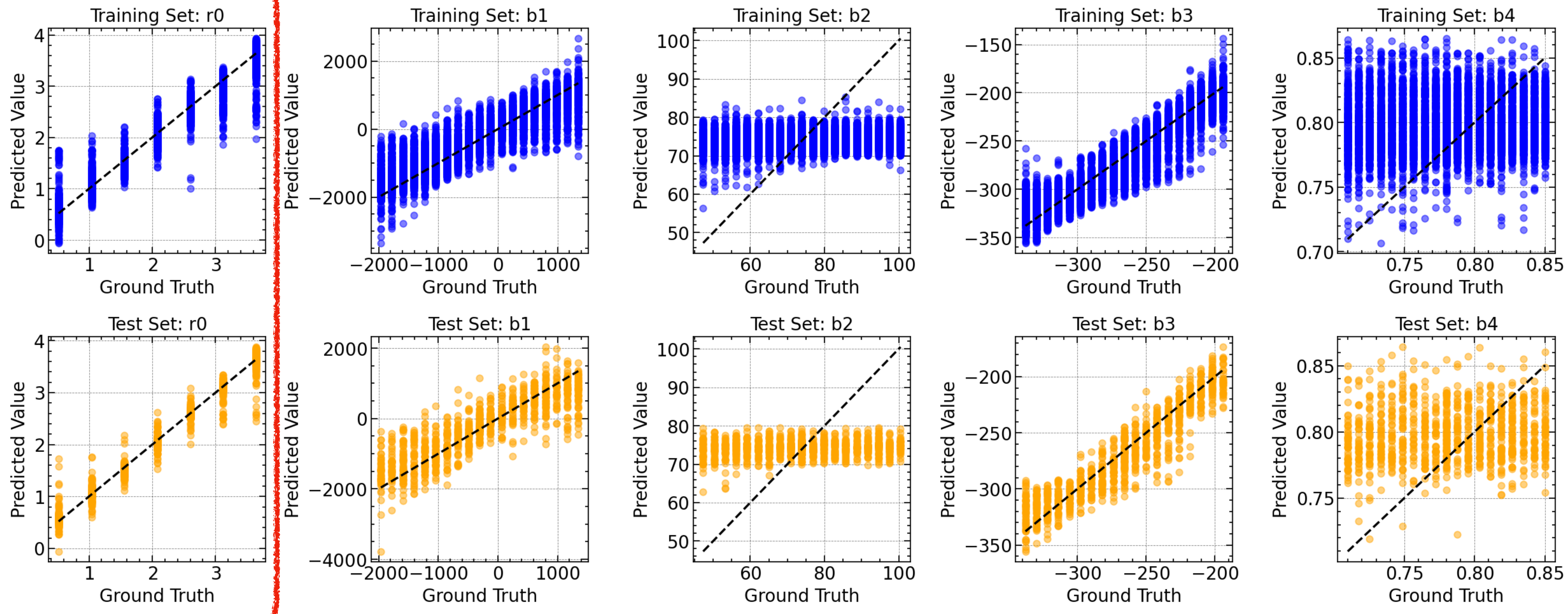
$r_0 = 1.3$ **fm**, $b_2 = 73.9, b_4 = 0.78, n_\pi = 2$

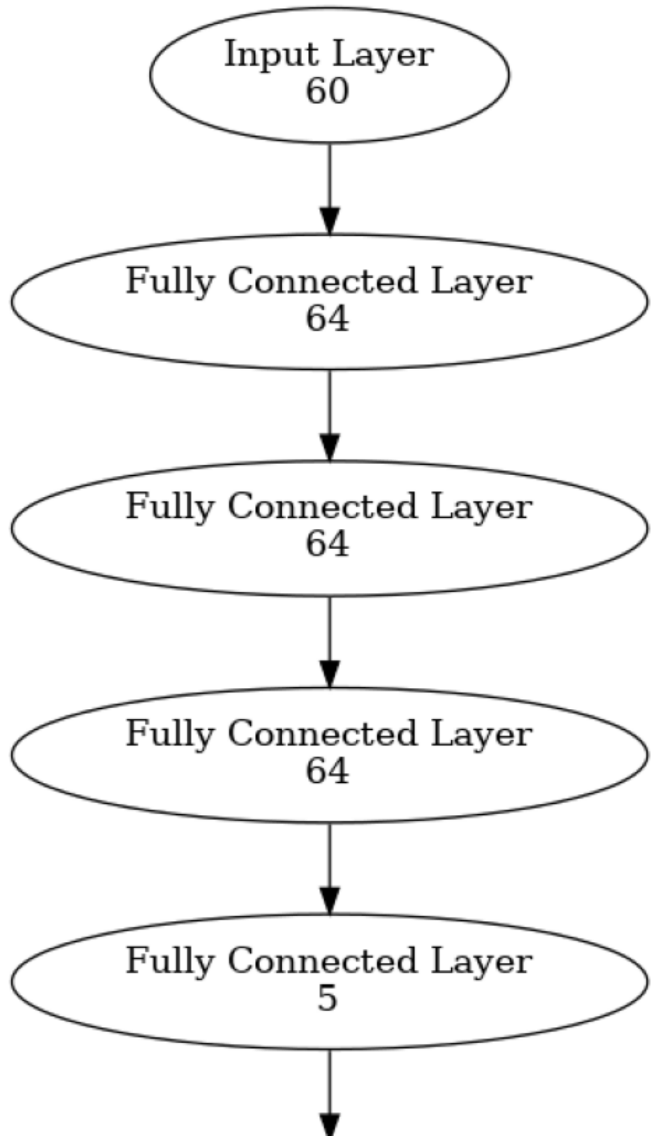
$N_c = 10000$

Inverse Femtoscopy

$N_c = 10000$ $n_\pi = 2$

in Preparation **with Liang Zhang, Jiaxing Zhao, etc.**





60 points

5 parameters

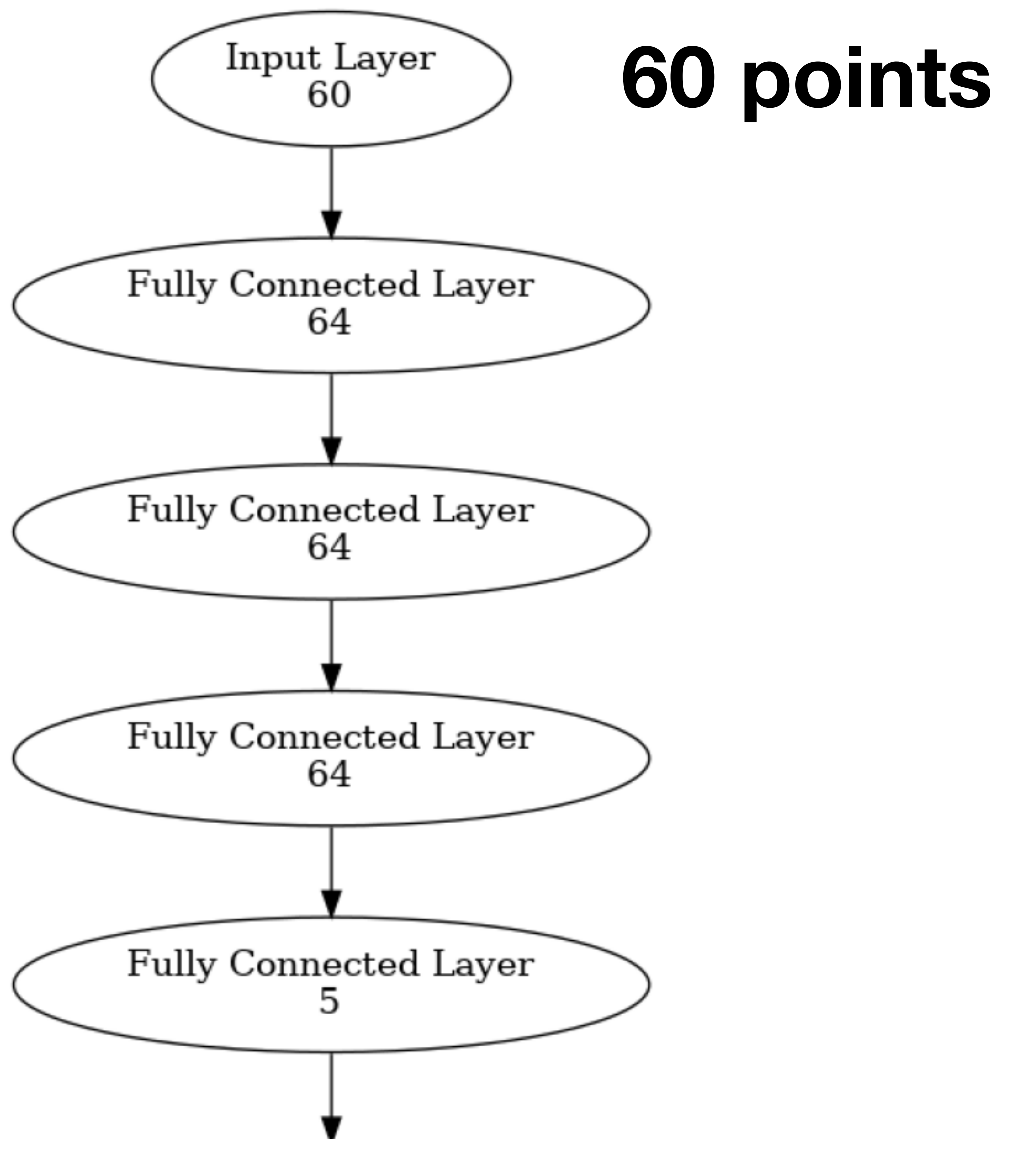
R-squared	r0	b1	b2	b3	b4
Training	0.99	0.86	0.02	0.90	0.00
Testing	0.99	0.86	0.02	0.90	0.00

Potential Functions

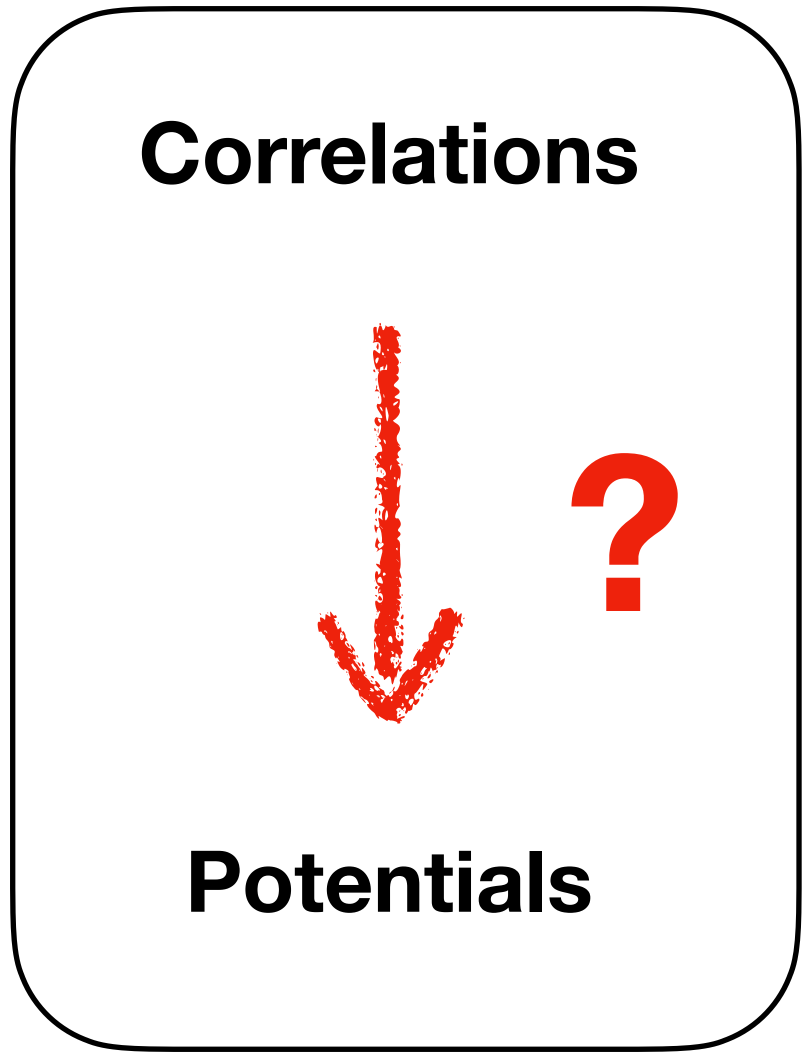
$$V(r) = b_1 e^{-b_2 r^2} + b_3(1 - e^{-b_4 r^2})(\frac{e^{(-m_\pi r)}}{r})^{n_\pi}$$

Summary I

- **Take-Home Messages**
  - PCA gives ≤ 3 main components
  - DNNs learn the inverse mapping successfully
  - Source size can also be identified
- Future Works
  - Flexible input length ✅
  - Different parameterizations ✅
  - Experimental data 💪
  - Bayesian inference 💪
  - Input beyond correlation functions 💭



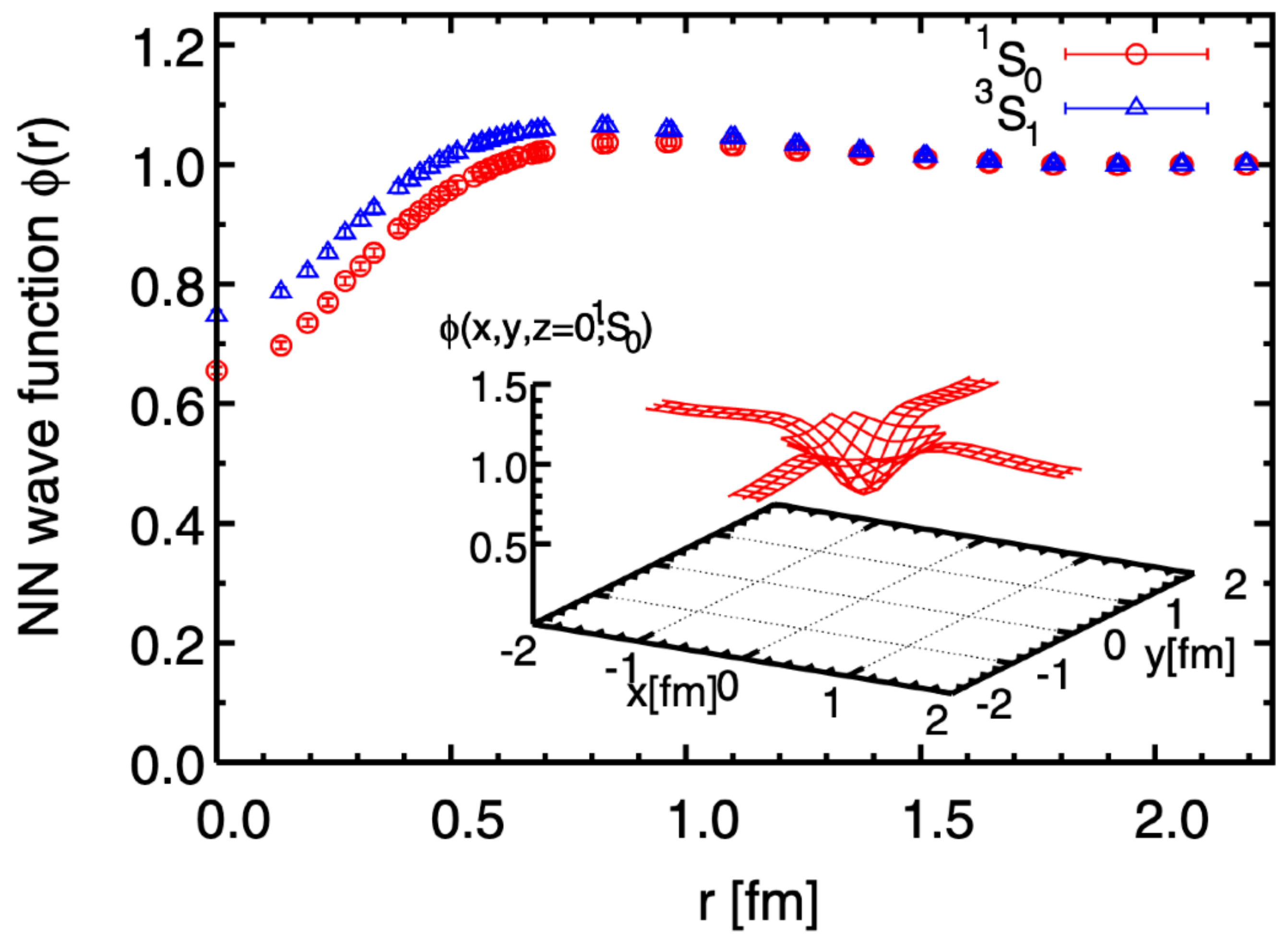
Parameters of potential functions

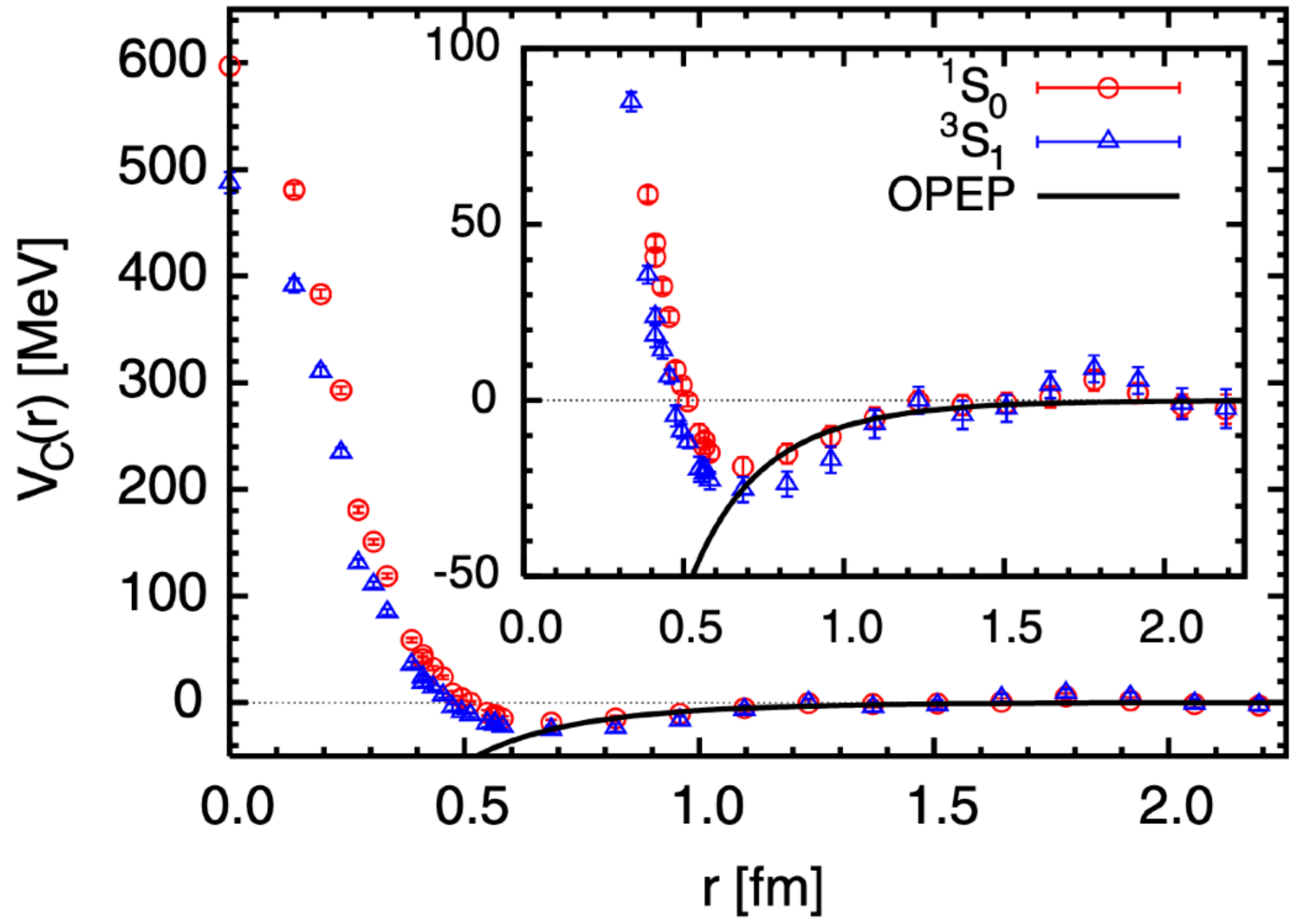


HAL QCD
meets DNNs

HAL QCD method

N. Ishii, S. Aoki, and T. Hatsuda, Phys. Rev. Lett. 99, 022001 (2007)
S. Aoki, T. Hatsuda, and N. Ishii, Prog. Theor. Phys. 123, 89 (2010)
Aoki, S., Doi, T., Front. in Phys. 8, 307 (2020)
S. Aoki and T. Doi, in Handbook of Nuclear Physics(2023), pp. 1–31





Local Approx.
Derivative Expansion

Nambu-Bethe-Salpeter (NBS) wave function

$$\psi_{NBS}(\vec{r}) = \langle 0|N(\vec{r})N(\vec{0})|N(\vec{k})N(-\vec{k}), in\rangle$$
$$\simeq e^{i\delta_l(k)} \sin(kr - l\pi/2 + \delta_l(k))/(kr)$$

(at asymptotic region)

HAL QCD method

Nulcear Force

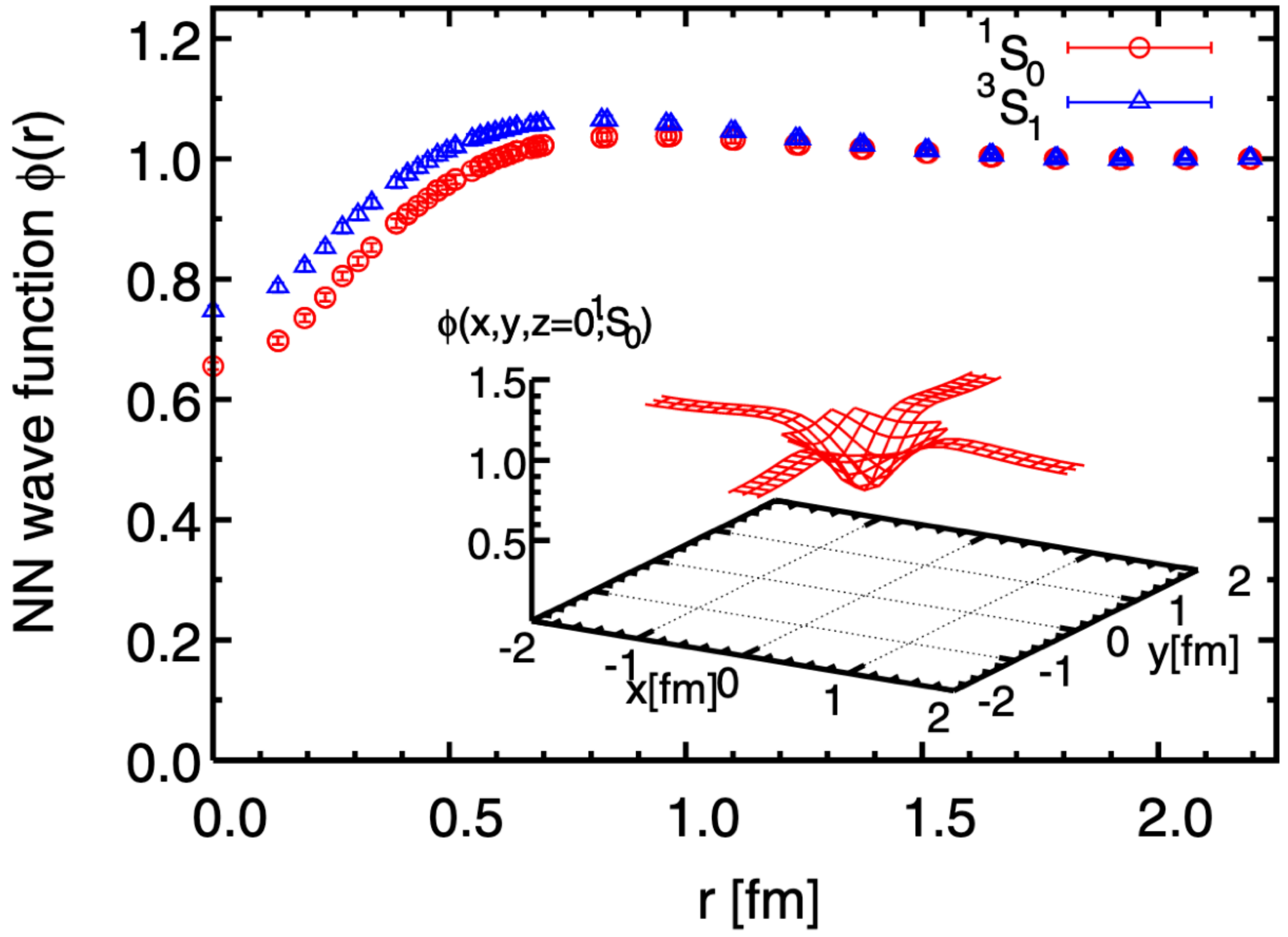
$$(k^2/m_N - H_0)\,\psi_{NBS}(\vec{r}) = \int d\vec{r}'\, U(\vec{r},\vec{r}')\psi_{NBS}(\vec{r}')$$

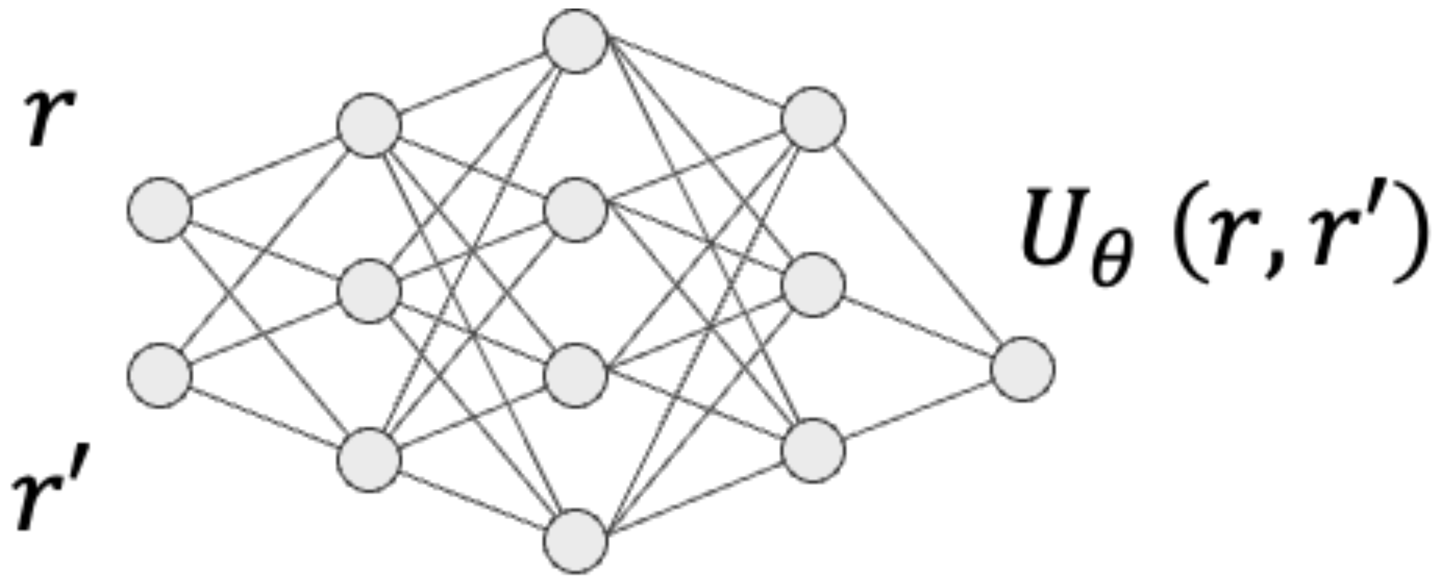
(Schrodinger eq.)

Kotaro's talk on Monday
Separable potential: L. Meng & Epelbaum (2023); K. Murakami and S.Aoki (2024).

HAL QCD: Inverse Problem Perspective

arXiv:2410.03082 (with HAL QCD)





Universal Approximation Theorem (1989,1991)

$$\theta_{i+1} \to \theta_i + \frac{\partial \mathcal{L}}{\partial U_\theta(r,r')} \frac{\partial U_\theta(r,r')}{\partial \theta}$$

Gradient Decent

NBS wave function → **Potential Function**

Data(Observations) ← Physics Properties

Maximize Likelihood Estimation

$$\mathbf{min}_\theta \mathcal{L} = \sum_k \int d^3\mathbf{r} \left[\left(E_k - H_0\right) \phi_\mathbf{k}(\mathbf{r}) - \int d^3\mathbf{r}' U_\theta(\mathbf{r},\mathbf{r}')\phi_\mathbf{k}(\mathbf{r}') \right]^2$$

Physics-Driven Deep Learning

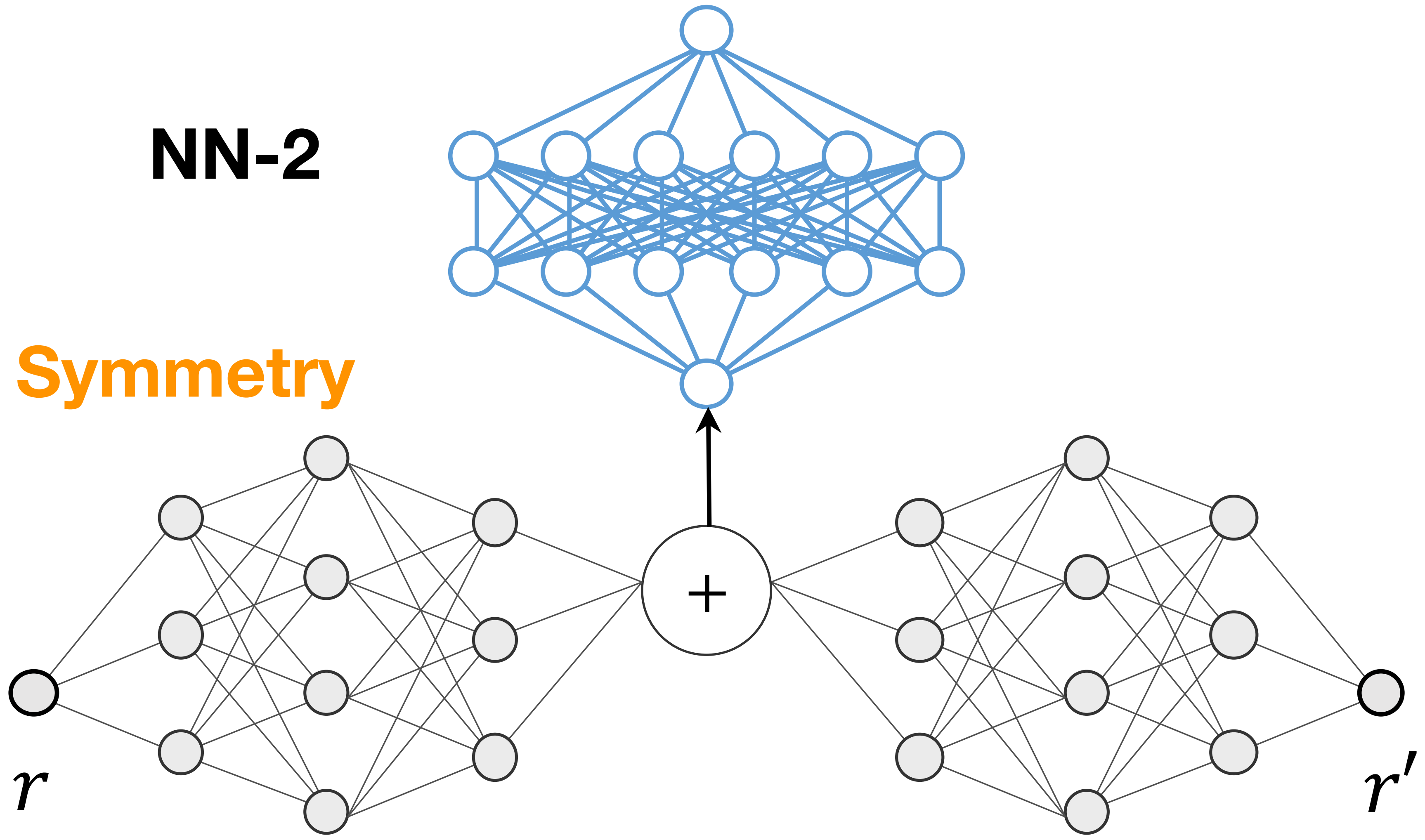
arXiv:2410.03082 (with HAL QCD)

Two particle interactions

$$U_\theta(r, r') \equiv g(f(r) + f(r'))$$

$U_\theta\,(r, r')$

NN-2



NN-1 **NN-1**

a. Permutation Symmetry

b. Asymptotic Behaviour as regulator

$$\lim_{\mathbf{r}>R, \mathbf{r}'>R} U_\theta(\mathbf{r}, \mathbf{r}') \to 0$$

Back-Propagation

Residual of Schrödinger Eq.

$$(E_k - H_0)\phi_{\mathbf{k}}(\mathbf{r}) = \int d^3 r' U(\mathbf{r}, \mathbf{r}')\phi_{\mathbf{k}}(\mathbf{r}')$$

$$E_k = \frac{k^2}{2m}, \quad H_0 = -\frac{\nabla^2}{2m}, \quad m = \frac{m_N}{2}$$

$\phi_k\,(r)$

or

$R(t, r)$

Phys. Lett. B 712, 437 (2012)

$$\left\{ \frac{1}{4m_N}\frac{\partial^2}{\partial t^2} - \frac{\partial}{\partial t} - H_0 \right\} R(t, r) = \int 4\pi r'^2 dr'\, U(r, r') R(t, r')$$

Mock Test: Separable Potential

As a numerical example, we take $\mu = 1.0, \omega = -0.017\mu^4, m = 3.30\mu, R = 2.5/\mu$

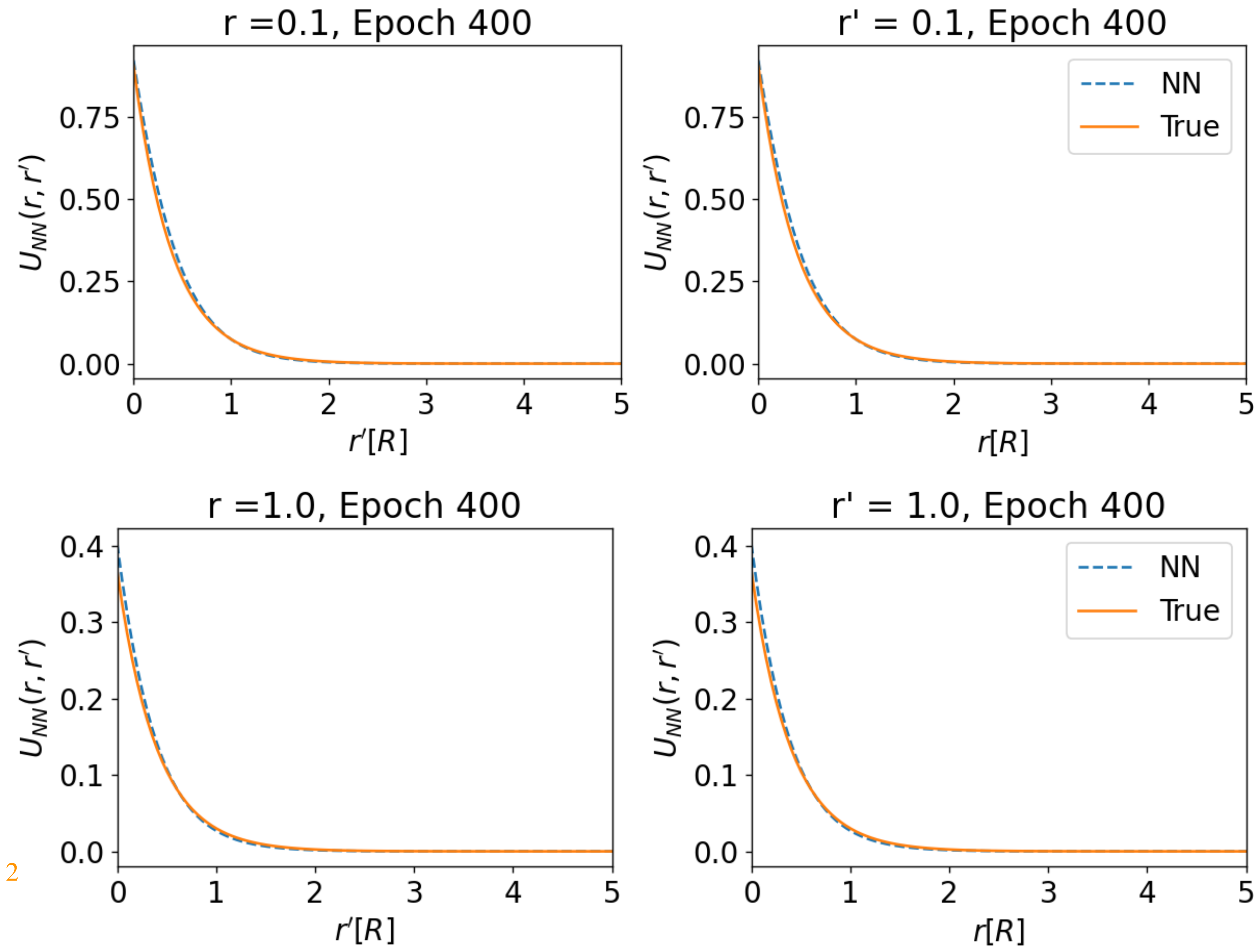
arXiv:2410.03082 (with HAL QCD)

$$U(\mathbf{r},\mathbf{r}') \equiv \omega \nu(\mathbf{r})\nu(\mathbf{r}'), \quad \nu(\mathbf{r}) \equiv e^{-\mu r}$$

$$U_{\mathbf{NN}}(r,r') = \omega f_\theta(r,r')$$

$$\phi_k^0(r) = \frac{e^{i\delta_0(k)}}{kr}\left[\sin\{kr+\delta_0(k)\} - \sin\delta_0(k)e^{-\mu r}\left(1+\frac{r(\mu^2+k^2)}{2\mu}\right)\right]$$

$$k\cot\delta_0(k) = -\frac{1}{4\mu^2}\left[2\mu(\mu^2-k^2) - \frac{3\mu^2+k^2}{4\mu^3}(\mu^2+k^2)^2 + \frac{(\mu^2+k^2)^4}{8\pi m\omega}\right]$$



Nambu-Bethe-Salpeter (NBS) wave function

$$\phi_{\mathbf{k}}(\mathbf{r})e^{-W_{\mathbf{k}}t} \equiv \langle 0 | N(\mathbf{x}+\mathbf{r},t)N(\mathbf{x},t) | NN, W_k\rangle$$

BP

$$(E_k - H_0)\phi_{\mathbf{k}}(\mathbf{r}) = \int d^3r' U(\mathbf{r},\mathbf{r}')\phi_{\mathbf{k}}(\mathbf{r}')$$

$$E_k = \frac{k^2}{2m}, \quad H_0 = -\frac{\nabla^2}{2m}, \quad m = \frac{m_N}{2}$$

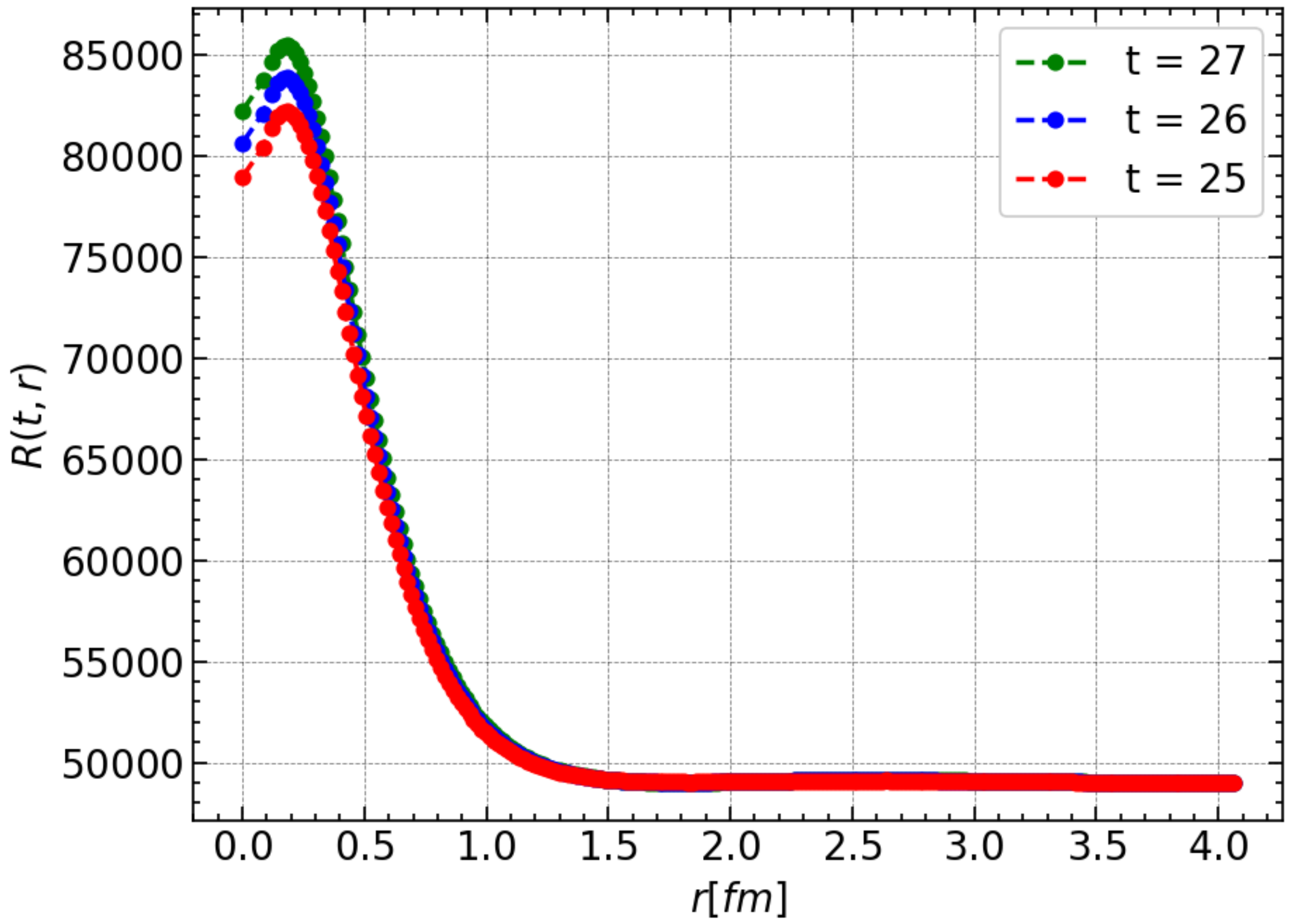
$$\mathscr{L} = \sum_k \int d^3r \left[\left(E_k - H_0\right)\phi_{\mathbf{k}}(\mathbf{r}) - \int d^3r' U_\theta(\mathbf{r},\mathbf{r}')\phi_{\mathbf{k}}(\mathbf{r}')\right]^2$$

$$U(r > 3R, r' > 3R) \to 0$$

$\Omega_{ccc}\Omega_{ccc}$ Interaction: ${}^1S_0$

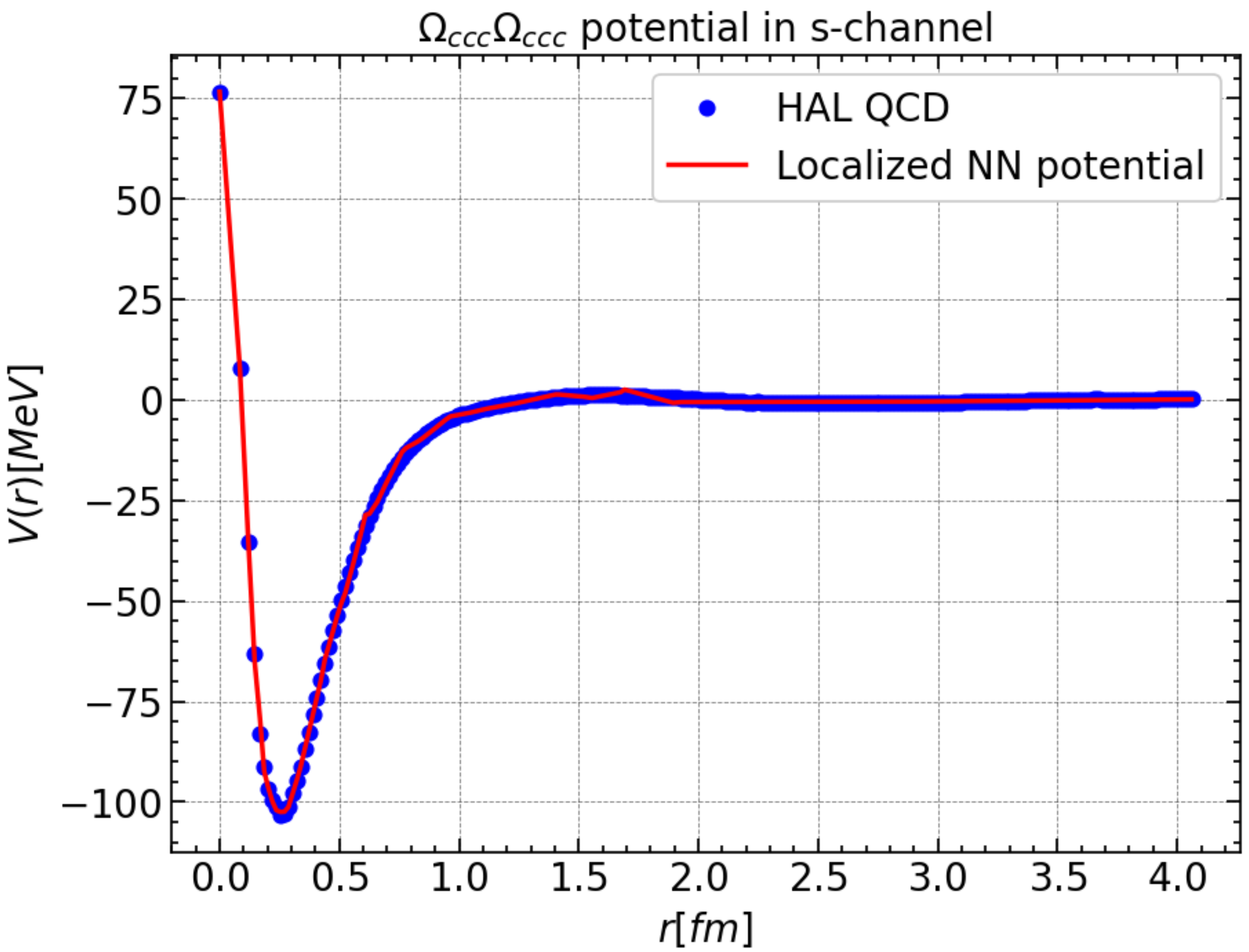
Y. Lyu, etc., Phys. Rev. Lett. 127, 072003 (2021)

$m_N = 2.073, a^{-1} = 2333.0$**MeV**



arXiv:2410.03082 (with HAL QCD)

Neural Network Hadron Force



$$\left\{ \frac{1}{4m_N}\frac{\partial^2}{\partial t^2} - \frac{\partial}{\partial t} - H_0 \right\} R(t,r) = \int 4\pi r'^2 dr'\, U(r,r') R(t,r')$$

$$R2 = R_{t+1} - 2R_t + R_{t-1}, R1 = (R_{t+1} - R_{t-1})/2, Rr = \nabla^2 R(t,r)$$

BP

Nambu-Bethe-Salpeter (NBS) wave function

$$\mathcal{L} = \sum_t \left\{ \frac{1}{4m_N} R2(t,r) - R1(t,r) + \frac{1}{m_N} Rr(t,r) - \int 4\pi r'^2 dr'\, U_\theta(r,r') R(t,r') \right\}$$

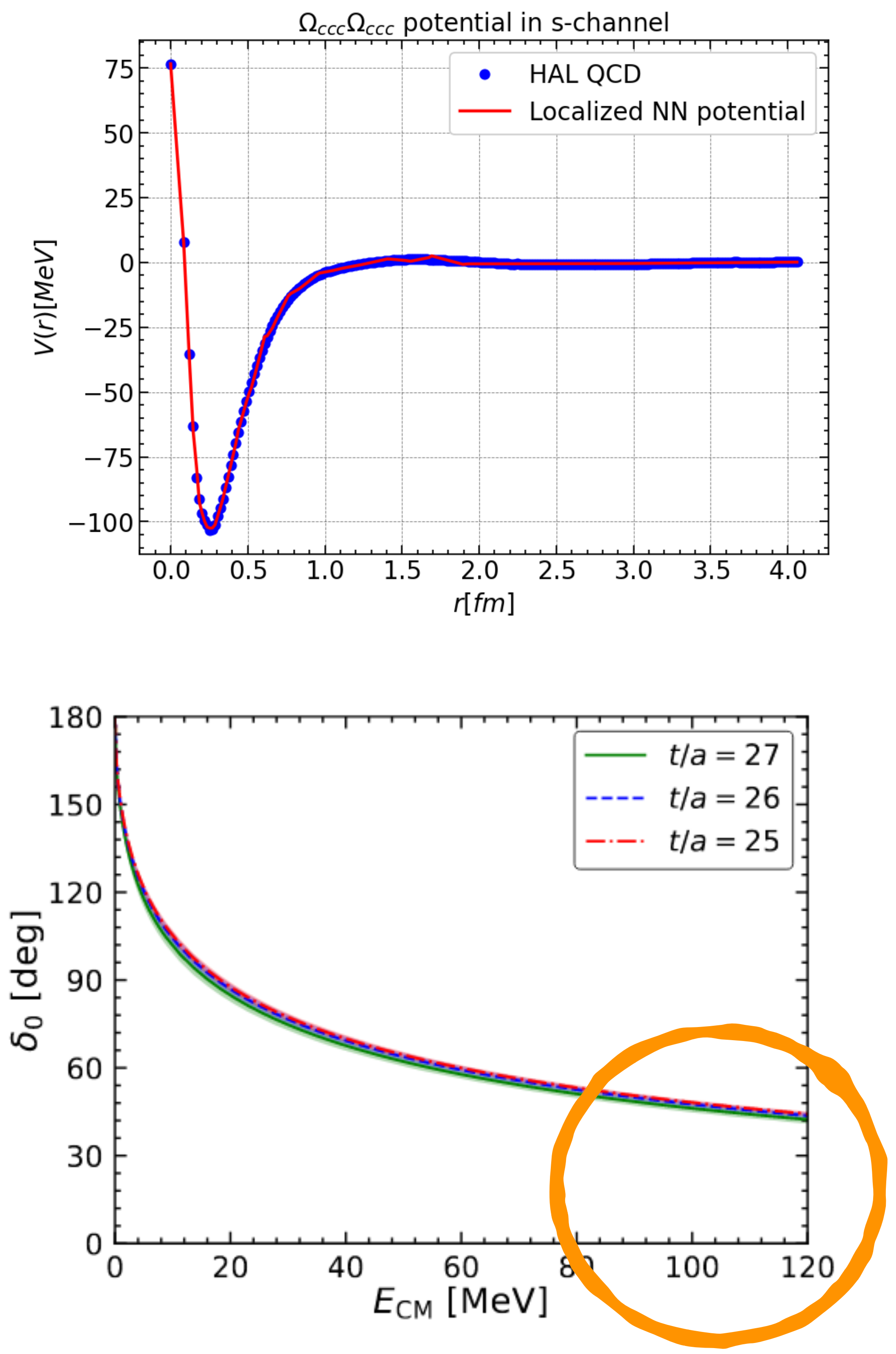
$$U(r > 3\textbf{fm}, r' > 3\textbf{fm}) \to 0$$

$$V_\theta(r) \equiv \frac{\sum_{r'} \Delta r' 4\pi r'^2 U_\theta(r,r') R(t,r')}{R(t,r)}$$

$\Omega_{ccc}\Omega_{ccc}$ Interaction: Non-Local Potential

arXiv:2410.03082 (with HAL QCD)

Y. Lyu, etc., Phys. Rev. Lett. 127, 072003 (2021)



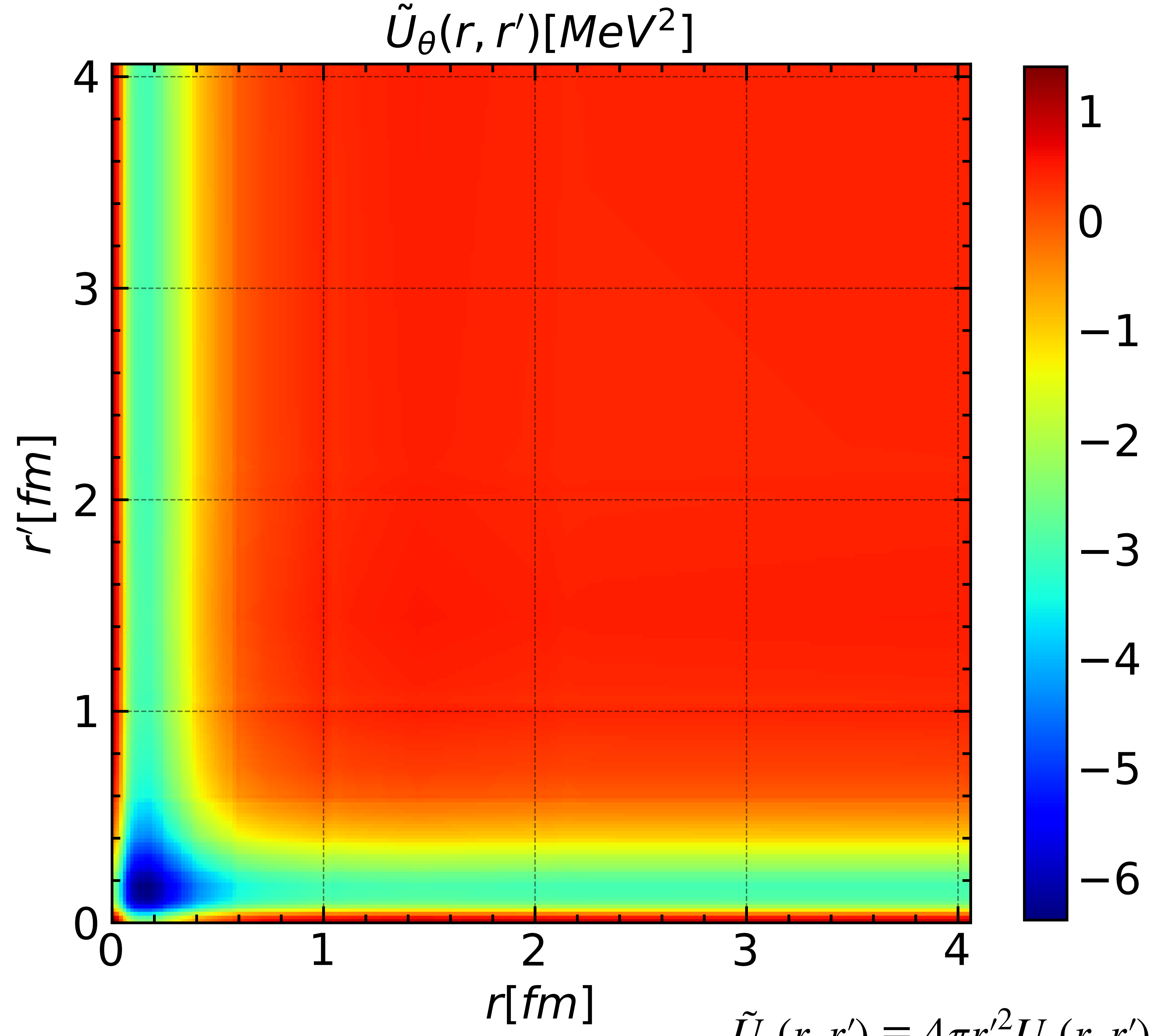
"3D Map"

$$(E_k - H_0)\phi_{\mathbf{k}}(\mathbf{r}) = \int d^3 r' U_\theta(\mathbf{r}, \mathbf{r}')\phi_{\mathbf{k}}(\mathbf{r}')$$

$$E_k = \frac{k^2}{2m}, \quad H_0 = -\frac{\nabla^2}{2m}, \quad m = \frac{m_N}{2}$$

Affect ?

Neural Network Hadron Force



$\tilde{U}_\theta(r, r') \equiv 4\pi r'^2 U_\theta(r, r')$

Summary II

- **Take-Home Messages**
  - Non-Local Potential As A “Map”
    - Neural Network Representations
  - Embed Physics Priors Explicitly
    - Permutation Symmetry
    - Asymptotic Behavior
- **Roadmap**
  - Separable Potential ✅
  - $\Omega_{ccc}\Omega_{ccc}(^1S_0)$ ✅
  - Phase Shifts ✅
  - More Cases 💪
    (B-B, N-B, N-M, N-N, …)
  - Joint Learning with Femtoscopy 💭

in preparation
[Review]

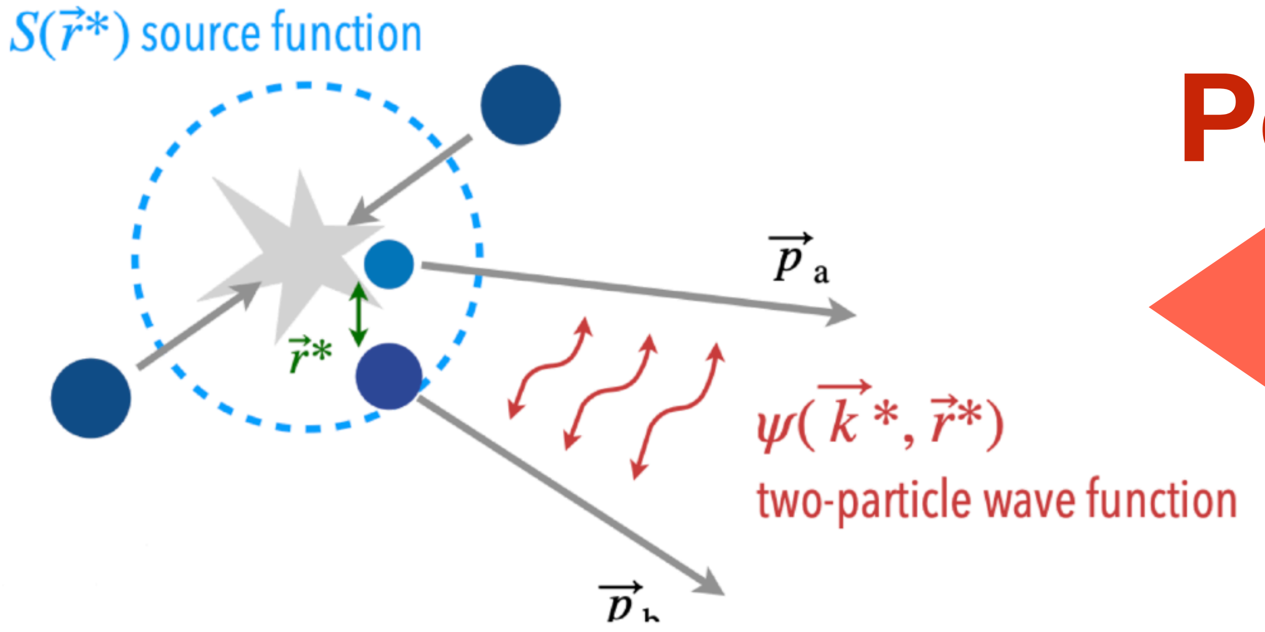
Physics-Driven Learning for Solving Inverse Problems towards QCD Physics

Gert Aarts[1], Kenji Fukushima[2], Tetsuo Hatsuda[3], Andreas Ipp[4], Shuzhe Shi[5], Lingxiao Wang[3,*], and Kai Zhou[6,7]

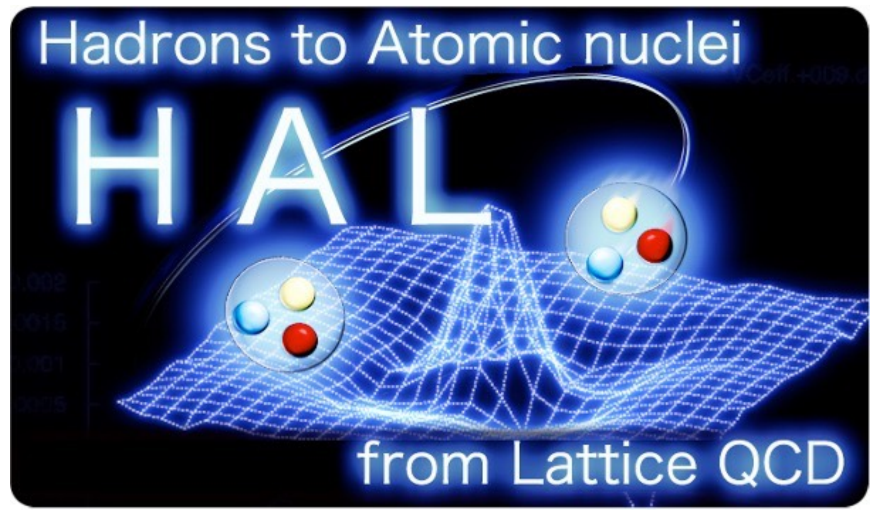
[1]Department of Physics, Swansea University, SA2 8PP, Swansea, United Kingdom
[2]Department of Physics, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan
[3]Interdisciplinary Theoretical and Mathematical Sciences Program (iTHEMS), RIKEN, Wako, Saitama 351-0198, Japan
[4]Institute for Theoretical Physics, TU Wien, Wiedner Hauptstraße 8-10/136, A-1040 Vienna, Austria
[5]Department of Physics, Tsinghua University, Beijing 100084, China
[6]School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen (CUHK-Shenzhen), Guangdong, 518172, China
[7]Frankfurt Institute for Advanced Studies, Ruth Moufang Strasse 1, D-60438, Frankfurt am Main, Germany
[*]e-mail: lingxiao.wang@riken.jp

ABSTRACT

The integration of deep learning techniques and physics-driven designs is reforming the way we address inverse problems, in which accurate physical properties are extracted from complex data sets. This is particularly relevant for quantum chromodynamics (QCD), the theory of strong interactions, with its inherent limitations in observational data and demanding computational approaches. This perspective highlights advances and potential of physics-driven learning methods, focusing on predictions of physical quantities towards QCD physics, and drawing connections to machine learning(ML). It is shown that the fusion of ML and physics can lead to more efficient and reliable problem-solving strategies. Key ideas of ML, methodology of embedding physics priors, and generative models as inverse modelling of physical probability distributions are introduced. Specific applications cover first-principle lattice calculations, and QCD physics of hadrons, neutron stars, and heavy-ion collisions. These examples provide a structured and concise overview of how incorporating prior knowledge such as symmetry, continuity and equations into deep learning designs can address diverse inverse problems across different physical sciences.



Potential



Thank You !

ML meets Physics, Opportunities and Challenges

Kouzou Sakai @Quanta Magazine

Learn to Sample

$$p(\phi) = e^{-S(\phi)}/Z$$

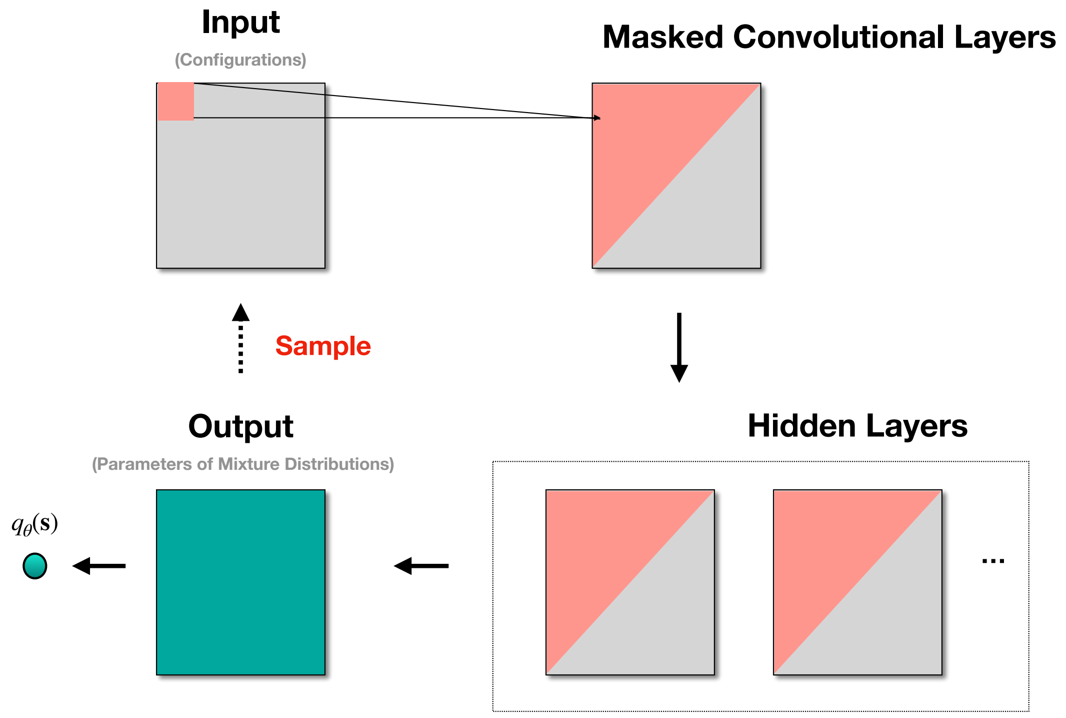
$$\langle O \rangle \approx \frac{1}{N} \sum_i O_i$$

→ **Physical Distribution, Sampling** via Generative Models

Global Sampling

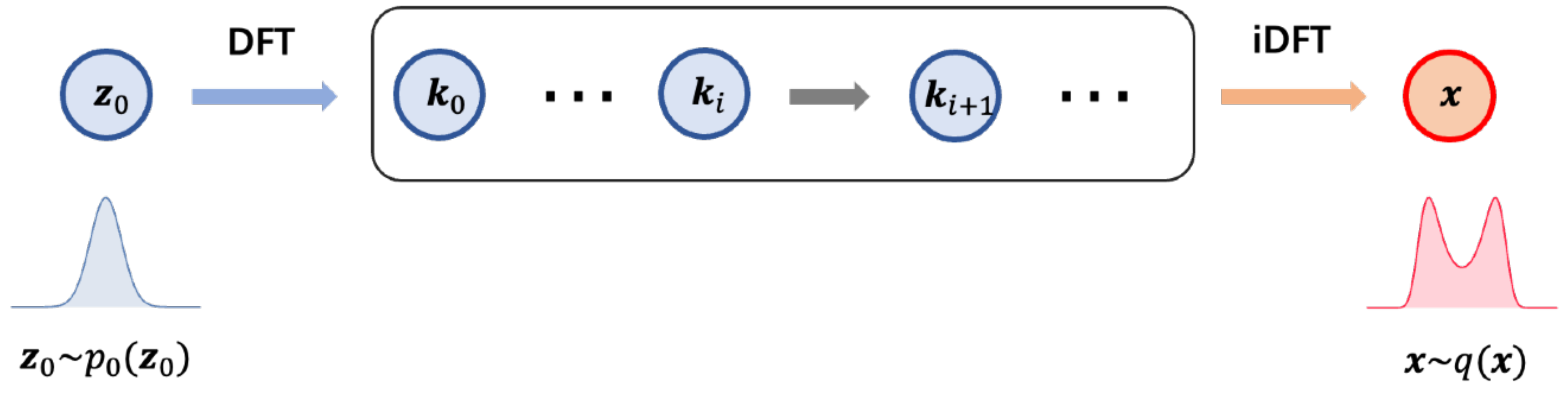
Fast and Independent Sampler

Prog.Part.Nucl.Phys. 104084(2023)



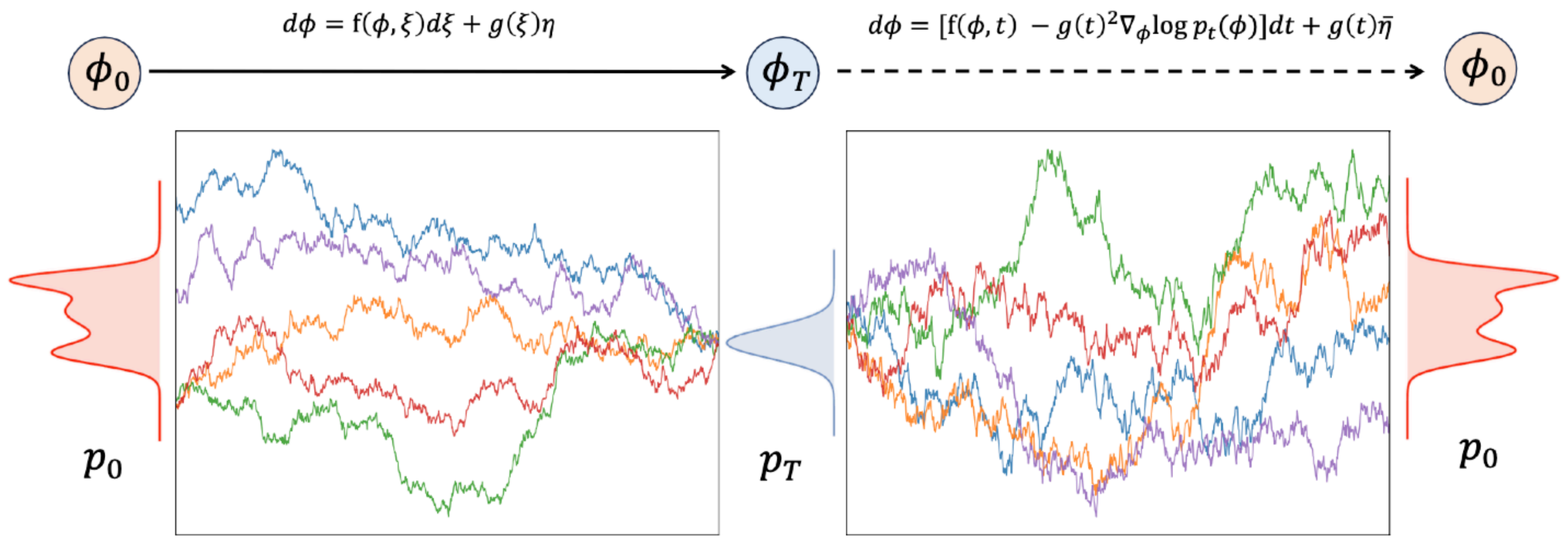
Continuous Autoregressive Models

Chin. Phys. Lett. 39, 120502 (2022)
Chin.Phys.C 48 (2024) 10, 103101



Fourier Flow-based Model

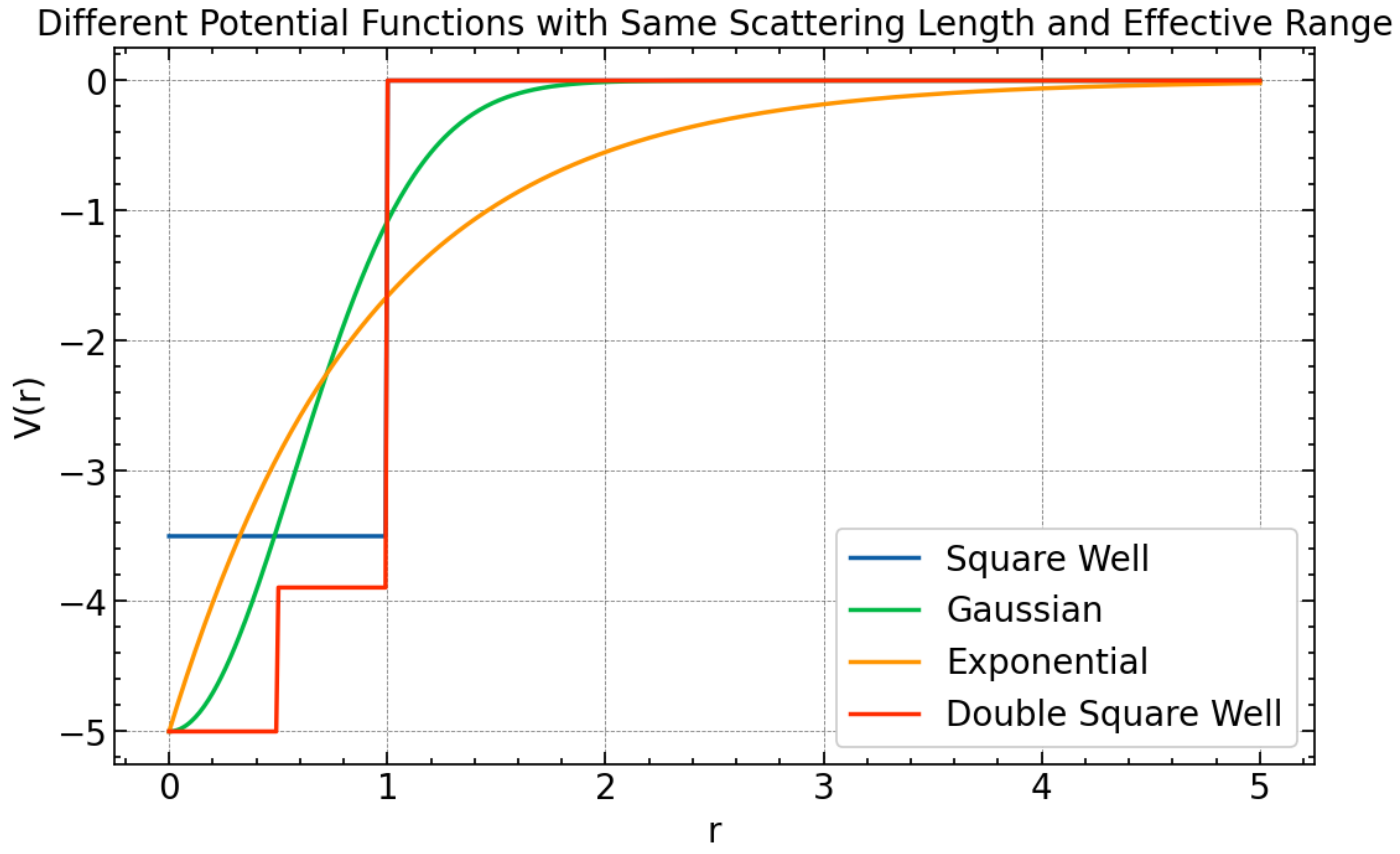
Phys. Rev. D 107, 056001

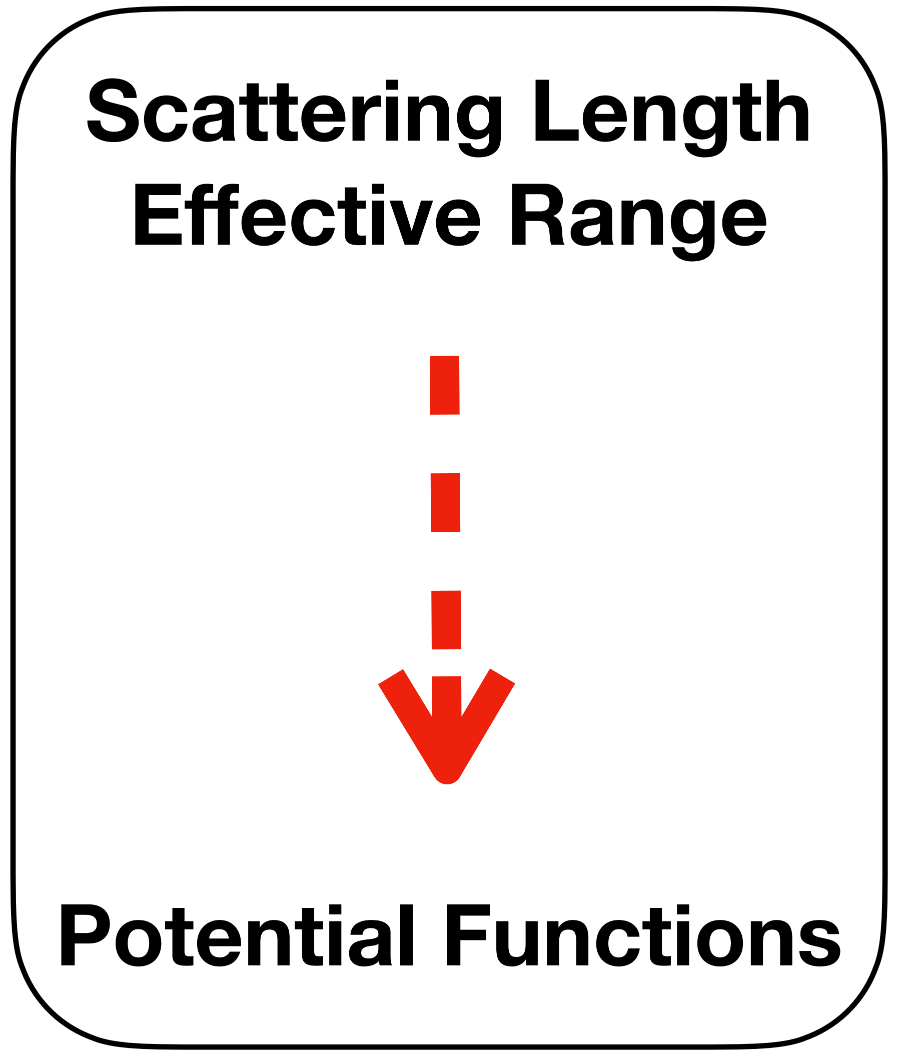


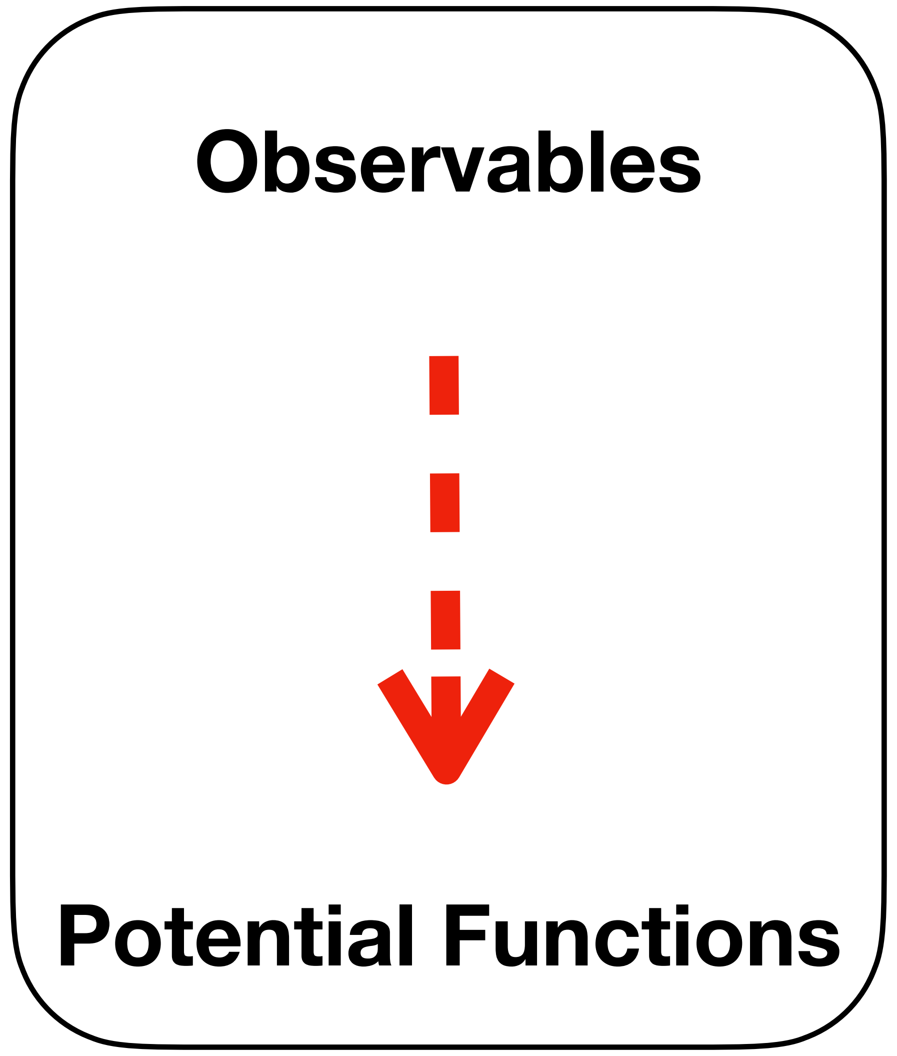
Diffusion Models

JHEP 05(2024)060
ArXiv:2410.21212
ArXiv:2410.19602

Backups: Existence and Uniqueness



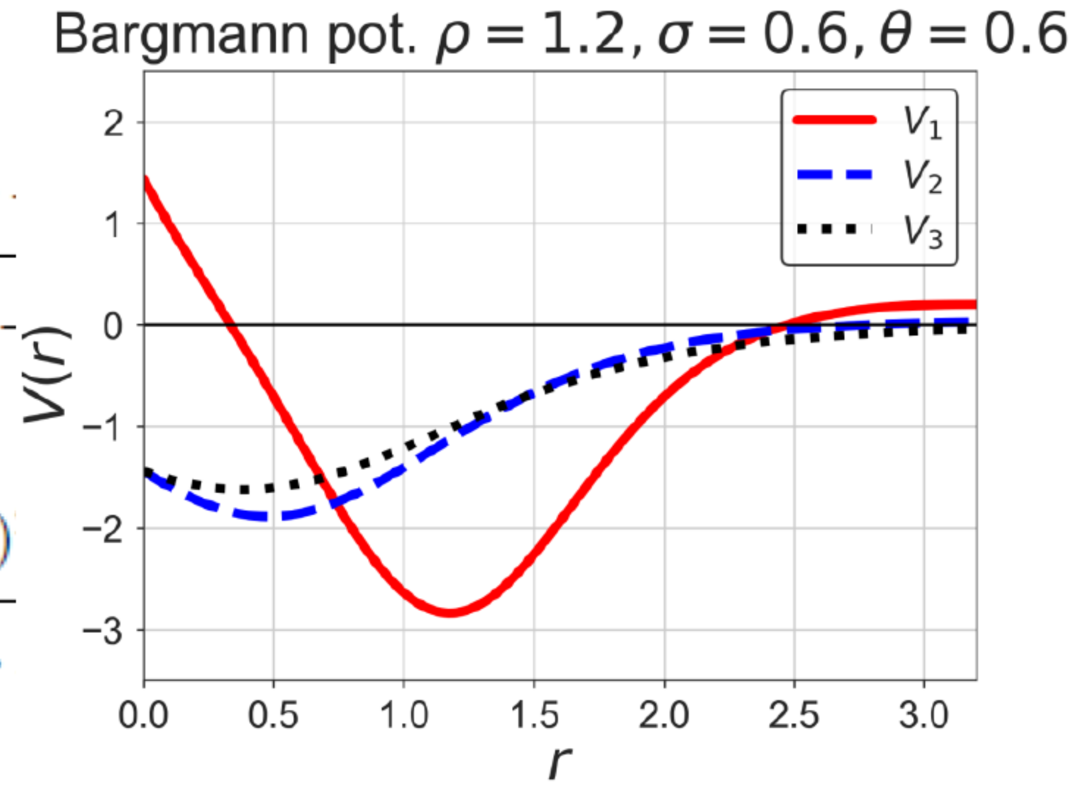




V. Bargmann, Phys.Rev.75(1949)301; Rev.Mod.Phys.21(1949)488

$$V_1(r) = \frac{\rho\sigma\left\{4\rho\sigma + (\rho-\sigma)^2\cosh\left[(\rho+\sigma)r - 2\theta\right] - (\rho \cdots\right.}{\left\{\sigma\sinh\left[\rho r - \theta\right] - \rho\sinh\left[\sigma r \cdots\right.\right.}$$

$$V_2(r) = \frac{\rho\sigma\left\{4\rho\sigma + (\rho-\sigma)^2\cosh\left[(\rho+\sigma)r\right] - (\rho+\sigma)\right.}{\left\{\sigma\sinh\left[\rho r + \theta\right] - \rho\sinh\left[\sigma r\right.\right.}$$

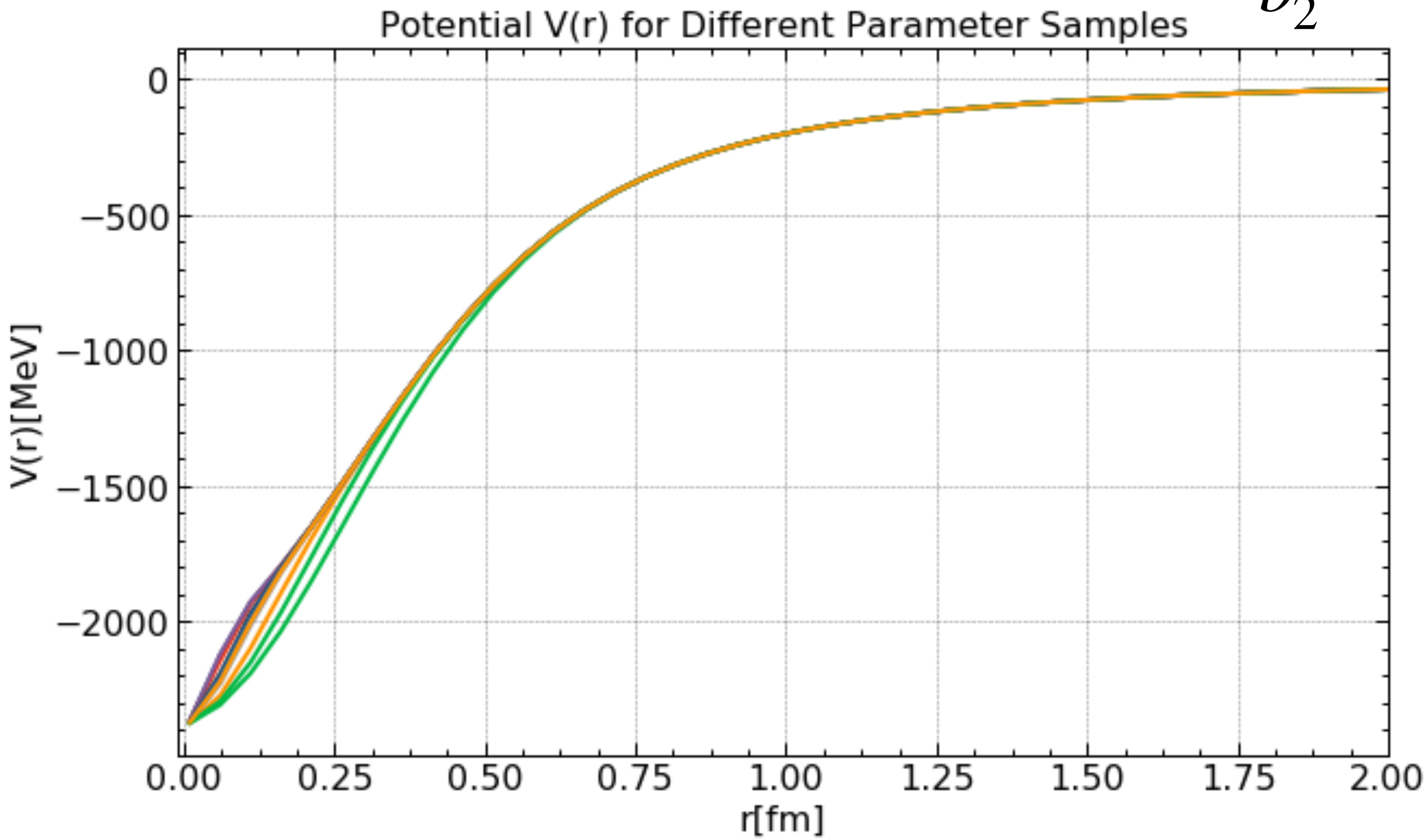


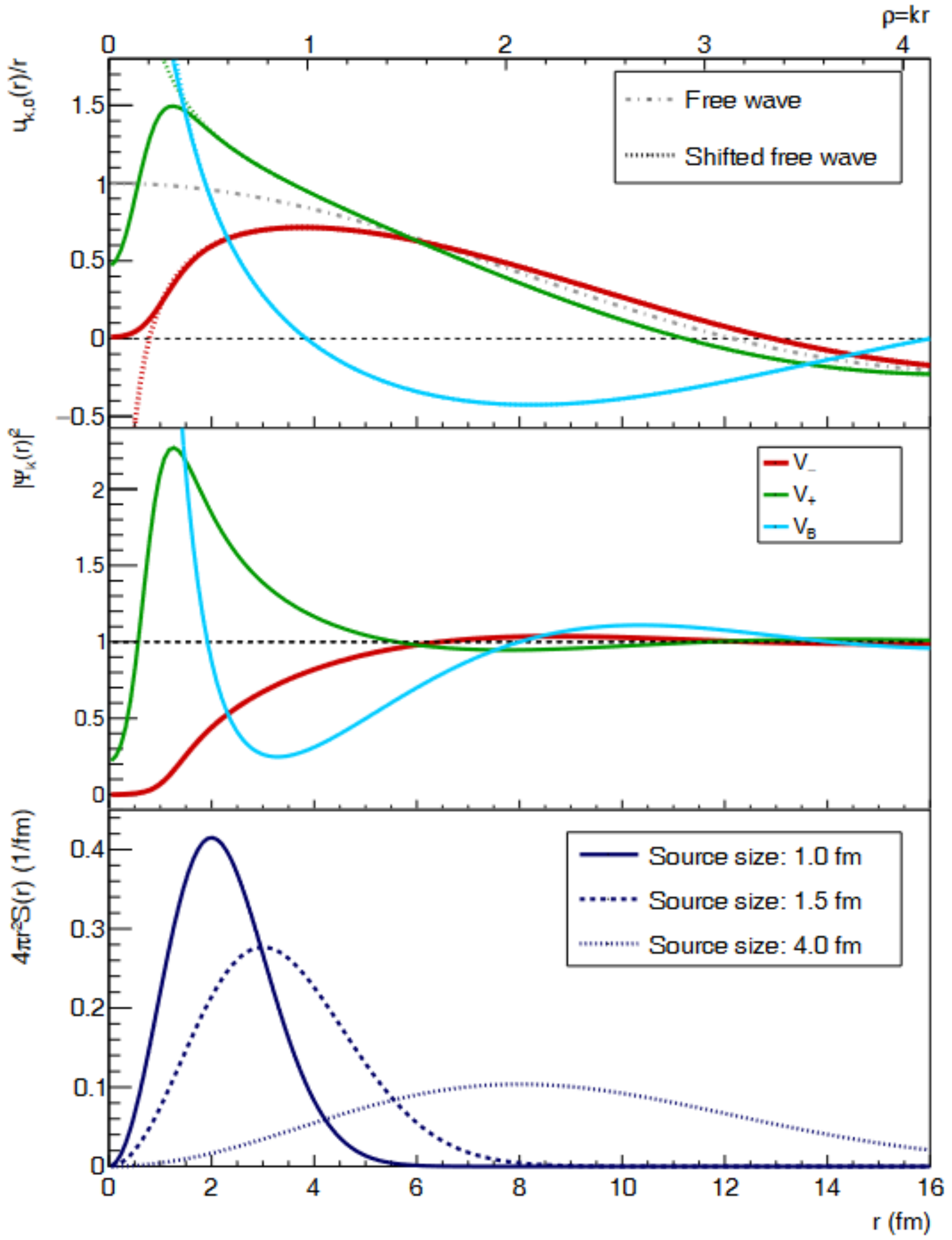
Yan Lyu's slides

$$V_1(r), V_2(r) \propto e^{-(\rho-\sigma)r} \quad (r \to \infty)$$

Backups: Potential vs. Source Function

b_2





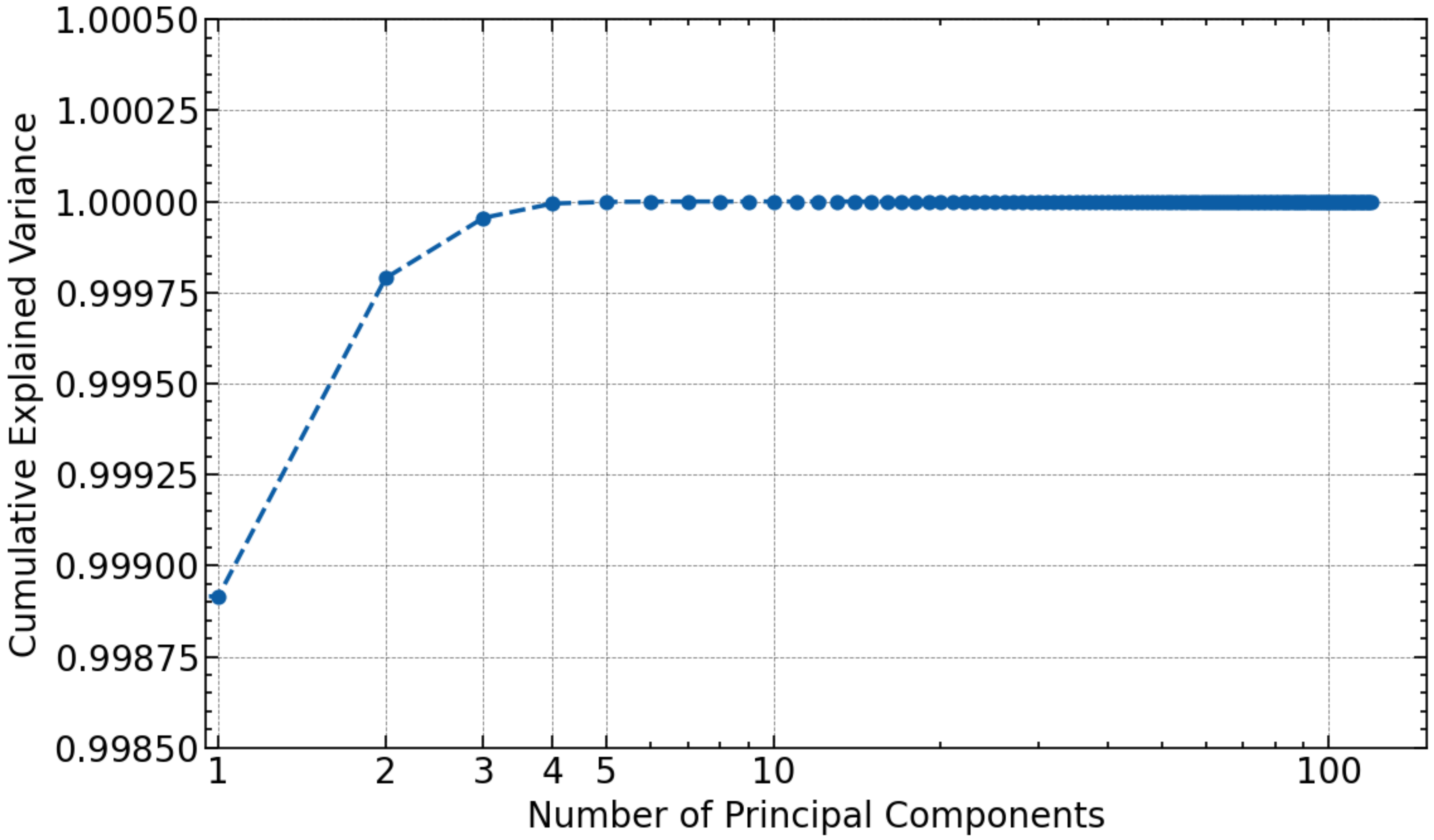
Potential Functions

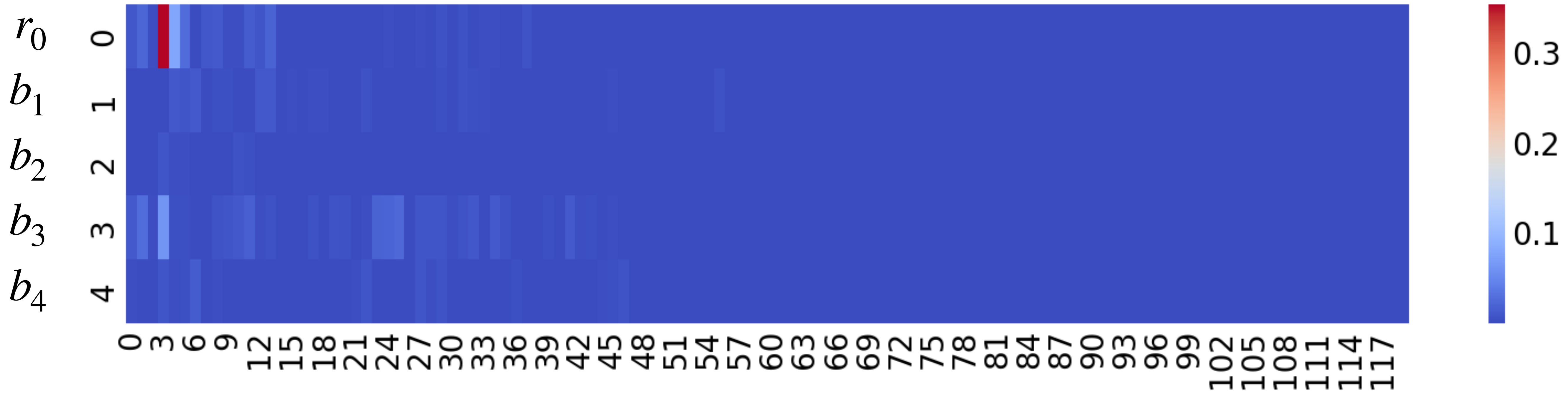
$$V(r) = b_1 e^{-b_2 r^2} + b_3(1 - e^{-b_4 r^2})(\frac{e^{(-m_\pi r)}}{r})^{n_\pi}$$

Source Function

$$S(r) = (4\pi r_0^2)^{-3/2} e^{-\frac{r^2}{4r_0^2}}$$

Backups: PCA and Correlation Matrix





in Preparation **with Liang Zhang, Jiaxing Zhao, etc.**

Data Standardization:

$$Z = \frac{X - \mu}{\sigma}$$

Covariance Matrix:

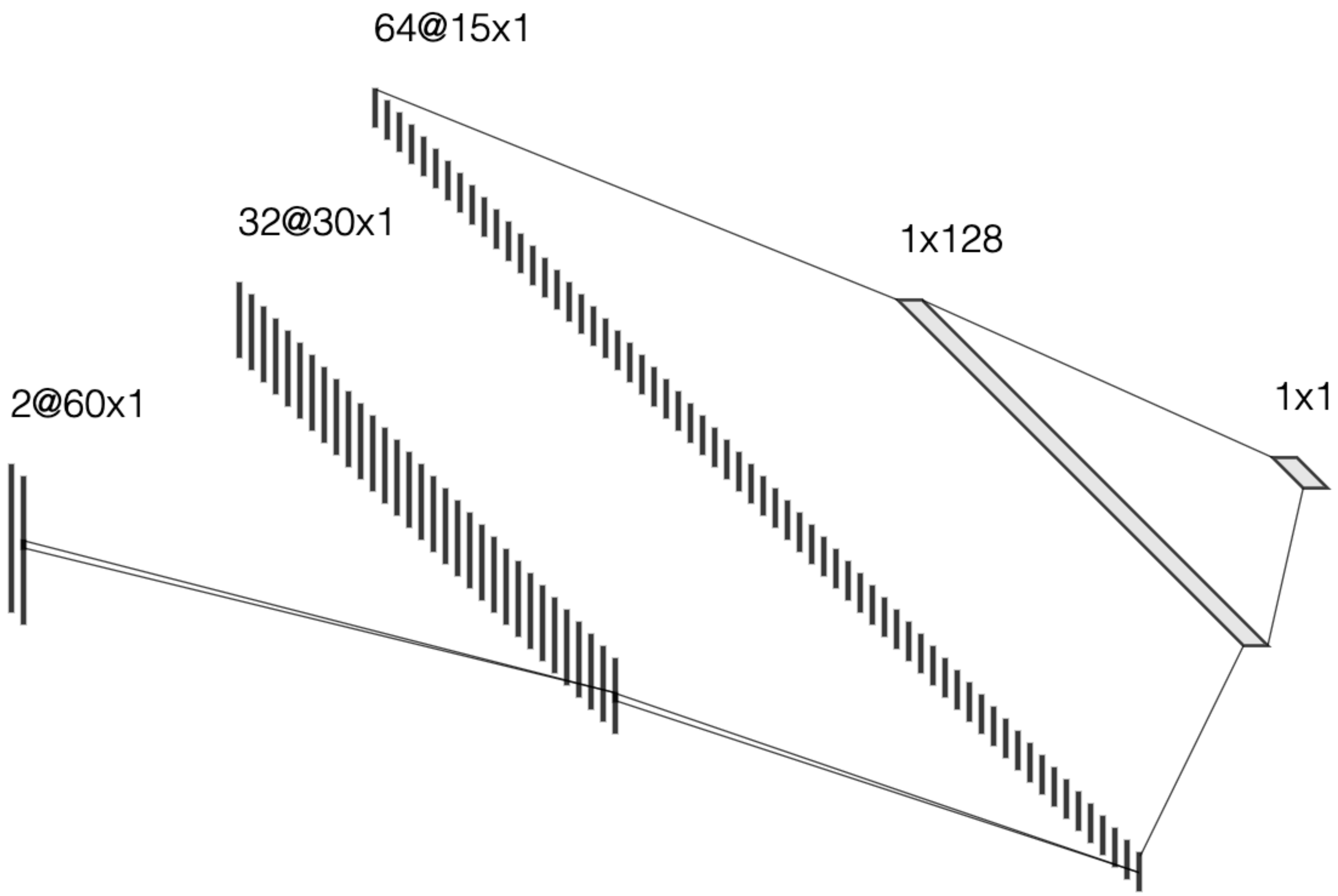
$$C = \frac{1}{n-1} Z^T Z$$

Eigenvalues and Eigenvectors:

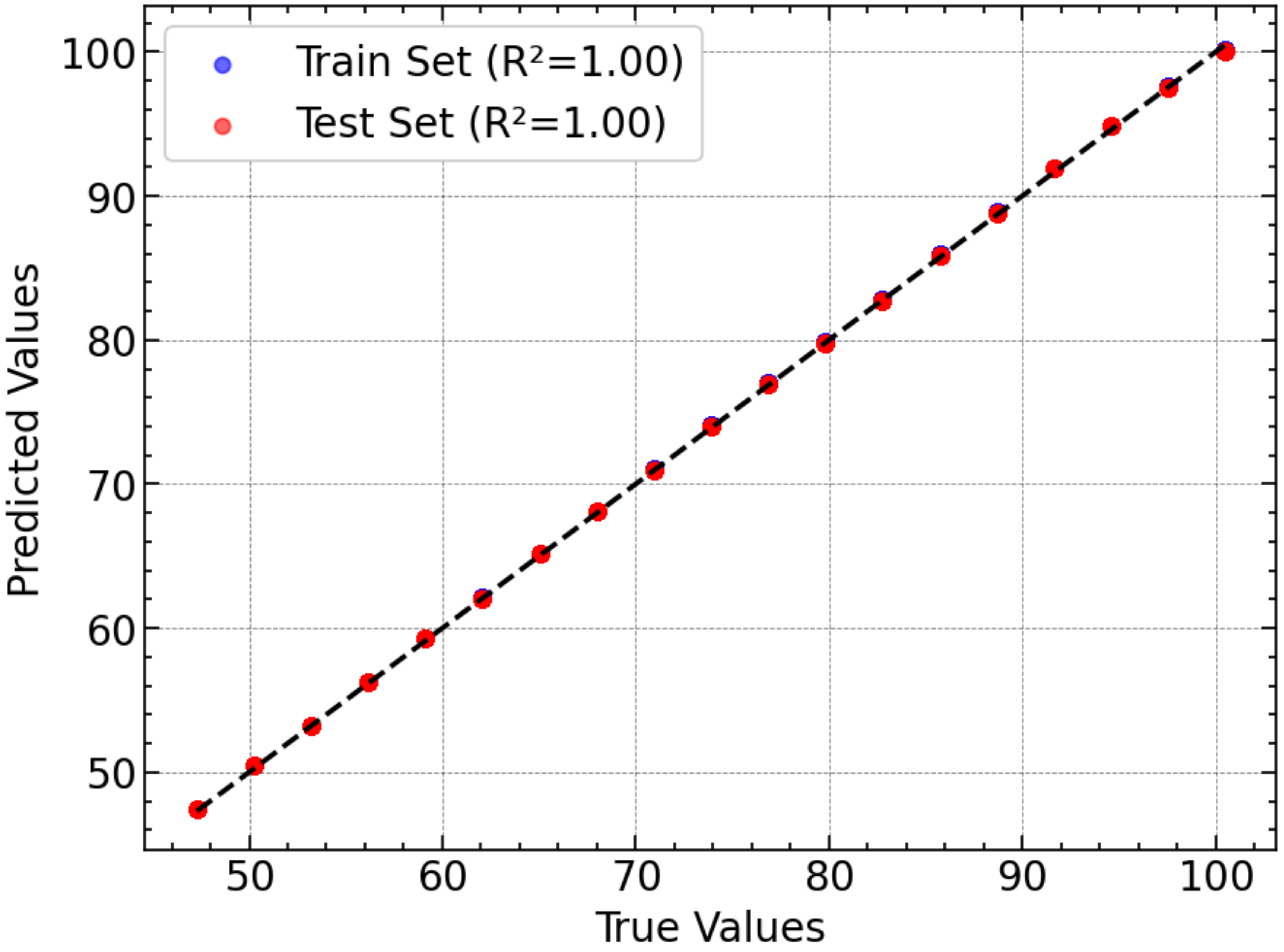
$$Cv = \lambda v$$

Backups: Flexible Input Length

Fully Convolutional Network

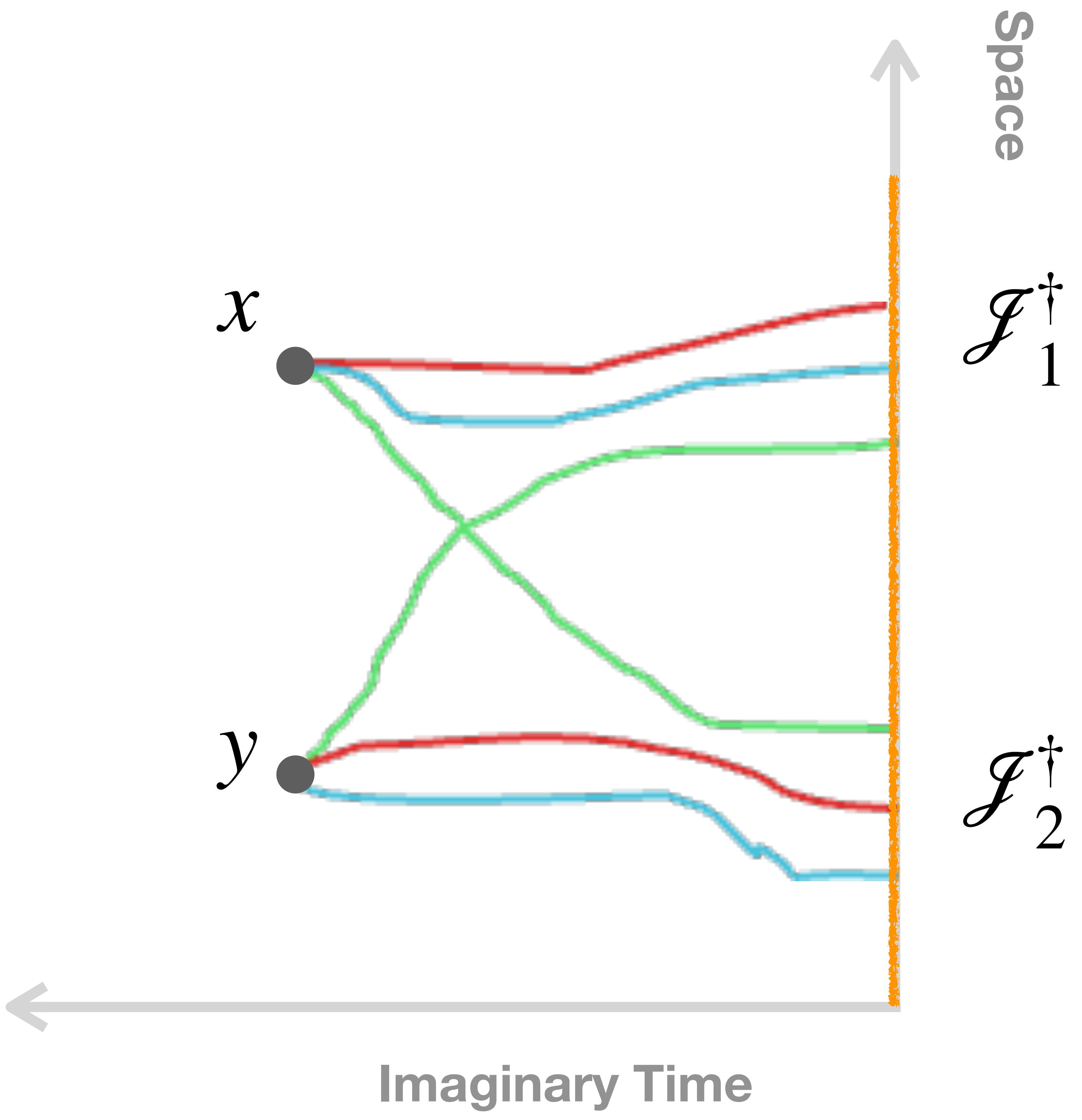


b_2



Backups: HAL QCD method

N. Ishii, S. Aoki, and T. Hatsuda, Phys. Rev. Lett. 99, 022001 (2007)
S. Aoki, T. Hatsuda, and N. Ishii, Prog. Theor. Phys. 123, 89 (2010)
Aoki, S., Doi, T., Front. in Phys. 8, 307 (2020)
S. Aoki and T. Doi, in Handbook of Nuclear Physics(2023), pp. 1–31

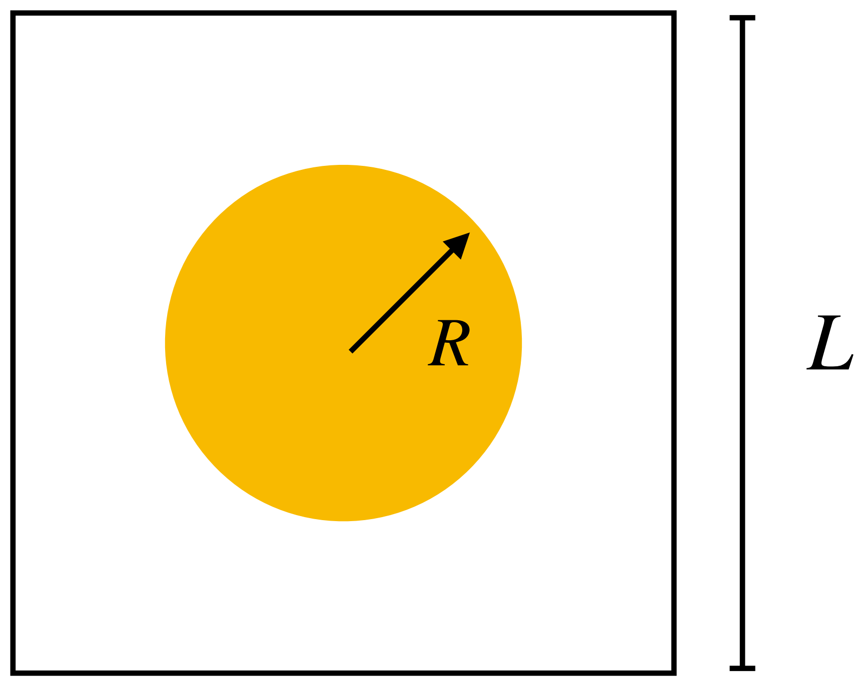


$$\langle N_1(\mathbf{x},t)N_2(\mathbf{y},t)\mathcal{J}_1^\dagger(0)\mathcal{J}_2^\dagger(0)\rangle$$

$$= \sum_n \langle 0|N_1(\mathbf{x})N_2(\mathbf{y})|n\rangle a_n e^{-E_n t}$$

$$\xrightarrow{t>t^*} \phi(\mathbf{r},t) = \sum_{n<n^*} b_n \phi_n(\mathbf{r}) e^{-E_n t}$$

$$(E_k - H_0)\phi_{\mathbf{k}}(\mathbf{r}) = \int d^3 r' U(\mathbf{r},\mathbf{r}')\phi_{\mathbf{k}}(\mathbf{r}'), \quad r < R$$



Consider the wave function at "interacting region"⟶Phase shift, Binding energy

$\phi(\mathbf{r},t) \to$ **2 PI Kernel**

Backups: Uncertainty Estimation

Consistent with Bayesian Inference

Phys. Rev. D 107, 083028

- x : reconstructed potential given a sample
- $O(x)$: observables, δ, E, etc.
  wave-functions in this case

Variance $\sigma(O)^2 = \langle \hat{O}^2 \rangle - \bar{O}^2$

Mean $\bar{O} = \langle \hat{O} \rangle = \sum_j^N w^{(j)} O^{(j)}$

Recall: Importance Sampling

- x : random variables
- $f(x)$: observables
- n : number of samples
- $p(x)$: original(true) distribution
- $q(x)$: reference distribution

$$E[f(x)] = \int f(x)p(x)dx = \int f(x)\frac{p(x)}{q(x)}q(x)dx$$

Weights $$w^{(j)} = \frac{\text{Posterior}(\boldsymbol{\theta}^{(j)}\,|\,\text{data})}{\text{Proposal}(\boldsymbol{\theta}^{(j)})}$$

$$\propto \frac{p(\text{data}\,|\,\boldsymbol{\theta}^{(j)}) \cdot \text{Prior}(\boldsymbol{\theta}^{(j)})}{p(\boldsymbol{\theta}^{(j)}\,|\,\text{samples}^{(j)}) \cdot p(\text{samples}^{(j)}\,|\,\text{data}) \cdot \text{Prior}(\text{data})}$$

L2 : $exp(-\lambda||\theta||_2^2) \sim 1, \quad \lambda \to 0$

$$p(\text{data}\,|\,\boldsymbol{\theta}^{(j)}) \propto \exp(-\chi^2(R_{\boldsymbol{\theta}^{(j)}}(t,r)))$$

$$p(\boldsymbol{\theta}^{(j)}\,|\,\text{samples}^{(j)}) = 1$$

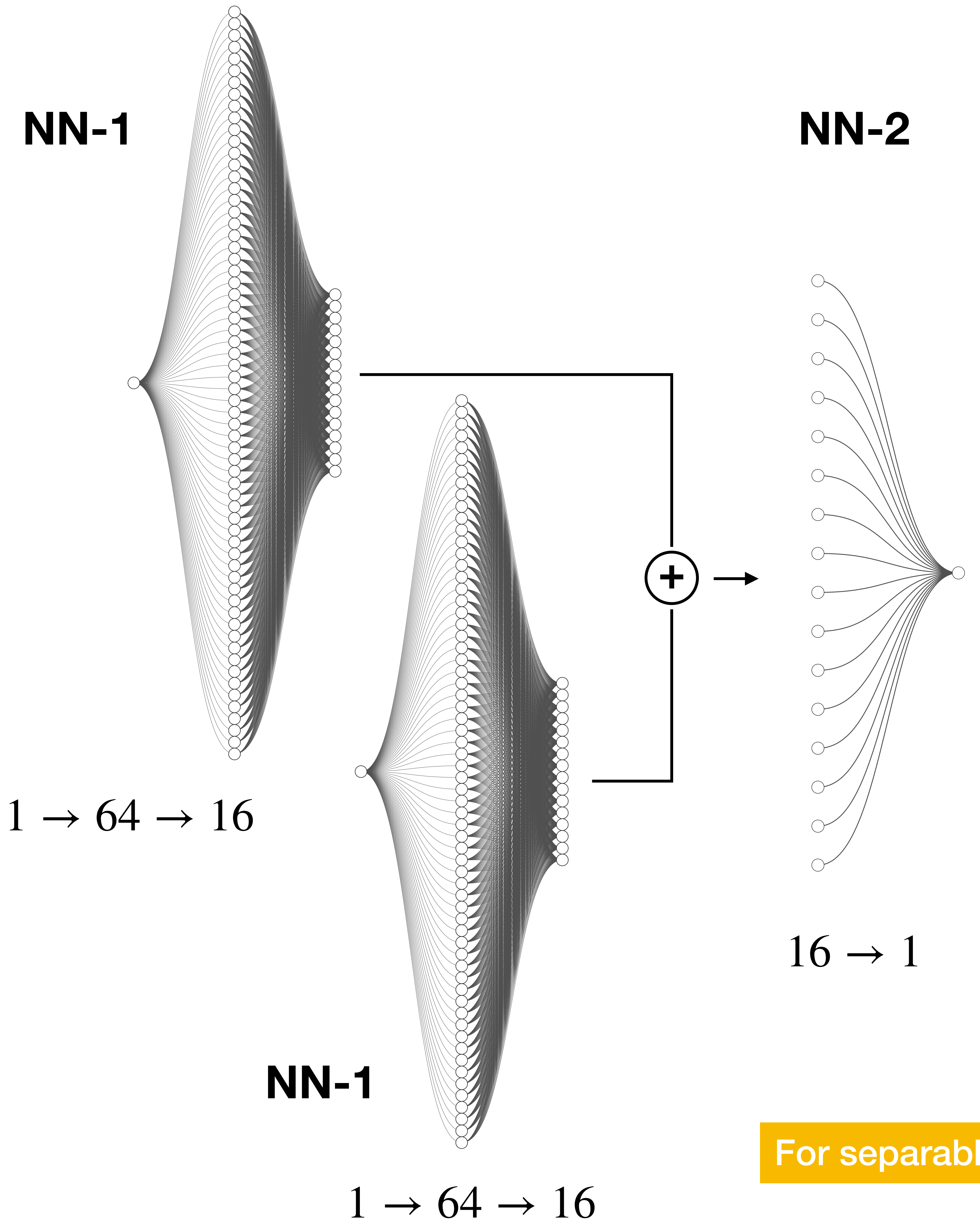
deterministic

$$p(\text{samples}\,|\,\text{data}) \sim \mathcal{N}(\mu, \sigma)$$

$$\tilde{w}^{(j)} = w^{(j)} / \sum_j w^{(j)}$$

normalization

Backups: Setup of DNNs



For separable potential

Regularization

$\lambda = 10^{-5}$

Gradient-based Optimization

Adam : $\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t} + \epsilon}\hat{m}_t$

arXiv:2410.03082 (with HAL QCD)

NN-1

$1 \to 32 \to 64 \to 128 \to 64 \to 32 \to 16$

NN-2

$16 \to 1$

For $\Omega_{ccc}\Omega_{ccc}$ $({}^1S_0)$

Backups: Why DNN helps

Regularization

- **Maximizing Bayesian Posterior**

$$P(\rho|D,I) = \frac{P(D|\rho,I)P(\rho|I)}{P(D|I)}$$

  - Likelihood, $P(D|\rho,I) = e^{-\chi^2/2}$
  - Prior, $P(\rho|I) = e^{\mathcal{S}[\rho]}$

- **Minimization the loss function**

$$J \equiv \frac{\chi^2}{2} - \mathcal{S}[\rho]$$

  - **Chi-square term**

$$\chi^2 \equiv \sum_{i,j=1}^{N} C_{ij}^{-1}(D_i^{\text{obs}} - D_i)(D_j^{\text{obs}} - D_j)$$

the inverse covariance matrix, C^{-1}

$$\Delta_i \equiv -\frac{\delta\chi^2/2}{\delta D(k_i)} = \sum_j C_{ij}^{-1}(D_j^{\text{obs}} - D(k_j))$$

  - **"Entropy" term serves as a regulator**

Three typical "Entropy" terms

$$\mathcal{S}_{\text{TK}} = -\frac{\alpha}{2}\sum_{a=1}^{N_\omega}(\rho_a - \text{DM}_a)^2\Delta\omega\,,$$

$$\mathcal{S}_{\text{MEM}} = \alpha\sum_{a=1}^{N_\omega}\left(\rho_a - \text{DM}_a - \rho_a\ln\frac{\rho_a}{\text{DM}_a}\right)\Delta\omega\,,$$

$$\mathcal{S}_{\text{BR}} = \alpha\sum_{a=1}^{N_\omega}\left(1 - \frac{\rho_a}{\text{DM}_a} + \ln\frac{\rho_a}{\text{DM}_a}\right)\Delta\omega\,.$$

α, a hyper parameter; defaulted model (DM); $\Delta\omega$, step length

$$\frac{\delta J}{\delta\rho(\omega)} = 0$$

$$\rho_a^{\text{TK}} - \text{DM}_a = \frac{1}{\alpha}\sum_i \Delta_i^{\text{TK}} K_{ia}\,,$$

$$\ln\frac{\rho_a^{\text{MEM}}}{\text{DM}_a} = \frac{1}{\alpha}\sum_i \Delta_i^{\text{MEM}} K_{ia}\,,$$

$$\frac{1}{\text{DM}_a} - \frac{1}{\rho_a^{\text{BR}}} = \frac{1}{\alpha}\sum_i \Delta_i^{\text{BR}} K_{ia}\,.$$

the optimal solution exists

Comput. Phys. Commun. 282, 108547

- **Neural Networks (e.g., NN representation)**

$$\rho_a \equiv \rho(\omega_a)$$

  - Output layer, $\rho_a = \text{DM}_a\,\sigma^{(l)}(f_a^{(l)})$
  - Activation functions, $f_a^{(n)} = \sigma^{(n)}(x_a^{(n)})$
  - Hidden layers, $x_a^{(n)} = \sum_b W_{ab}^{(n)} f_b^{(n-1)}$

$a = 1,2,\cdots,N^{(n)};\ n = 1,2,\cdots,l$

- **Set-ups**

$$N^{(0)} = 1, N^{(l)} = N_\omega$$

  - **Input layer,** $a_1^{(0)} = 1$
  - **Hidden layer, no activation functions**
  - **Output layer,** $\sigma^{(l)}(x) = \sigma(x),\ f_a \equiv f_a^{(l)}$
  - **L2 regulation**, $L_2 \equiv \alpha\Delta\omega\sum_{l,a,b}\left(W_{ab}^{(l)}\right)^2$

$$\frac{\delta J}{\delta\rho(\omega)} = 0$$

the optimal solution exists!

$$\frac{f_a/\sigma'(f_a)}{\left(\sum_b f_b^2\right)^{\frac{l-1}{l}}} = \frac{\text{DM}_a}{\alpha}\sum_i \Delta_i K_{ia}$$

non-local constraints from NN!

$$\frac{(1+e^{-f_a})f_a}{\left(\sum_b f_b^2\right)^{\frac{l-1}{l}}} = \frac{\text{DM}_a}{\alpha}\sum_i \Delta_i K_{ia}$$ for *Softplus* activation function

$$\alpha\, f_a(1+e^{-f_a}) \equiv \text{DM}_a\left(\sum_b f_b^2\right)^{\frac{l-1}{l}}\sum_i \Delta_i K_{ia}$$