Methods and results for spectral functions - why do we care about inverse Laplace transform?

Chris Allton [Swansea University]

FASTSUM Collaboration

Firenze March 2022

UK Academic Strike

FASTSUM Collaboration's NRQCD Project

Gert Aarts, CRA, Tim Burns, Sam Offler, Benjamin Page, Thomas Spriggs
Swansea University

Benjamin Jäger
Southern Denmark University

Seyong Kim
Sejong University, Korea

Maria Paola Lombardo
INFN, Florence

Sinead Ryan
Trinity College, Dublin

Jon-Ivar Skullerud
National University of Ireland, Maynooth, Ireland

Overview

- Mathematical Limitations

- Towards chiral & continuum limits
  - $M_\pi = 392, 236, 140\ MeV$
  - $a_\tau = 33, 17\ am$

- Spectral Reconstruction from 7 Methods
  - *Max.Likelihood (x2)*
  - *Moments*
  - *Bayesian (x2)*
  - *Machine Learning*
  - *Backus-Gilbert*

NRQCD and the Laplace transform

- QCD expansion in p^2
- Long history in phenomenology from late 1970's
- Agrees surprisingly well with experiment

$$G(\tau) = \int \frac{d\omega}{2\pi} K(\omega, \tau)\, \rho(\omega)\, d\omega = \text{Euclidean Correlation Function}$$

$$K(\omega, \tau) = e^{-\omega\tau} = \text{Kernel}$$

Spectral Functions

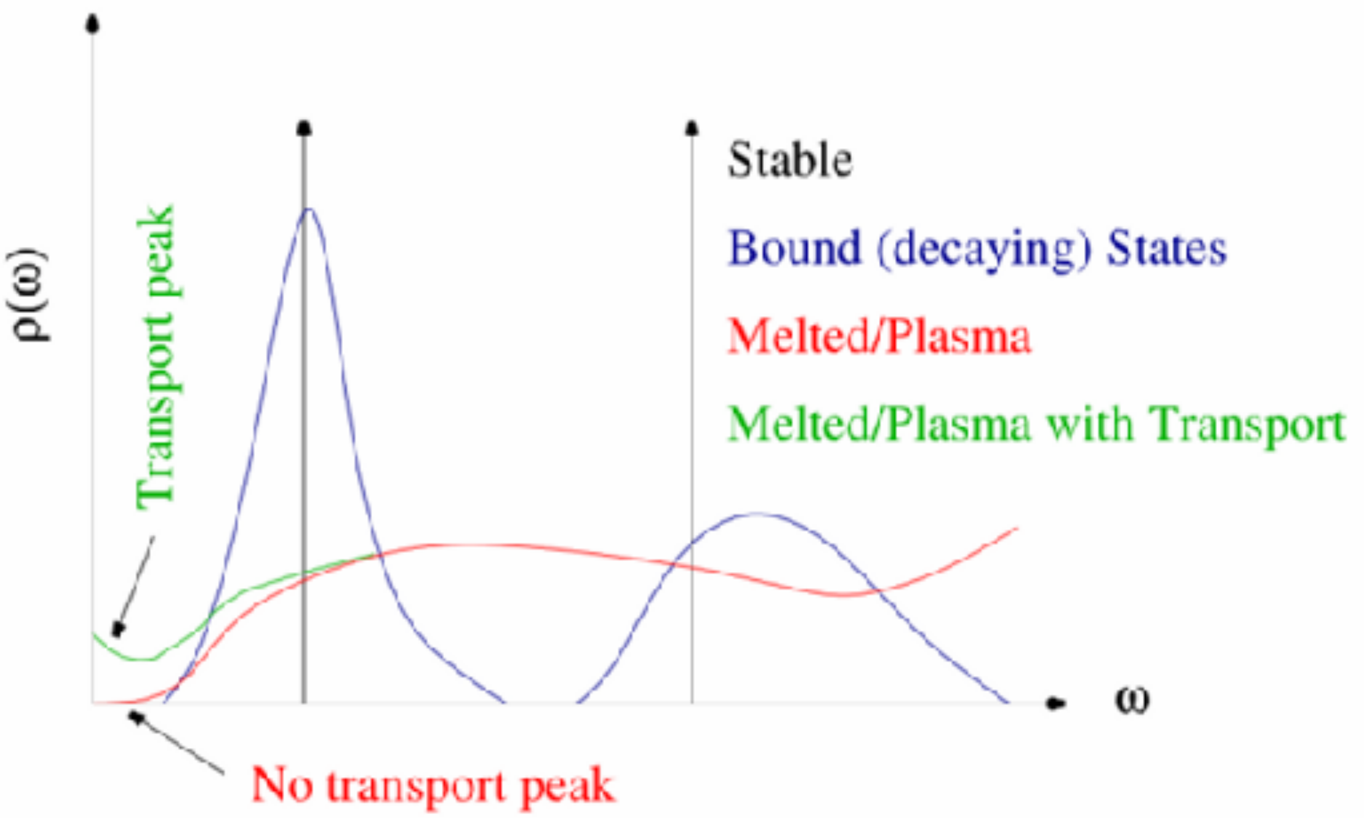
$$G(\tau) = \int \frac{d\omega}{2\pi} K(\omega,\tau)\, \rho(\omega)\, d\omega$$

$\uparrow$ Euclidean Correlator

$\nearrow$ (Lattice) Kernel

$\downarrow$ Spectral Function

$K(\omega,\tau) \sim e^{-\omega\tau}$ for NRQCD



Why bother?

Extracting $\rho(\omega)$ crucial for:

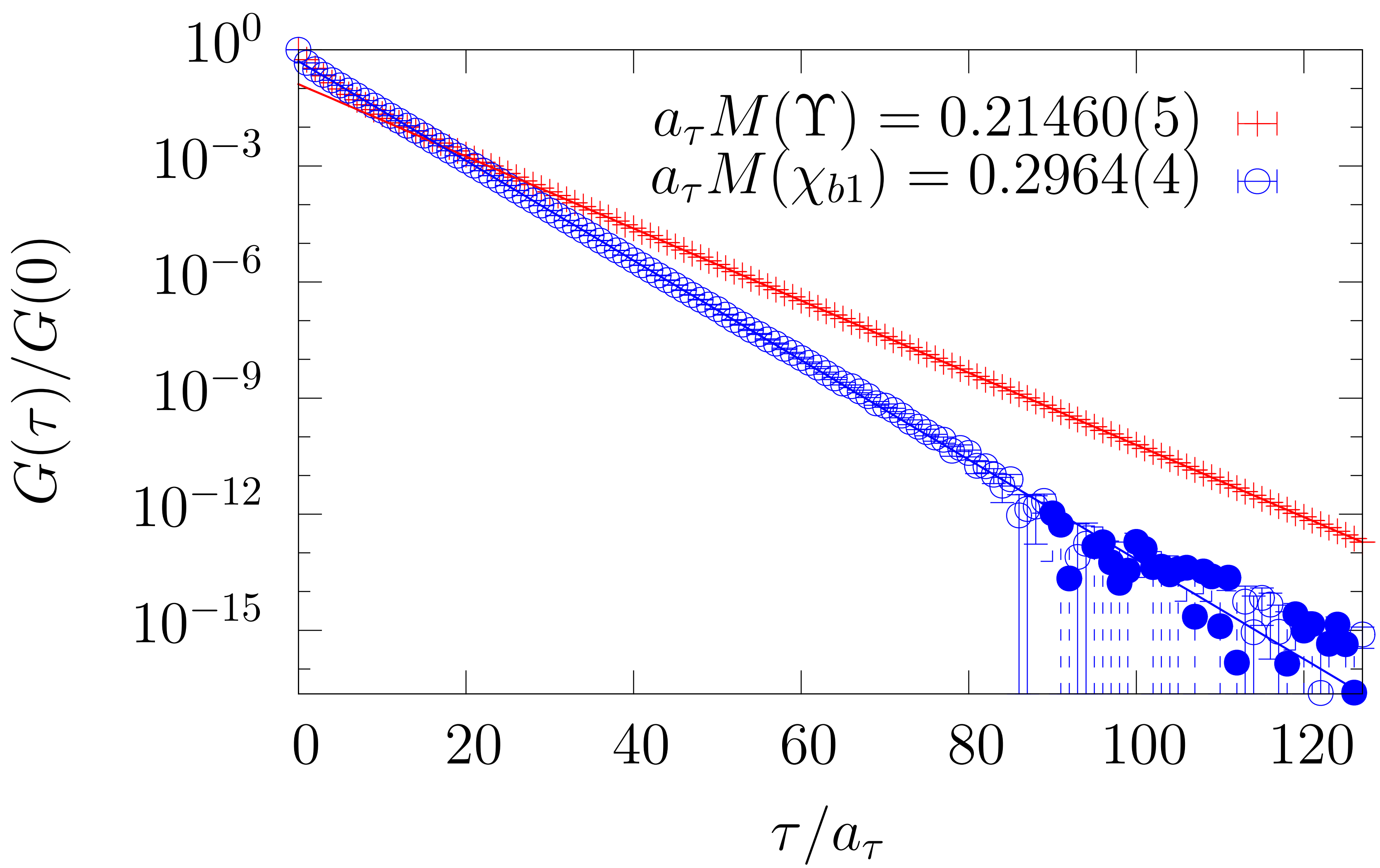
- masses
- widths
- “unbinding”

You may think it’s easy, but…

T=0 Correlators

FASTSUM Generation 2

$$G(\tau) \equiv \sum_x \langle 0|J(x,\tau)J^\dagger(0,0)|0\rangle \quad \xrightarrow{\tau\to\infty} \quad \frac{|\langle 0|J|\text{gnd}\rangle|^2}{2M}\, e^{-M\tau}$$



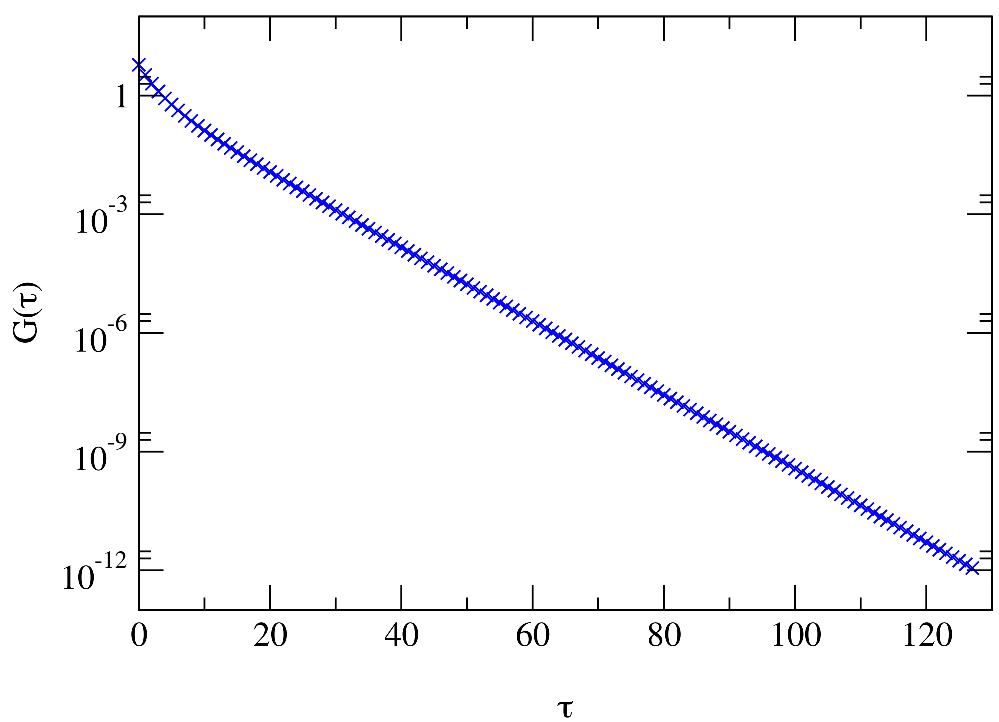
Lattice Determinations of Quarkonia Width

Extracting Spectral F'ns

Euclidean Lattice Correlator → $$G(\tau) = \int d\omega \, K(\omega, \tau) \, \rho(\omega)$$ ← Spectral F'n

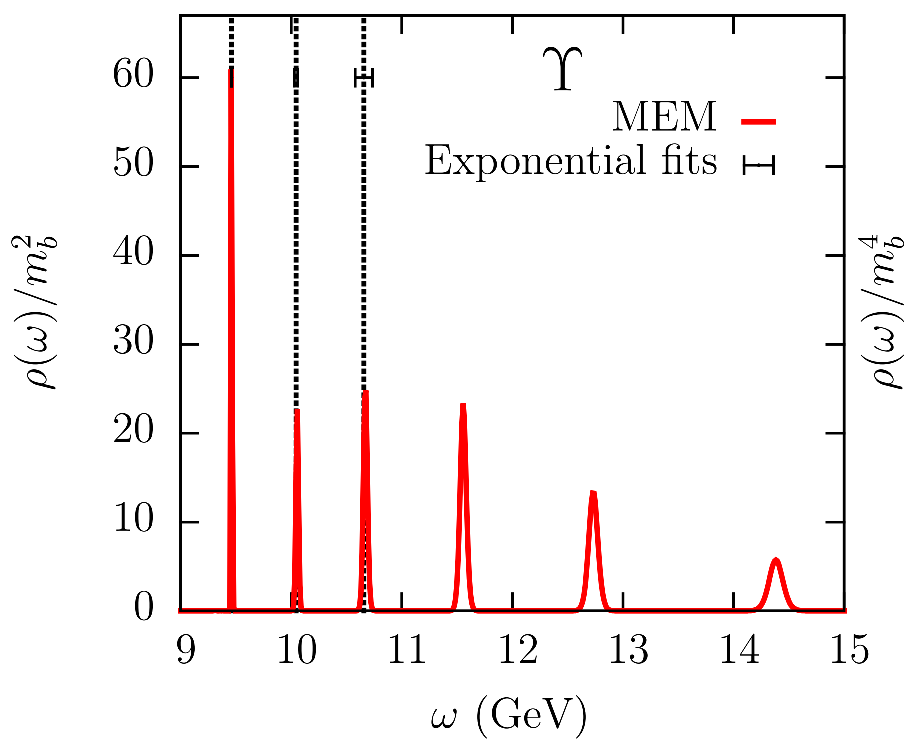
Input Data:

$G_{\pm}(\tau)$

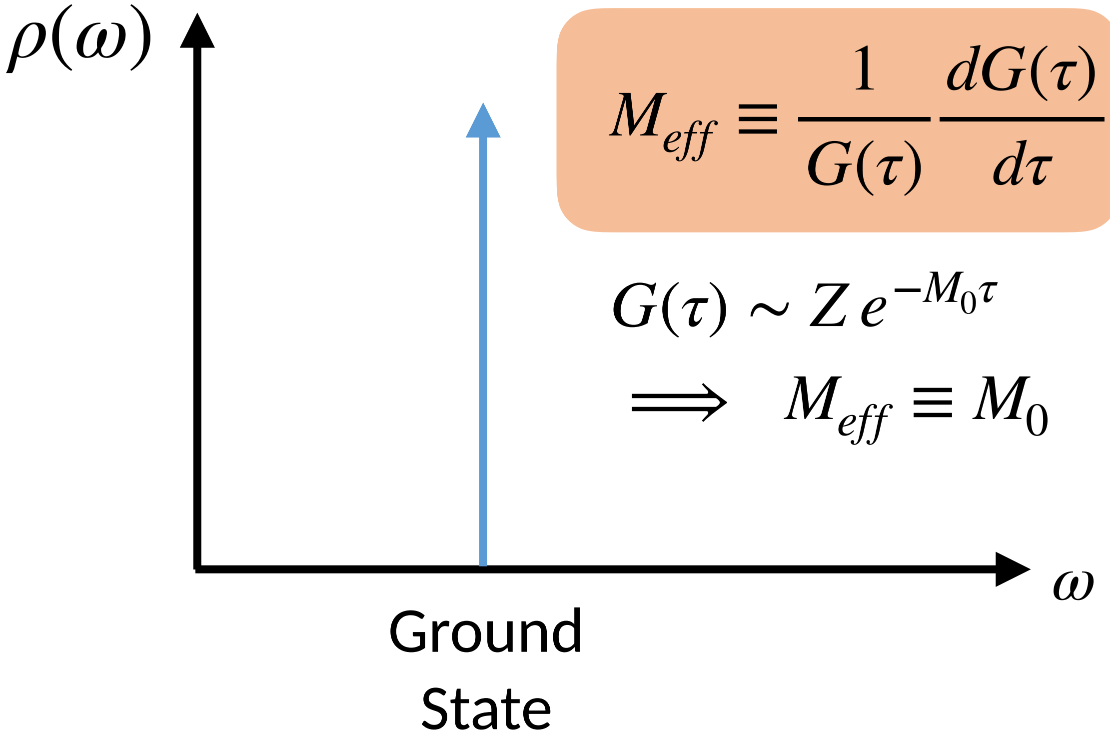


Output Data:

$\rho_{\pm}(\omega)$



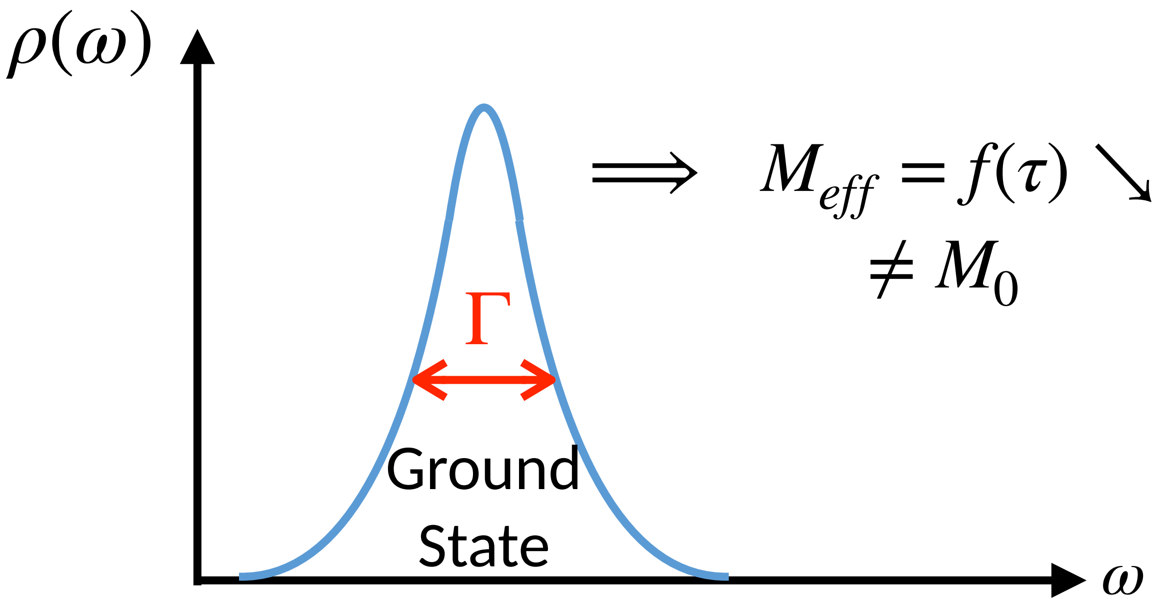
Thermal Case



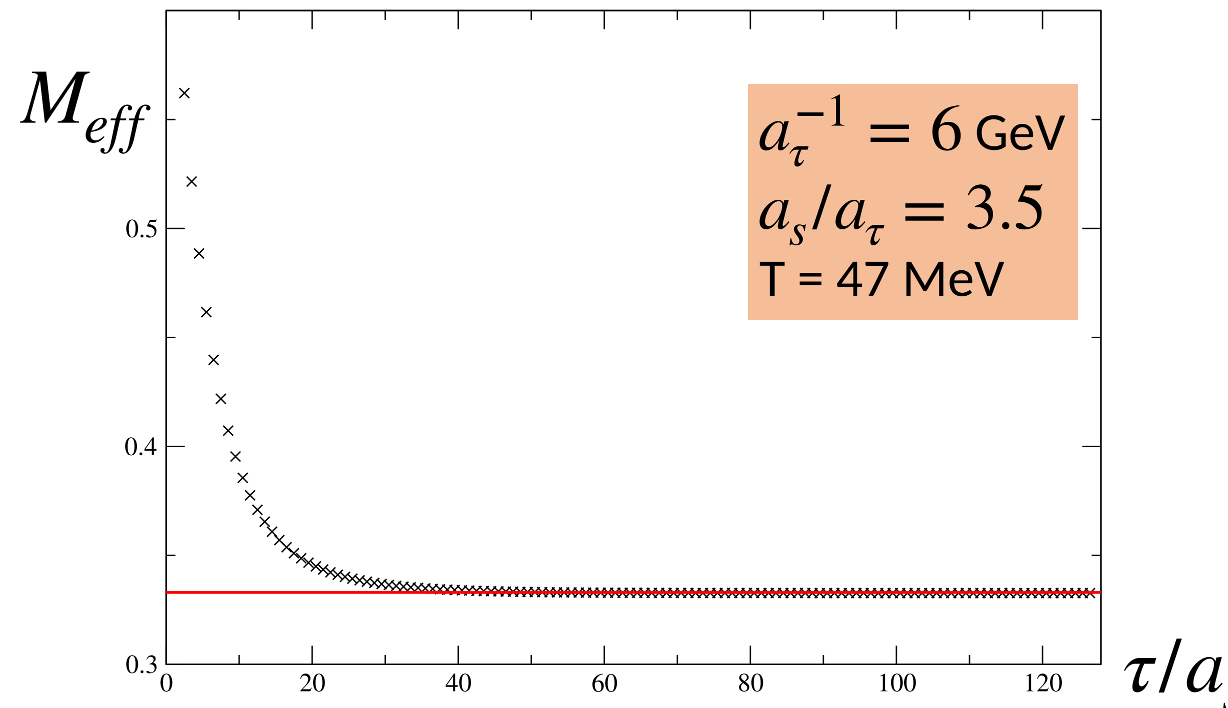
$$M_{eff} \equiv \frac{1}{G(\tau)} \frac{dG(\tau)}{d\tau}$$

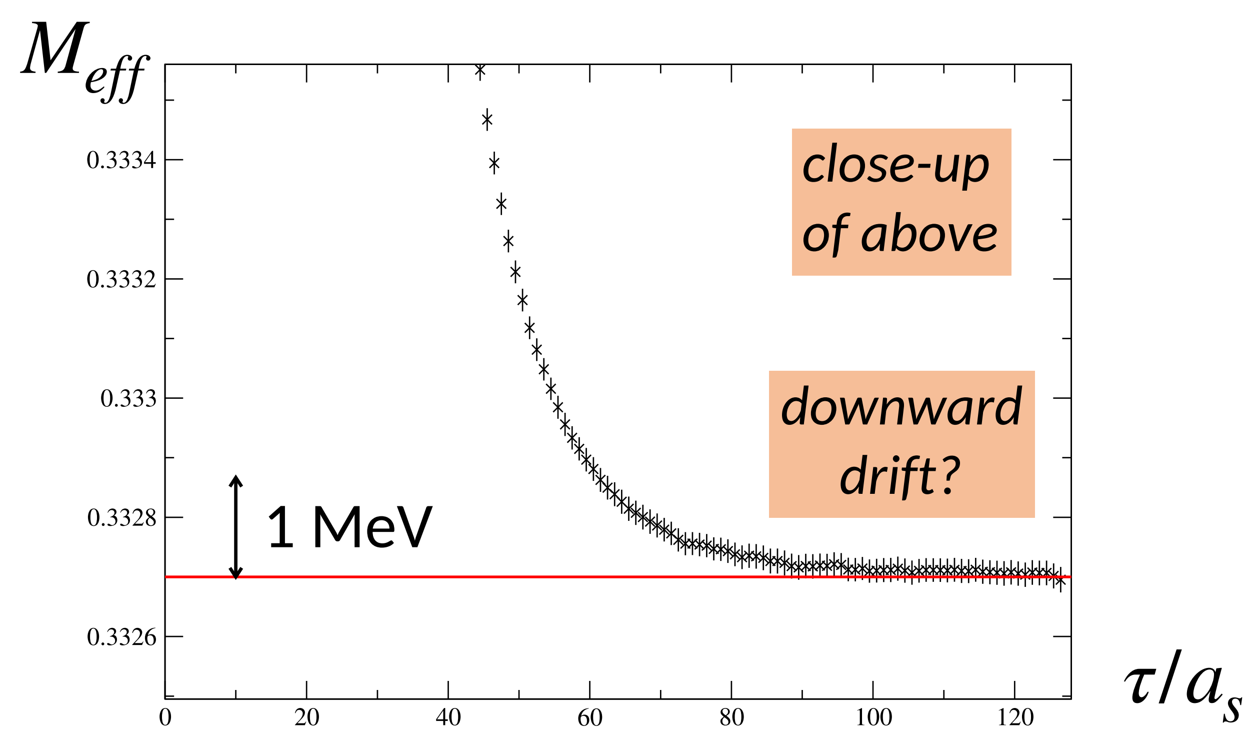
$$G(\tau) \sim Z\, e^{-M_0 \tau}$$

$$\Longrightarrow \quad M_{eff} \equiv M_0$$



$$\Longrightarrow \quad M_{eff} = f(\tau) \searrow$$
$$\neq M_0$$





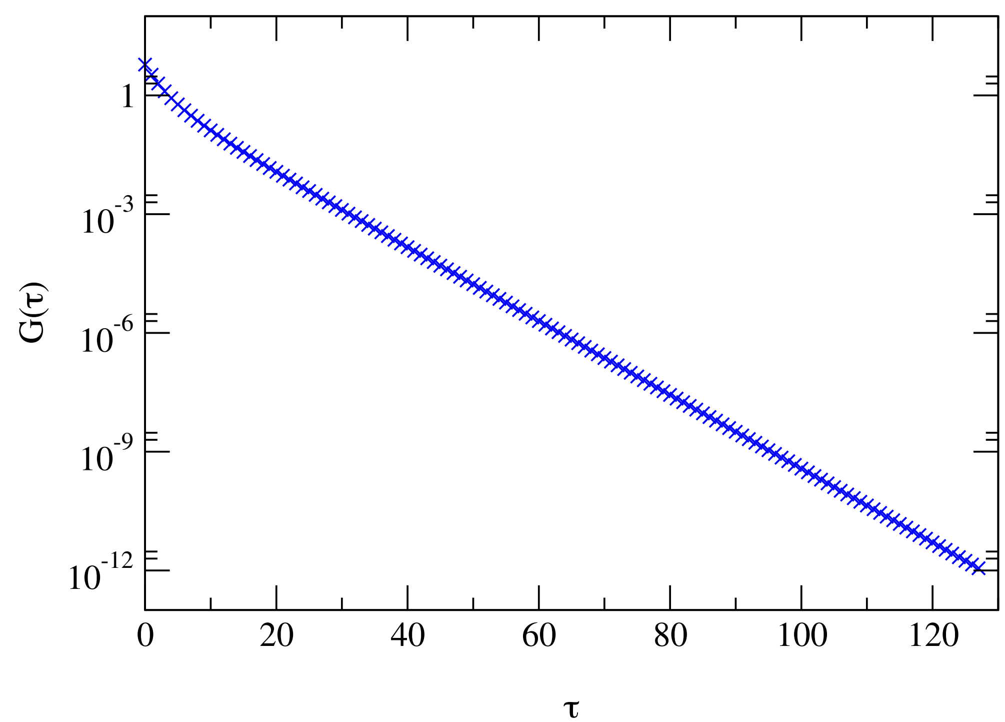
Lattice Determinations of Quarkonia Width

Extracting Spectral F'ns

Euclidean Lattice Correlator → $G(\tau) = \int d\omega\, K(\omega, \tau)\, \rho(\omega)$ ← Spectral F'n

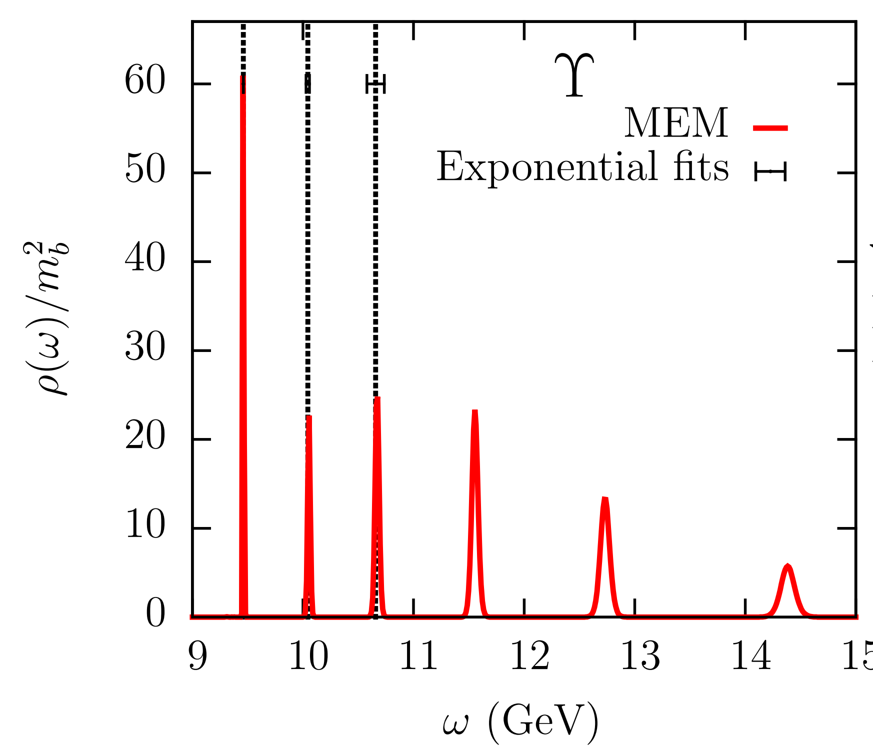
Input Data: *AND they are correlated!*

$G(\tau)$, $\tau = 1, \ldots, \mathcal{O}(10 - 100)$



Output Data:

$\rho_{\pm}(\omega)$, $\omega \sim 1, \ldots, \mathcal{O}(1000)$



ill-posed ! ***i.e. ∞ solutions with*** $\chi^2 = 0$

An allegory of life: You can't get more out than you put in.

Mathematical Limitations on Inverse Problem

McWhirter and Pike, J.Phys.A (1978) 1729
On the numerical inversion of the Laplace transform…

Cuniberti, De Micheli and Viano, Commun. Math. Phys. **216** (2001), 59-83
(courtesy of Mikko Laine)

Shuzhe Shia, Lingxiao Wang, Kai Zhou arXiv:2201.02564

Why is it difficult to extract $\rho(\omega)$ from $G(\tau)$?

- Is it because

$$G(\tau),\ \ \tau = 1,\ldots,\mathcal{O}(10-100) \qquad \rho_{\pm}(\omega),\ \ \omega \sim 1,\ldots,\mathcal{O}(1000)$$

- (and) or something else?

Mathematical Limitations on Inverse Problem

McWhirter and Pike, J.Phys.A (1978) 1729

On the numerical inversion of the Laplace transform…

Consider: $G(\tau) = \int_0^\infty d\omega\, K(\omega\tau)\, \rho(\omega)$ ***Note the product*** $\omega \times \tau$

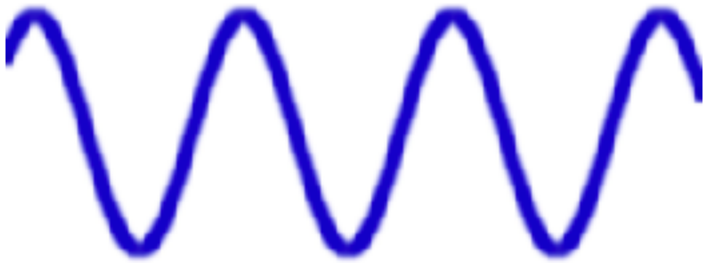
Note that Laplace transform has kernel of form: $K(\omega\tau) = e^{-\omega\tau}$

Note relativistic kernel is not in this form: $K(\omega, \tau) = \dfrac{\cosh(\omega(\tau - 1/2T))}{\sinh(\omega/2T)}$

Require that $K(\omega\tau)$ is bounded: $\int_0^\infty |K(x)|\, x^{-1/2} dx < \infty$

Mathematical Limitations on Inverse Problem

View $G(\tau) = \int_0^\infty d\omega \, K(\omega\tau) \, \rho(\omega)$ as a transformation



Fourier Transform

$G(\tau)$

time space

$\rho(\omega)$

frequency space

Mathematical Limitations on Inverse Problem

Consider a perturbation (error) in $\rho(\omega)$: $\delta\rho_{\Omega}(\omega) = \sin(\Omega\omega)$

This leads to a corresponding perturbation in $G(\tau)$:

$$\delta G(\tau) = \int K(\omega\tau)\ \sin(\Omega\omega)\, d\omega$$

Since $K(\omega\tau)$ is integrable: $\int K(\omega\tau)\ \sin(\Omega\omega)\, d\omega \longrightarrow 0$ as $\Omega \to \infty$

ie. $\delta G(\tau)$ can be made negligible!

Hence there are in ∞ number of $\rho(\omega)$ possible, i.e. $\rho(\omega)$ is ***not unique!***

Mathematical Limitations on Inverse Problem

McWhirter and Pike, J.Phys.A (1978) 1729

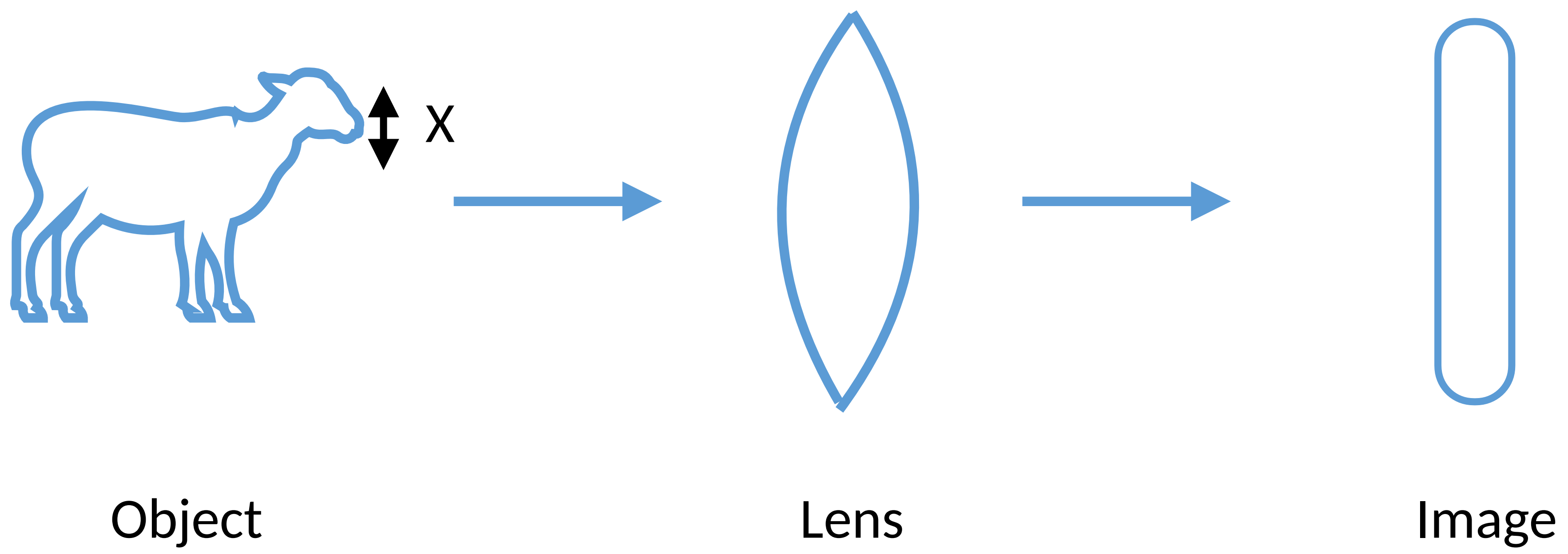
On the numerical inversion of the Laplace transform…

"… This difficult numerical problem, which is frequently encountered by physicists and engineers, is still the subject of much attention"

"… need to consider information content in order to avoid obtaining meaningless results."

Mathematical Limitations on Inverse Problem

Consider an optics example



$$I(x') = \frac{1}{2\pi} \int_{-\Omega}^{\Omega} d\omega \, e^{-i\omega x'} \int_{-X/2}^{X/2} e^{i\omega x} \, O(x) \, dx$$

where Ω is he highest spatial frequency mode transmitted by lens

Mathematical Limitations on Inverse Problem

$$I(x') = \frac{1}{2\pi}\int_{-\Omega}^{\Omega} d\omega\, e^{-i\omega x'} \int_{-X/2}^{X/2} e^{i\omega x}\, O(x)\, dx$$

$$= \int_{-X/2}^{X/2} \frac{\sin[\Omega(x'-x)]}{\pi(x'-x)}\, O(x) dx$$

Lens's highest spatial frequency mode has associate resolution limit of π/Ω

$\rightarrow$ Concept of "Information Content" = No. of π/Ω in Object

Hence Object of size X gives an Image with

$S = \dfrac{X}{\pi/\Omega}$ independent d.o.f. = ***Shannon Number*** of Information Theory

Making Mathematical Progress on Inverse Problem

McWhirter and Pike, J.Phys.A (1978) 1729

Key concept is to cast inverse problem as eigenvalue problem

$$\int_{-X/2}^{X/2} \frac{\sin[\Omega(x' - x)]}{\pi(x - x')} \phi_n(x')\, dx' = \lambda_n \phi_n(x)$$

ϕ_n form complete orthogonal set.

$$I(x') = \sum_{n=0}^{\infty} b_n \phi_n(x') \quad \text{and} \quad O(x) = \sum_{n=0}^{\infty} a_n \phi_n(x')$$

Then $O(x) = \sum_{n=0}^{\infty} \frac{b_n}{\lambda_n} \phi_n(x)$

Looks easy… BUT….

Making Mathematical Progress on Inverse Problem

McWhirter and Pike, J.Phys.A (1978) 1729

On the numerical inversion of the Laplace transform…

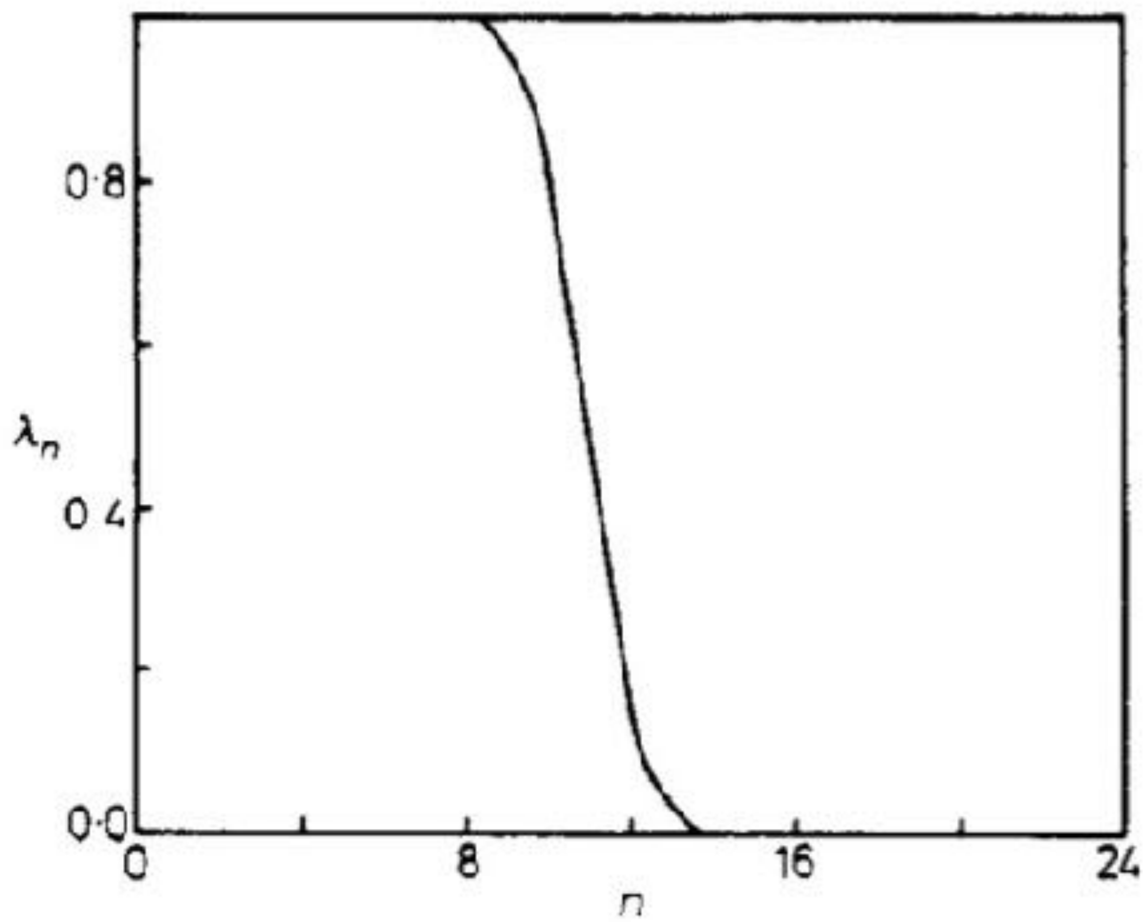


Figure 2. Eigenvalues λ_n of equation (2.7) as a function of n.

$$\frac{X\Omega}{\pi} = 11.5$$

For other kernels/integral transforms, λ_n has similar features (but doesn't fall quite as fast)

Making Mathematical Progress: Laplace Case

Laplace Case: $\lambda_n \sim \sqrt{\dfrac{\pi}{\cosh(\pi n)}}$

Recall $O(x) = \sum_{n=0}^{\infty} \dfrac{b_n}{\lambda_n} \phi_n(x)$

In general, we can write $\rho(\omega)$ as

$$\rho(\omega) = \sum_{n=0}^{N} a_n \phi_n(\omega) + \sum_{n=N+1}^{\infty} \theta_n \phi_n(\omega)$$

"Knowable"
$\lambda \gg 0$

"Unknowable"
$\lambda \sim 0$

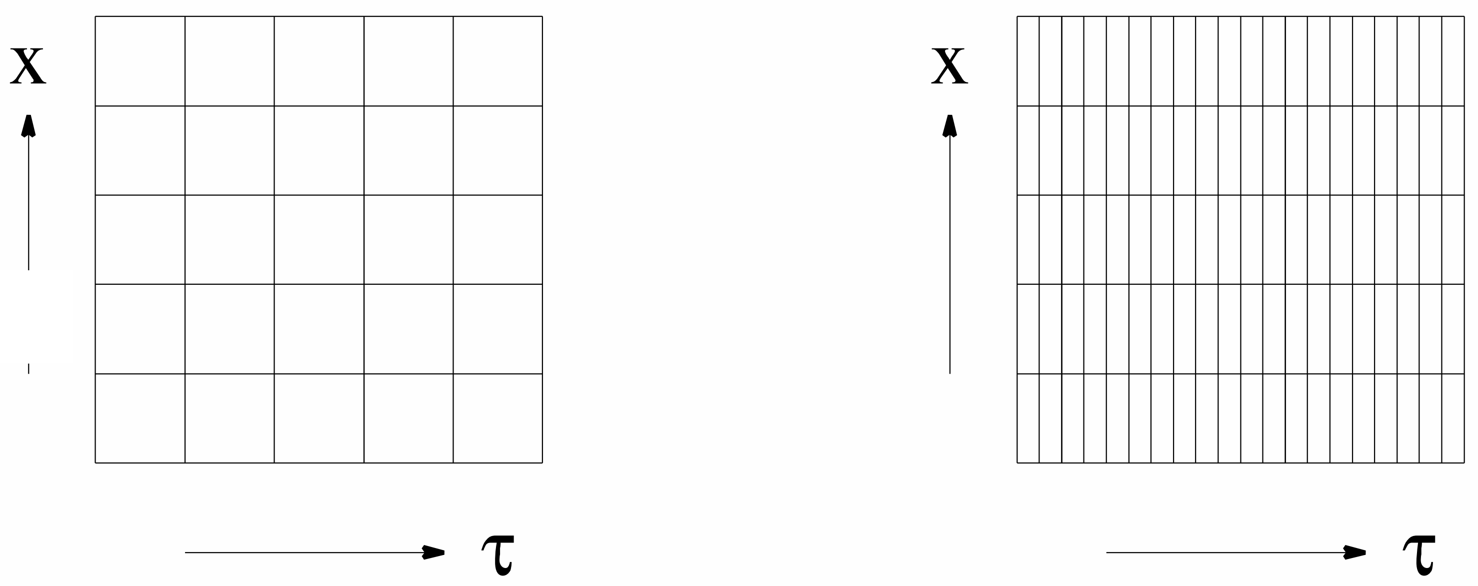
FASTSUM setup

Anisotropic Lattice:

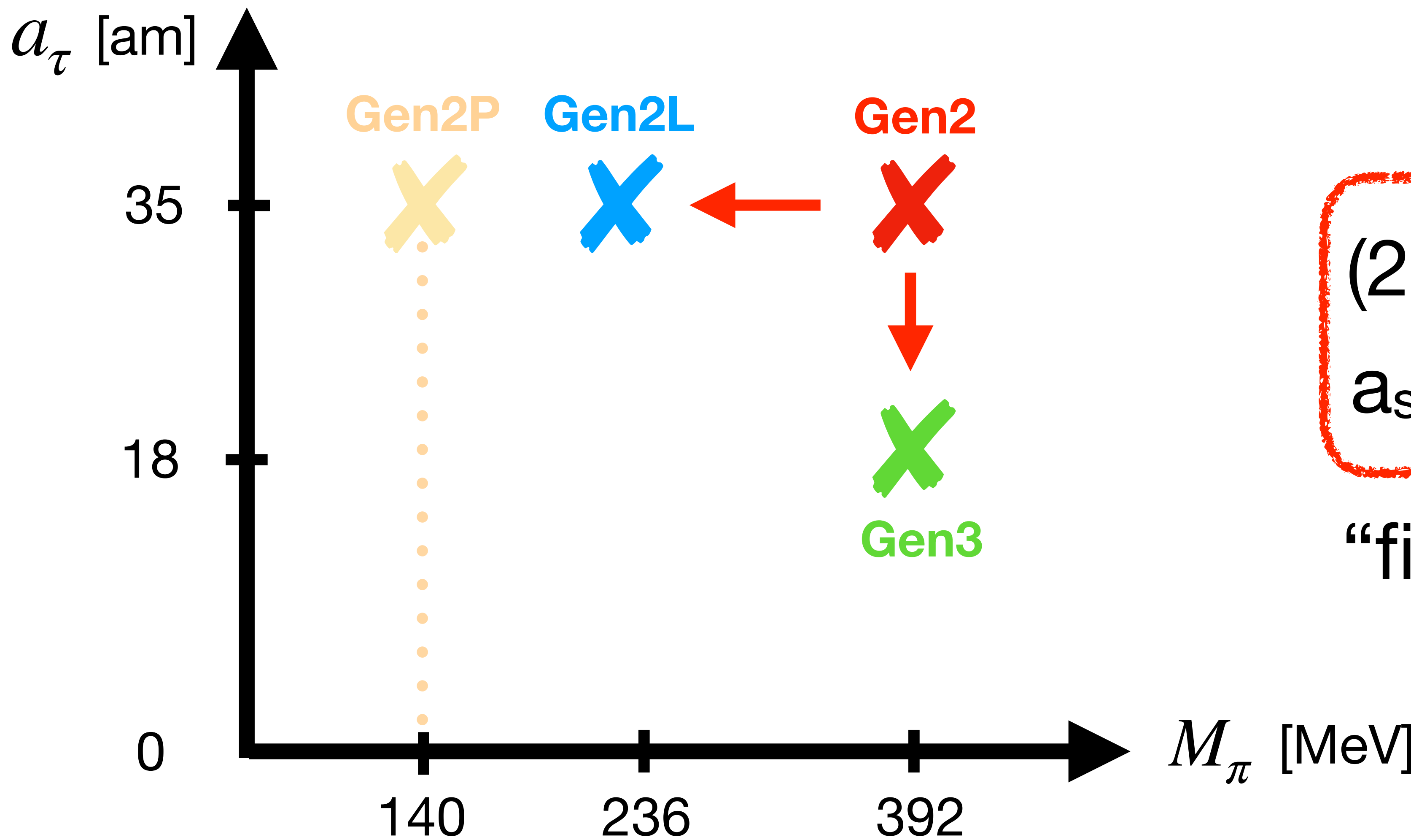
$$a_\tau < a_s$$

allowing for better resolution, particularly at finite temperatures, since

$$T = \frac{1}{L_\tau} = \frac{1}{N_\tau a_\tau}$$



Lattice Parameters



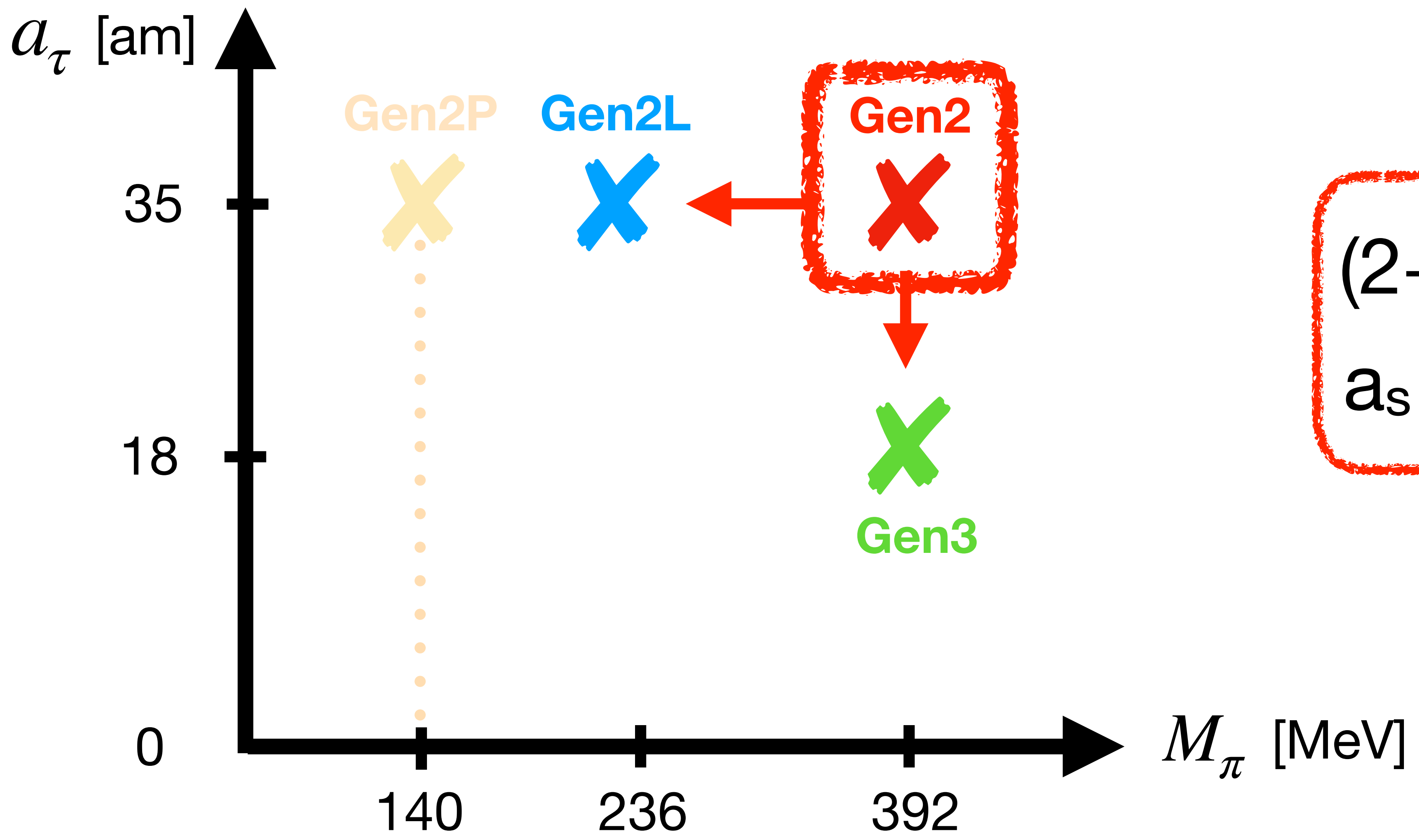
(2+1) flavour

a_s ~ 0.12 fm

"fixed scale"

Aarts *et al*, JHEP 07 (2014) 097

Lattice Parameters



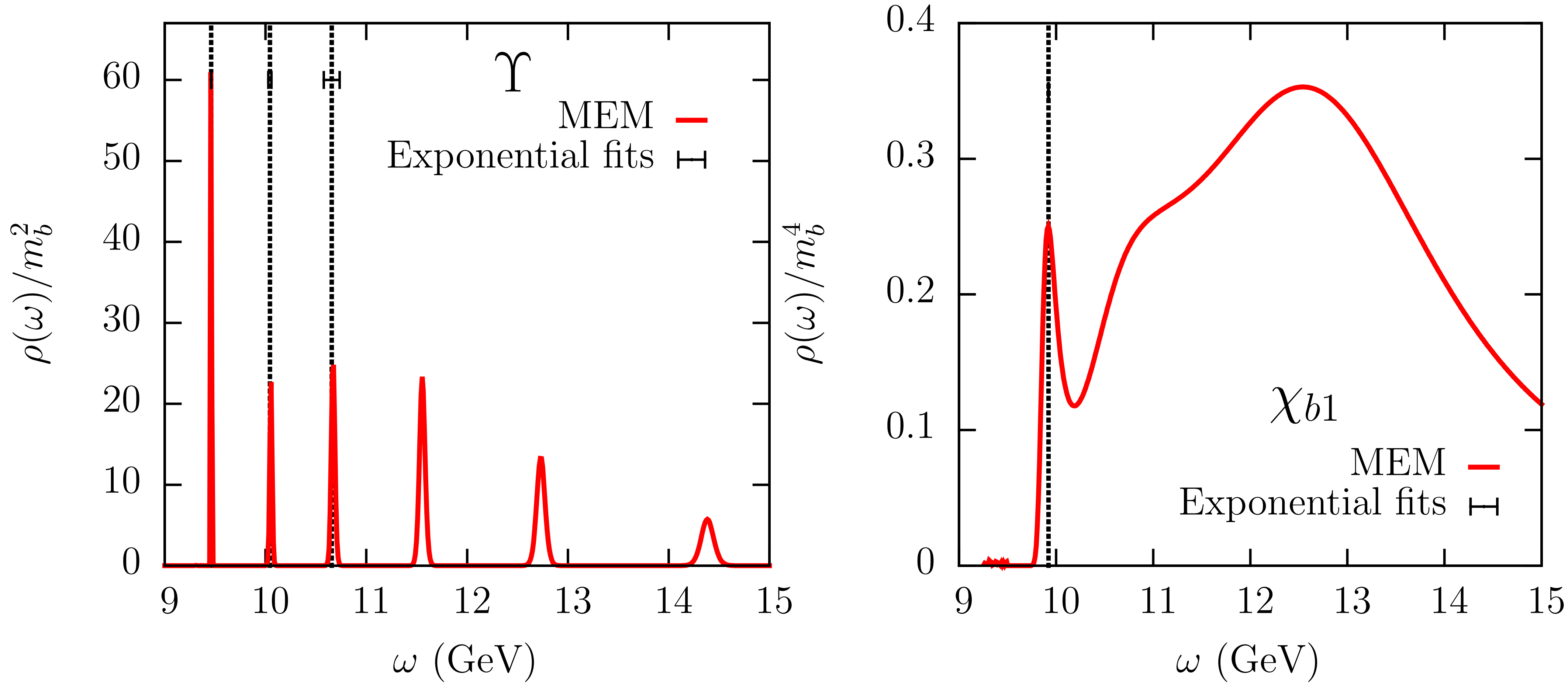
(2+1) flavour

$a_s \sim 0.12$ fm

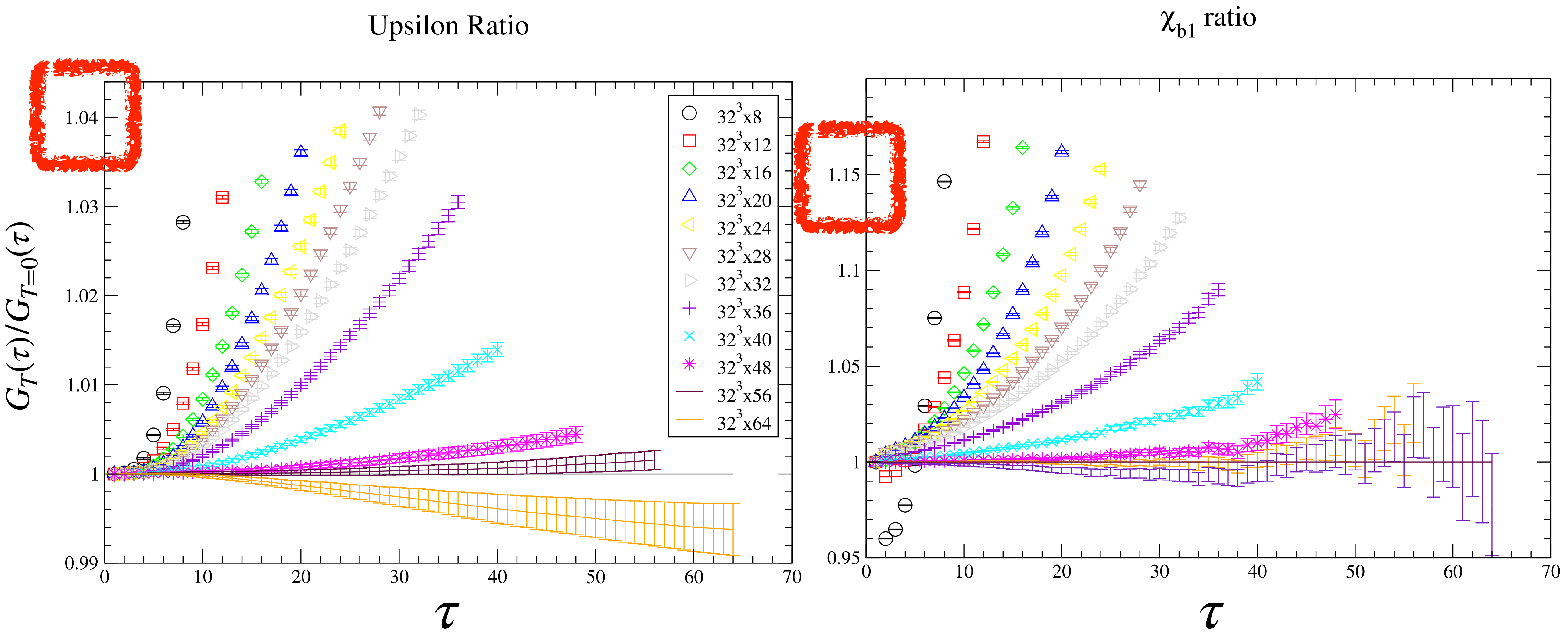
T=0 spectral functions

Generation 2

$$G(\tau) = \int_{\omega_{\min}}^{\omega_{\max}} \frac{\mathrm{d}\omega}{2\pi} K(\tau,\omega)\,\rho(\omega), \qquad K(\tau,\omega) = e^{-\omega\tau}.$$

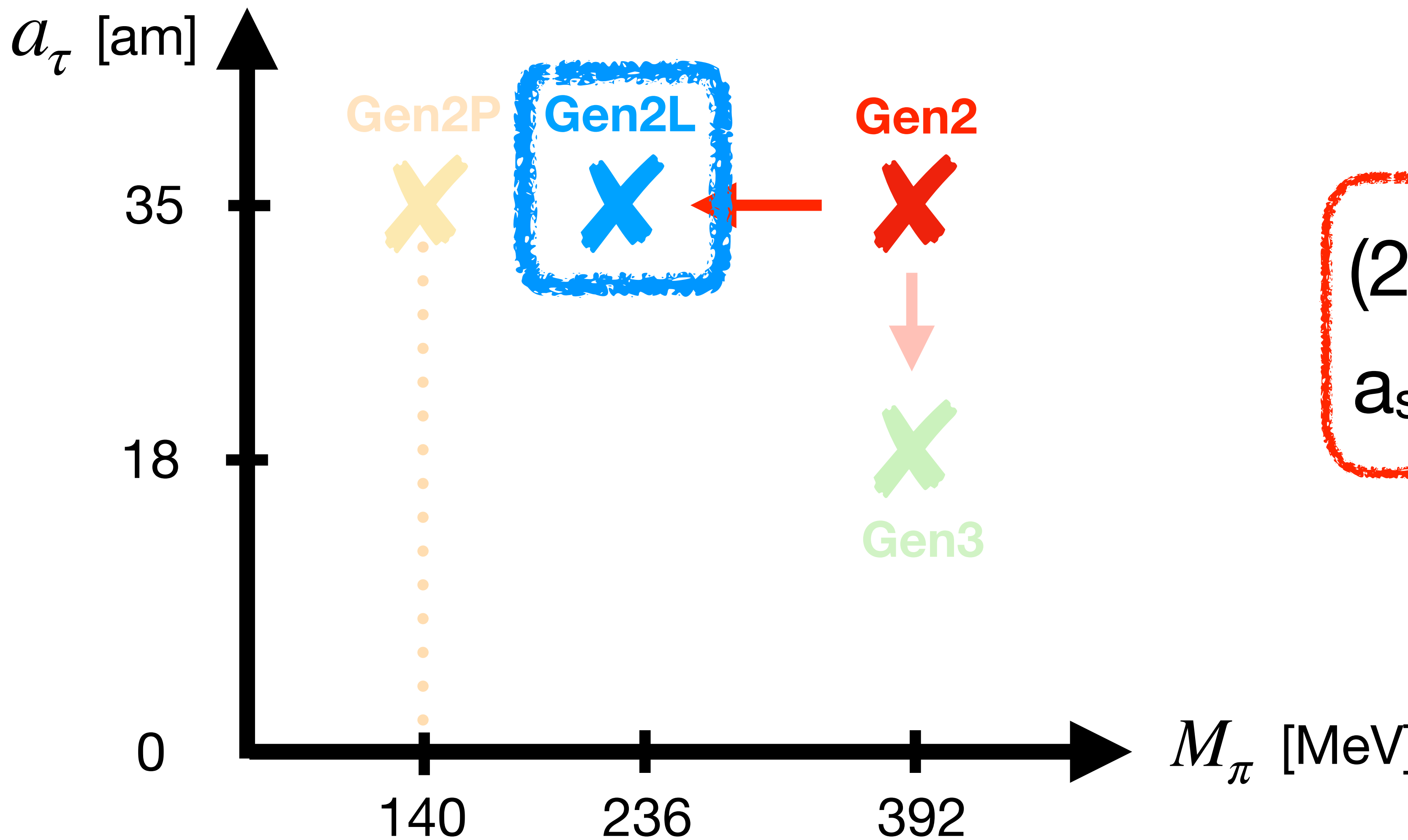


Correlation Ratios: Upsilon & χ_{b1}



Thermal effects stronger in χ_{b1}

Lattice Parameters

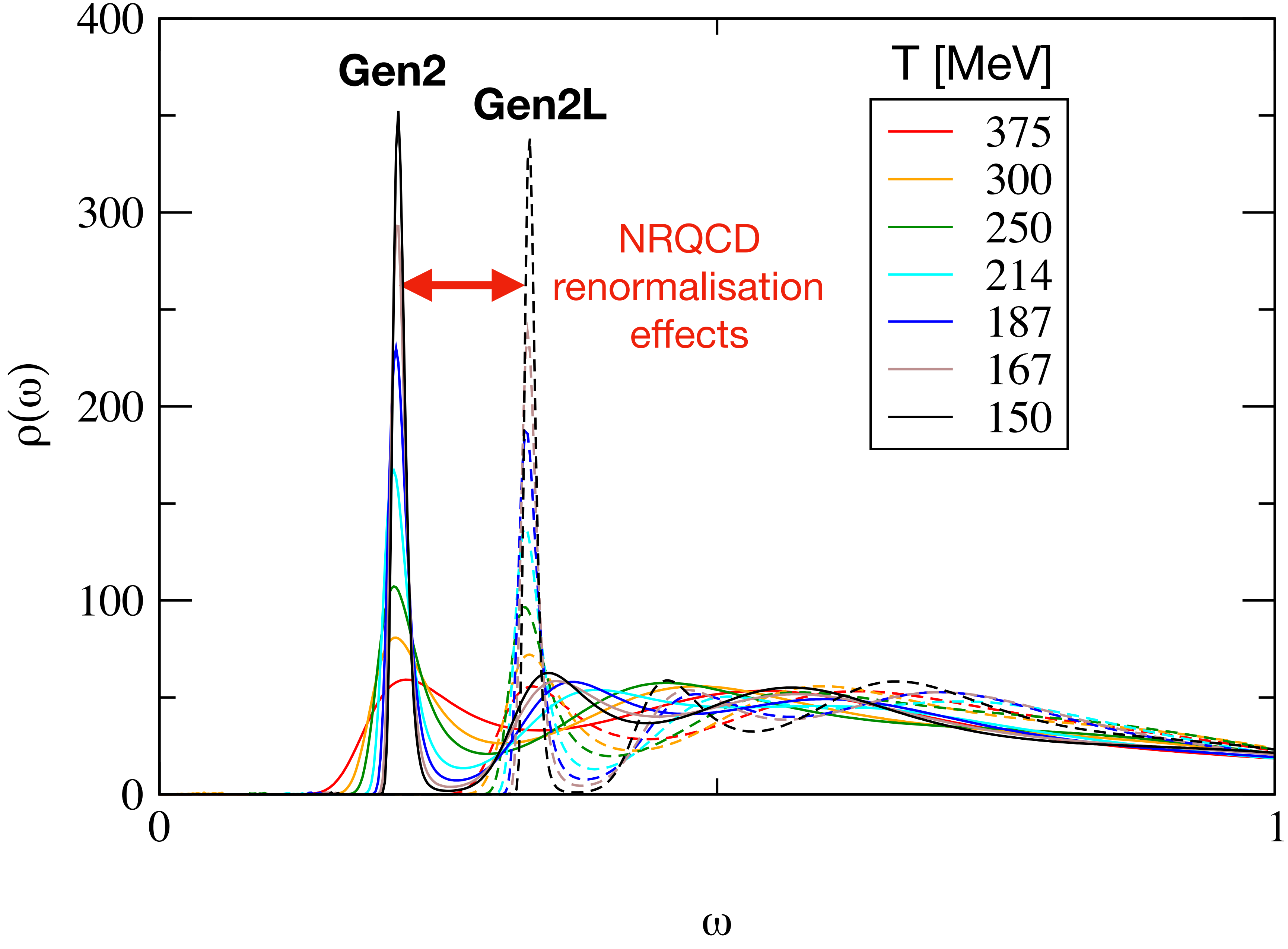


(2+1) flavour

$a_s \sim 0.12$ fm

Upsilon: Generation2 vs Generation2L

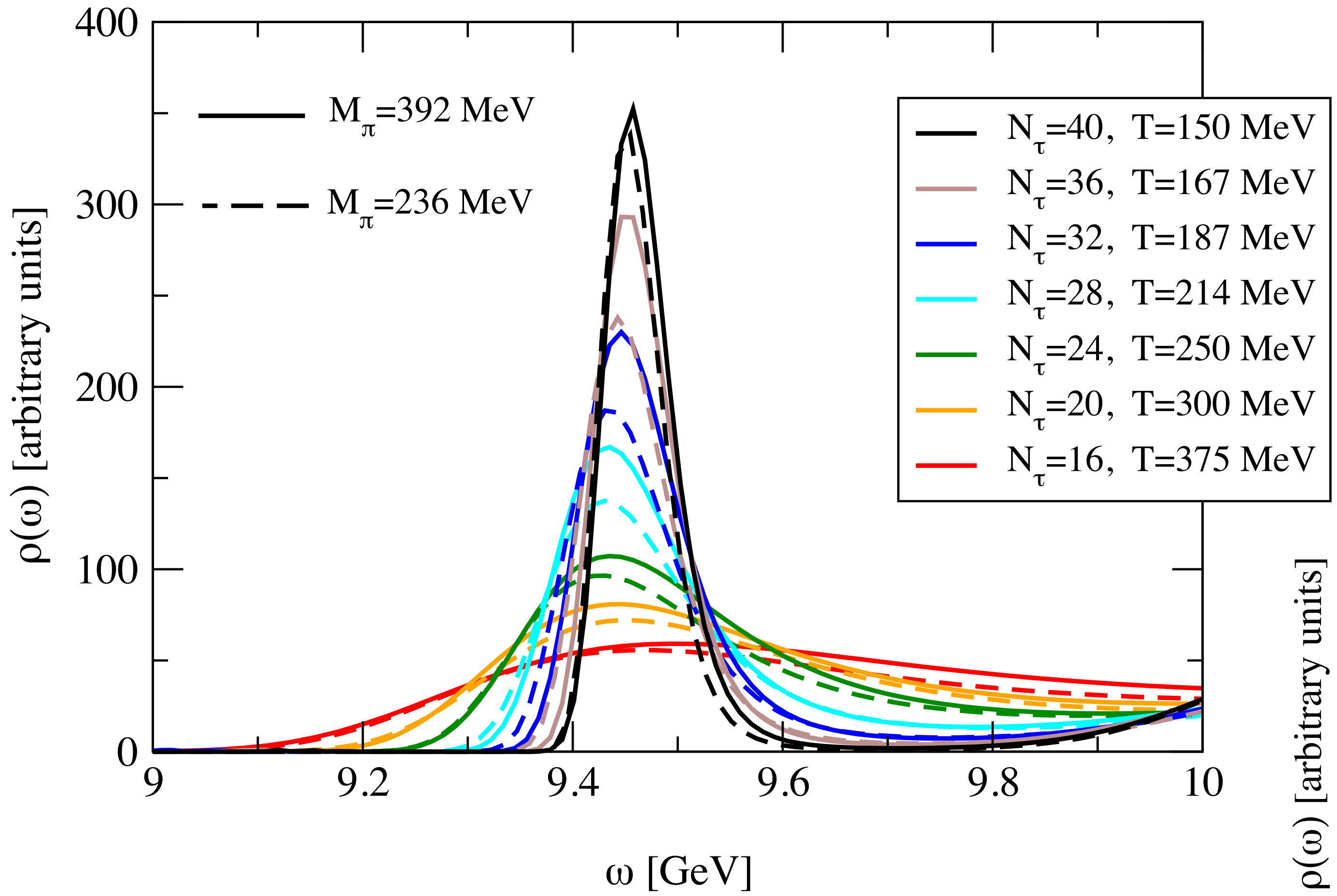
Going lighter

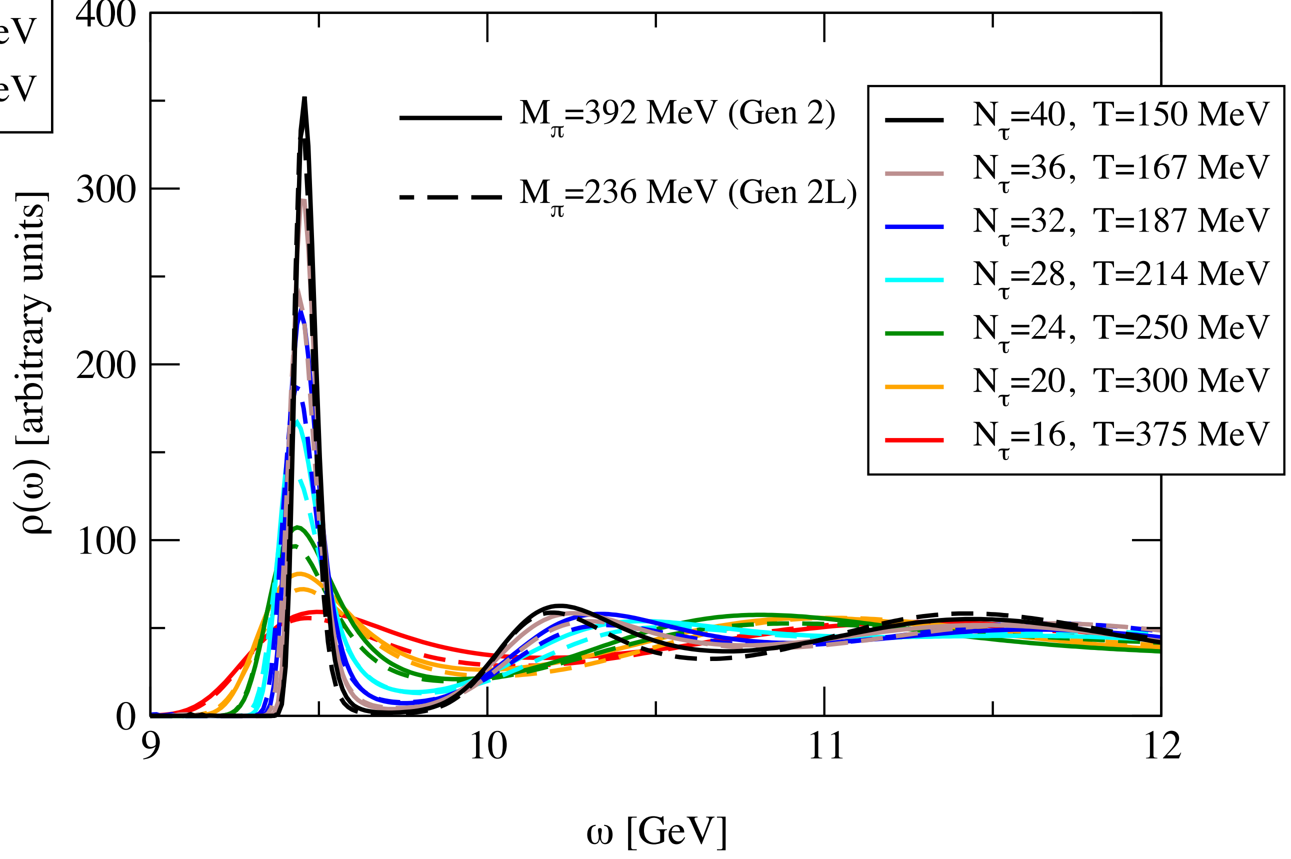


Preliminary

Upsilon: Generation2 vs Generation2L

Going lighter

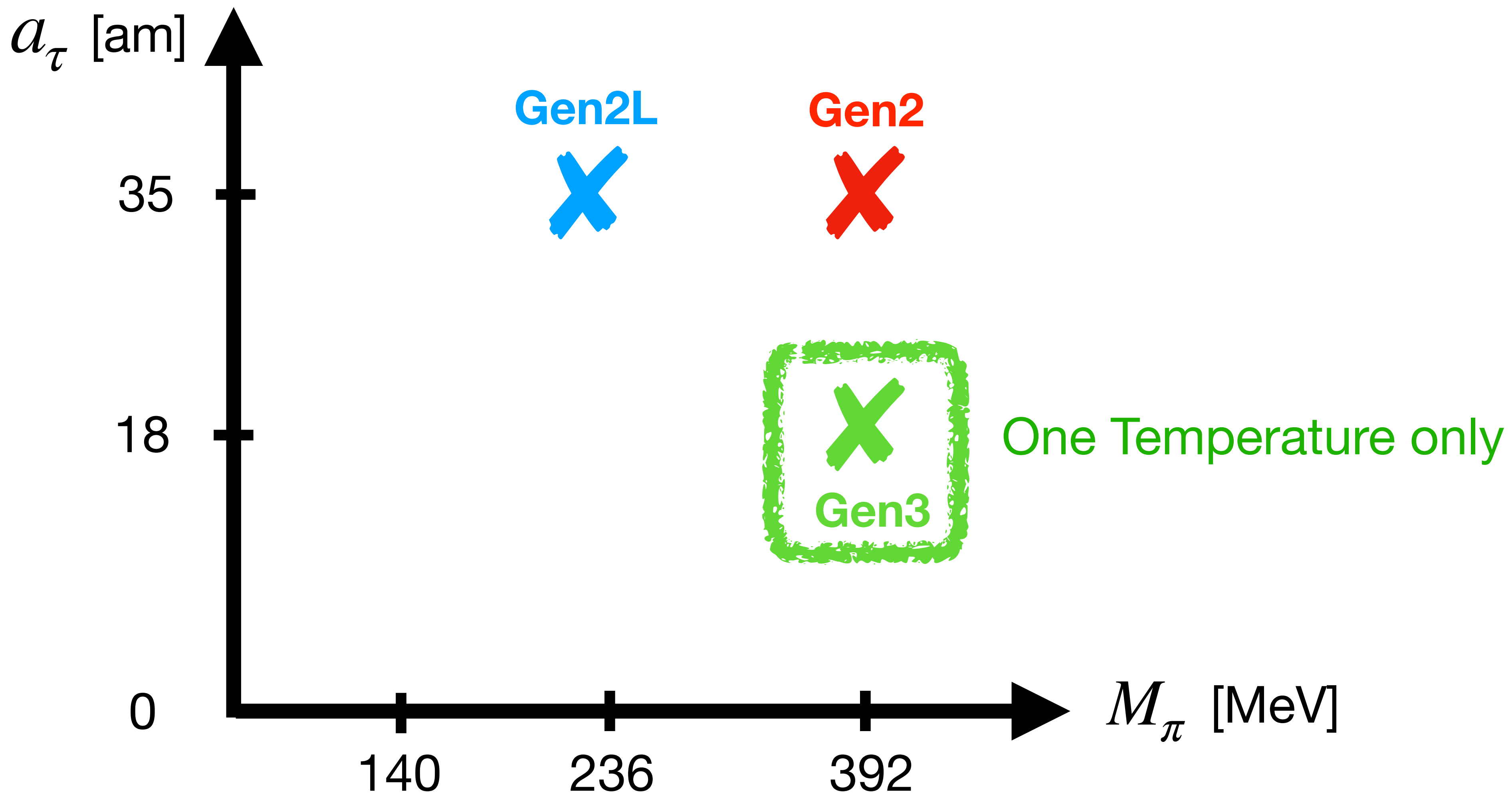




Preliminary

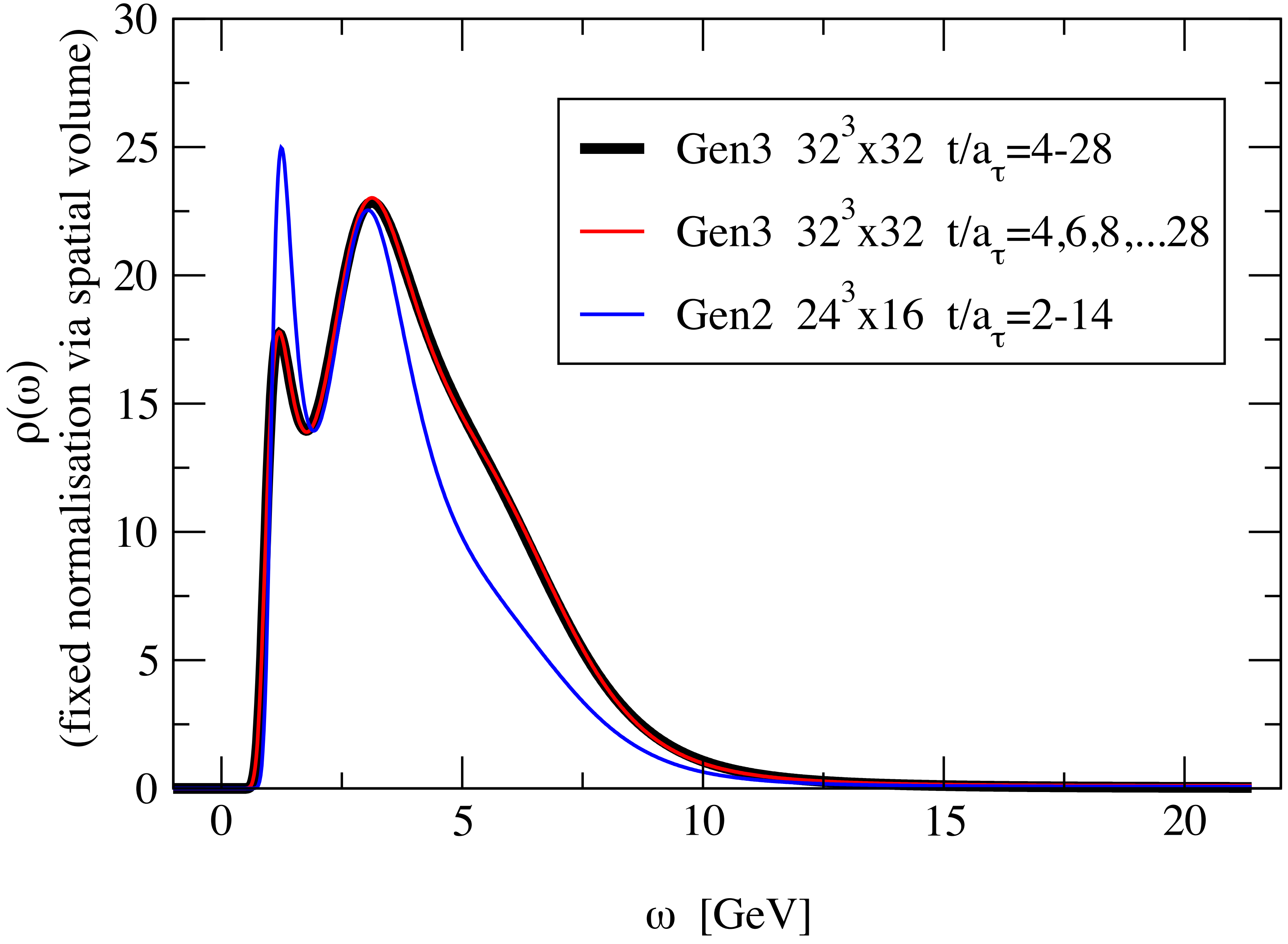
Generation3 Results

Going finer



Upsilon: Generation2 vs Generation3

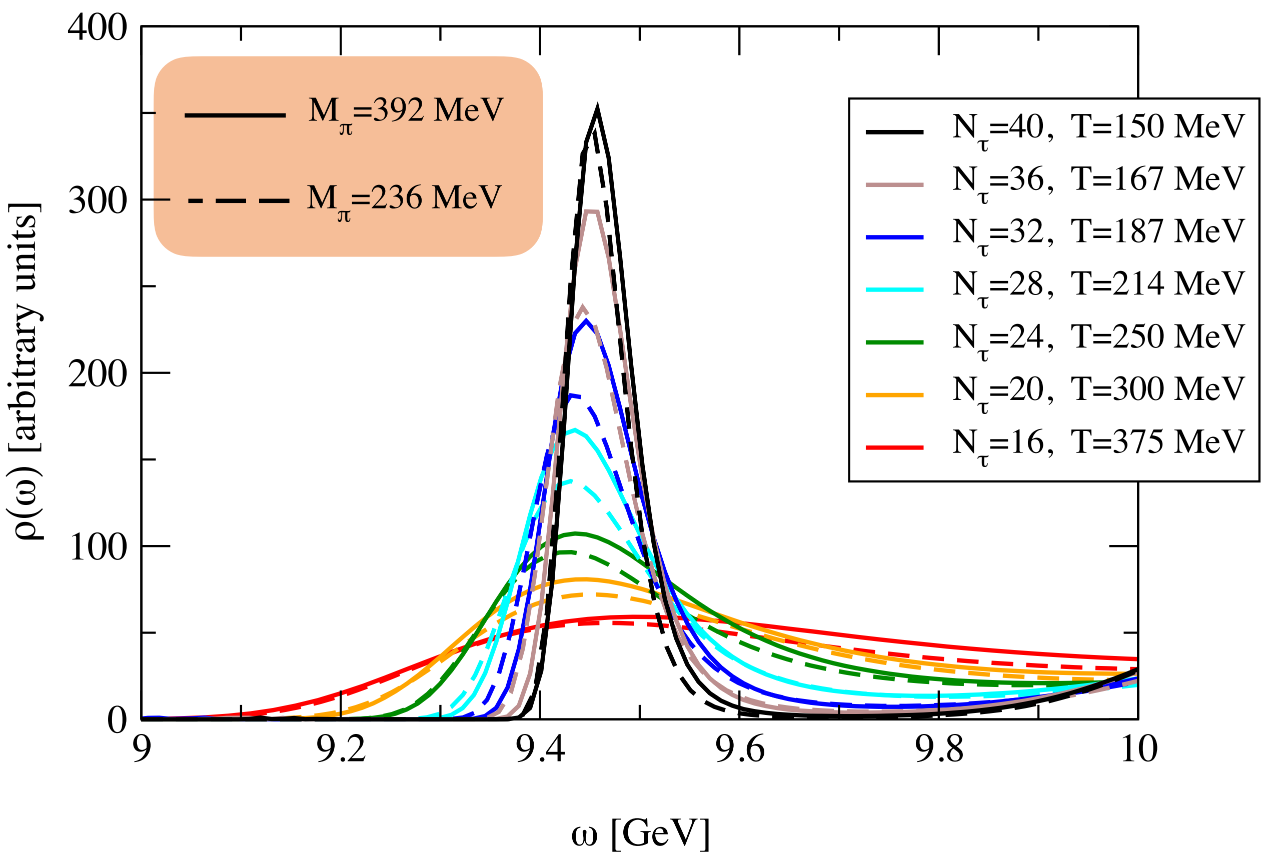
Going finer



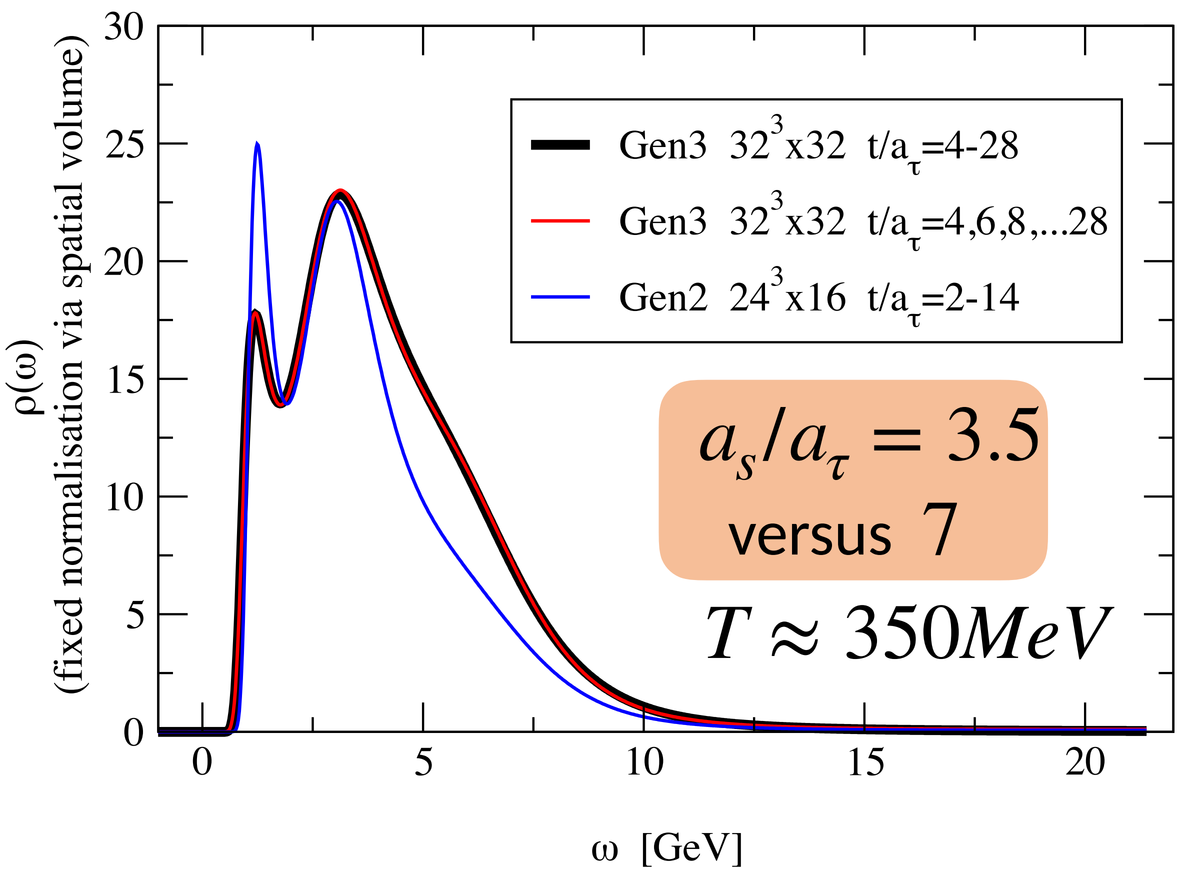
$T \approx 350MeV$

Lattice systematics - are "small"

Going lighter $m_q \searrow$



Going finer $a_\tau \searrow$



w/o NRQCD additive constant

Preliminary

Study of Numerical Methods

1. Exponential (Conventional δ f'ns)
2. Gaussian Ground State (+ δ f'n excited)

} Maximum Likelihood (Minimise χ^2)

3. Moments of Correlation F'ns

Direct Method - "no" fit

4. BR Method
5. Maximum Entropy Method

} Bayesian Approaches

6. Kernel Ridge Regression

Machine Learning

7. Backus Gilbert

from Geophysics

Moments

$$G(\tau) = \int e^{-\omega\tau}\rho(\omega)\, d\omega \quad \longrightarrow \quad \frac{dG(\tau)}{d\tau} = \int \omega\, e^{-\omega\tau}\, \rho(\omega)\, d\omega$$

$$-\frac{1}{G(\tau)}\frac{dG(\tau)}{d\tau} \;=\; M_{eff}(\tau) \;=\; \frac{1}{G(\tau)} \int \omega\, e^{-\omega\tau}\, \rho(\omega)\, d\omega \;=\; \langle \omega \rangle_{e^{-\omega\tau}\rho(\omega)}$$

Similarly, we can take a 2nd derivative to calculate

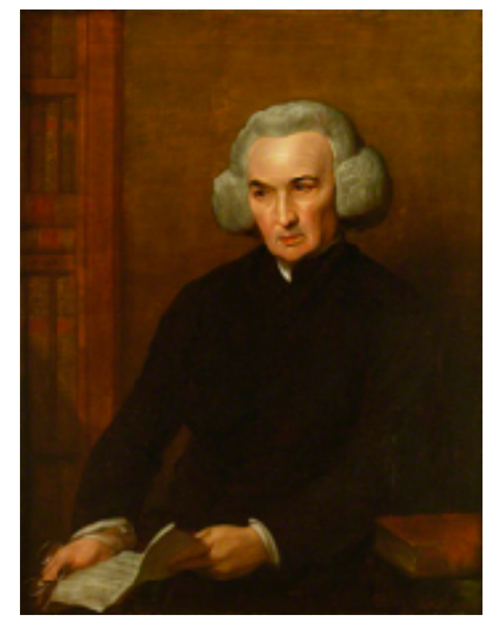
Variance (i.e. width):

$$\Gamma^2 = \frac{1}{G(\tau)}\frac{d^2G(\tau)}{d\tau^2} \;-\; M_{eff}^2 \;=\; \langle\, (\omega - \langle \omega \rangle)^2 \,\rangle$$

Thomas Bayes 1701 - 1761

- Religious Minister
- Did not publish Bayes Theorem

Richard Price 1723 - 1791



- Born in Wales
- Educated in Neath
- Also a religious minister
- Published Bayes Theorem after Bayes death
- Friends and proponent of American Independence Leaders, Benjamin Franklin, Thomas Jefferson, George Washington

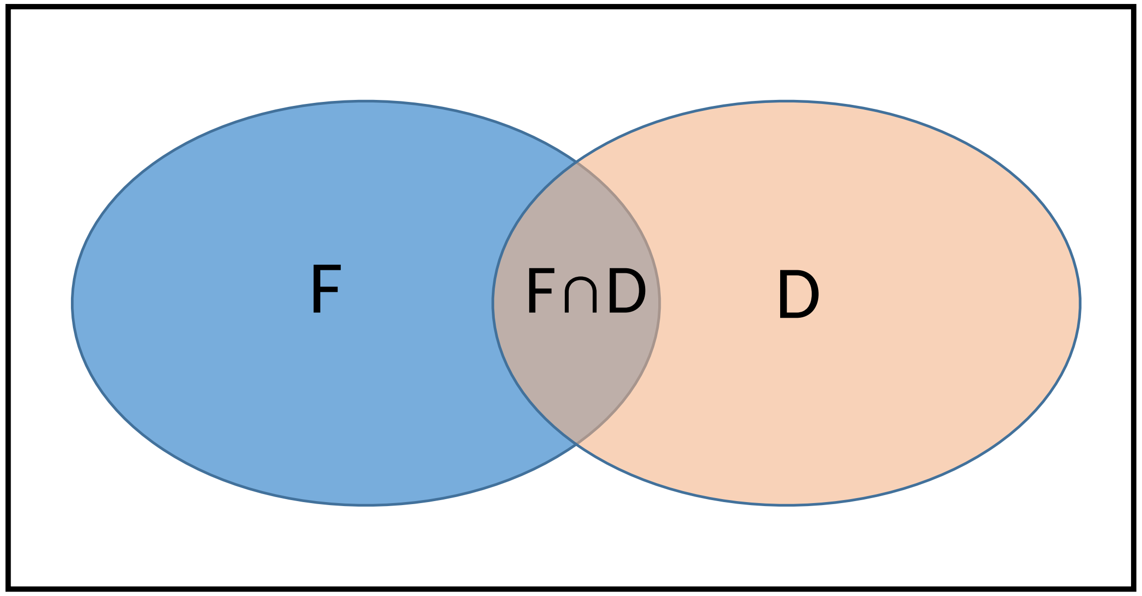
Bayesian Approaches

Need to maximise $P(F \mid D)$

Bayes Theorem :

$$P(F \cap D) = P(F \mid D)\, P(D) = P(D \mid F)\, P(F)$$

i.e. $P(F \mid D) = \dfrac{P(D \mid F)\, P(F)}{P(D)}$



Note $P(D \mid F) \sim \chi^2$

So we should always include $P(F)$ = "Priors"

$P(F)$ is encoded as an *Entropy*

BR and MEM use different Entropy definitions

Bayesian Approaches

Bayes's Theorem: $P(F\,|\,D) = \dfrac{P(D\,|\,F)\,P(F)}{P(D)}$

Is "Maximum Likelihood" method wrong because we don't include $P(F)$ = "Priors" ?

In fact we ***do*** include priors in Maximum Likelihood method in our choice of fitting f'n

e.g. $f(\tau) = Z\,e^{-M\tau}$

So we are always including prior information…

The Prior is a way of regulating the Inverse Problem and removing degeneracies.

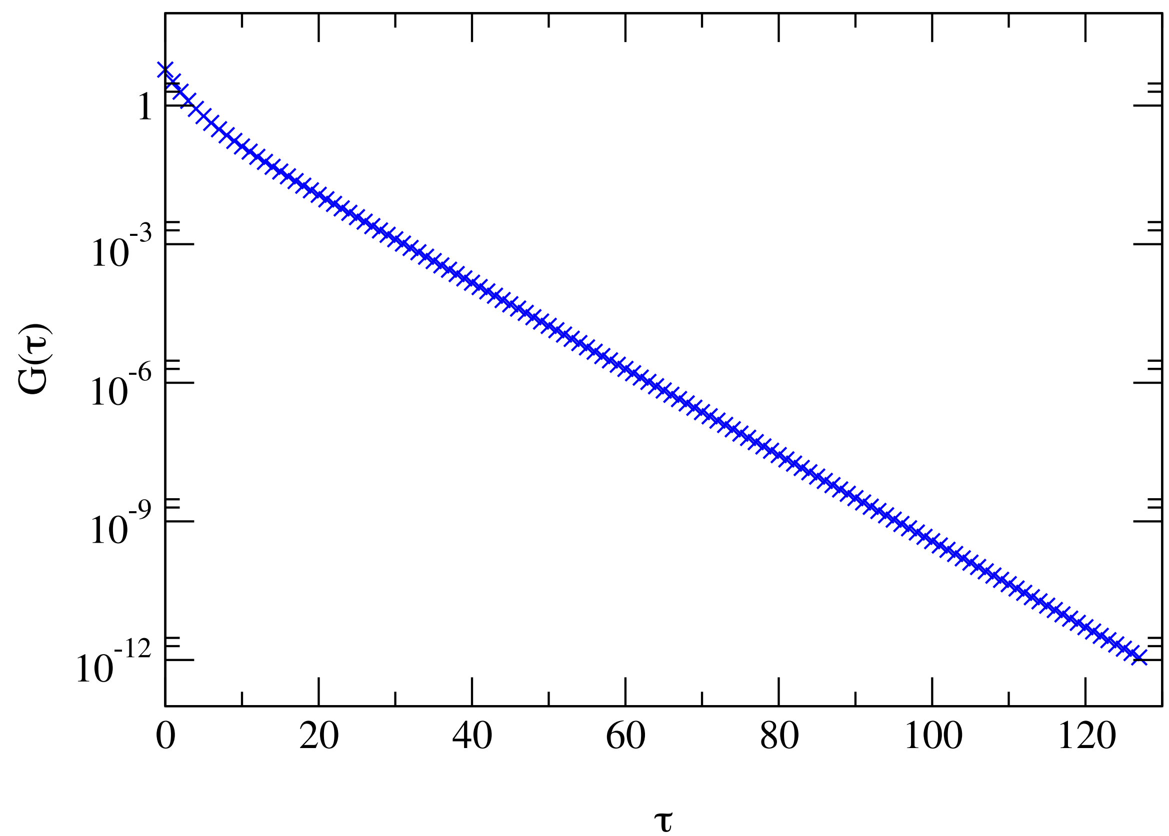
Entropy

	No Data	Data
No Prior	$\mathcal{I}(F) \equiv 0$	F from min χ^2
Prior	$F \equiv$ prior	F from max $P(F\|D)$

Extracting Spectral Functions

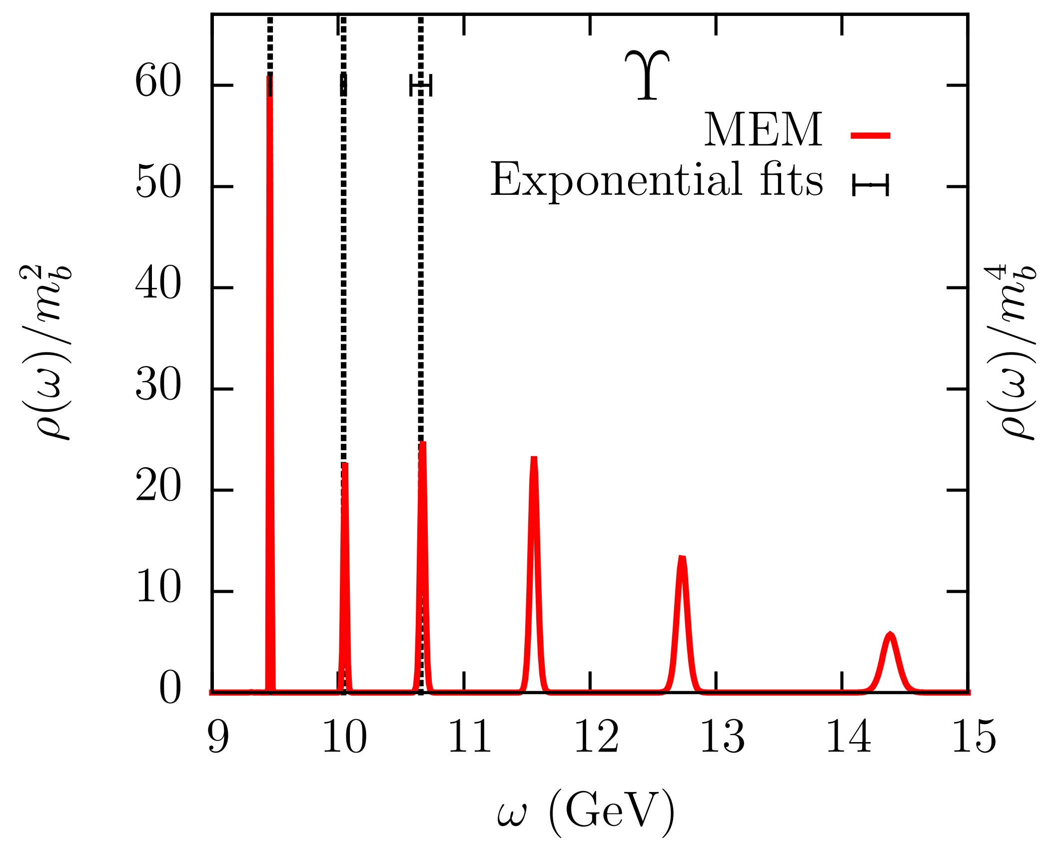
Input Data:

$G_{\pm}(\tau), \ \ \tau = 1,...,\mathcal{O}(10-100)$



Output Data:

$\rho_{\pm}(\omega), \ \ \omega \sim 1,...,\mathcal{O}(1000)$



ill-posed ! ***i.e.*** ∞ ***solutions with*** $\chi^2 = 0$

"Entropy" Factor $P(F)$ breaks this degeneracy

Choice of Entropy Term

$P(F) \sim e^{S}$ $\quad$ S = *Entropy*

Maximum Entropy Method:

Shannon-Jaynes Entropy:

$$S = \int_0^\infty d\omega \left[\rho(\omega) - m(\omega) - \rho(\omega) \ln \frac{\rho(\omega)}{m(\omega)} \right]$$

Asakawa, Hatsuda, Nakahara, Prog.Part.Nucl.Phys. 46 (2001) 459

BR Method:

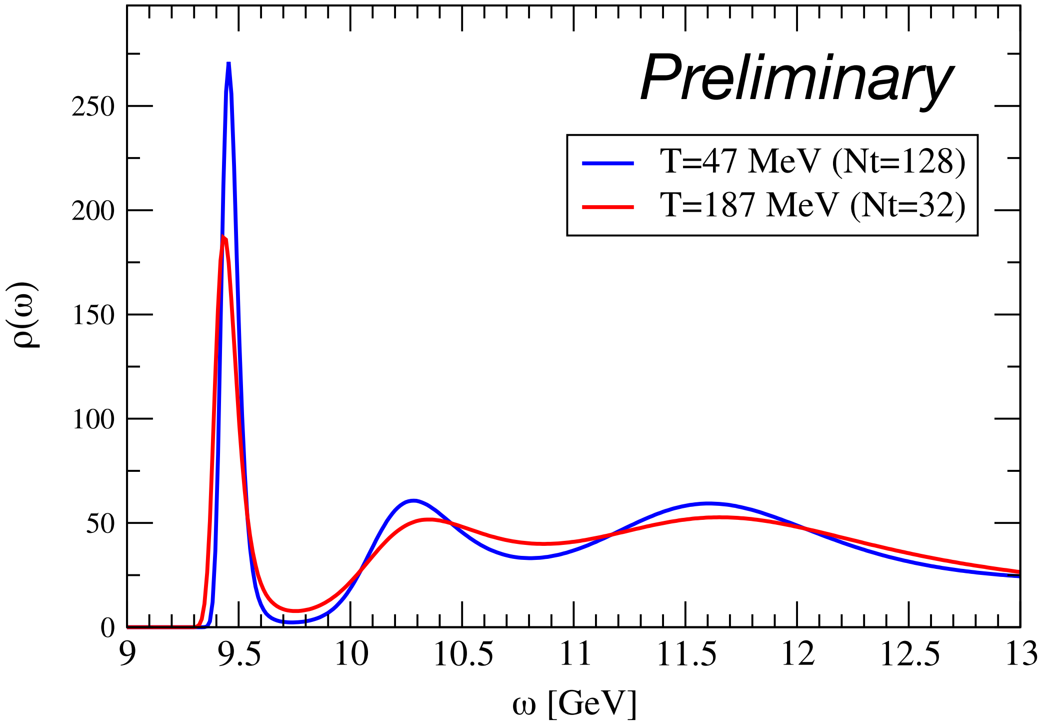
$$S = \int_0^\infty d\omega \left[1 - \frac{\rho(\omega)}{m(\omega)} + \ln \frac{\rho(\omega)}{m(\omega)} \right]$$

Burnier & Rothkop Phys.Rev.Lett. 111 (2013) 182003

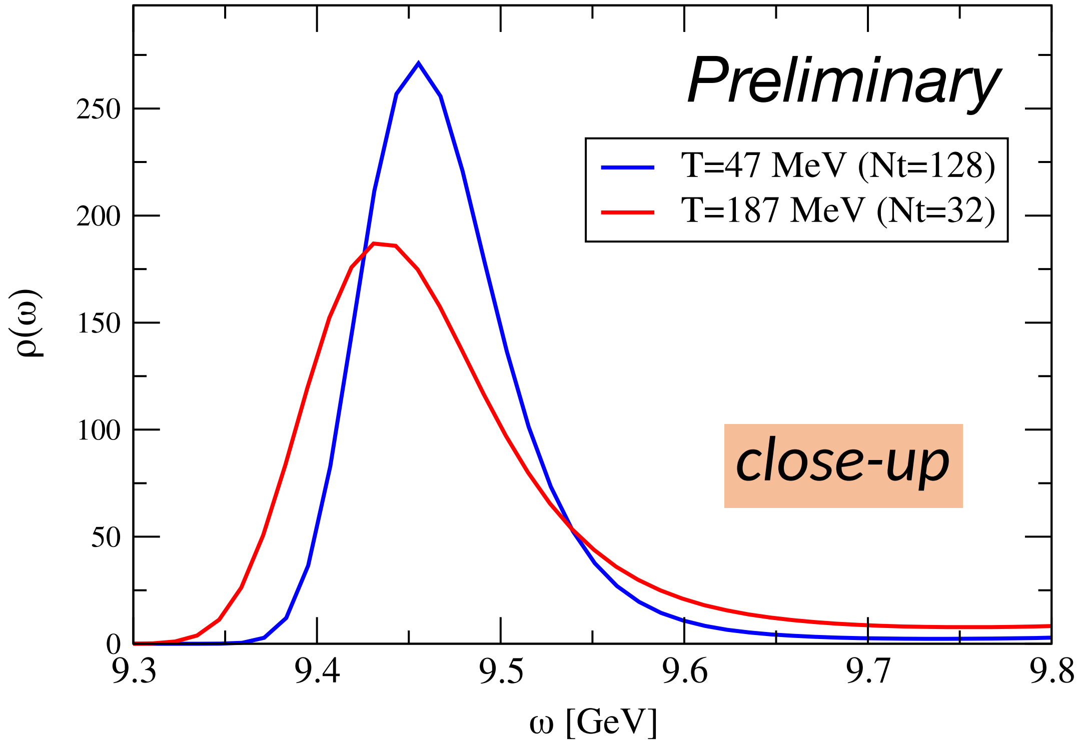
Apples and Apples

Systematic effects in T (MEM)

Fitting: $\tau = [2,30]$ for *both* T



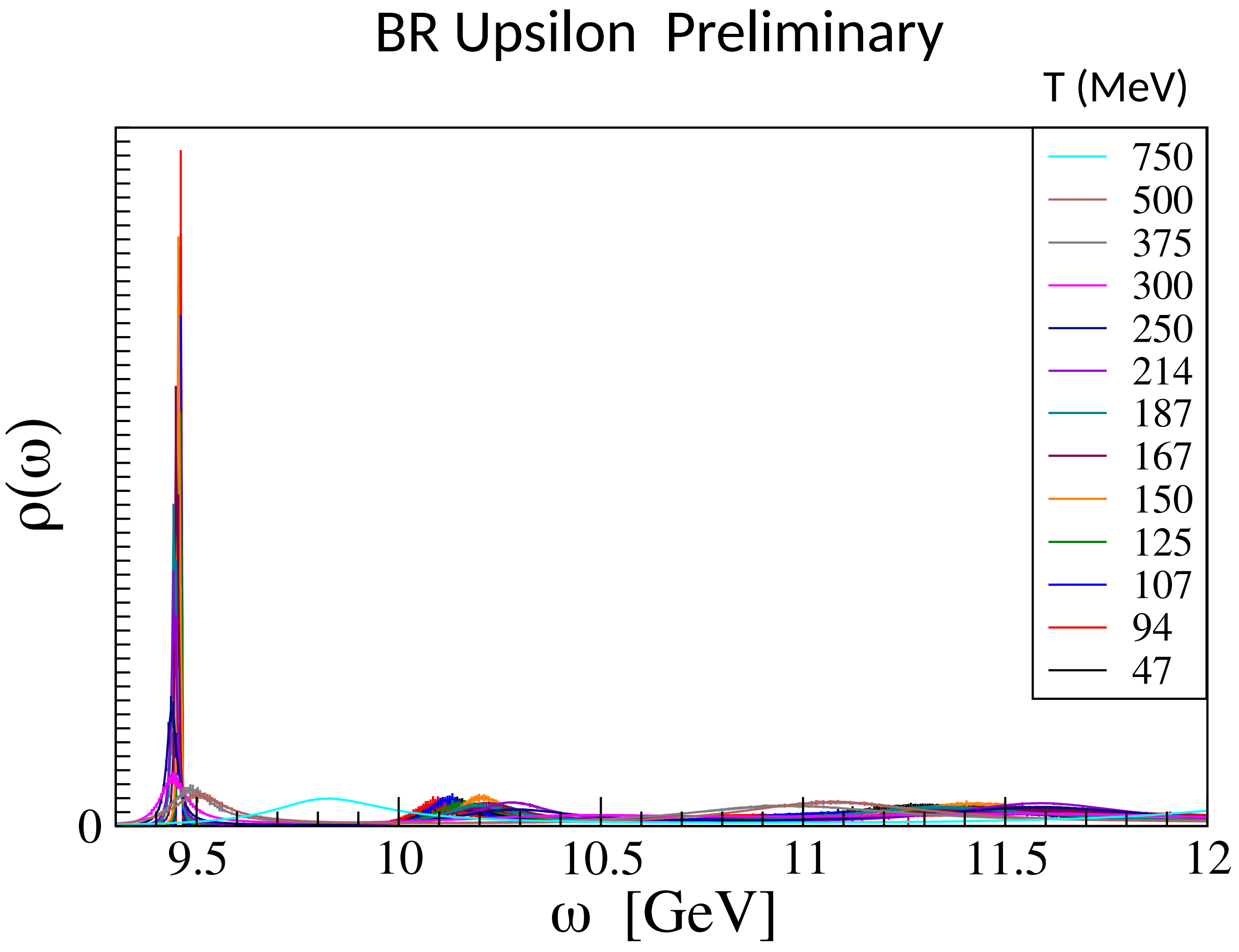
Sequential suppression

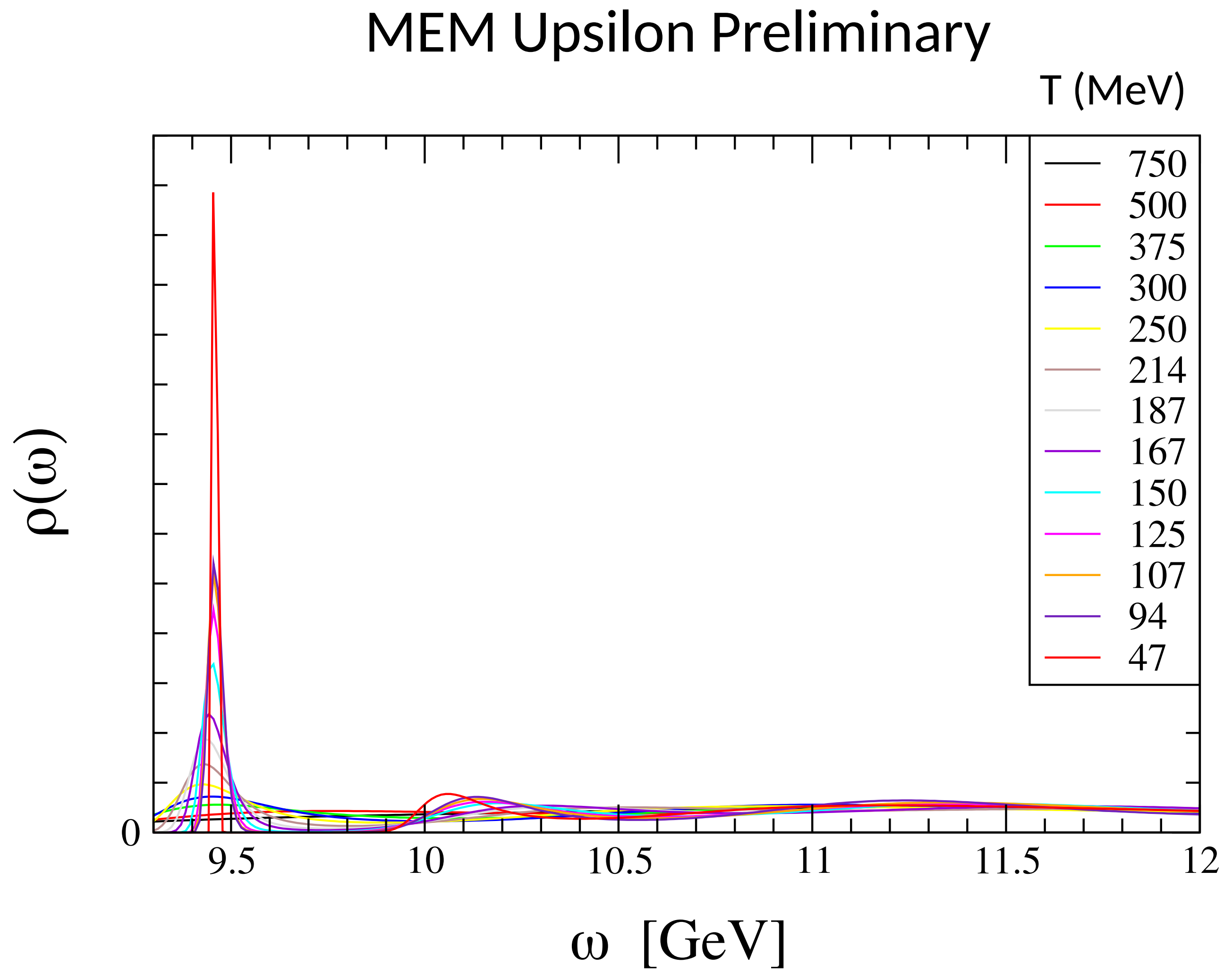


$\Gamma \nearrow$ as $T \nearrow$

Although Γ is *upper bound*, *we can resolve thermal trends*

Direct comparison of Bayesian Approaches





Kernel Ridge Regression

Machine Learning

- uses training data to determine an *alpha matrix* of parameters determined analytically using a cost function
- cost function includes a term to prevent overfitting
- training data set is $\mathcal{O}(10^4)$ mock data with 5 Gaussians
- difficult to produce systematic error estimate

Backus Gilbert

Take $G(\tau) = \int \rho(\omega)\, e^{-\omega\tau}\, d\omega = \int \rho(\omega)\, K(\omega, \tau)$

Generate *averaging functions*: $A(\omega, \omega_0) = \sum_\tau c_\tau\, K(\omega, \tau)$

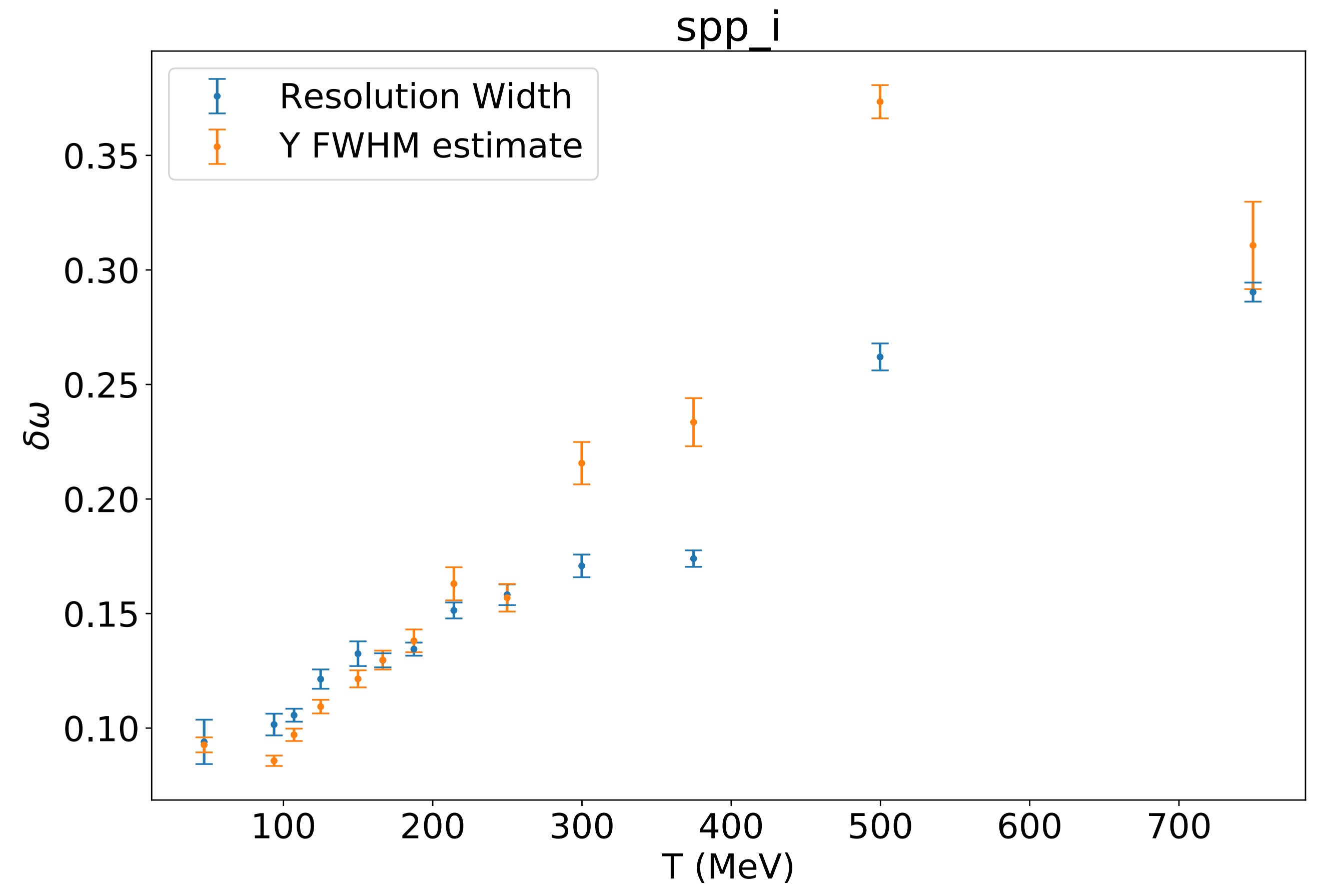
(an approximation to the δ f'n), such that

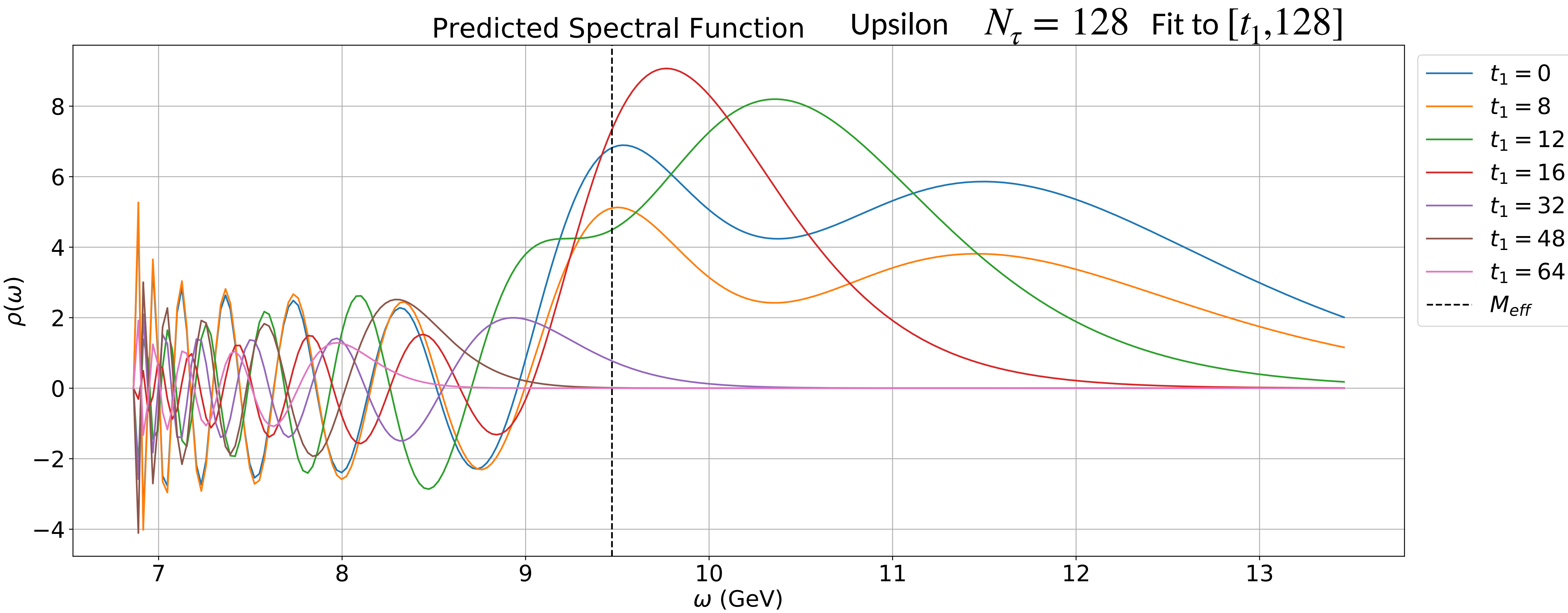
$$
\begin{aligned}
\hat{\rho}(\omega_0) &= \int A(\omega, \omega_0)\, \rho(\omega)\, d\omega \\
&= \sum_\tau c_\tau\, G(\tau) \\
&\approx \rho(\omega_0)
\end{aligned}
$$

Averaging coeffts c_τ determined by *minimising the width of* $A(\omega, \omega_0)$

Backus Gilbert Systematics

- Intrinsic Resolution
- Time window systematics



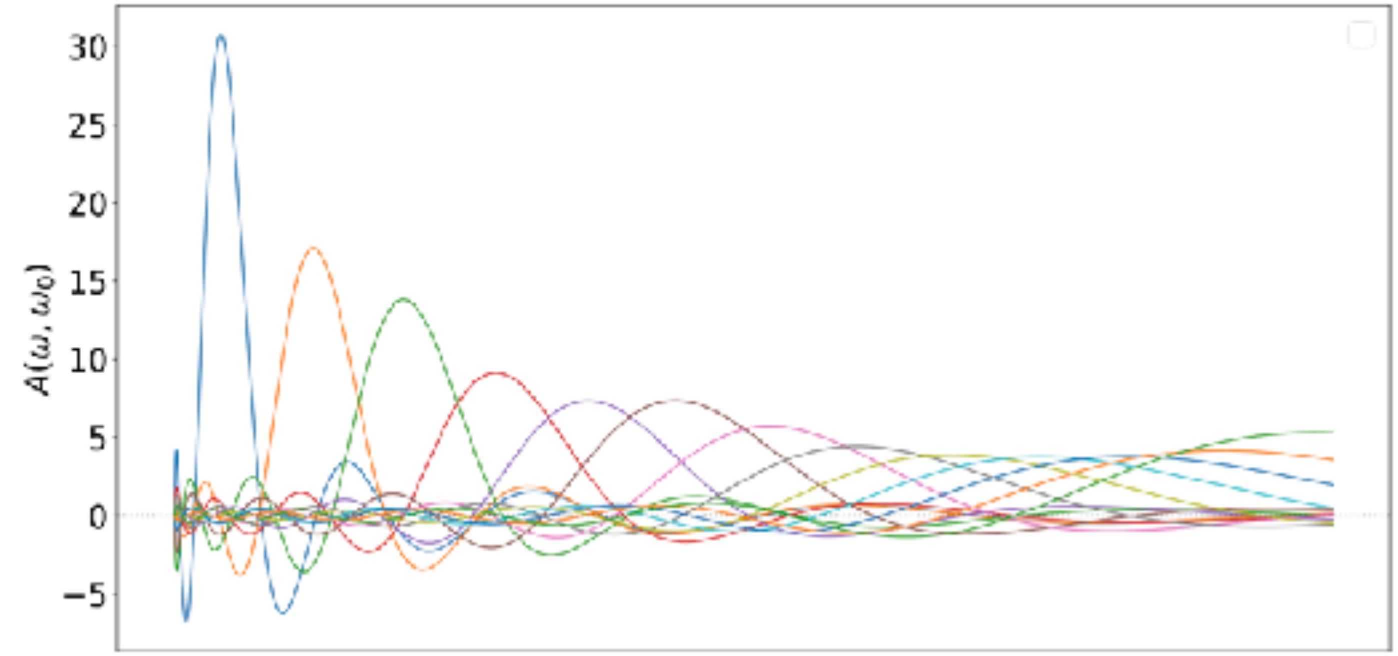


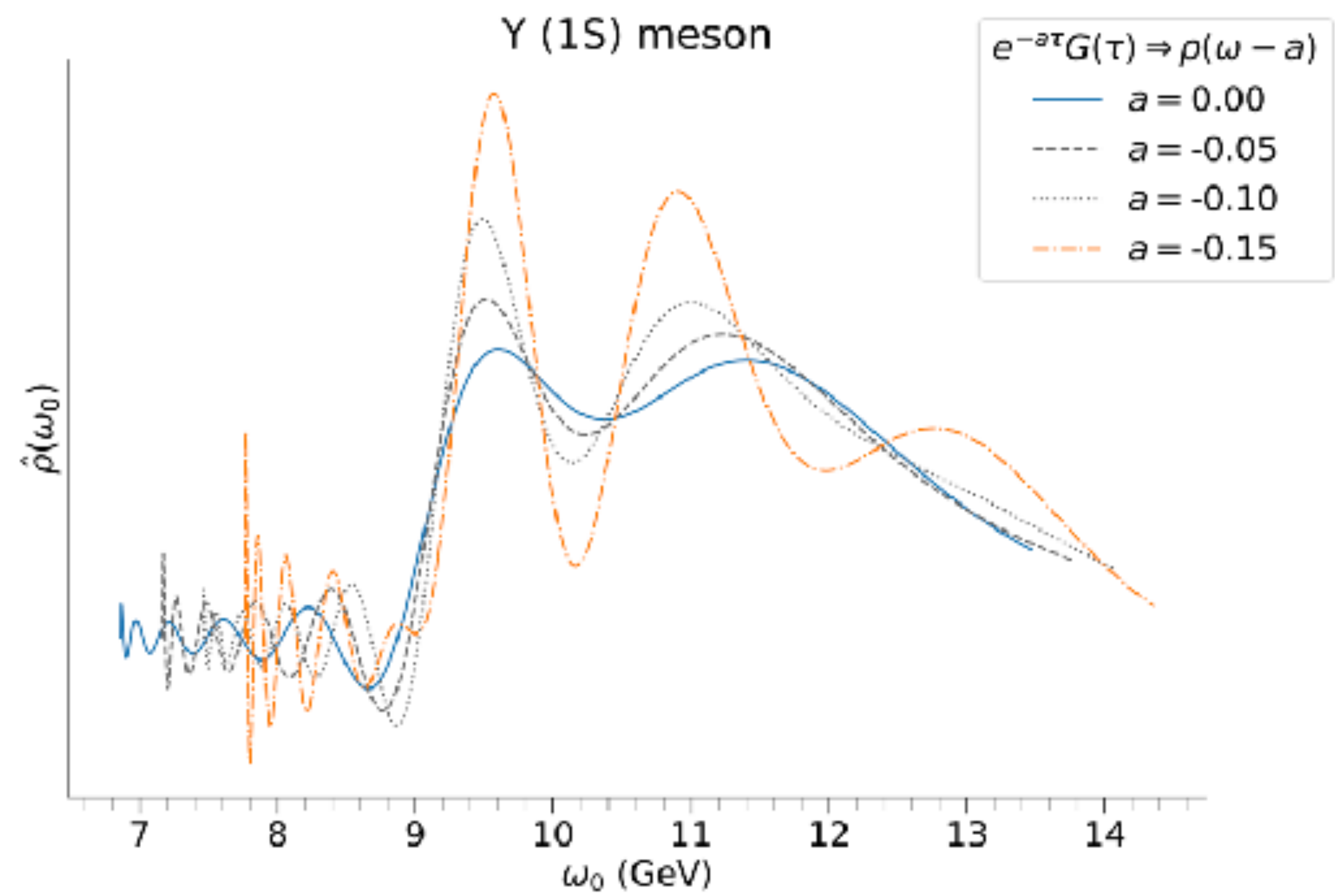
Backus Gilbert Laplace Shift

- Increased sensitivity near origin
- Shift spectral features towards origin using Laplace transform properties:

$$G(\tau) = \int \rho(\omega) e^{-\omega\tau}$$

$$e^{a\tau} G(\tau) = \int \rho(\omega - a) e^{-\omega\tau}$$



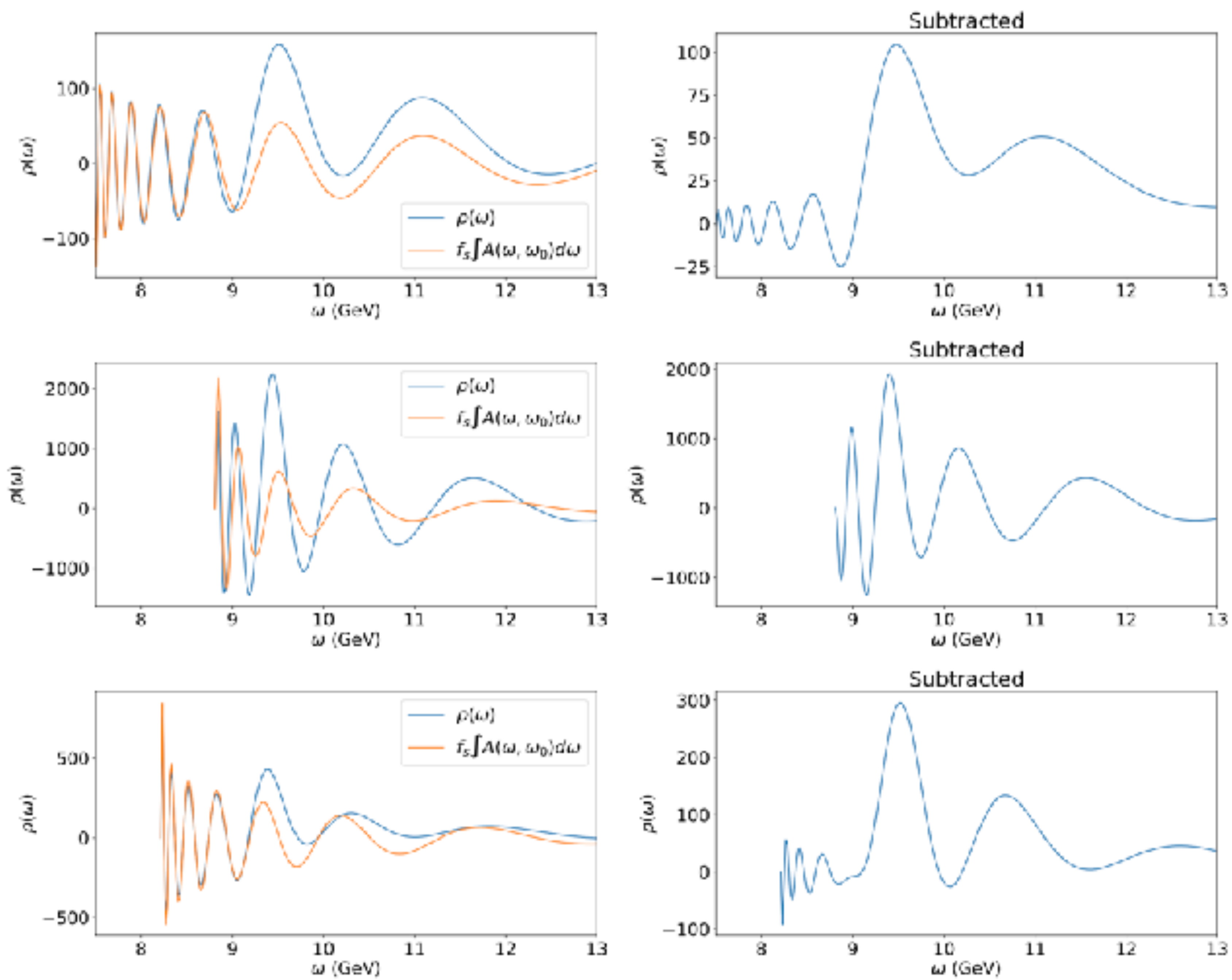


Backus Gilbert Noise Subtraction

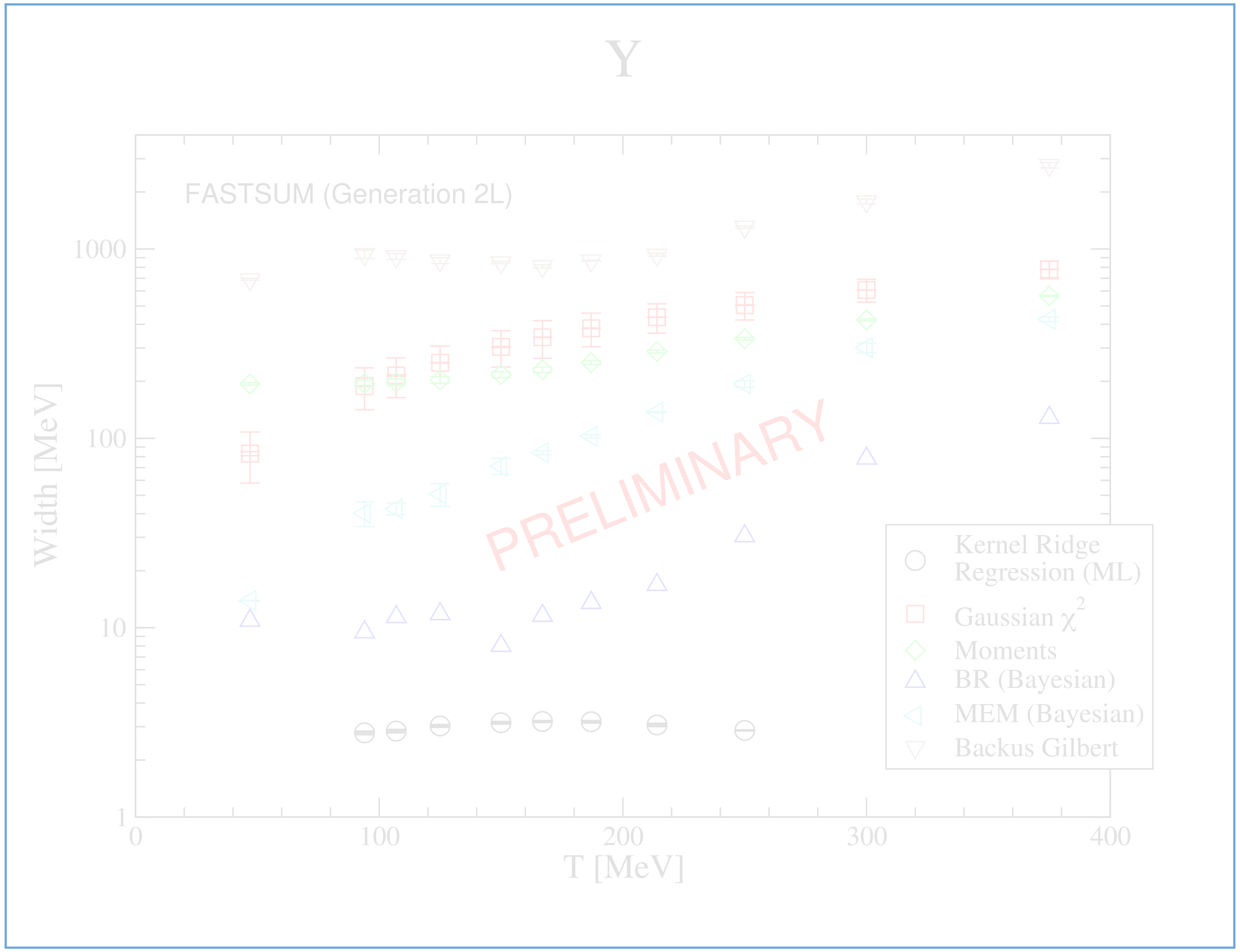
Can approximate noise/systematics by subtracting

$$\int A(\omega, \omega_0) d\omega \sim \text{BG spectrum of constant}$$

from BG spectrum



Comprehensive Study of Systematics from Analysis Techniques



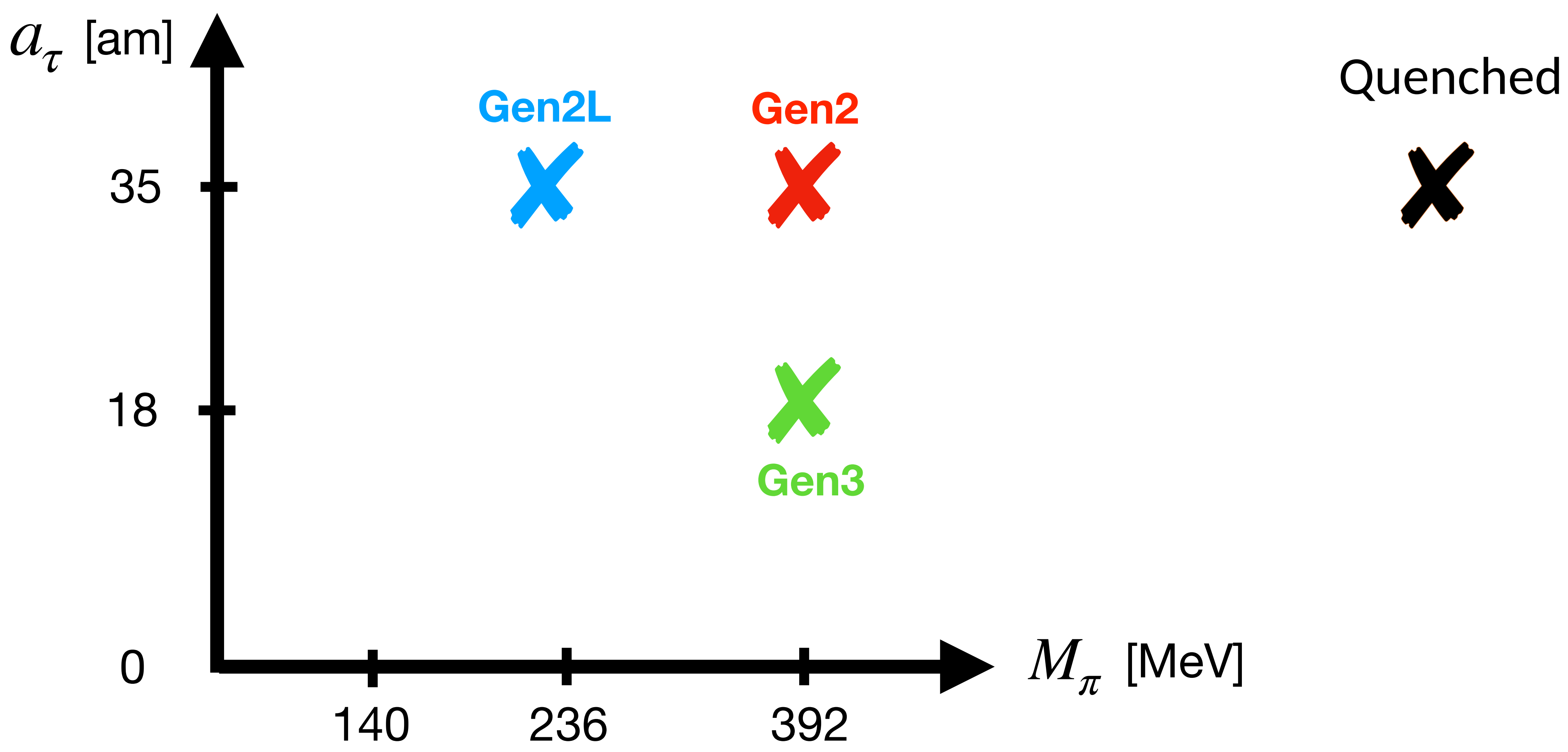
Quantity	Order of Difficulty
M_0	*Easy O(1)*
Γ	*Difficult O(2)*
Line Shape	Very Difficult! *O(3)*

FASTSUM Quenched

Large Volume

High Statistics

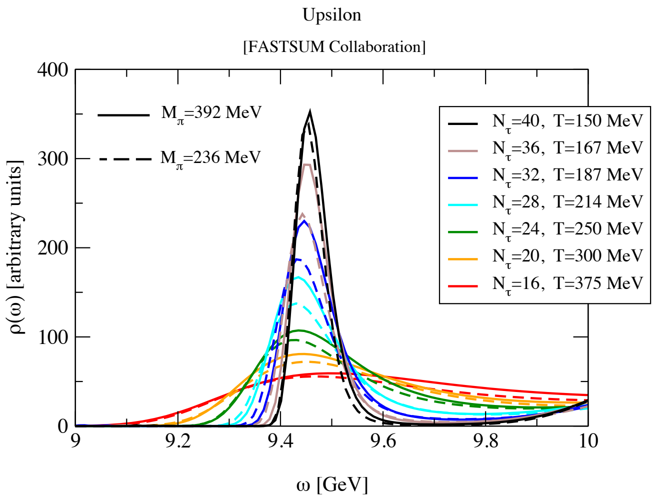
- We will apply all our Inverse techniques

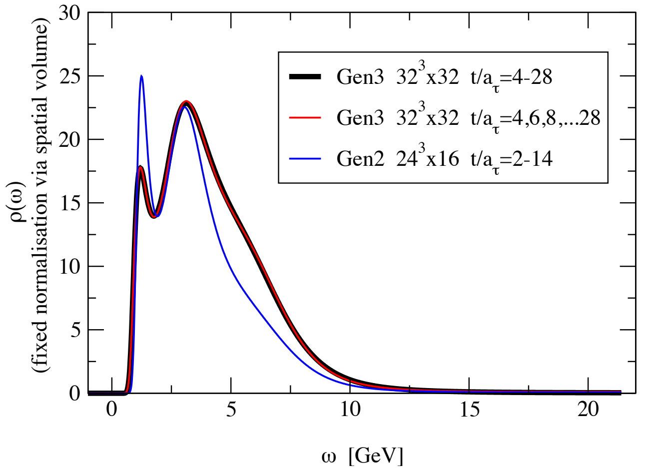


Summary

- Mathematical Limitations of Laplace Transform

- Towards chiral & continuum limits
  - $M_\pi = 392, 236, 140\ MeV$
  - $a_\tau = 33, 17\ am$





- Spectral Reconstruction from 7 Methods
  - *Max.Likelihood (x2)*
  - *Moments*
  - *Bayesian (x2)*
  - *Machine Learning*
  - *Backus-Gilbert*

Towards Systematic Understanding of Bottomonium Spectrum from the Lattice