Application of radial basis functions networks in spectral functions

Zhou Meng

Peking University

November 26, 2021

The 7th China LHC Physics Workshop

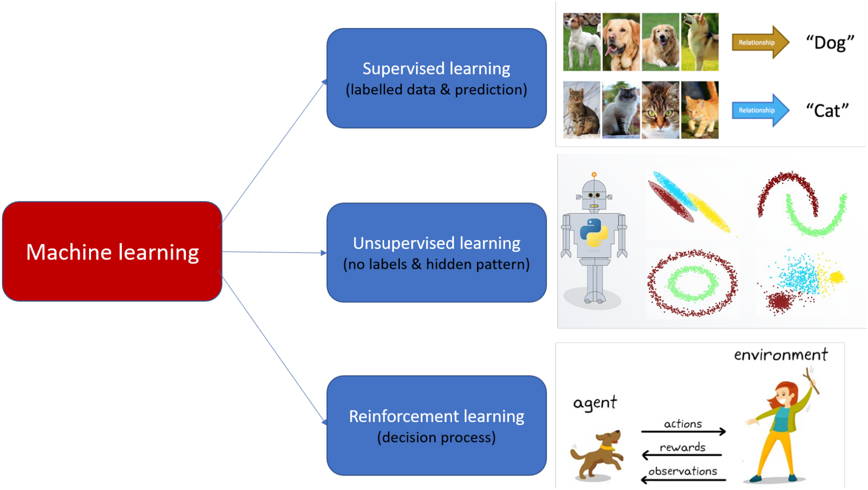
Outline

1. Introduction
   - Introduction to Machine Learning
   - ML in heavy ion physics
2. Machine Learning Spectral Functions
   - the Spectral Functions
   - Traditional Method
   - Our Method
3. Summary

Outline

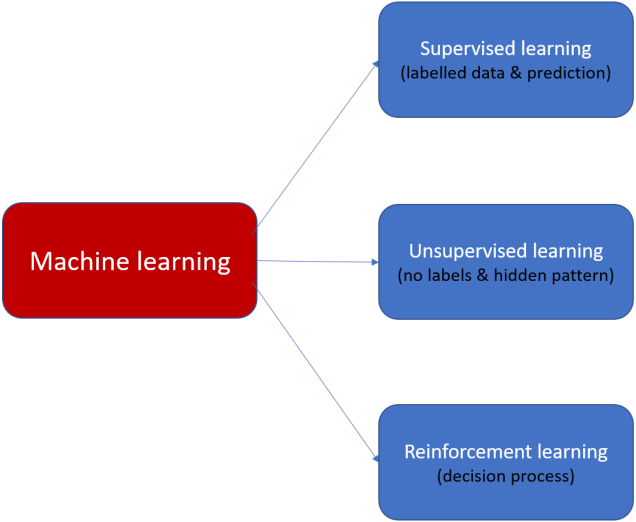
Types of machine learning

Machine learning is a collection of algorithms that let the computer find patterns from data by themselves.



Types of machine learning

Machine learning is a collection of algorithms that let the computer find patterns from data by themselves.

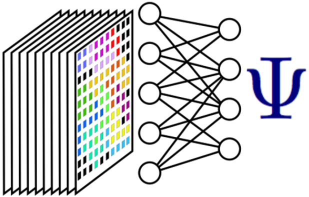


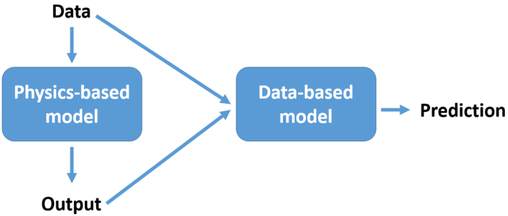
Classification
Regression
...

Dimension reduction
Clustering
...

Machine learning and Physics

- Learning knowledge:
  - Human: learn physics laws from observed variables
  - Machine: learn hidden patterns from big data
- Understanding Machine learning:
  - Hidden patterns = physics law?
- Use ML to solve physics problem (Comparing to traditional method):



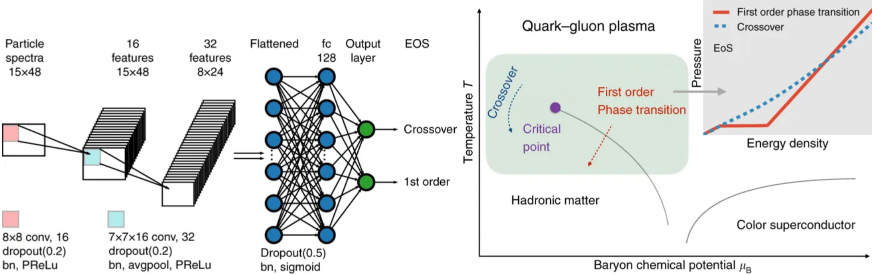


Data driven science

Outline

ML in heavy ion physics

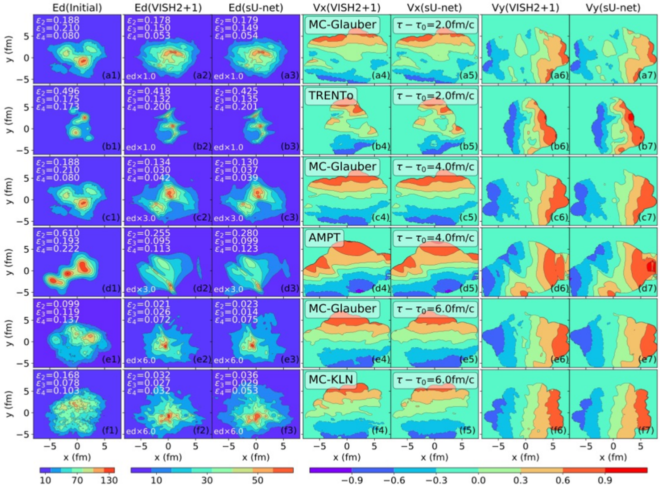
Deep learning for nuclear phase transition



$$l(\theta) = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log \hat{y}_i + (1-y_i)\log(1-\hat{y}_i)] + \lambda||\theta||_2^2$$

Nature Communications 2018, LG. Pang, K.Zhou, N.Su, H.Petersen, H. Stoecker, XN. Wang.

Stacked U-net for relativistic hydrodynamics

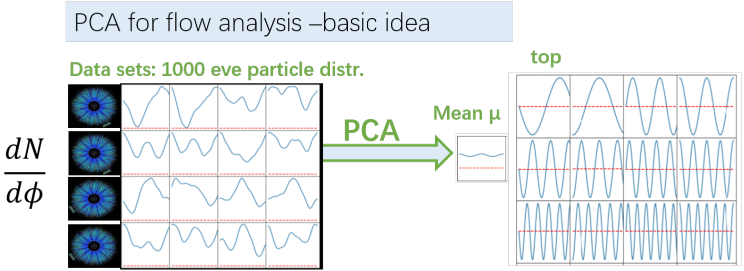


Huang H, Xiao B, Xiong H, Song H,et al. Nuclear Physics A, 2019, 982: 927-930.

Definition of flow harmonics by human

$$\frac{dN}{d\phi} = \frac{1}{2\pi}(1 + 2\sum_{n=1}^{\infty} v_n e^{-in(\phi-\Psi_n)}) \tag{1}$$

But why Fourier expansion?



Liu Z, Zhao W, Song H. The European Physical Journal C volume 79, Article number: 870 (2019)

Outline

the Spectral Functions(SPFs)

- Closely related to correlator $G(\tau)$ by definition

$$G(\tau) = \int_0^\infty \frac{d\omega}{2\pi} K(\tau,\omega)\rho(\omega) \tag{2}$$

Why spectral function?

- Theoretical understanding of some experimental observables.
  - In medium modification of hadrons
  - heavy quark/quarkonium survival in QGP
  - Transport coefficients
  - Photon emission rate
  - ...
- Conductivity as an example:

$$\sigma \propto \lim_{\omega\to 0} \frac{\rho(\omega)}{\omega} \tag{3}$$

Why hard?

- By definition

$$G(\tau) = \int_0^\infty \frac{d\omega}{2\pi} K(\tau, \omega) \rho(\omega) \tag{4}$$

- $G(\tau)$: lattice QCD calculations or other non-perturbative methods.
- Reconstruction of SPF:

$$\rho(\omega_i) = K_{ij}^{-1} \circ G(\tau_j) \tag{5}$$

- ill-posed problem
  - Large number of degrees of freedom required while fewer are given
  - Unavoidable noises
  - Smoothing effect of K_{ij}

Outline

Traditional method

Maximum Entropy Method (MEM)

- Bayes' theorem

$$P[\rho|GH] = \frac{P[G|\rho H]P[\rho|H]}{P[G|H]} \tag{6}$$

- Prior probability $P[\rho|H] \propto e^{\alpha S}$
- the Shannon-Jaynes Entropy

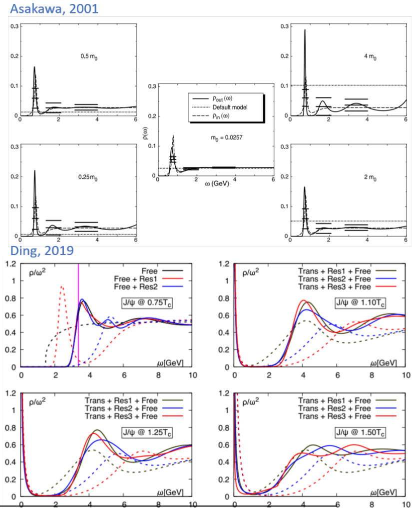
$$S = \int_0^\infty [\rho(\omega) - m(\omega) - \rho(\omega)\log(\frac{\rho(\omega)}{m(\omega)})] \tag{7}$$

- Find the most probable $\rho(\omega)$ (an optimization problem)[1]

$$\frac{\delta P[\rho|GH]}{\delta \rho} = 0 \tag{8}$$

[1] Using Newton's Method

Results using Maximum Entropy Method



Examples using MEM in Fig.

Advantages of MEM:

- Prior info encoded
- Sharp peaks can be reconstructed

Defects of MEM:

- Dependence on default model.
- Unstable results.

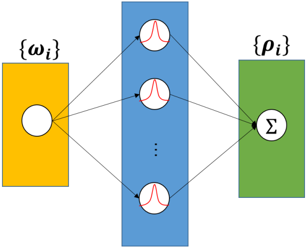
We need comparison between prior information, and also between different methods.

Ding H T, Kaczmarek O, Kruse A L, et al. Nuclear Physics A, 2019, 982: 715-718.

Asakawa, M., T. Hatsuda, and Y. Nakahara. Progress in Particle and Nuclear Physics 46, no. 2 (2001): 459–508.

Outline

Radial Basis Functions(RBF) method



To approximate a target function

$$y^{target}(x) \approx \sum_{i=0}^{N} a_i \phi(||x - x_i||)$$

A single hidden layer feed-forward networks

- RBF as activation functions

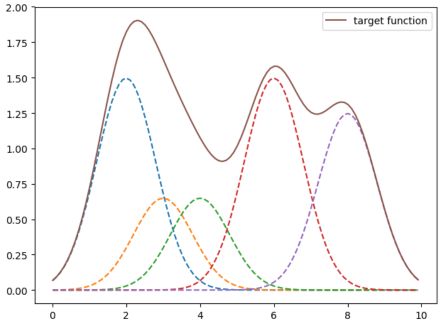
Universal approximation has been proven[2]

- Commonly used Radial Basis Functions

$$\begin{aligned} \text{Thin-plate spline:} \quad & \phi(r) = r^2 ln(r) \\ \text{Gaussian:} \quad & \phi(r) = e^{-\frac{r^2}{2a^2}} \\ \text{Multiquadric:} \quad & \phi(r) = (r^2 + a^2)^{\frac{\alpha}{2}} \end{aligned}$$

[2] Park J, Sandberg I W. Neural computation, 1991, 3(2): 246-257. Any continuous function on a closed, bounded set with arbitrary precision.

Radial Basis Functions(RBF) method



Applications of RBFs

- multivariate function interpolation and approximation[3,4]
- solution of differential and integral equations[5,6,7]

[3] David S Broomhead and David Lowe. Technical report, DTIC Document, 1988.

[4] Wang J G, Liu G R. International Journal for Numerical Methods in Engineering, 2002, 54(11): 1623-1648.

[5] Kansa EJ. Computers and Mathematics with Applications 1990; 19:127–161.

[6] C. Franke , R. Shaback, Applied Mathematics and Computation, Volume 93, Issue 1 (1998) , pp. 73-82.

[7] A. Golbabai , S. Seifollahi, Applied Mathematics and Computation, 174 (2006), pp.877-883.

Reconstruction of Weights A_k

Reformulation of the problem

- Using RBF networks, the SPF is parametrized as, ($\phi, \bar{m}_j$ are given)

$$\rho(\omega_j) = \sum_k \phi(|\omega_j - \bar{m}_k|) A_k \tag{9}$$

- Now the SPF is expressed as a weighted summation of RBFs ! Weights A_j are to be solved.
- The discreted integral equation now reads (Matrix multiplication)

$$G(\tau_i) = \sum_j K(\tau_i, \omega_j) \rho(\omega_j) = K \cdot \phi \cdot A \tag{10}$$

- And we want A (although $K \cdot \phi$ is irreversible)
- Method: using TSVD. Fast, stable, and shape parameter.

Mock SPFs

- Breit-Wigner type SPF:

$$\rho_{Mock}(\omega) = \sum_i \rho_{BW}(A_i, \Gamma_i, M_i, \omega) \tag{11}$$

with

$$\rho_{BW}(A_i, \Gamma_i, M_i, \omega) = \frac{4A_i\Gamma_i\omega}{(M_i^2 + \Gamma_i^2 - \omega^2)^2 + 4\Gamma_i^2\omega^2}, \tag{12}$$

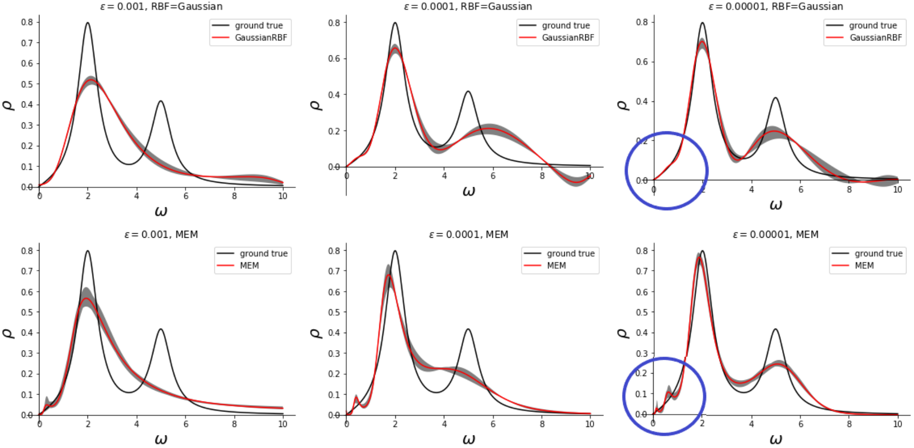
where A_i is the normalization parameter, M_i denotes the mass of the particle, carrying the location of the peak, and Γ_i is the width.

- SPF associated with diffusion coefficient η_D

$$\begin{aligned} \rho_V(\omega) =& \frac{6\chi_{00}T}{M_0}\frac{\omega\eta_D}{\omega^2 + \eta_D^2} + \frac{3}{2\pi}\Theta(\omega^2 - 4M_0^2) \\ & \times \omega^2 \tanh(\omega/4T)\sqrt{1 - 4M_0^2/\omega^2} \\ & \times [1 + 4M_0^2/\omega^2]. \end{aligned} \tag{13}$$

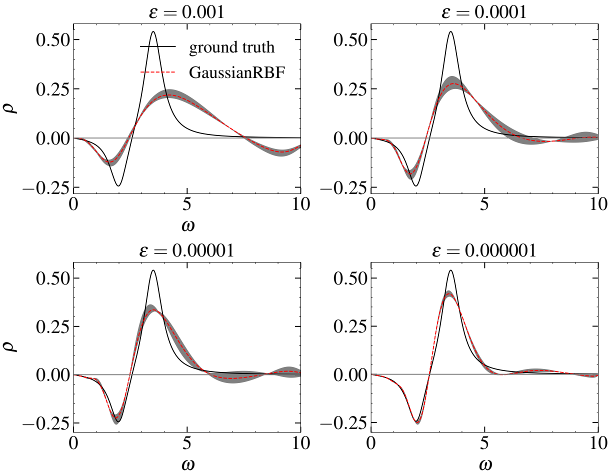
Results of RBF

- Prediction using correlators with Gaussian Noise(width=ϵ)
  - Mock SPF of Breit-Wigner type.
- Comparison between RBF and Maximum Entropy method.



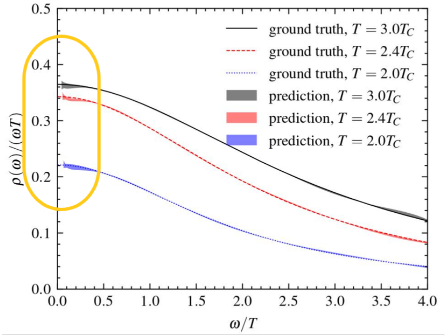
Results of RBF

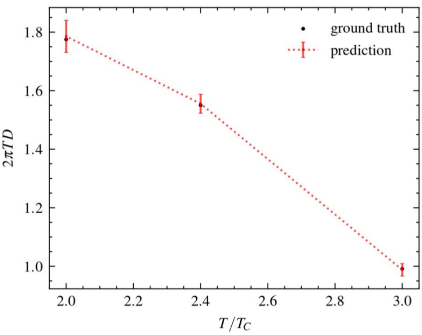
- Prediction using correlators with Gaussian Noise(width=ϵ)
  - Mock SPF of Breit-Wigner type.
- SPF with negative part.



Results of RBF

- Prediction using correlators
  - Associated with the heavy quark diffusion
  - Uncertainty comes from noises in parameter
- Low frequency part of SPF and predicted coefficients.





Summary

- RBF representation is continuous and analytic.
- TSVD method is fast and stable, and we need shape parameter to consider prior information.
- Our results in general show better behaviors compared to other traditional schemes, especially in low frequency region.
- RBF method can cope with more complex situations such as negative SPFs.

Thank You