Chapter 5

Correlator Analysis: Determining the Finite-Volume Spectrum

5.1 Correlator Matrices

In order to compute the low-lying finite-volume stationary states of QCD, we compute temporal correlator matrices of the form,

$$\mathcal{C}_{ij}(t) = \langle 0 | \mathcal{O}_i (t + t_0) \overline{\mathcal{O}}_j (t_0) | 0 \rangle , \tag{5.1}$$

where $\mathcal{O}_i(t + t_0)$ are the operators which annihilate the states of interest at a time $t + t_0$ and $\overline{\mathcal{O}}_j(t_0)$ are the corresponding operators which create the states of interest at an earlier time t_0. The operators are engineered such that the correlator matrix is Hermitian.[1] Since the energies are determined from the exponential decay rates of these two-point correlators, we can rescale the operators without changing the energy spectrum. In order to dampen the effects of differing normalizations among the operators, we rescale $\mathcal{C}(t)$ to form the actual correlator matrix $C(t)$ which we analyze as follows,

$$C_{ij}(t) = \frac{\mathcal{C}_{ij}(t)}{\sqrt{\mathcal{C}_{ii}(\tau_N)\mathcal{C}_{jj}(\tau_N)}}, \tag{5.2}$$

where the *normalization time* τ_N is taken at a very early time. Echoing the argument in Sec. 2.3, we can perform a spectral decomposition of $C(t)$ to obtain

$$C_{ij}(t) = \sum_n Z_i^{(n)} Z_j^{(n)*} e^{-E_n t}, \tag{5.3}$$

---

[1]This follows from the requirement that $\langle 0 | O_i | n \rangle = \langle n | \overline{O}_i | 0 \rangle^*$.

where the $E_n < E_{n+1}$, the energies have been shifted such that $E_0 \equiv 0$, and we have defined the *overlap factors*,

$$Z_j^{(n)*} = \langle n | \bar{O}_j | 0 \rangle, \quad Z_j^{(n)} = \langle 0 | O_j | n \rangle. \tag{5.4}$$

The overlap factors offer us some insight into how much the state produced by a particular operator overlaps onto a particular energy eigenstate. This information can be used to make qualitative assessments of the contents of a particular energy state, e.g. if it is single- or multi-hadron dominated. Note that invariance under the a shift in phase

$$Z_j^{(n)} \rightarrow Z_j^{(n)} e^{i\phi_n} \tag{5.5}$$

implies that we can only determine the magnitude $|Z_j^{(n)}|$ of these overlap factors.

5.2 The Generalized Eigenvalue Problem

In the limit of infinite statistics, one could in principle fit Eq. (5.3) and determine the entire energy spectrum. In practice, however, we are only able to fit the lowest one or two energies in the sum. It would be better to devise a method that allows us to fit more than just the lowest one or two energies in the spectrum. Ref. [32] proves a theorem that is the cornerstone for our approach to extracting energies from temporal correlators:

Theorem: For every $t \geq 0$, let $\lambda_n(t)$ be the eigenvalues of an $N \times N$ Hermitian correlation matrix $C(t)$ ordered such that $\lambda_0 \geq \lambda_1 \geq \cdots \geq \lambda_{N-1}$, then

$$\lim_{t \to \infty} \lambda_n(t) = b_n e^{-E_n t} \left[1 + \mathcal{O}(e^{-t\Delta_n}) \right], \quad b_n > 0, \quad \Delta_n = \min_{m \neq n} |E_n - E_m|. \tag{5.6}$$

This tells us that a *principal correlator* $\lambda_n(t)$ tends to a decaying exponential whose decay rate is the nth energy of the spectrum. This theorem is still insufficient for use in practice, however. At large t such that the $\mathcal{O}(e^{-t\Delta_n})$ correction is negligible, the determination of $C(t)$ has large uncertainties due to a decreasing signal-to-noise ratio [36]. At small t such that $C(t)$ is well determined, the $\mathcal{O}(e^{-t\Delta_n})$ is not negligible. Fortunately, it is shown in Ref. [32] that the correction term can be brought down to $\mathcal{O}(e^{-t(E_N - E_n)})$ by instead solving the following generalized eigenvalue problem (GEVP):

$$C(t)v_n(t, \tau_0) = \lambda_n(t, \tau_0) C(\tau_0) v_n(t, \tau_0), \quad n = 1, \cdots, N-1, \quad \frac{t}{2} \leq \tau_0 < t. \tag{5.7}$$

In practice, we find that a good rule of thumb to keep the $\mathcal{O}(e^{-t(E_N - E_n)})$ correction term low is to demand $N \geq \frac{3n}{2}$, where n is the number of energies we wish to extract. Solving

Eq. (5.7) is equivalent to diagonalizing $G(t) = C^{-1/2}(\tau_0) C(t) C^{-1/2}(\tau_0)$, and the eigenvectors tend to [37]

$$\lambda_n(t) \to |Z_n'|^2 e^{-E_n t}, \quad t \to \infty. \tag{5.8}$$

The overlaps are determined by

$$Z_j^{(n)} \approx C_{jk}(\tau_0)^{1/2} V_{kn}(t) Z_n' \quad \text{(no sum over n)}, \tag{5.9}$$

where V is the unitary matrix whose columns are the eigenvectors of $G(t)$.

5.3 Pivots and Operator Pruning

Two important assumptions about $C(t)$ are that it is Hermitian as well as positive-definite, which also implies that $G(t)$ is Hermitian and positive-definite. The assumption of positive-definiteness can be violated by operators that produce noisy correlators, or by an operator set that is linearly dependent. Noise can cause diagonal elements of the correlator to become zero or negative, and linearly dependent operators will produce eigenvalues of $C(t)$ that are zero. The odds of this occurring increases as we increase the number of operators in our correlator matrix. We call the process of judiciously choosing operators to ensure our correlator matrices remain well-conditioned *pruning*. Pruning is necessary, but not sufficient for ensuring well-conditioned correlator matrices. It is still necessary to monitor for ill-conditioned matrices and take corrective action.

The condition number ξ^{cn} of a matrix is defined as the magnitude of the ratio of its maximum eigenvalue $\lambda_{\max}$ to its minimum eigenvalue $\lambda_{\min}$:

$$\xi^{\text{cn}} = \left| \frac{\lambda_{\max}}{\lambda_{\min}} \right|. \tag{5.10}$$

When a matrix is ill-conditioned, its condition number will be high (and in some cases it may have negative eigenvalues). In order to prevent statistical noise from ruining our extraction of the energy spectrum, it is important to modify our diagonalization procedure to correct for high condition numbers, which is essentially introducing a singular value decomposition. These methods choose a threshold for the maximal allowed condition number $\xi^{\text{cn}}_{\text{threshold}}$ and project the correlator matrix onto the subspace spanned by the eigenvectors whose eigenvalues are larger than $\lambda_{\text{threshold}} = \lambda_{\max}/\xi^{\text{cn}}_{\text{threshold}}$. (The eigenvalues and condition numbers are all functions of t.)

Our entire method, which we call a pivot procedure, can be summarized as follows. Choose values for $\xi^{\text{cn}(0)}_{\max}$, the maximal acceptable condition number for the $N \times N$ matrix

$C(\tau_0)$. Let the $N_0 \leq N$ eigenvectors whose eigenvalues are greater than $\lambda_{\max}^{(0)}/\xi_{\max}^{\text{cn}(0)}$ form the columns of the $N \times N_0$ matrix P_0, where $\lambda_{\max}^{(0)}$ is the largest magnitude of the eigenvalues of $C(\tau_0)$. Then define the $N_0 \times N_0$ matrices

$$
\begin{aligned}
\widetilde{C}(\tau_0) &= P_0^\dagger C(\tau_0) P_0, \\
\widetilde{C}(t) &= P_0^\dagger C(t) P_0, \\
\widetilde{G}(t) &= \widetilde{C}(\tau_0)^{-1/2} \widetilde{C}(t) \widetilde{C}(\tau_0)^{-1/2},
\end{aligned}
\tag{5.11}
$$

for $t \neq t_0$. We then solve for the eigenvectors and eigenvalues of $\widetilde{G}(t)$. Let the $N_P \leq N_0$ eigenvectors whose eigenvalues are greater than $\lambda^{(t)}/\xi^{(t))\max}$ form the columns of the $N_0 \times N_P$ matrix $\widetilde{V}(t)$, where $\lambda_{\max}^{(t)}$ is the largest magnitude of the eigenvalues of $\widetilde{G}(t)$. The $N_P \times N_P$ matrix

$$
\widetilde{\Lambda}(t) = \widetilde{V}^\dagger(t) \widetilde{G}(t) \tilde{V}(t)
\tag{5.12}
$$

will satisfy

$$
\begin{aligned}
\widetilde{\Lambda}(\tau_0) &= I, \\
\widetilde{\Lambda}(t) &= \text{diag}(\lambda_n(t)),
\end{aligned}
\tag{5.13}
$$

where $\lambda_n \to \left|\widetilde{Z}'_n\right|^2 e^{-E_n t}$ for large t and $Z_i^{(n)} \approx P_{0ij} \widetilde{C}_{jk}(\tau_0)^{1/2} \widetilde{V}_{kn}(t) \widetilde{Z}'_n$. We can therefore fit the different $\lambda_n(t)$ for energies and overlap factors of the spectrum. The pivot method presented above involves rotating the correlator at every time slice, but in practice, this is not strictly necessary. We can instead take a simpler approach of diagonalizing $\widetilde{G}(t)$ at only one time τ_d chosen such that $\tau_d > \tau_0$, and using the same matrix to diagonalize at every other time, provided that the off-diagonal elements are statistically consistent with zero for all time slices $t \geq \tau_d$. Using the same transformation at every time slice is an ingredient in the *single pivot* method, and is the procedure used in this work. Note that choosing $\tau_d \approx 2\tau_0$ ensures minimal contamination from higher lying states [32].

5.3.1 The Zero Temperature Assumption

An important caveat to everything presented thus far is that we work in the assumption of a zero-temperature (i.e. infinite time extent) limit. Recall that the expression relating two-point correlators to a sum of decaying exponentials, given by Eq. (2.49) rests on the assumption that $t \ll T$, where T is the time extent of the lattice. When calculating correlator

matrix elements, this should be corrected to

$$
\begin{aligned}
\mathcal{C}_{ij} &= \left\langle \mathcal{O}_i(t)\overline{\mathcal{O}}_j(0)\right\rangle_T \\
&= \frac{1}{Z_T}\operatorname{Tr}\left[e^{-HT}\mathcal{O}_i(t)\overline{\mathcal{O}}_j(0)\right] \\
&= \frac{1}{Z_T}\sum_n \langle n|\, e^{-H(T-t)}\mathcal{O}_i(0)e^{-Ht}\overline{\mathcal{O}}_j(0)\,|n\rangle \\
&= \frac{1}{Z_T}\sum_{n,m} e^{-E_n(T-t)}e^{-E_m t}\,\langle n|\,\mathcal{O}_i(0)\,|m\rangle\,\langle m|\,\overline{\mathcal{O}}_j(0)\,|n\rangle\,,
\end{aligned}
\tag{5.14}
$$

where

$$
\begin{aligned}
Z_T &= \operatorname{Tr} e^{-HT} \\
&= \sum_n \langle n|\, e^{-HT}\,|n\rangle \\
&= \sum_n e^{-E_n T}.
\end{aligned}
\tag{5.15}
$$

Only in the limit $T \to \infty$ does Eq. (5.14) agree with Eq. (2.49) and Eq. (5.3). The thermal effects of finite time extent will manifest as backwards propagating modes which will be briefly discussed in Sec. 5.4. Fortunately, on the $32^3 \times 256$ anisotropic lattice used in this work, these thermal effects (or temporal wrap-around effects) are negligible, and even on smaller lattices, we only find it necessary to account for backwards propagating modes for the lightest states.

5.4 Fit Forms

Fitting the diagonalized correlators is accomplished by forming fit ansatzes based on decaying exponentials, or symmetric exponentials when we wish to account for thermal effects. One such fit form is the single exponential,

$$
C(t) = Ae^{-Et}.
\tag{5.16}
$$

This form is strictly incorrect for finite t due to excited state contamination (i.e. subleading exponential contributions from higher-lying states), but it is approximately correct for sufficiently large t. The single exponential has the advantage that it has few fit parameters and is thus less sensitive to statistical noise, but it has the disadvantage that it is highly sensitive to the minimum time in the used in the fit domain. The fit domain must only include times at which the subleading contributions are negligible.

We can account for the effects of higher-lying states by using a two-exponential fit,

$$C(t) = Ae^{-Et}\left(1 + Be^{-\Delta^2 t}\right), \tag{5.17}$$

where the term Δ^2 is used to ensure that the gap to the next energy is positive. We can also attempt to account for higher-lying states by using an ansatz that mocks up these states as a sum of equally-spaced levels above the fit energy E:

$$C(t) = \frac{Ae^{-Et}}{1 - Be^{-\Delta^2 t}} = Ae^{-Et}\left(\sum_{n=0}^{\infty} B^n e^{-\Delta^{2n} t}\right). \tag{5.18}$$

These forms have the advantage that they are less sensitive to the fit domain and more accurately model the theory, but they have the disadvantage that they contain more parameters and are more sensitive to statistical noise. For this reason, we must often resort to using a single-exponential fit for noisier correlators.

If we wish to account for thermal effects for mesons, we can also make any fit form symmetric, e.g.

$$C(t) = A\left(e^{-Et} + e^{-E(T-t)}\right) \tag{5.19}$$

for the single exponential fit form. For baryons, things become more complicated, since the backwards propagating mode actually corresponds to a baryon's parity partner, which would necessitate fitting in two symmetry channels simultaneously. Fortunately, baryons are heavy enough that accounting for temporal wrap-around effects is usually not necessary. Note that no extra parameters are needed for this, so it is easy to verify if such fit forms are necessary. Additionally, a constant can be added to any of the fit forms to help account for vacuum expectation values and to account for thermal effects on smaller lattices.

5.4.1   The Effective Energy

A very useful tool in fitting correlators is the so-called *effective energy* or *effective mass* which is given by

$$E_{\text{eff}}(t) = -\frac{d}{dt}\ln C(t), \tag{5.20}$$

which can be discretized and estimated as

$$-\frac{1}{\Delta t}\left(\ln C(t + \Delta t) - \ln C(t)\right). \tag{5.21}$$

The effective energy is designed such that it tends to the fit energy E as $t \to \infty$. When performing our fits, we are usually visually guided by effective energy plots rather than the
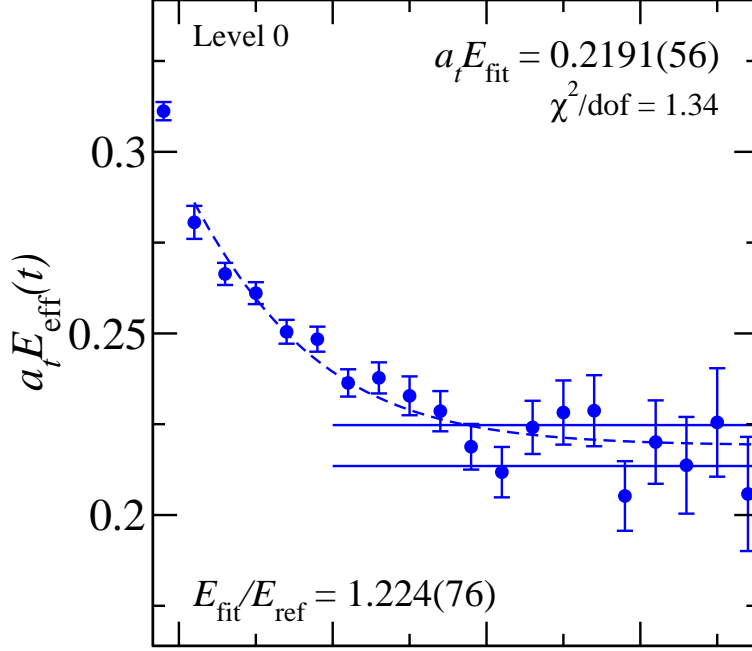
Figure 5.1: Example of an effective energy plot. A curve obtained from a two-exponential correlator fit is overlaid, along with horizontal bars above and below the plateau, which depict the confidence interval of the energy of the leading exponential.

correlators themselves, but it is crucial to emphasize that we do not fit the effective energy itself. An example of an effective energy plot is given in Fig. (5.1).

5.5 Error Analysis and Resampling

The observables we estimate through direct Monte Carlo measurements on gauge configurations are called *simple* observables. Estimating the variance of simple observables is straightforward via calculating the population variance using the Monte Carlo method. All other observables are called *non-simple* observables. An example of a non-simple observable is a model parameter obtained when we fit a correlator to data points which are themselves averages over many gauge configurations. The way to approach error analysis for non-simple observables is through *resampling* techniques. Resampling involves forming new datasets from an initial dataset in order to estimate properties of how that dataset is distributed. For example, if we wanted to estimate the variance of the mean of some variable x, we would need several sets of measurements $\{x\}_i$ from which we can estimate the variance of the mean as

$$\sigma_{\overline{x}}^2 = \langle \overline{x}^2 \rangle - \langle \overline{x} \rangle^2, \tag{5.22}$$

where $\overline{x}$ denotes the population mean of a single set, and $\langle\rangle$ denotes the average over all populations $\{x\}_i$. In general, we wish to inquire about any arbitrary property of a set $\{x\}$. When only one dataset $\{x\}$ is available, resampling can be used to form new sets $\{x\}_i$ using only the original set. The dataset we consider is a set of gauge configurations $\{U\}$. Here we outline two resampling techniques, namely the *jackknife* and the *bootstrap* [38].

Jackknife resampling proceeds as follows. For a set of N_c configurations $\{U\}$, form new sets $\{U\}_i$ which consist of every configuration in $\{U\}$ *except* for the ith configuration. There will be N_c such new sets, and we expect that because only one configuration has been removed from each, that any properties we measure of the set will be close to those of the original set.

Bootstrap resampling proceeds as follows. For a set of N_c configurations $\{U\}$, form N_b new sets $\{U_i\}$ by choosing N_c configurations from $\{U\}$ randomly with replacement. Since configurations can be sampled multiple times or not at all, we do not expect that properties we measure of the resampled sets will necessarily be distributed closely to that of the original set, as we do with jackknife resampling.

The properties we wish to estimate are the covariances of our model fit parameters, such as the energies and overlap factors. For jackknife resampling, we estimate the covariance as

$$\text{cov}(f_i, f_j) = \frac{N_c - 1}{N_c} \sum_{i=1}^{N_c} \left(\langle f\rangle_i - \langle f\rangle^J\right) \left(\langle f\rangle_j - \langle f\rangle^J\right), \tag{5.23}$$

where $\langle f\rangle_i$ denotes the average of f over the ith jackknife resampling, and $\langle f\rangle^J$ denotes the average of f over all resamplings. For bootstrap resampling, we estimate the covariance as

$$\text{cov}(f_i, f_j) = \frac{1}{N_b - 1} \sum_{i=1}^{N_b} \left(\langle f\rangle_i - \langle f\rangle^B\right) \left(\langle f\rangle_j - \langle f\rangle^B\right), \tag{5.24}$$

where $\langle f\rangle_i$ denotes the average of f over the ith bootstrap resampling, and $\langle f\rangle^B$ denotes the average of $\langle f\rangle_i$ over all bootstrap resamplings.

5.6   Correlated χ^2 Minimization

For uncorrelated data, it is usually sufficient to fit the model parameters $\boldsymbol{\alpha}$ of a function $f(t; \boldsymbol{\alpha})$ correlater data $C(t)$ by minimizing the uncorrelated χ^2 metric defined by

$$\chi^2 = \sum_t \frac{(C(t) - f(t; \boldsymbol{\alpha}))^2}{\sigma_t^2}, \tag{5.25}$$

where σ_t^2 is the variance of the data at each time point. Our measurements are not statistically independent, however, because each measurement is taken on the same set of gauge configurations. We must therefore fit our correlators by instead minimizing the correlated χ^2, defined by

$$\chi^2 = \sum_{t,t'} = \left(C(t) - f(t; \boldsymbol{\alpha}) \right) \mathrm{cov}^{-1} \left(C(t), C(t') \right) \left(C(t') - f(t'; \boldsymbol{\alpha}) \right). \tag{5.26}$$

The model parameters are then fit by minimizing χ^2 on each resampling and calculating their uncertainties using the relevant covariance formula for the given resampling scheme.