NOVEMBER 2, 2020

Argonne NATIONAL LABORATORY

Machine Learning-based Inversion of Nuclear Responses

Krishnan Raghavan[1]

- **Prasanna Balaprakash**[1]**, Alessandro Lovato**[2,3]**, Noemi Rocco**[2] **and Stefan Wild**[1]

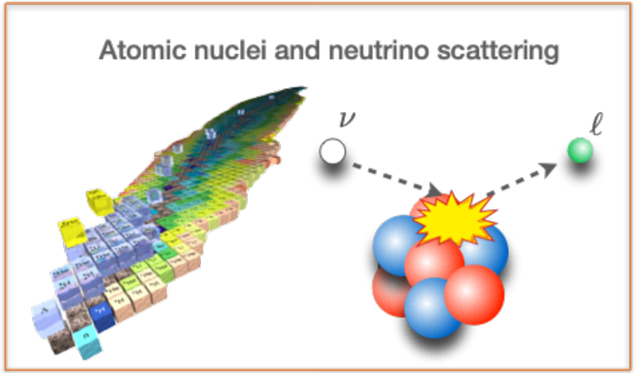
Mathematics and Computer Science Division[1]**, Physics Division**[2]

[1,2]Argonne National Laboratory, [3]Istituto Nazionale di Fisica Nucleare (INFN)

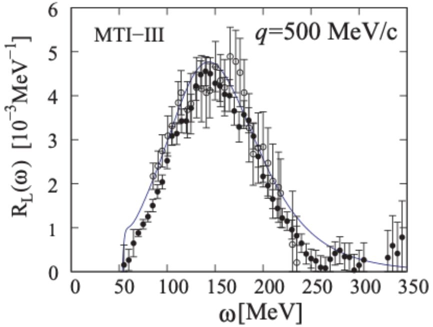
Outline

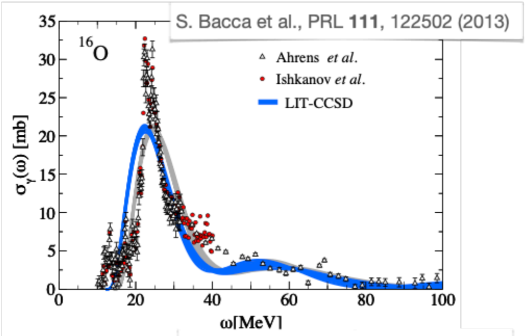
- Problem statement
- Challenges
- Inversion using Neural Networks
- Dataset
- Results
- Conclusion and future works

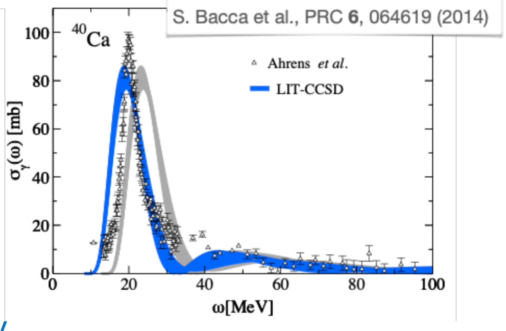
Introduction



The response functions contain all information on the structure and dynamics of the target







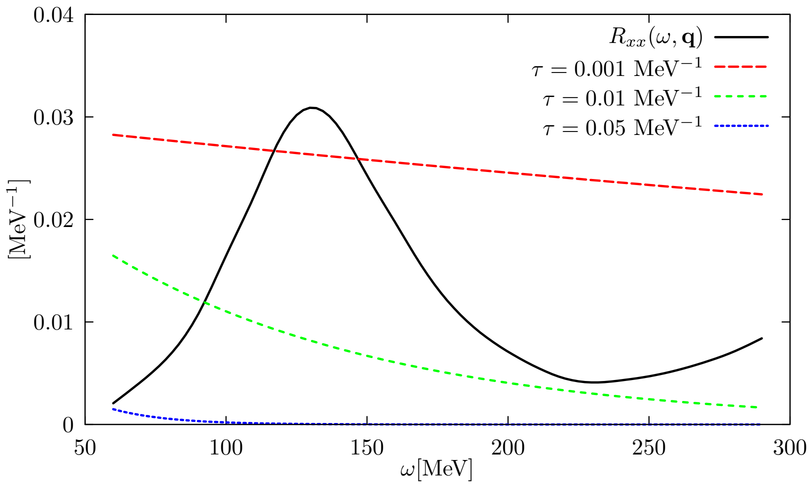
Measured by electron-scattering experiments

$$R_{\alpha\beta}(\omega,\mathbf{q}) = \sum_f \langle\Psi_0|J^\dagger_\alpha(\mathbf{q})|\Psi_f\rangle\langle\Psi_f|J_\beta(\mathbf{q})|\Psi_0\rangle\delta(\omega - E_f + E_0)$$

Laplace Transform

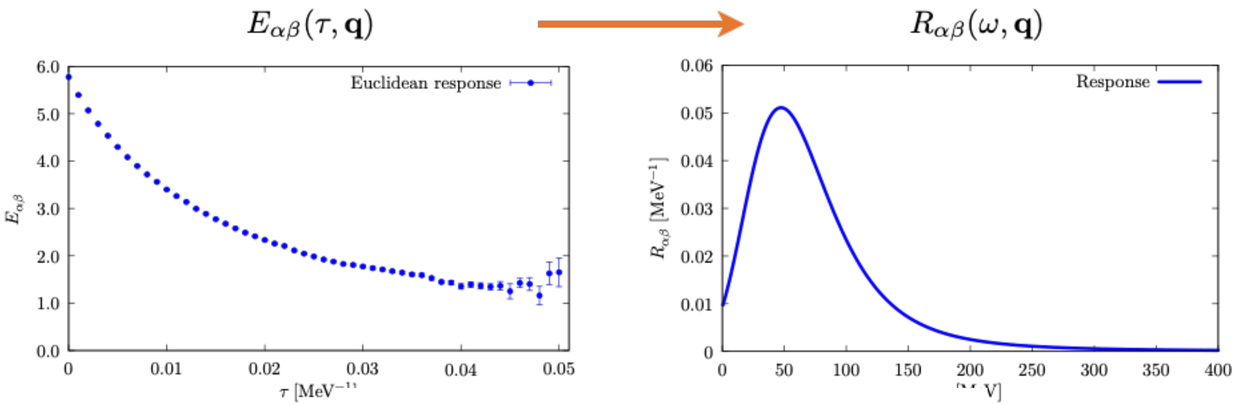
Valuable information on the energy dependence of the response functions can be inferred from their **Laplace transforms**

The system is first heated up by the transition operator.
Its cooling determines the **Euclidean response** of the system



$$\underbrace{E_{\alpha\beta}(\tau, \mathbf{q})}_{\text{Euclidean Response}} \equiv \int d\omega e^{-\omega\tau} \underbrace{R_{\alpha\beta}(\omega, \mathbf{q})}_{\text{Response Function}}$$

Laplace Transform

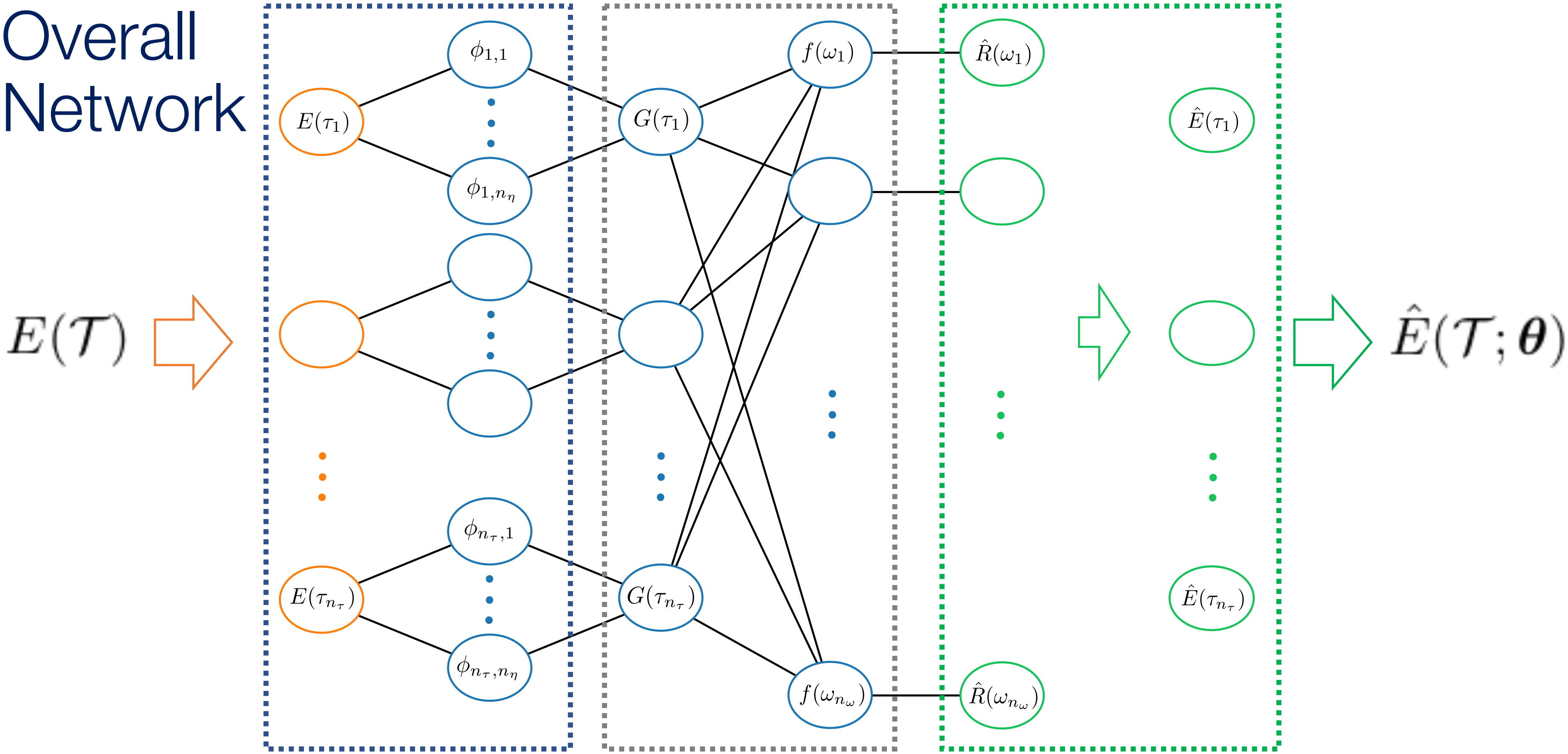


$$R(\Omega) = K(\Omega, \mathcal{T})^{-1} E(\mathcal{T}).$$

- The inverse is ill-posed.
  - Multiple response functions can have the same Euclidean response (within errors).
- High noise in the Euclidean response results in unstable inversions.

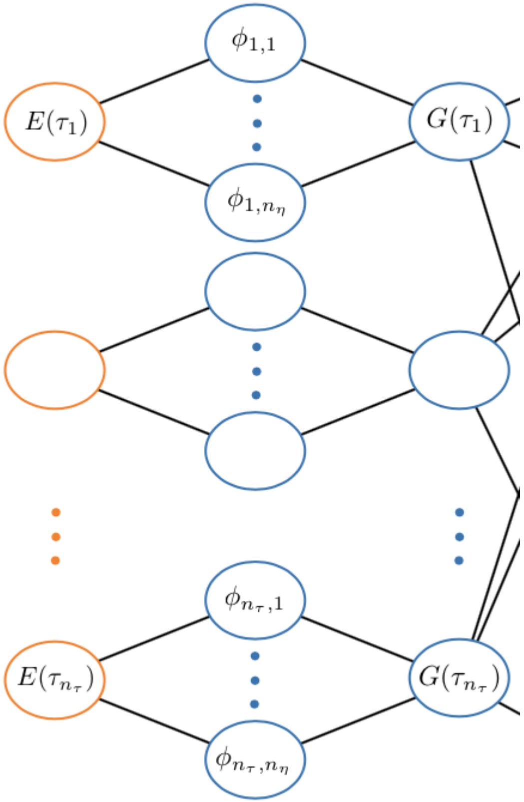
$$\hat{R}(\Omega; \boldsymbol{\theta}) = \frac{1}{\mathcal{N}_0} e^{f(E(\mathcal{T}); \boldsymbol{\theta})}$$

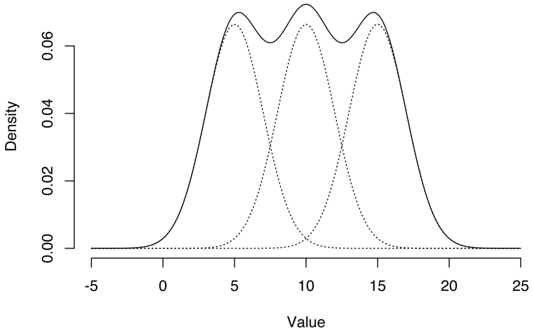
Overall Network

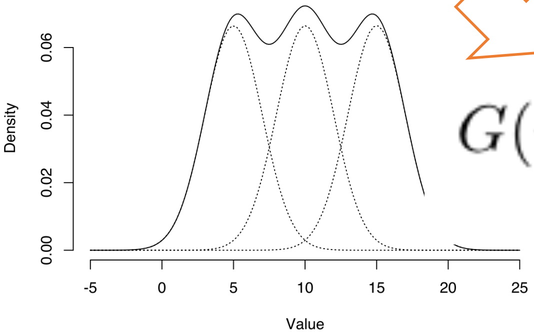


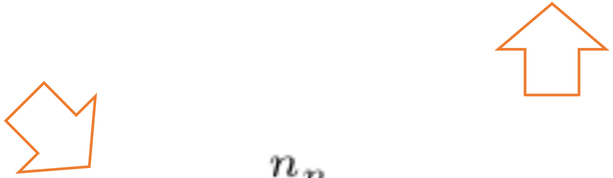
Mixture of Gaussians Neural Network Laplace Transform

Mixture of Gaussians







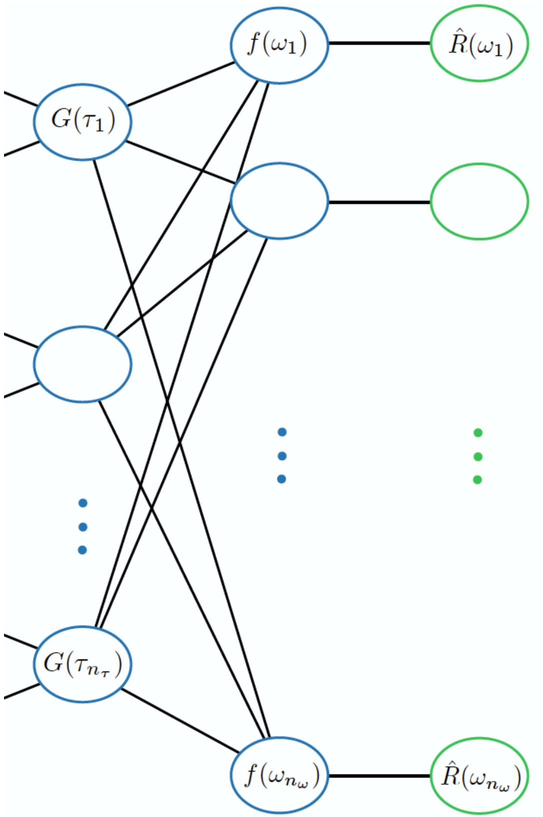


$$\begin{bmatrix} G(\tau_0) \\ \vdots \\ G(\tau_{n_\tau}) \end{bmatrix}$$

$$G(\tau_j) = \sum_{k=1}^{n_\eta} \phi(E(\tau_j), \mu_{j,k}, \sigma_{j,k})$$

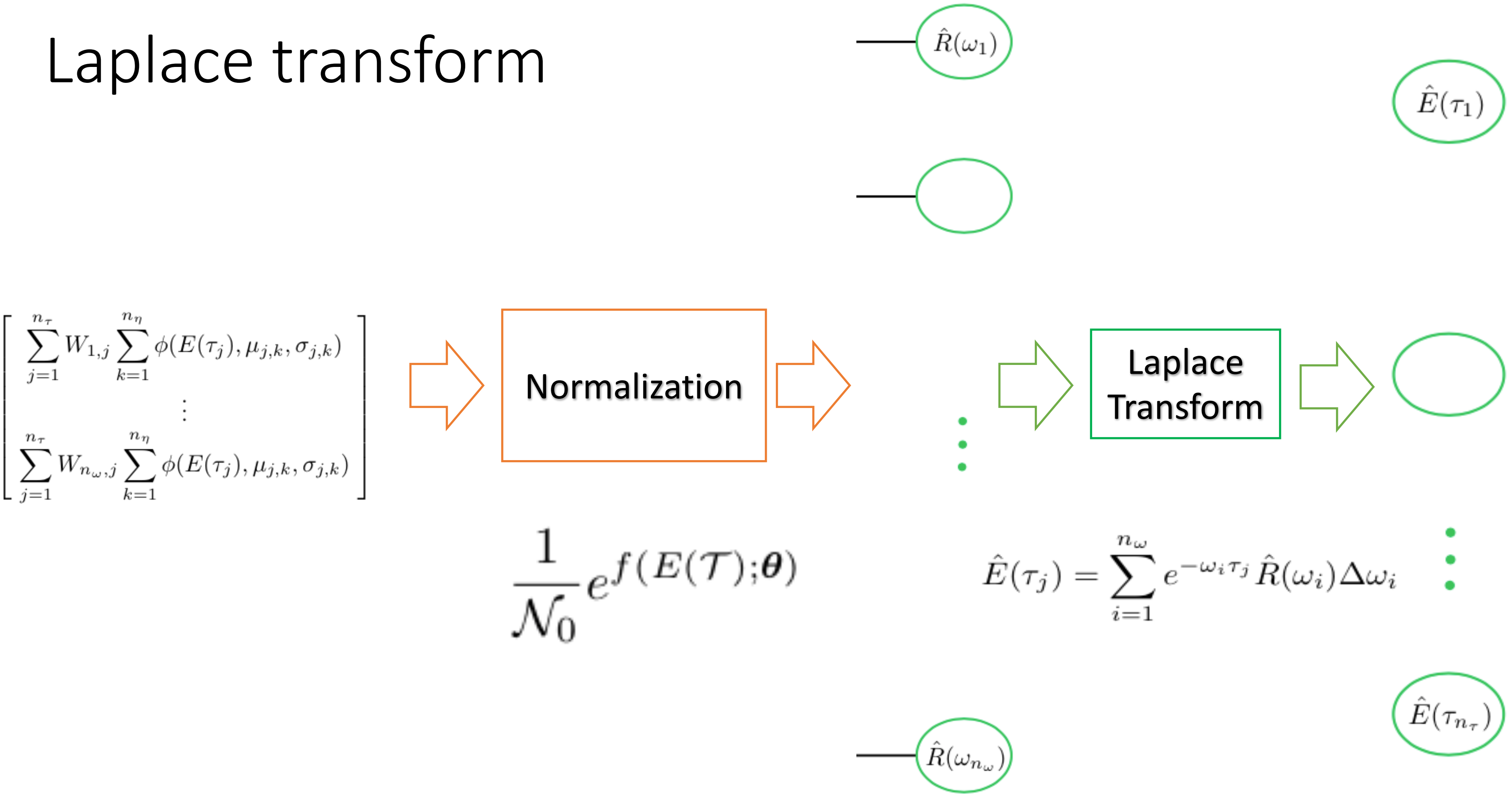
The Neural Network

$$\begin{bmatrix} G(\tau_0) \\ \vdots \\ G(\tau_{n_\tau}) \end{bmatrix}$$

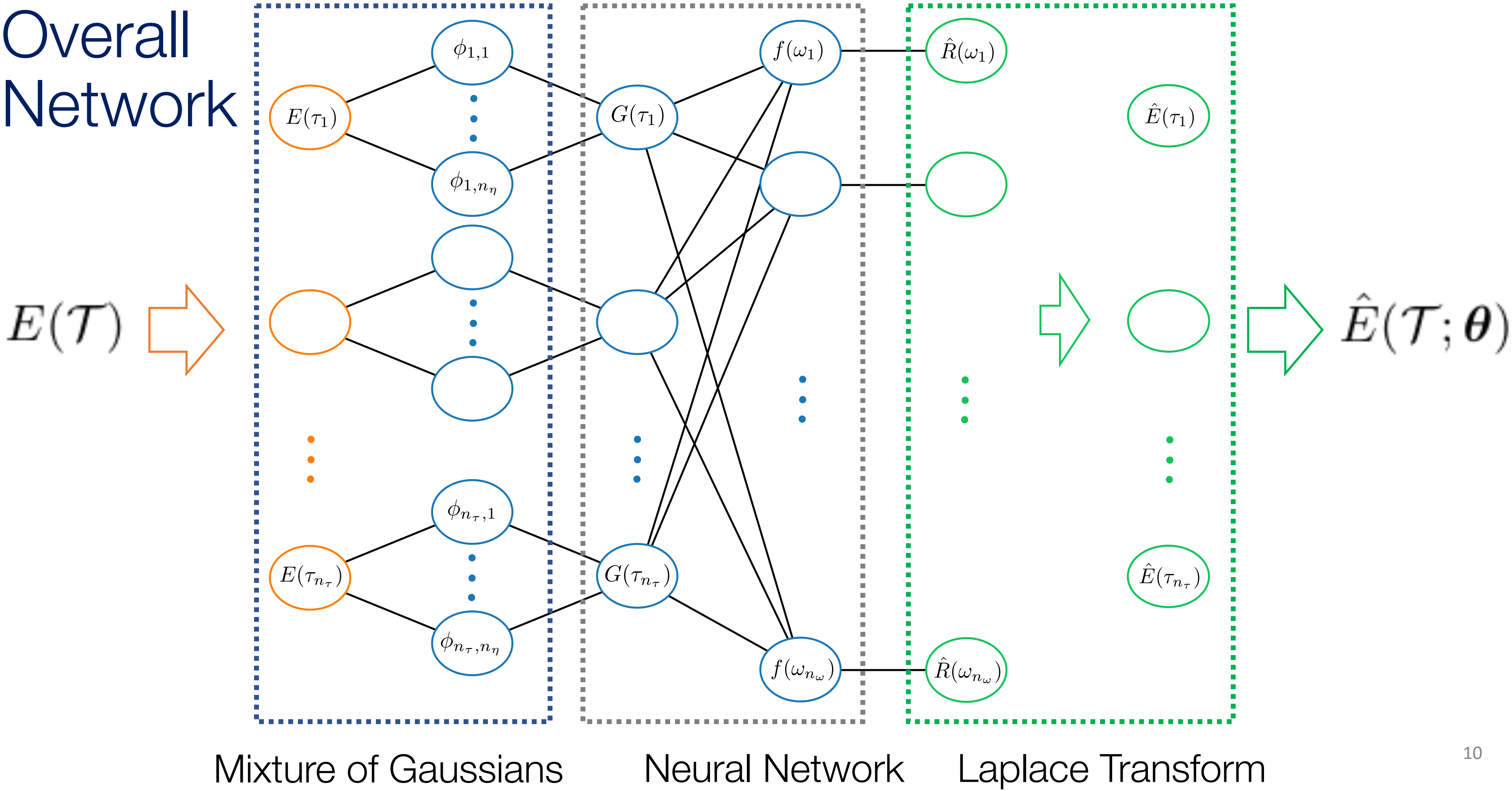


$$\begin{bmatrix} \sum_{j=1}^{n_\tau} W_{1,j} \sum_{k=1}^{n_\eta} \phi(E(\tau_j), \mu_{j,k}, \sigma_{j,k}) \\ \vdots \\ \sum_{j=1}^{n_\tau} W_{n_\omega,j} \sum_{k=1}^{n_\eta} \phi(E(\tau_j), \mu_{j,k}, \sigma_{j,k}) \end{bmatrix}$$

Laplace transform



Overall Network



Learning Problem

$$\min_{\boldsymbol{\theta}} \frac{1}{|\mathbb{T}|} \sum_{k \in \mathbb{T}} \ell\left(E_k(\mathcal{T}), R_k(\Omega), \hat{R}_k(\Omega; \boldsymbol{\theta})\right) \quad \boldsymbol{\theta} = (\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{W})$$

$$\ell(E_k, R_k, \hat{R}_k) = \gamma_R S_R(R_k, \hat{R}_k) + \gamma_E \chi^2_E(E_k, \hat{R}_k)$$

$$\sum_{i=1}^{n_\omega} \left(R(\omega_i) - \hat{R}(\omega_i) - R(\omega_i) \ln \left(\frac{R(\omega_i)}{\hat{R}(\omega_i)} \right) \right) \Delta\omega_i$$

Entropy

$$\frac{1}{n_\tau} \sum_{j=1}^{n_\tau} \frac{1}{\sigma_j^2} \left(E(\tau_j) - \hat{E}(\tau_j) \right)^2$$

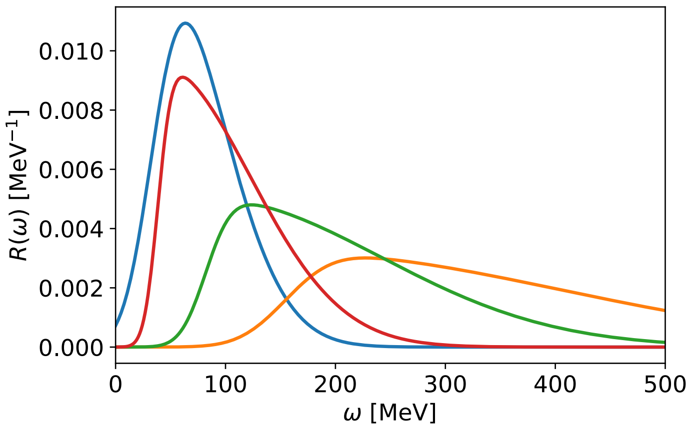
χ^2

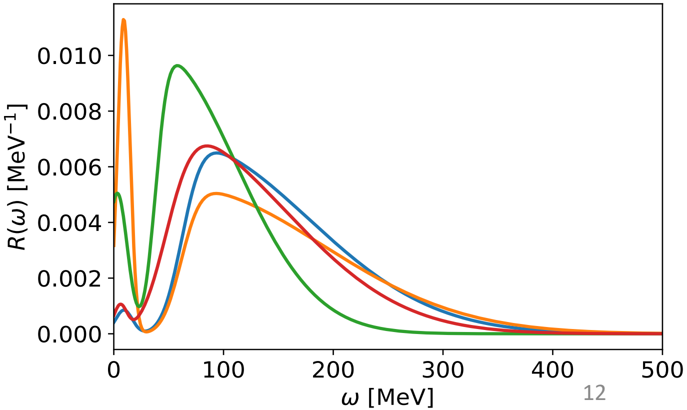
The Dataset and the Metrics

- We use two distinct data-sets of (E, R).
  - Two hundred thousand responses corresponding to one peak (Fig., top).
  - Two hundred thousand responses corresponding to two peak (Fig., bottom).

- Metrics– R^2, Entropy (S_R) and chi-squared (χ^2).
  - We run our simulations for thirty initial conditions and the mean and standard deviations are summarized.

- We compare our approach to MaxEnt.



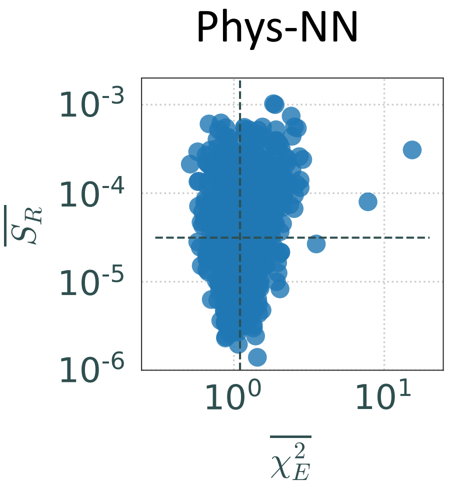


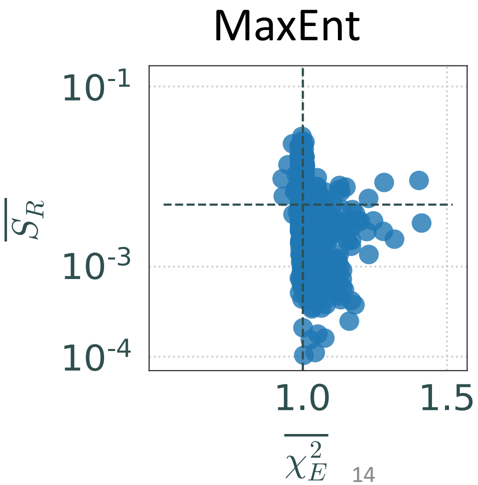
Overall Results

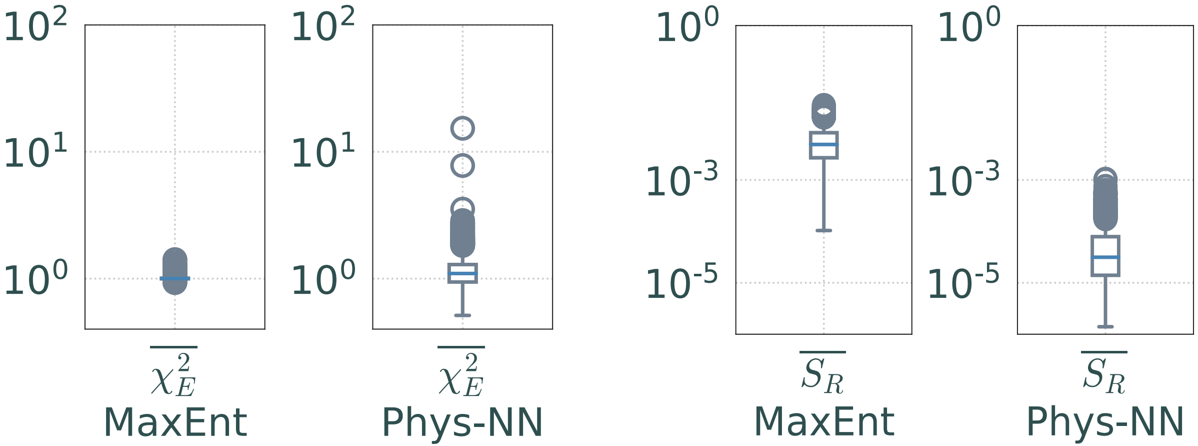
	$1-\overline{R^2_R}$ $\times 10^{-4}$	$\overline{\chi^2_E}$	$\overline{S_R}$ $\times 10^{-4}$
		Phys-NN	
One-peak	0.42	1.171(13)	0.72
Two-peak	9.04	3.220(87)	9.16
Combined	0.61	2.335(14)	3.66
		MaxEnt	
One-peak	29.7	1.015 (1)	60.4
Two-peak	84.8	1.016 (1)	107
Combined	57.2	1.015 (1)	83.7

One Peak Dataset

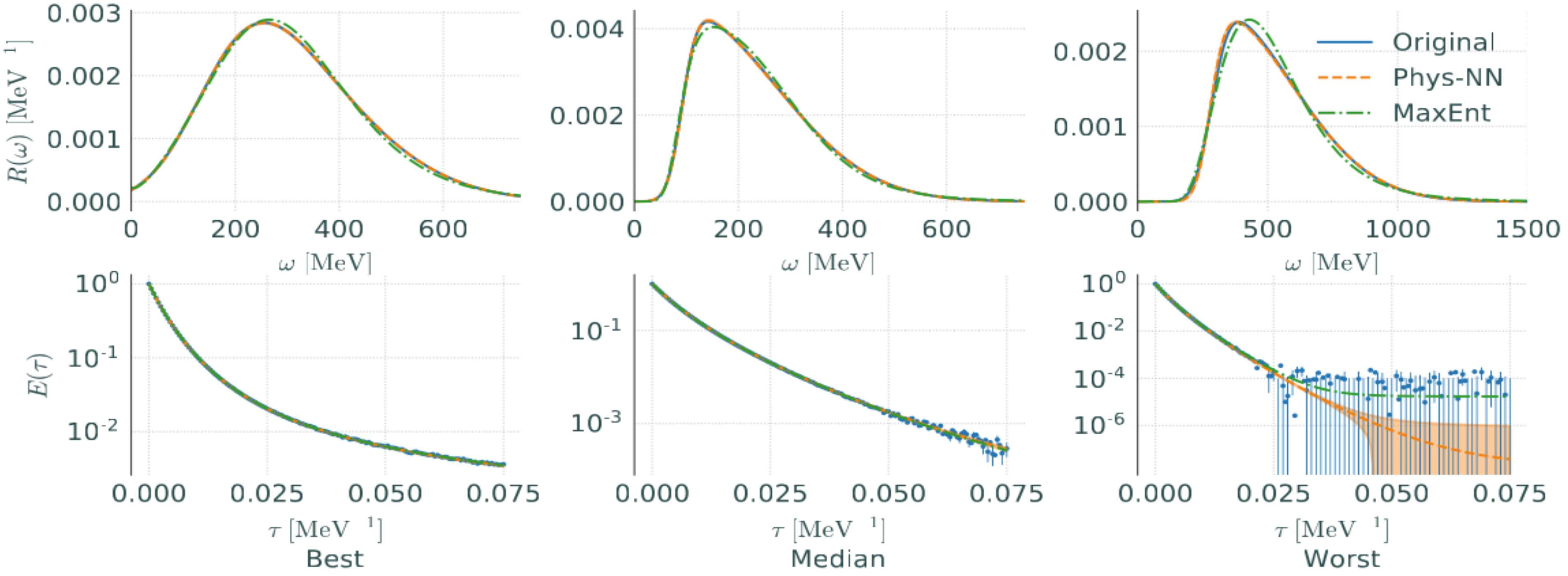
- First demonstrate performance with just the one peak data.
  - On the bottom, metrics on the test data
  - On the right, correlation between S_R and χ^2







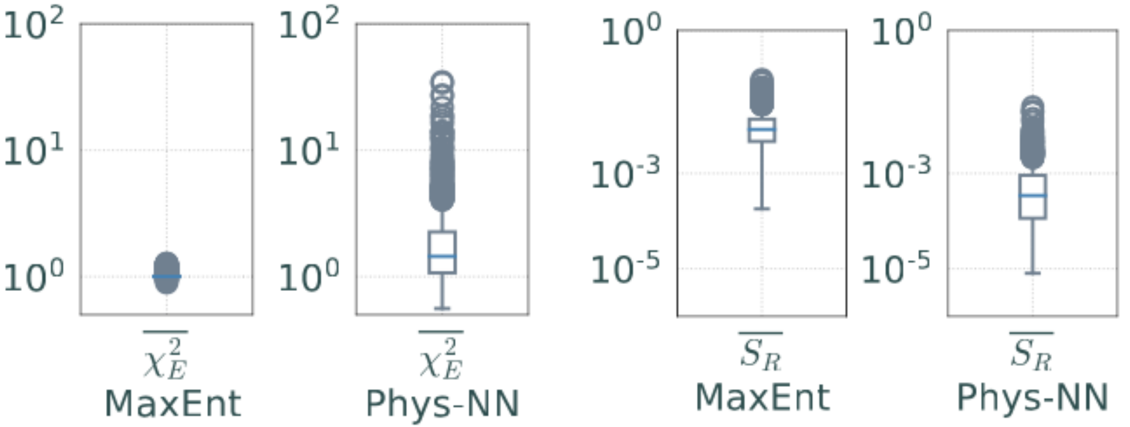
One Peak Reconstruction

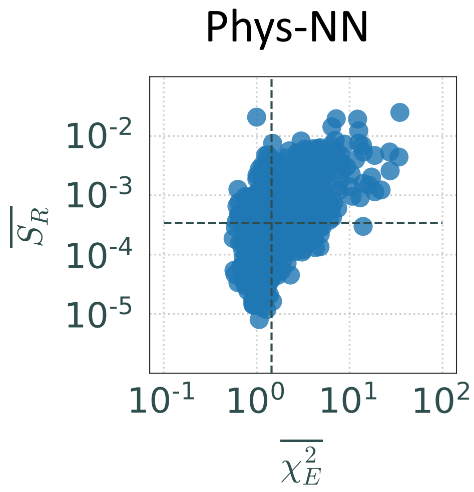


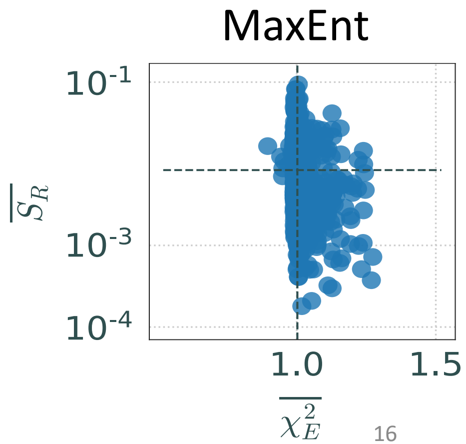
- The observed vs predicted plots for best and the worst responses for the one peak data.

Two Peak Dataset

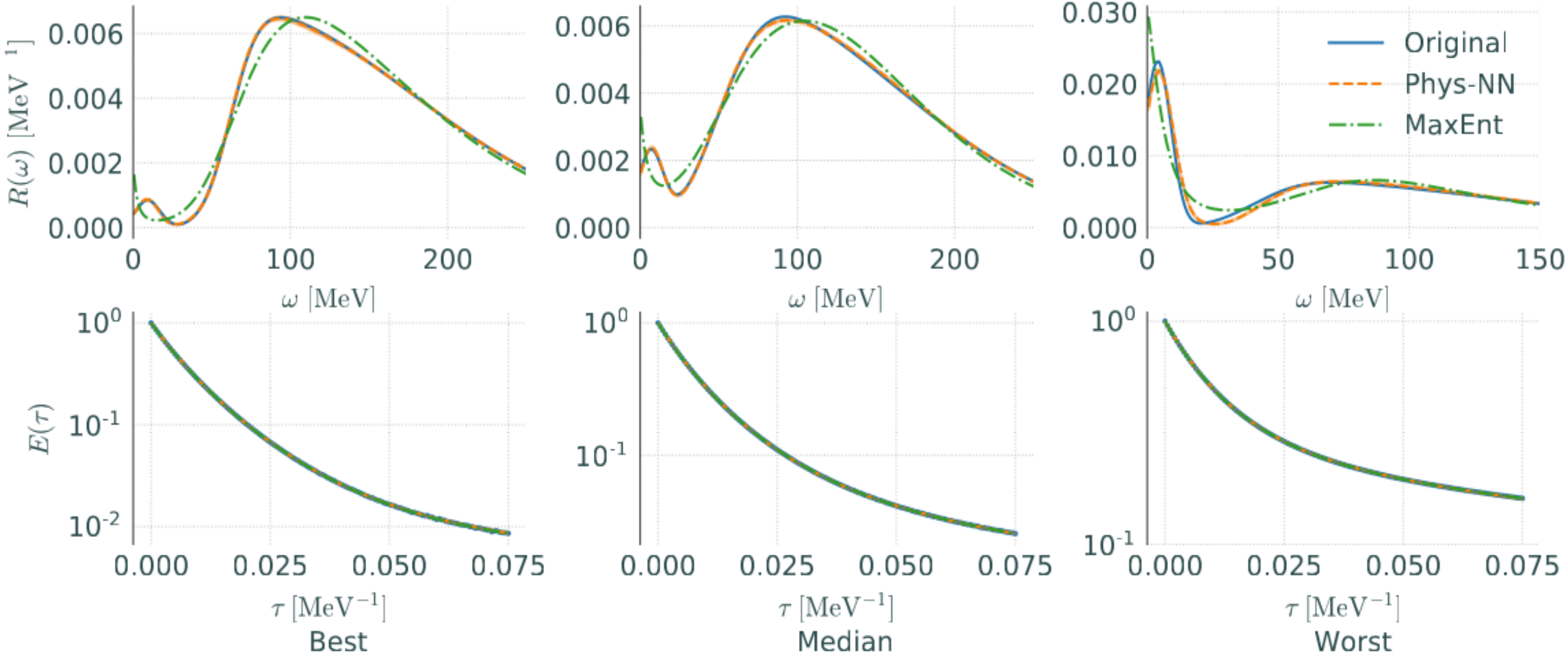
- Next, the two peak data
  - On the bottom, metrics on the test data.
  - On the right, correlation between S_R and χ^2.





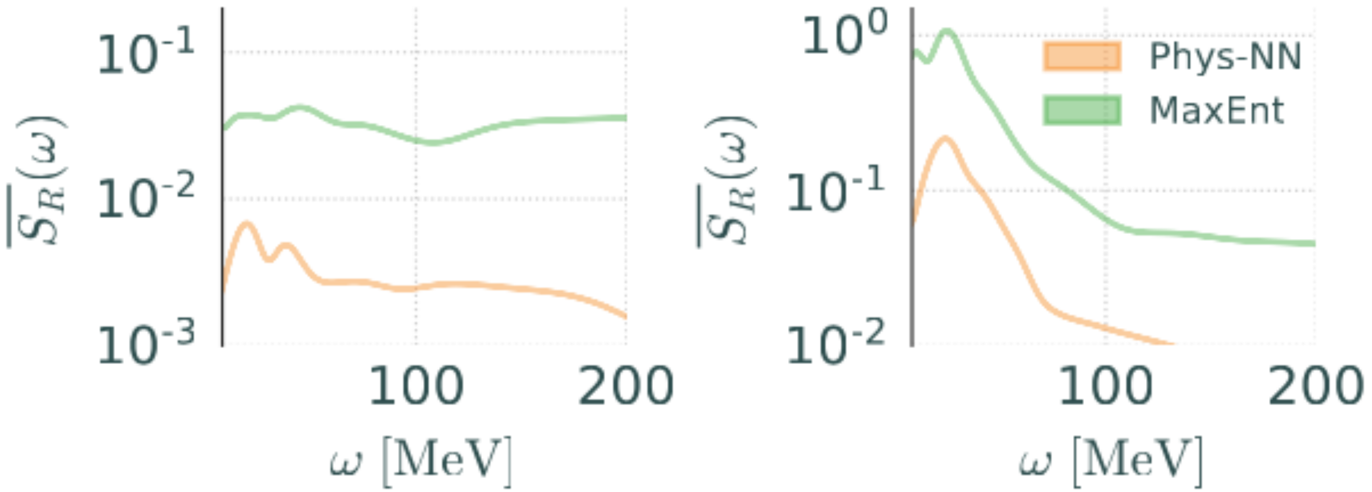


Two Peak Reconstruction

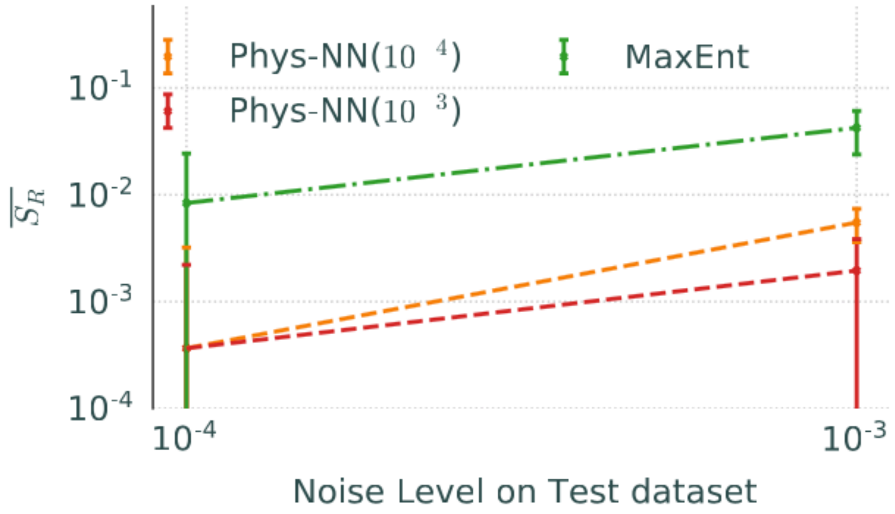


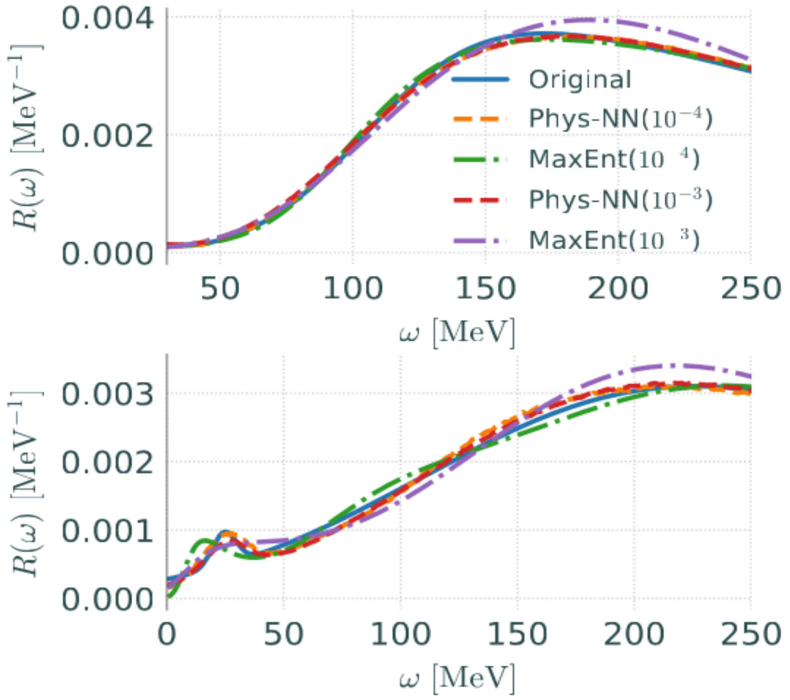
- The observed vs predicted plots for best and the worst responses for the two peak data.

Performance in Low omega



Resilience to Noise





Conclusion and Future Work

- Using Gaussian mixture models for capturing the structure of the response functions is very accurate.
- Outperforms Maximum Entropy and captures the structure in low omega region of the response function.
- Better than MaxEnt in terms of noise.
- In the future, we would provide the capability of propagating the errors in the Euclidean responses into the future.

U.S. DEPARTMENT OF ENERGY

DOE Early Career Research Program, ASCR