Nuclear Science Computing Center at CCNU

Machine learning Hadron Spectral Functions in Lattice QCD

Feiyi Liu

Based on arXiv:2110.13521 in collaboration with
Shiyang Chen, Hengtong Ding, Gábor Papp, Chunbin Yang

12th GPU Day, June 20-21, 2022

Partner institute

Nuclear Science Computing Center

at Central China Normal University(CCNU)

N: Nuclear **S**: Science **C³**: Color 3 -> QCD

Built in September 2018

Theoretical Peak (Rpeak):

2.7 PFlop/s (peta flops per second)

GPUs: 306 Nvidia V100 + 30 A100

Word Rand	2018	2019	2020
TOP 500 The List.	466	389	432
THE GREEN 500	268	13	17

Cooperation with ELTE: Machine learning Hadron Spectral Functions in Lattice QCD

S.-Y Chen, H.-T Ding, F.-Y Liu, G. Papp and C.-B. Yang, arXiv:2110.13521

An ill-posed problem in hadron spectral function

Hadron spectral functions: Carry all information about hadrons

1. Quarkonia dissociation temperature T. Matsui & H. Satz, PLB 178, 416 (1986)

2. Heavy Quark diffusion coefficient P. Petreczky & D. Teaney, PRD 73,1649 (2006)

Reconstruction

Correlator (Discretized) → *Inverse problem!* / *ill-posed problem!* → **Spectral function** (Continuous)

$\mathcal{O}(10)$ $\qquad$ $\mathcal{O}(1000)$

$$G(\tau,T) = \sum_{x,y,z} \langle J_H(0,\vec{0}) J^+(\tau,\vec{x})\rangle_T = \int_0^\infty \frac{d\omega}{2\pi} K(\tau,\omega,T)\rho(\omega)$$

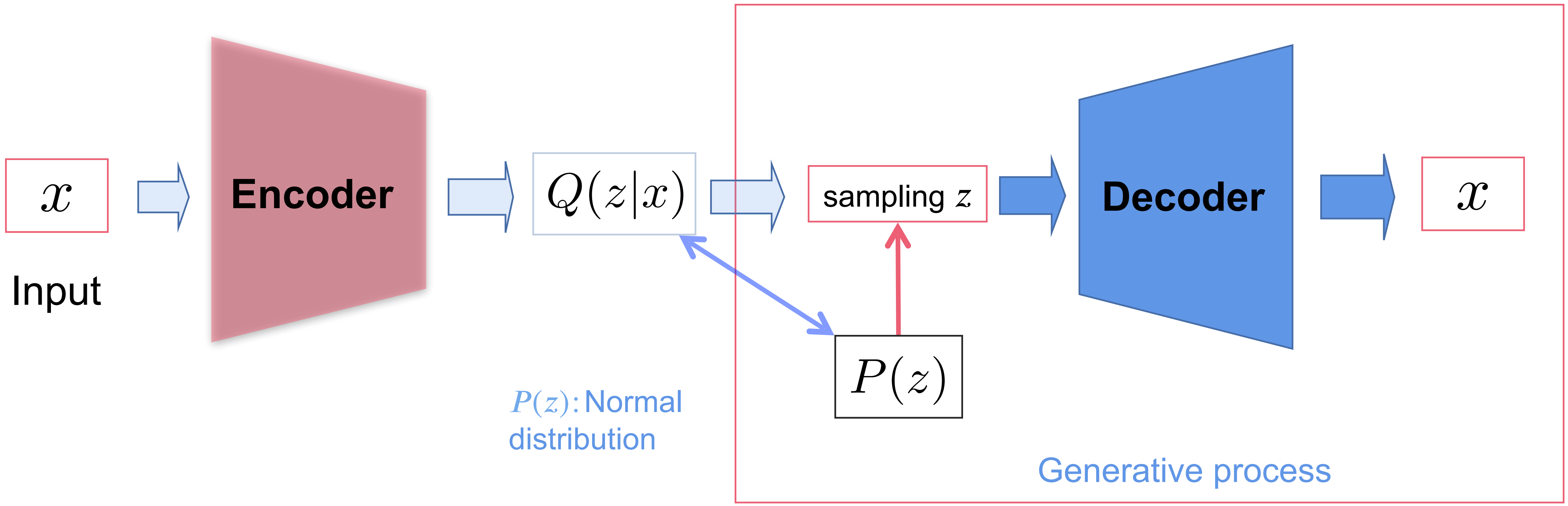
$$K(\omega,\tau,T) = \frac{\cosh(\omega(\tau - \frac{1}{2T}))}{\sinh(\frac{\omega}{2T})}$$

Common methods:

1. Maximum Entropy Method (MEM) Asakawa, Hatsuda & Nakahara, (2001)
2. Backus Gilbert method G. Backus et al, Geophysical Journal International 16, 169(1968)
3. New Bayesian Method (improved MEM) Y. Burnier et al, PRL 111,182003 (2013)
4. Stochastic Approaches H.-T. Ding et al., PRD 97, 094503 (2018)

……

Variational AutoEncoder (VAE) C.Doersch, arXiv:1606.05908



Natural tool to tackle the inverse problem.

Our Goal:

1. **Find certain constraints on latent space Z.**
2. **Design a neural network that can learn to balance the prior & likelihood.**
3. **Train the neural network using a general input**

A newly structure of VAE

To reconstruct ρ from G → $P(\rho|G)$ (Conditional probability)

$$\log P(\rho\,|\,G) = \log\left[\int P(\rho\,|\,z, G) P(z\,|\,G)\, dz\right]$$

$$= \log\left[E_{Q(z|\rho_{gt},G_{gt})}\left(P(\rho\,|\,z, G)\frac{P(z\,|\,G)}{Q(z\,|\,\rho_{gt}, G_{gt})}\right)\right]$$

$$\geq E_{Q(z|\rho_{gt},G_{gt})}\left(\log P(\rho\,|\,z, G)\frac{P(z\,|\,G)}{Q(z\,|\,\rho_{gt}, G_{gt})}\right)$$

$$E_{Q(z|\rho_{gt},G_{gt})}\big(f(z)\big) = \int dz\, f(z) Q(z\,|\,\rho_{gt}, G_{gt})$$

Jensen's inequality & the concave feature of the logarithm function

- Define loss function = - Evidence lower bound

$$\mathscr{L} = -E_{Q(z|\rho_{gt},G_{gt})}\left(\log P(\rho\,|\,z, G)\frac{P(z\,|\,G)}{Q(z\,|\,\rho_{gt}, G_{gt})}\right)$$

$$= -E_{Q(z|\rho_{gt},G_{gt})}\big(P(\rho\,|\,z, G)\big) + KL(Q(z\,|\,\rho_{gt}, G_{gt})\,||\,P(z\,|\,G))$$

$$\int dz Q(z\,|\,\rho_{gt}, G_{gt}) \log\left(\frac{Q(z\,|\,\rho_{gt}, G_{gt})}{P(z\,|\,G)}\right)$$

KL divergence gives the difference between 2 probabilities expected to denoise

$$P(\rho\,|\,z, G) = P(\rho\,|\,G)\; P(G\,|\,\rho, z)/P(G\,|\,z)$$

Prior information | Likelihood | Evidence

SVAE

$$\mathscr{L} = -E_{Q(z|\rho_{gt},G_{gt})}\left(\frac{P(\rho\,|\,G)P(G\,|\,\rho,z)}{P(G\,|\,z)}\right) + KL(Q(z\,|\,\rho_{gt},G_{gt})\,||\,P(z\,|\,G))$$

Prior information:

$$P(\rho\,|\,G) = \frac{1}{Z_s}e^{S}, \qquad S = \int d\omega\ \rho_{gt} - \rho(z) - \rho(z)log\frac{\rho(z)}{\rho_{gt}}.$$

Shannon-Jaynes entropy term: ground truth

ρ_{gt}: ground truth value of the spectral function, best prior information of ρ!

Likelihood:

$$P(G\,|\,\rho,z) = \frac{1}{Z_L}e^{-L}, \qquad L = L_j = \sum_j \frac{(\hat{G}_j[\rho(z)] - G_j)^2}{\alpha^2(z)G_j^2}.$$

G: correlator data

$\hat{G}$: computed from ρ(z)

From Maximum Entropy Method (MEM)

$\alpha(z)$: controls the relative weight of the entropy (having ρ close to the prior) and the likelihood (having $\rho(z)$ close to the data G)

Evidence:

$$P(G\,|\,z) = \int \mathscr{D}\rho(z)P(\rho\,|\,G)P(G\,|\,\rho,z) = \int \mathscr{D}\rho(z)\,\frac{1}{Z_sZ_L}e^{-L+s}$$

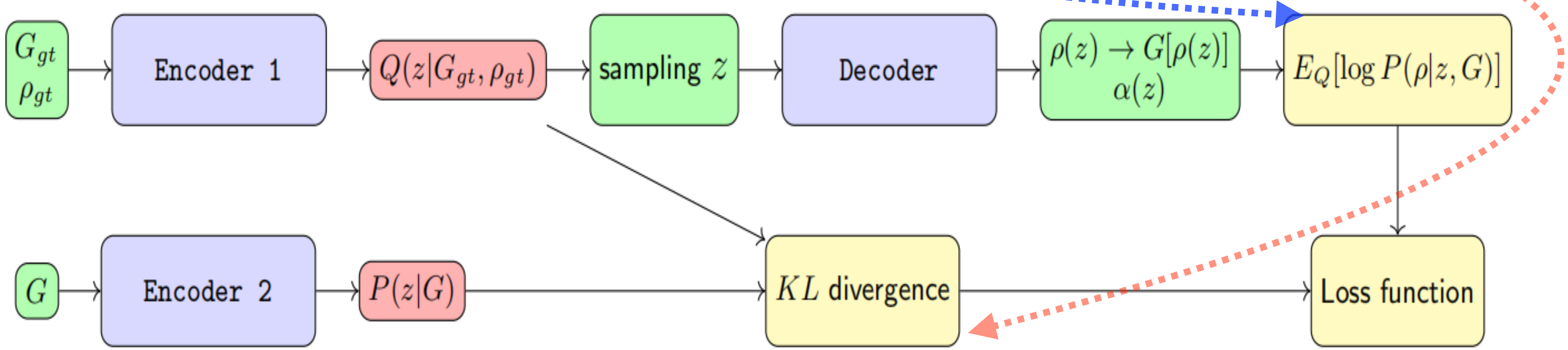
Topology of SVAE

Decoder | Encoder1 | Encoder2

$$\mathscr{L} = -E_{Q(z|\rho_{gt},G_{gt})}\left(\frac{P(\rho\,|\,G)P(G\,|\,\rho, z)}{P(G\,|\,z\,)}\right) + KL(Q(z\,|\,\rho_{gt}, G_{gt})\,||\,P(z\,|\,G))$$

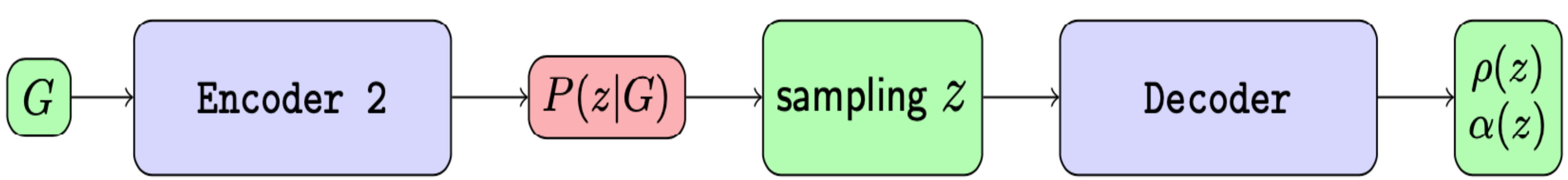
Epoch = 5 million

Training:



1. feed ground truth value G_{gt} and ρ_{gt} into Encoder 1
2. feed G generated stochastically to Encoder2;
   Noisy G is sampled from a Gauss distribution $N(G_{gt},\ \sigma)$ with $\sigma = b \times G_{gt}$ and $b = \sigma_{latt}/G_{latt}$
3. Perform the feeding in steps 1 and 2 simultaneously
4. For each repeat steps 1, 2 & 3.

Reconstruction:



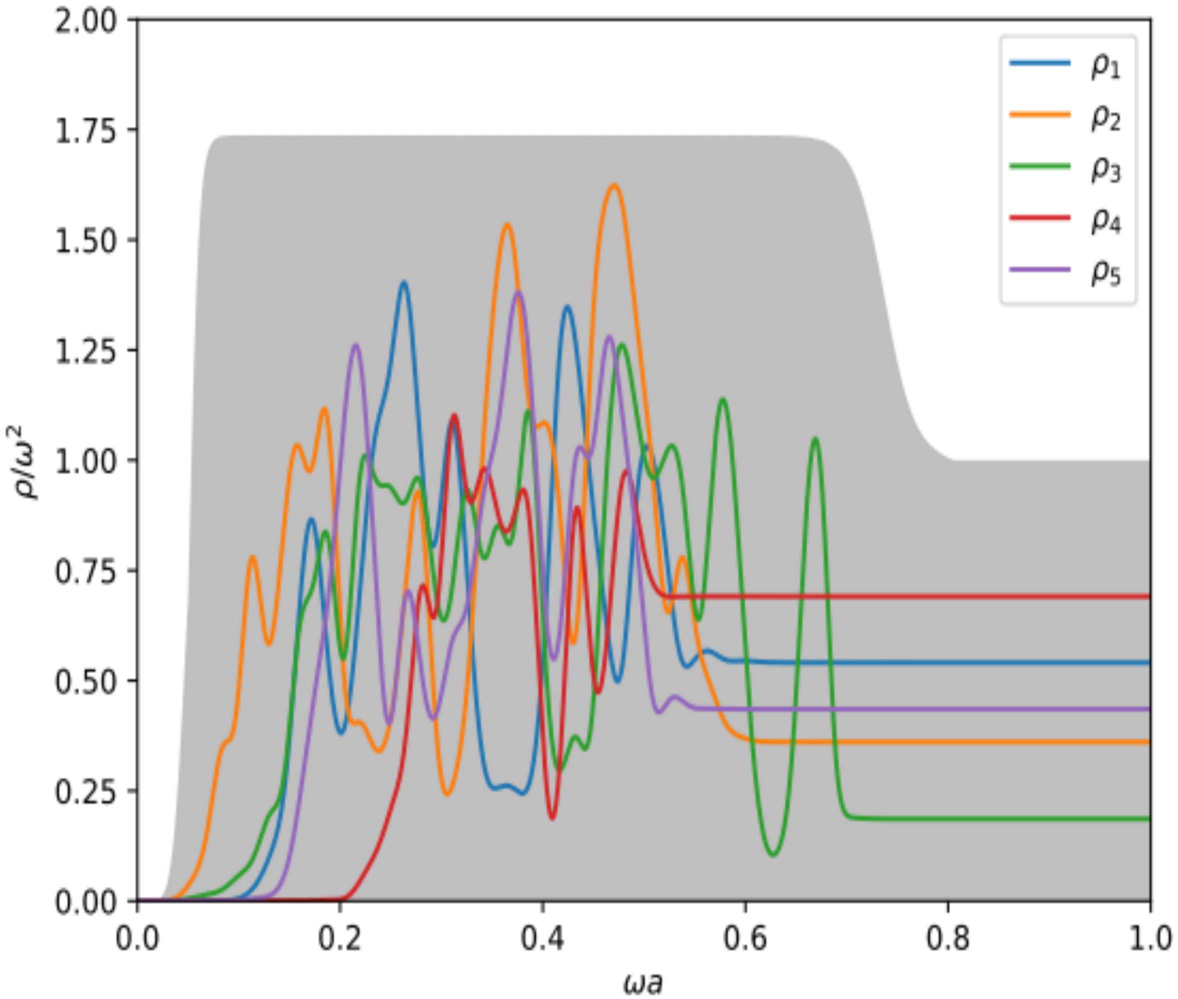
A general spectral function used for the training

$$\frac{\rho_{train}}{\omega^2} = \hat{\theta}(\omega, \delta_1,\zeta_1)\left(\sum_{i=1}^{\hat{N}_g} C_i e^{-\left(\frac{\omega - M_i}{\gamma}\right)^2}\left(1 - \hat{\theta}(\omega, \delta_2,\zeta_2)\right) + C_0\hat{\theta}(\omega, \delta_2,\zeta_2)\right)$$

Gaussian peaks — continuum part

M_i: peak location with fixed interval γ: fixed peak width $C_0 \in [0, 1]$

$\hat{N}_g = 50$, number of Gauss peaks



Training samples

Step functions: $\hat{\theta}(\omega, \delta_i, \zeta_i) = \frac{1}{1 + \exp(-\frac{\omega-\delta_i}{\zeta_i})}$

Parameters	interval	Parameters	interval
M_i	$[0.05, 0.8]$	$\zeta_2(\sim [0.1, 0.2]m_c)$	$[0.005, 0.02]$
C_i	$[0, 1]$	$\delta_1(\sim [1, 3]m_c)$	$[0.05, 0.3]$
$\zeta_1(\sim [0.1, 0.2]m_c)$	$[0.005, 0.02]$	$\delta_2(\sim 8m_c)$	$[\delta_1 + 0.1, \delta_1 + 0.6]$

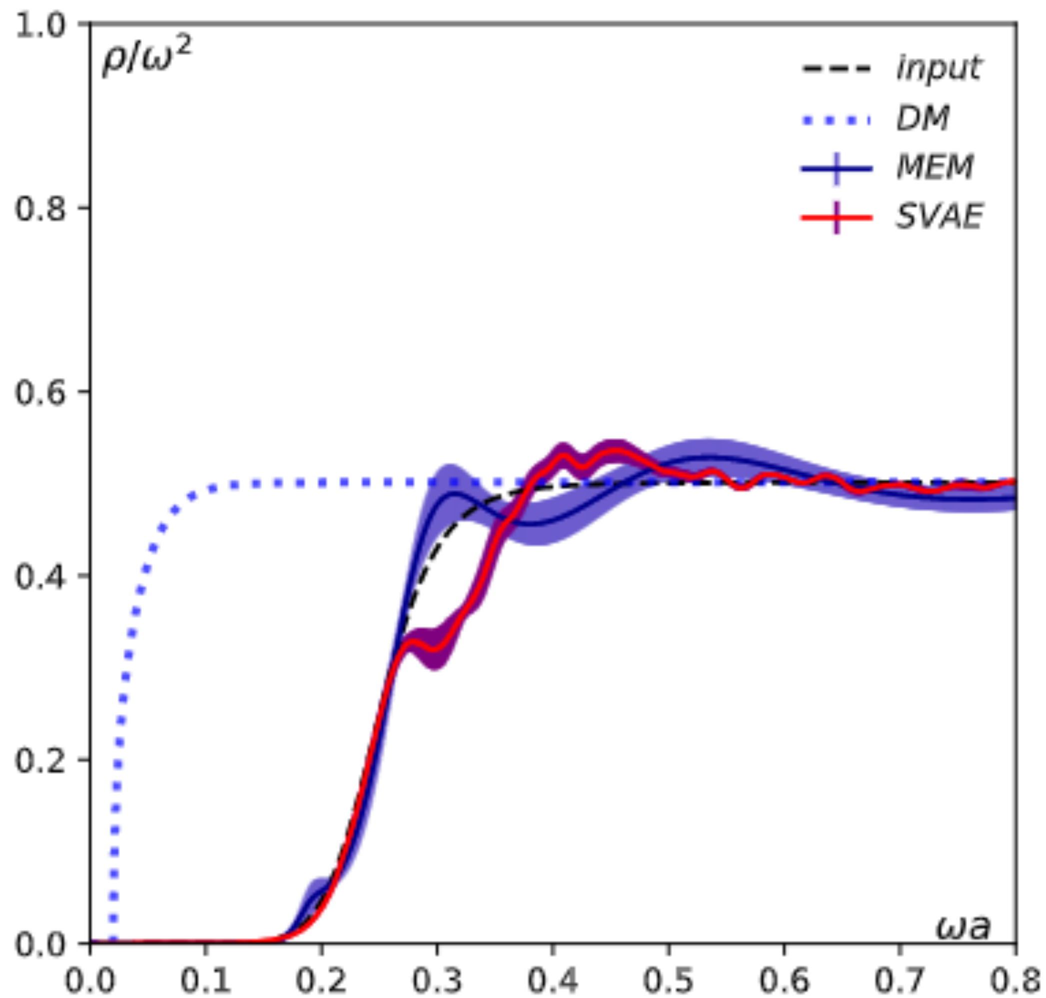
The parameters are set according to the theoretically range of physics.

H. T. Ding et al, Phys. Rev. D 86, 014509 (2012)

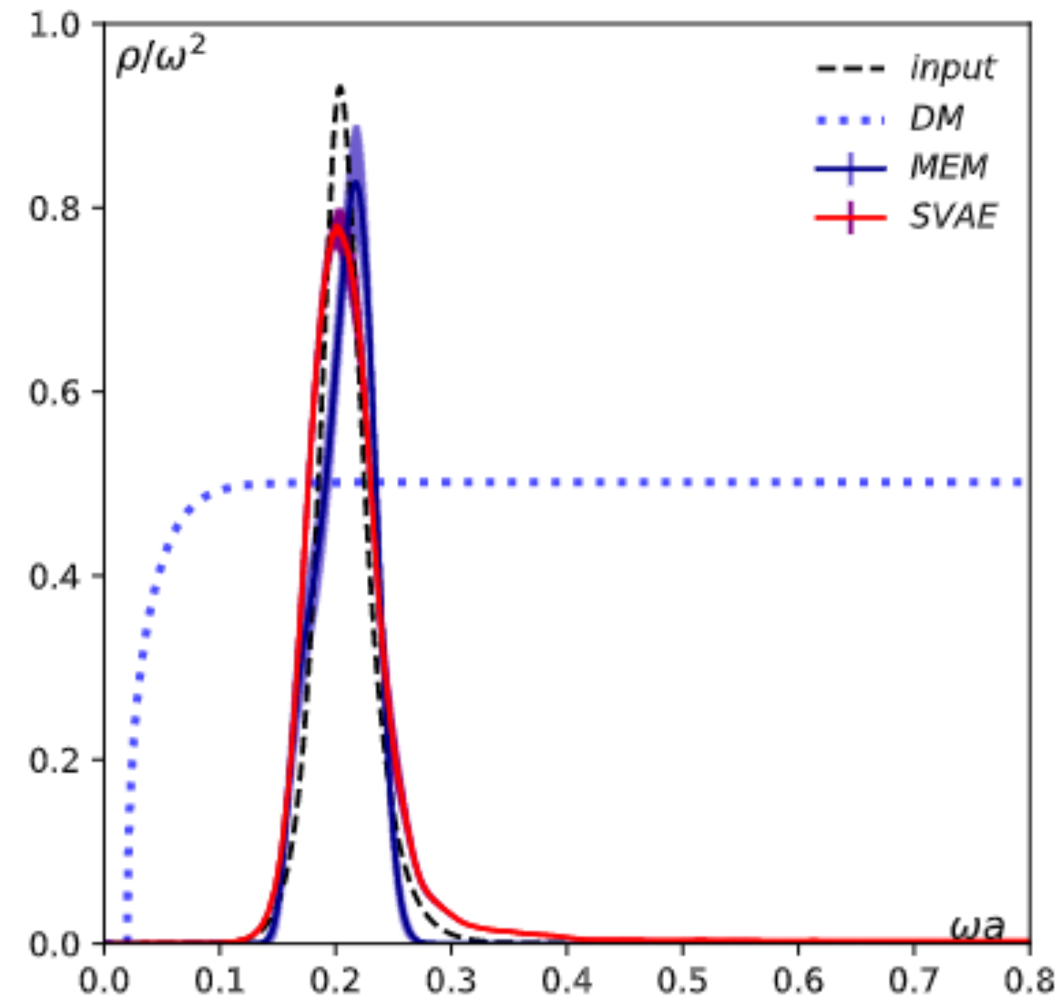
Mock data test I: simplest cases

Temporal extent $N_\tau = 96$

MEM results depend on the chosen of default model (DM)



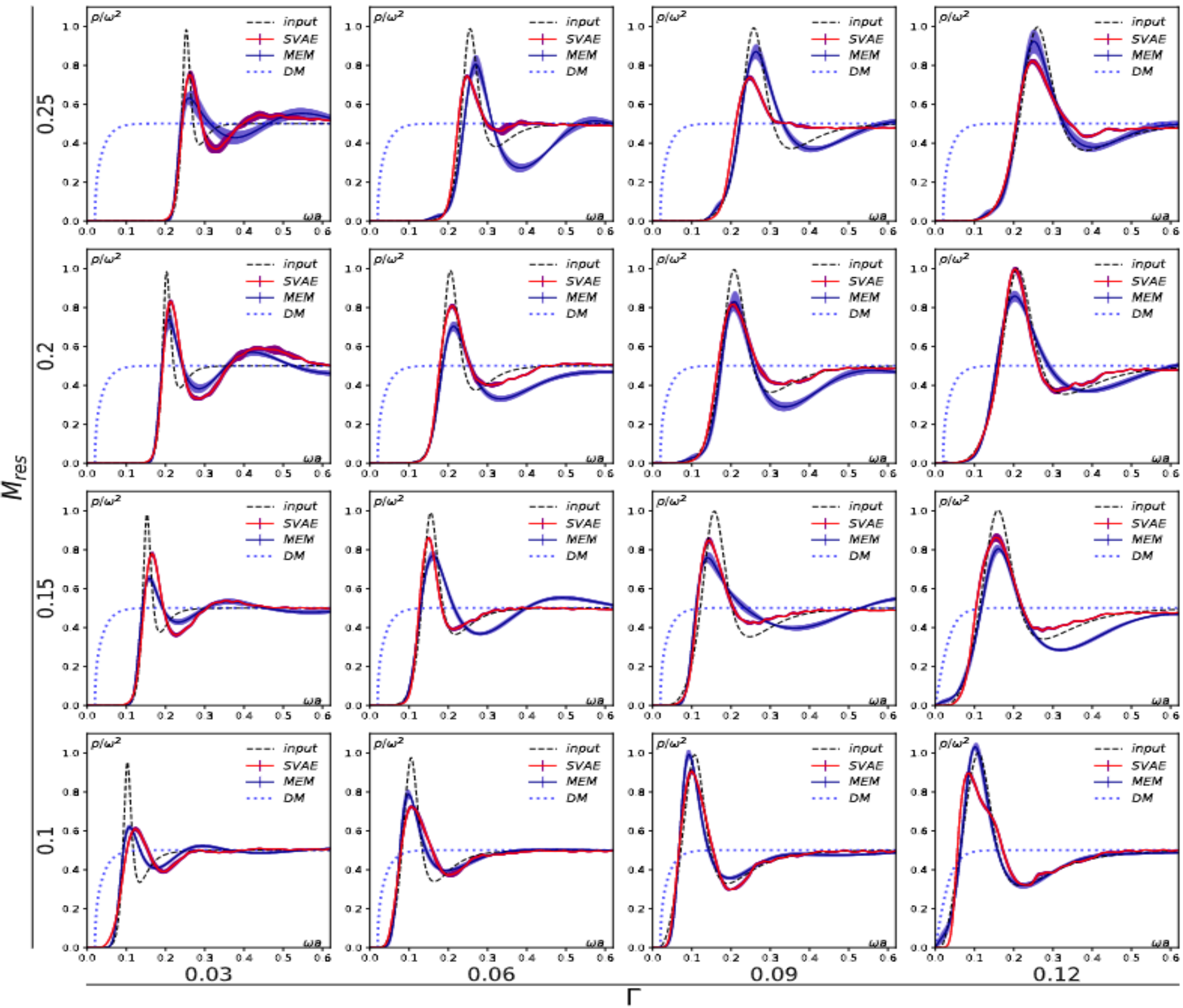
Only a continuum part



Only a resonance peak

Both output spectral functions from the SVAE and MEM can more or less reproduce the input spectral functions

Mock data test II: A resonance peak + continuum part



$N_\tau = 96$

Charm quark

Peak location

$M_{res} = 0.1, 0.15, 0.2, 0.25$

Peak width

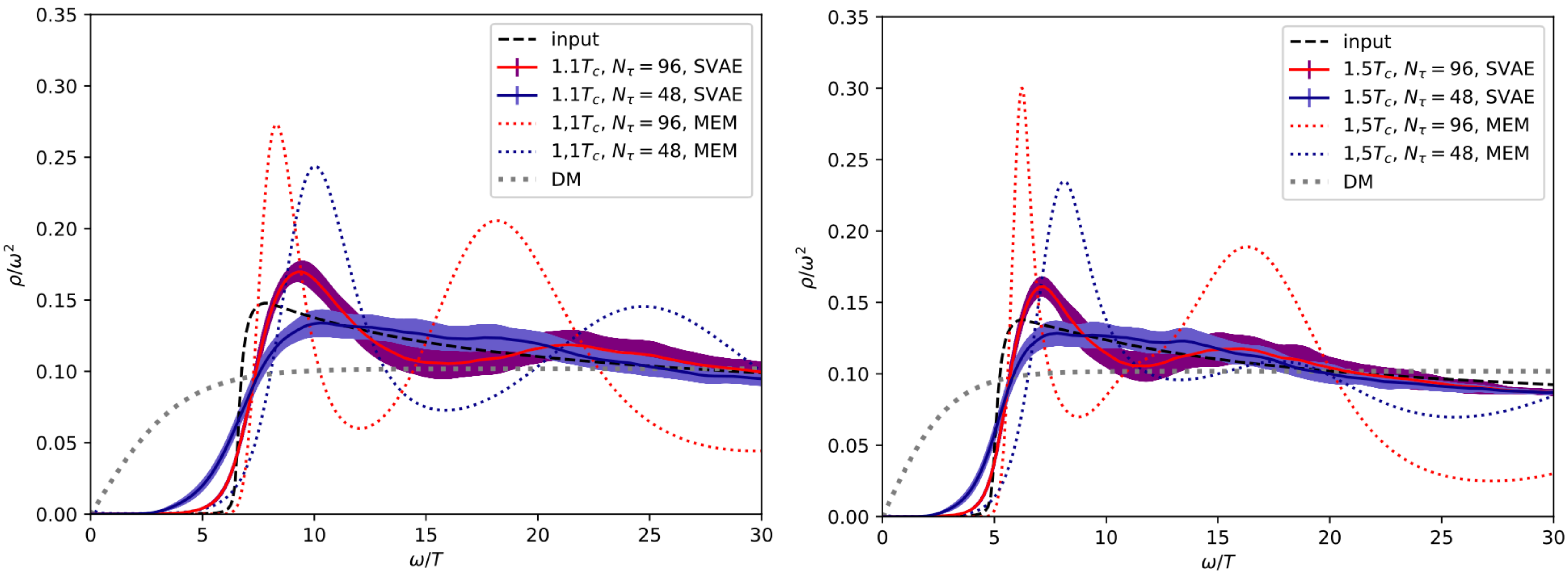
$\Gamma = 0.03, 0.06, 0.09, 0.12$

Reconstruction quantity depends more on Γ rather than M_{res}

Mock data test III: Non-relativistic QCD

Input ρ obtained from the Non-relativistic QCD (NRQCD)

Y. Burnier et al, JHEP 2017 (11), 206

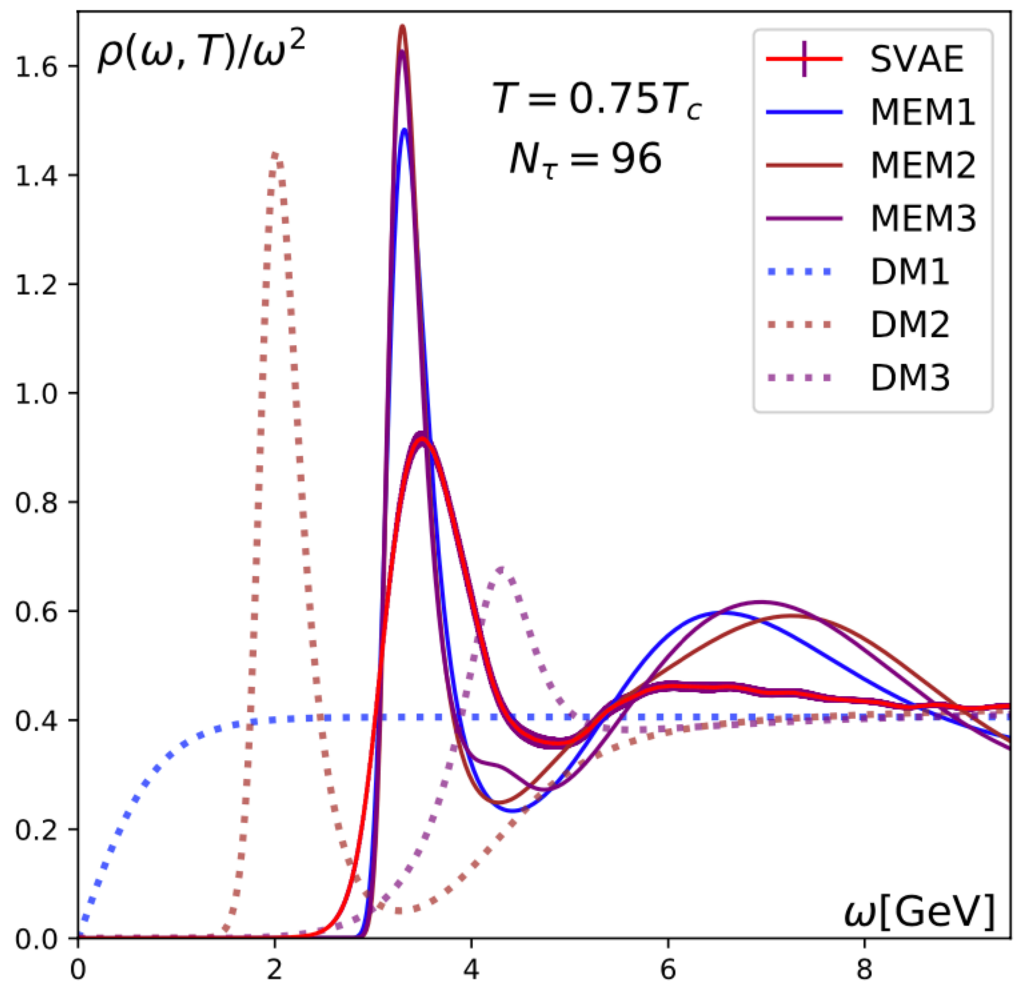


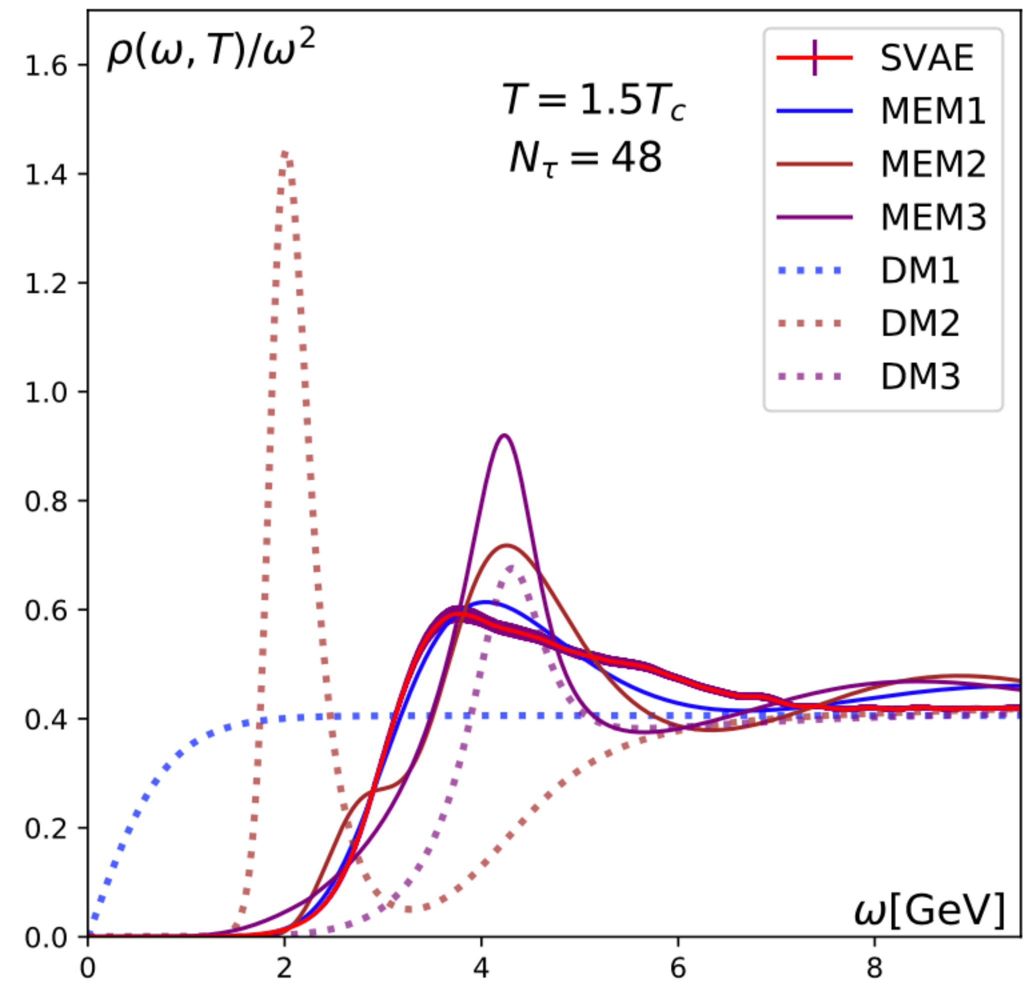
Although the SVAE results are closer to the mock data, both the SVAE and MEM cannot reproduce the input NRQCD spectral function in a satisfactory way.

Application to Charmonium correlator in the pseudo-scalar(η_c) channel

Correlators computed on the finest lattices $128^3 \times 96(0.75T_c)$ and $128^3 \times 48(1.5T_c)$ with a fixed scale approach.

H. T. Ding et al, Phys. Rev. D 86, 014509 (2012)





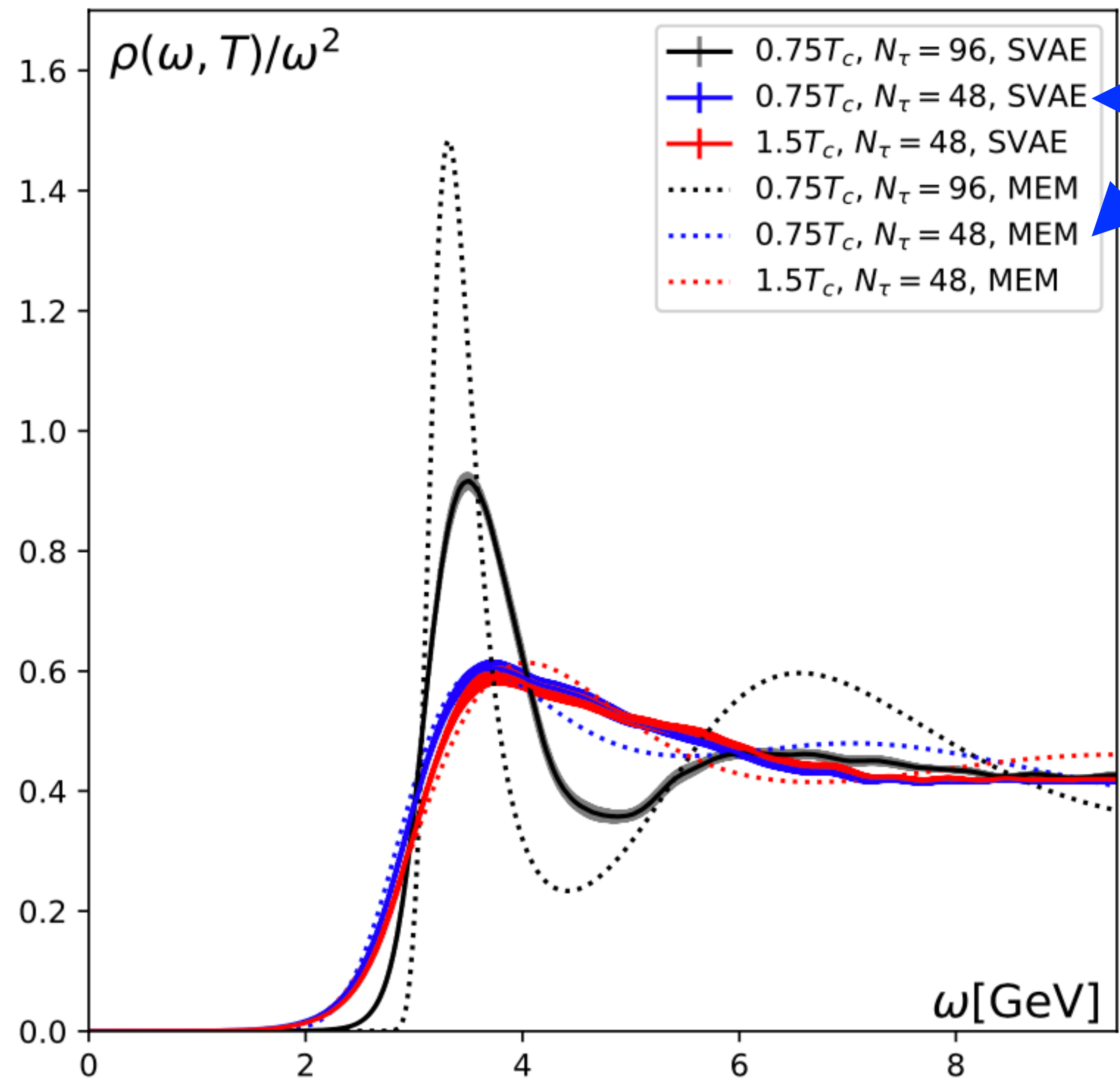
- Peak locations reconstructed from these two methods are more reliable compared to the peak height

- Based on the results from both methods it seems that the resonance peak for ηc is substantially modified at 1.5Tc

Application to Charmonium correlator in the pseudo-scalar(η_c) channel

Correlators computed on the finest lattices $128^3 \times 96(0.75T_c)$ and $128^3 \times 48(1.5T_c)$ with a fixed scale approach.

H. T. Ding et al, Phys. Rev. D 86, 014509 (2012)



Reconstruct correlator with N_τ= 96 to that with N_τ= 48 of ρ fixed at 0.75 Tc

$$G_{rec}(\tau, T; T') = \int d\omega K(\tau, T; T')\rho(\omega, T')$$

$$= \sum_{\tau'=\tau}^{N'_\tau - N_\tau + \tau} G(\tau', T')$$

$$T = 1.5T_c,\ T' = 0.75T_c,\ N_\tau = 48,\ N'_\tau = 96$$

- $N_\tau = 48$: SVAE is consistent with MEM at both 0.75 Tc and 1.5Tc.
- Output $\rho(\omega)$ larger depends on N_τ.

S.-Y Chen, H.-T Ding, F.-Y Liu, G. Papp and C.-B. Yang, arXiv:2110.13521

Summary& Outlook

- We propose a novel neural network, SVAE, which can be trained to obtain the most probable image of the spectral function
- The loss function of SVAE includes an entropy and a likelihood term which are balanced by a weight $\alpha(z)$
- For training: A general ρ is used with corresponding correlator having the error of correlators mimic to LQCD
- Mock tests shows that SVAE is comparable to MEM and even outperforms MEM in certain cases
- For the application to lattice QCD data of charmonium correlator in the pseudo-scalar channel, SVAE is consistent with MEM results.
- Output $\rho(\omega)$ larger depends on N_τ

➤ Trial with a different kernel K = $\omega^2(\omega^2 + p^2)$ or $e^{-\tau\omega}$

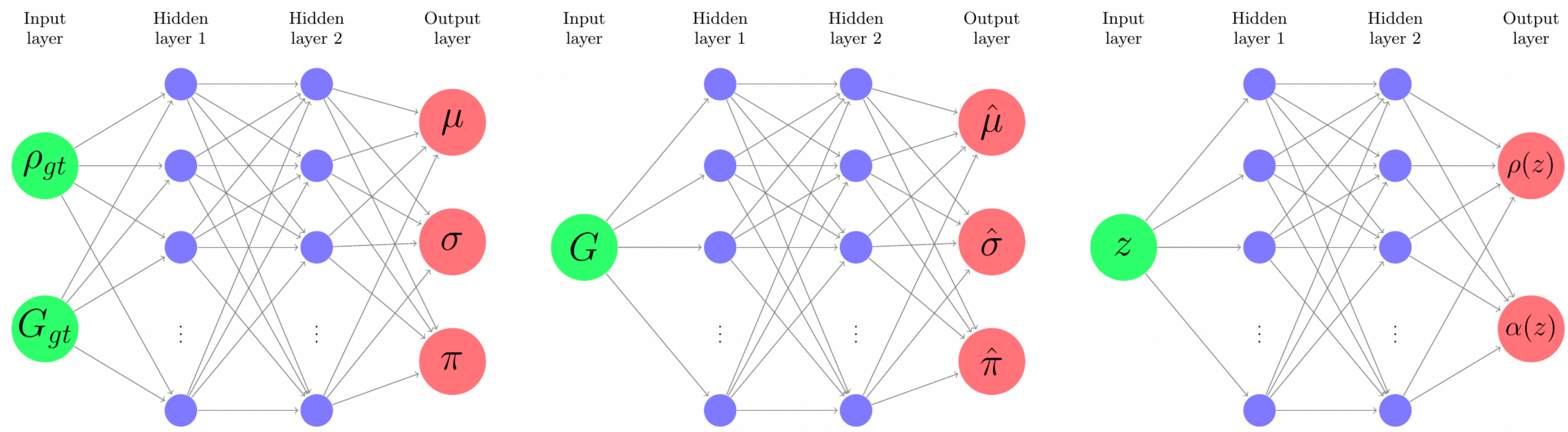
$$G(\tau, T) = \int_0^\infty \frac{\mathrm{d}\omega}{2\pi} \, K(\omega, \tau, T)\rho(\omega, T)$$

Current: $K(\omega, \tau, T) = \frac{\cosh(\omega(\tau - \frac{1}{2T}))}{\sinh(\frac{\omega}{2T})}$

Thank you!

Back up

Networks



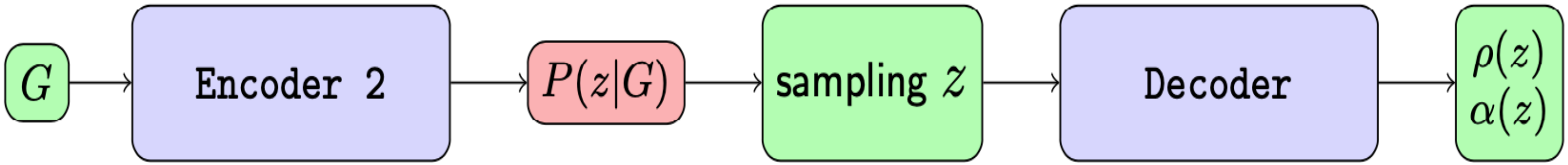
(a) Encoder 1 (b) Encoder 2 (c) Decoder

$$Q(z|\rho_{gt}, G_{gt}) = \sum_i^{N_g} \pi_i \prod_k^{N_z} \mathcal{N}(\mu_{i,k}, \sigma_{i,k}) \quad P(z|G) = \sum_i^{N_g} \hat{\pi}_i \prod_k^{N_z} \mathcal{N}(\hat{\mu}_{i,k}, \hat{\sigma}_{i,k}).$$

Q and P(z|G): parameterized based on the Gauss mixture model with $\mathcal{N}$ the Gauss function, outputs from Encoder 1 and 2

$\rho(z), \alpha(z)$: outputs of Decoder

Topology of the sVAE for reconstruction



$$\tilde{\rho}_{\Theta} = \frac{1}{N_c}\sum_{m=1}^{N_c}\int dz \rho(z) P(G_m|\rho, z) P(z|G_m)$$

$$\to \frac{1}{N_c}\sum_{m=1}^{N_c}\sum_{n=1}^{N_s}\frac{\rho(z_n)P(G_m|\rho, z_n)}{\sum_{k=1}^{N_s}P(G_m|\rho, z_k)} \equiv \frac{1}{N_c}\sum_{m=1}^{N_c}\tilde{\rho}_{\Theta,m}\,.$$

N_s: Number of samples in Z space

N_c: Number of bootstrap samples in configurations space

Bootstrap method: 1. Firstly randomly choose an epoch Θ (uniformly)

2. At this certain epoch we sample N_c times of $\tilde{\rho}_{\Theta,m}$ uniformly: $\tilde{\tilde{\rho}}_i$

Final mean value and its uncertainty:

$$\overline{\rho} = \frac{1}{N_{btp}}\sum_{i=1}^{N_{btp}}\tilde{\tilde{\rho}}_i \qquad \sigma_\rho = \sqrt{\frac{1}{N_{btp}}\sum_{i=1}^{N_{btp}}(\tilde{\tilde{\rho}}_i - \overline{\rho})^2}$$

N_{btp} : Number of bootstrap samples in model parameter space

Application to Charmonium correlator in the pseudo-scalar(η_c) channel

Correlators computed on the finest lattices $128^3 \times 96(0.75T_c)$ and $128^3 \times 48(1.5T_c)$ with a fixed scale approach.

H. T. Ding et al, Phys. Rev. D 86, 014509 (2012)

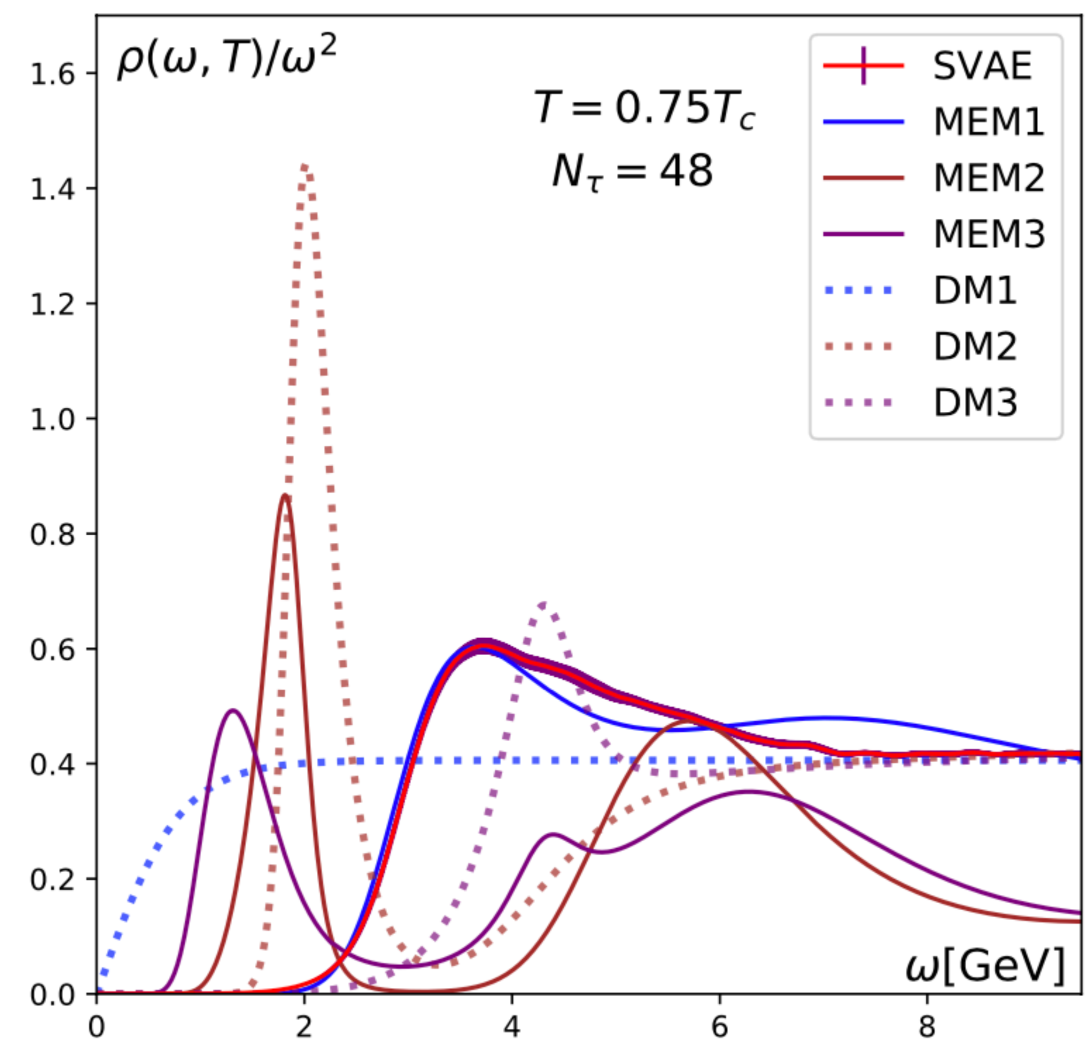
- At higher temperatures, the information of the spectral function is compressed into the correlation function with a short temporal extent

Reconstruct correlator with N_τ=96 to that with N_τ= 48 with ρ fixed at 0.75 Tc

$$G_{rec}(\tau, T; T') = \int d\omega K(\tau, T; T')\rho(\omega, T')$$

$$= \sum_{\tau'=\tau}^{N'_\tau - N_\tau + \tau} G(\tau', T')$$

$$T = 1.5T_c,\ T' = 0.75T_c,\ N_\tau = 48,\ N'_\tau = 96$$



- Output $\rho(\omega)$ larger depends on N_τ.