$K_L - K_S$ mass difference from Lattice QCD

Jianglei Yu

Submitted in partial fulfillment of the

requirements for the degree

of Doctor of Philosophy

in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2014

Abstract

$K_L - K_S$ mass difference from Lattice QCD

Jianglei Yu

The $K_L - K_S$ mass difference is a promising quantity to reveal new phenomena which lie outside the standard model. A state-of-art perturbation theory calculation has been performed at next-to-next-to-leading order (NNLO) and a 40% error is quoted in the final result. We develop and demonstrate non-perturbative techniques needed to calculate the $K_L - K_S$ mass difference, ΔM_K, in lattice QCD and carry out exploratory calculations. The calculations are performed on a 2+1 flavor, domain wall fermion, $16^3 \times 32$ ensemble with a 421 Mev pion and a $24^3 \times 64$ lattice ensemble with a 330 MeV pion. In the 16^3 lattice calculation, we drop the double penguin diagrams and the disconnected diagrams. The short distance part of the mass difference in a 2+1 flavor calculation contains a quadratic divergence cut off by the lattice spacing. Here, this quadratic divergence is eliminated through the Glashow-Iliopoulos-Maiani (GIM) mechanism by introducing a quenched charm quark. We obtain a mass difference ΔM_K which ranges from $6.58(30) \times 10^{-12}$ MeV to $11.89(81) \times 10^{-12}$ MeV for kaon masses varying from 563 MeV to 839 MeV. On the 24^3 lattice, we include all the diagrams and perform a full calculation. Our result is for a case of unphysical kinematics with pion, kaon and charmed quark masses of 330, 575 and 949 MeV respectively. We obtain $\Delta M_K = 3.19(41)(96) \times 10^{-12}$ MeV, quite similar to the experimental value. Here the first error is statistical and the second is an estimate of the systematic discretization error. An interesting aspect of this calculation is the importance of the disconnected diagrams, a dramatic failure of the OZI rule.

Contents

List of Tables

List of Figures

ACKNOWLEDGMENTS

I would like to thank my advisor Prof. Norman Christ very much for his guidance, suggestions, encouragements, and help during my research at Columbia University. I really enjoyed the weekly discussions, and learned a lot.

I am grateful to Prof. Robert Mawhinney for all the insightful discussions, help on the machines and other suggestions. I would also like to thank Prof. Chris Sachrajda, Armarjit Soni, Taku Izubuchi for instructions and help on all kinds of problems. I greatly appreciate Ziyuan Bai's work on checking part of my calculation.

I enjoyed working and benefited a lot from all my colleagues: Tom Blum, Luchang Jin, Chris Kelly, Christoph Lehner, Qi Liu, Zhongjie Lin, Greg McGlynn, Hantao Yin, Daiqian Zhang, I would like to thank all of them for the discussions and help.

Chapter 1

Introduction

The standard model of particle physics is a theory using quantum filed theory to describe the dynamics of subatomic particles. The current formulation of the standard model was finalized in the 1970s. Then the discoveries of the $W^{\pm}$ and Z boson, the top quark, the tau neutrino and finally the Higgs boson all confirm the correctness of the standard model. Because the standard model can explain a wide variety of experimental results, it is considered to be a theory of almost everything.

There are three fundamental interactions in the standard model: electromagnetic (EM), weak and strong interactions. We believe that we have fully understood the EM interaction with the theory of quantum electrodynamics (QED), which has withstood some of the most stringent experimental tests in the history of physics. The excellent agreement between the experimental and theoretical values of the anomalous magnetic dipole moment of the electron is considered to be one of the most significant triumphs of physics during the last century.

The strong interaction is described by quantum chromodynamics (QCD) in the standard model. Although QCD is widely believed to be well understood, its analytic treatment at low energy is extremely difficult if not entirely impossible. Perturbation theory calculation fail at an energy scale of Λ_{QCD}, because the coupling constant becomes equal to or larger than

$O(1)$. Other analytical methods such as chiral perturbation theory are also unsatisfactory. These methods are usually not from first principles and require input parameters other than the parameters of the standard model. More importantly, these methods are usually not able to give precise enough predictions for low energy QCD effects.

The weak interaction is the least understood part of the standard model and is considered to be the place where new physics may be discovered. However, various tests of the weak interaction have been done and no significant discrepancy from the standard model has been found. Among the ongoing lattice calculations, the unitarity of the Cabibbo-Kobayashi-Maskawa (CKM) matrix and both the direct and indirect CP violation parameters may be the most interesting tests. These calculations are usually quite difficult. The weak interaction itself can be treated precisely with perturbation theory. However, many interesting weak interaction processes involve mesons and baryons. A full calculation of such process requires inputs from non-pertrubative QCD calculation.

Lattice QCD is the only known method to provide first-principal calculation of non-perturbative QCD effects in electroweak processes. Due to both theoretical and practical reasons, we usually don't simulate weak interaction processes directly on lattice. The stanard approach is to split the energy scale into high ($> \mu$) and low ($< \mu$). We start from the full standard model and integrate out the $W^{\pm}$ and Z bosons. The weak interaction process will be described by an effective Hamiltonian. We use perturbation theory to run the Wilson coefficients from the W meson mass scale to a lower scale μ. The physics above the scale μ is encoded in the Wilson coefficients of the effective Hamiltonian. The low-energy part of the weak matrix element is a pure non-perturbative QCD problem and can be solved using lattice QCD techniques.

At the early stage of lattice QCD, due to the limitation of computing resources, most of the lattice calculations were done at unphysical kinematics without dynamical fermions (quenched approximation). During the last few decades, the computing power of supercom-

puters has been improved by a few order of magnitudes. The development of the lattice algorithms has boosted the efficiency tremendously. Simulations at physical kinematics with full dynamical fermions became possible in recent years. Actually, the measurement of the standard quantities such as f_K, f_π have reached sub-percent level precision. At such precision, isospin breaking effects and QED corrections become important. Many efforts and computing resources are invested to further improve the precision of these measurements. However, we should be aware that there are some other more interesting quantities that can be calculated with lattice QCD. Among many such quantities, the $K_L - K_S$ mass difference is one of the most interesting ones.

The kaon mass difference ΔM_K with a value of $3.483(6) \times 10^{-12}$ MeV [4] plays a very important role in the history of particle physics. It led to the prediction of charm quark fifty years ago [5, 6, 7]. This extremely small mass difference is believed to arise from K^0-$\overline{K}^0$ mixing via second-order weak interaction. However, because it arises from an amplitude in which strangeness changes by two units, this is a promising quantity to reveal new phenomena which lie outside the standard model. A closely related quantity is the indirect CP violation parameter ϵ_K, which arises from the same process. The experimental value of ΔM_K and ϵ_K are both known very accurately, making the precise calculation of ΔM_K and ϵ_K an important challenge.

In perturbation theory calculations, the standard model contribution to ΔM_K is separately into short distance and long distance parts. The short distance part receives most contributions from momenta on the order of the charm quark mass. As pointed out in the recent next-to-next-to-leading-order (NNLO) calculation [8], the NNLO terms are as large as 36% of the leading order (LO) and next-to-leading order (NLO) terms, raising doubts about the convergence of the QCD perturbation series at this energy scale. The long distance part of ΔM_K can receive contributions from distance as large as $1/m_\pi$. So far there is no result with controlled uncertainty available since this is highly non-perturbative. How-

ever, an estimation given by Donoghue *et. al.* [9] suggest that there can be sizable long distance contributions. The calculation of ϵ_K is under much better control and the largest contribution involves momenta on the scale of top quark mass. However, the same NNLO difficulties in the predicting the charm quark contribution to ϵ_K enters at the 8% level. In addition, the long distance part of ϵ_K, which is estimated to be 3.6% by Buras *et al.* [10], also becomes increasing interesting and non-perturbative methods are required for a more reliable calculation.

It is customary when discussing the K_L-K_S mass difference to follow the convention of referring to distance scales at or below $1/m_c$ as short distance and those larger than $1/m_c$ as long distance. We will follow this convention here. However, since the perturbation theory calculation of ΔM_K at the charm quark mass scale converges badly, the inverse charm quark mass represents a somewhat large distance to act as a boundary between the short and long distance regions. Non-perturbative methods are needed for the proper treatment of short distance contributions at the charm quark mass scale and that it may be better to adopt a shorter distance demarcation between short and long distances in the future.

Here we propose a method to compute $K_L - K_S$ mass difference including the long distance effects on a Euclidean lattice [11]. We devise an Euclidean-space amplitude which can be evaluated in lattice QCD and which contains the second-order mass difference of interest. As explained in the following chapters, we perform a second-order integration of the product of two first-order weak Hamiltonians in a given space-time volume. The integration sums the contribution to the mass difference from all possible intermediate states.

A generalization of the Lellouch-Luscher method [12] is used to correct potentially large finite-volume effects coming from the two-pion state which can be degenerate with the kaon and the associated principal part appearing in the infinite volume integral over intermediate states [11]. This is an important part of this proposal. However, in the kinematic region studied in this paper, we are unable to resolve the two-pion intermediate state signal from

statistical fluctuations, so this last piece cannot be studied numerically. We therefore only give a theoretical discussion in this thesis.

Our first preliminary work is performed on a 2+1 flavor, domain wall fermion (DWF), $16^3 \times 32 \times 16$ lattice ensemble with a 421 MeV pion. We drop the double penguin diagrams and the disconnected diagrams in this first exploratory calculation. Although this is a non-unitary calculation at unphysical kinematics, the main purpose of this work is to show that the second order weak process can be evaluated using lattice methods. Then a full calculation is performed on a $24^3 \times 64 \times 16$ ensemble with a 330 Mev pion. Although the inclusion of the disconnected diagrams increased the noise substantially, we are able to get a good statistical error with more sophisticated measuring techniques.

This work is organized as follows. In Chapter 2, we review the $K^0 - \overline{K}^0$ mixing in the standard model and give a brief summary of the perturbation theory calculation. Chapter 3 summarizes all the building blocks of a lattice QCD calculation of the $K_L - K_S$ mass difference; In Chapter 4, we discuss the lattice measurement methods used in our calculation; Chapter 5 and 6 give the results from a $16^3 \times 32 \times 16$ lattice ensemble and a $24^3 \times 64 \times 16$ lattice ensemble respectively. In Chapter 7, we discuss our results and future prospects.

Chapter 2

Kaon Mixing in the Standard Model

In this chapter, we will discuss the neutral kaon mixing in the Standard Model. We start from basic quantum mechanics and give the standard formalism for the $K_L - K_S$ mass difference ΔM_K and the indirect CP violation parameter ϵ_K. After that we explain the perturbation theory calculation of $K^0 - \overline{K}^0$ mixing in the Standard Model. Then we summarized the NNLO results and discuss the limitations of these calculations.

2.1 $K^0 - \overline{K}^0$ Mixing

One of the most interesting features of the neutral K mesons is the mixing of the K^0 and $\overline{K}^0$. The K^0 meson with strangeness $+1$ and the $\overline{K}^0$ meson with strangeness -1 are eigenstates of the strong interaction Hamiltonian. They are antiparticle of each other and have identical mass. However, the weak interactions do not conserve strangeness and cause the mixing of the K^0 and $\overline{K}^0$. The strangeness of these two mesons are differ by two. Hence the $K^0 - \overline{K}^0$ mixing is a second order weak process and the amplitude of this process are extremely small.

The K^0 and $\overline{K}^0$ state are charge conjugates of each other

$$C|K^0\rangle = |\overline{K}^0\rangle, \quad C|K^0\rangle = |\overline{K}^0\rangle. \tag{2.1}$$

Here we use the most natural phase convention. Because the kaons are pseudoscalar mesons, we have

$$CP|K^0\rangle = -|\overline{K}^0\rangle, \quad CP|K^0\rangle = -|\overline{K}^0\rangle. \tag{2.2}$$

It is straightforward to construct eigenstates of CP

$$K_{1/2} = \frac{K^0 \mp \overline{K}^0}{\sqrt{2}}. \tag{2.3}$$

Here K_1 is CP even state and K_2 is CP odd state. If we neglect the effects of CP violation, which are at the 0.1% level, these two states are also mass eigenstates. From the experimental observation, we know the life times of these two states differ by a more than a factor 10^3. The explanation for this difference is also straightforward. The CP even state K_1 can decay to $\pi\pi$ states which have a much larger phase space than the three-pion states. So K_1 decays much faster than K_2 and has a much shorter life time,

$$K_S \approx K_1, \quad K_L \approx K_2. \tag{2.4}$$

This equation is exact only when CP is conserved. If CP violation is included, the $K_{S/L}$ mesons will be combinations of the $K_{1/2}$ states. The exact solution can be easily determined by using second order perturbation theory. The time evolution of the $K^0 - \overline{K}^0$ system is given by

$$i\frac{d}{dt}\begin{pmatrix} K^0(t) \\ \overline{K}^0(t) \end{pmatrix} = (M - \frac{i}{2}\Gamma)\begin{pmatrix} K^0(t) \\ \overline{K}^0(t) \end{pmatrix}, \tag{2.5}$$

where the matrices M and Γ are given by

$$M_{ij} = m_K\delta_{ij} + \mathcal{P}\sum_n \frac{\langle i|H_W|n\rangle\langle n|H_W|j\rangle}{m_K - E_n} \tag{2.6}$$

$$\Gamma_{ij} = 2\pi\sum_n \langle i|H_W|n\rangle\langle n|H_W|j\rangle\delta(m_K - E_n), \tag{2.7}$$

where the indices i and j take the values 0 and $\overline{0}$, H_W is the $\Delta S = 1$ weak effective Hamiltonian and $\mathcal{P}$ indicates that the principal part should be take when an integral with a vanishing energy denominator is encountered.

We will assume CPT symmetry which requires that $M_{00} = M_{\overline{0}\overline{0}}$ and $\Gamma_{00} = \Gamma_{\overline{0}\overline{0}}$. Because the M and Γ matrices are hermitian, we have $M_{0\overline{0}} = M^*_{\overline{0}0}$ and $\Gamma_{0\overline{0}} = \Gamma^*_{\overline{0}0}$ and the diagonal terms are real. If we also assume CP symmetry, $M_{0\overline{0}}$ and $\Gamma_{0\overline{0}}$ are both real. With all these constraints, we can rewrite M and Γ matrices

$$M - \frac{i}{2}\Gamma = \begin{pmatrix} A & p^2 \\ q^2 & A \end{pmatrix}, \tag{2.8}$$

where p and q are complex numbers which differ slightly due to CP violation. This matrix can be diagonalized and the difference in eigenvalues is given by

$$\begin{aligned} 2pq &= (m_L - m_S) - \frac{i}{2}(\Gamma_L - \Gamma_S) \\ &= 2\sqrt{(M_{0\overline{0}} - \frac{i}{2}\Gamma_{0\overline{0}})(M^*_{0\overline{0}} - \frac{i}{2}\Gamma^*_{0\overline{0}})}. \end{aligned} \tag{2.9}$$

If we neglect the effects of CP violation, the mass difference is given by

$$m_L - m_S = 2M_{0\overline{0}} = 2\mathcal{P}\sum_n \frac{\langle\overline{0}|H_W|n\rangle\langle n|H_W|0\rangle}{m_K - E_n}, \tag{2.10}$$

and the mass eigenstates are just $K_{1/2}$ states.

After the inclusion of CP violation effect, the new exact eigenstates are given by

$$\begin{aligned} |K_L\rangle &= \frac{|K_2\rangle + \overline{\epsilon}|K_1\rangle}{\sqrt{1 + |\overline{\epsilon}|^2}} \\ |K_S\rangle &= \frac{|K_1\rangle + \overline{\epsilon}|K_2\rangle}{\sqrt{1 + |\overline{\epsilon}|^2}}, \end{aligned} \tag{2.11}$$

where the quantity $\overline{\epsilon}$ is defined by

$$\overline{\epsilon} = \frac{p - q}{p + q} \approx \frac{i}{2}\frac{\mathrm{Im}M_{0\overline{0}} - \frac{i}{2}\mathrm{Im}\Gamma_{0\overline{0}}}{\mathrm{Re}M_{0\overline{0}} - \frac{i}{2}\mathrm{Re}\Gamma_{0\overline{0}}}, \tag{2.12}$$

where the final approximation is valid if the CP violation effects are small.

The quantity $\overline{\epsilon}$ gives the size of the CP even component in the K_L meson. So $\overline{\epsilon}$ is closely related to the indirect CP violation parameter ϵ_K, defined in terms of the $K \to \pi\pi$ amplitudes,

$$\epsilon_K = \frac{2\eta_{+-} + \eta_{00}}{3}, \quad \eta_{ij} = \frac{\mathcal{A}(K_L \to \pi^i\pi^j)}{\mathcal{A}(K_S \to \pi^i\pi^j)}. \tag{2.13}$$

Following the usual convention, we introduce the definite isospin amplitudes A_0 and A_2,

$$A_I e^{\delta_I} = A(K \to \pi\pi(I)), \tag{2.14}$$

where δ^I is the strong phase from the π-π interaction. The $\pi^+\pi^-$ and $\pi^0\pi^0$ states are related to the states with definite isospin by

$$\begin{aligned} |\pi^+\pi^-\rangle &= \frac{1}{\sqrt{3}}\left(|\pi\pi(I=2)\rangle + \sqrt{2}|\pi\pi(I=0)\rangle\right) \\ |\pi^0\pi^0\rangle &= \frac{1}{\sqrt{3}}\left(-\sqrt{2}|\pi\pi(I=2)\rangle + |\pi\pi(I=0)\rangle\right). \end{aligned} \tag{2.15}$$

Substituting Eq. 2.15 and 2.11 into Eq. 2.13 and dropping all the $(\mathrm{Re}A_2/\mathrm{Re}A_0)^2$ term gives

$$\epsilon_K = \bar{\epsilon} + i\xi, \tag{2.16}$$

where ξ is the weak phase of the $K^0 \to (\pi\pi)_{I=0}$ amplitude,

$$\xi = \frac{\mathrm{Im}A_0}{\mathrm{Re}A_0} \tag{2.17}$$

In order to evaluate ϵ_K more conveniently, we can simplify the expression by defining the superweak phase

$$\phi_\epsilon = \arctan\left(\frac{-2\mathrm{Re}M_{0\bar{0}}}{\Gamma_{0\bar{0}}}\right). \tag{2.18}$$

We can use the observation that the $|(\pi\pi)_{I=0}\rangle$ states dominate the sum in Eq. 2.7, which implies

$$\frac{\mathrm{Im}\Gamma_{0\bar{0}}}{\mathrm{Re}\Gamma_{0\bar{0}}} \approx -2\frac{\mathrm{Im}A_0}{\mathrm{Re}A_0} = -2\xi. \tag{2.19}$$

Combining all these together we arrive

$$\epsilon_K = e^{i\phi_\epsilon}\sin\phi_\epsilon\left(\frac{\mathrm{Im}M_{0\bar{0}}}{\Delta m_K} + \xi\right) \tag{2.20}$$

The experimental value for ϕ_ϵ is close to 45°. The ξ term contributes to ϵ_K at the few percent level.

2.2 Perturbation Theory Calculation for ΔM_K

In the standard model, the largest contribution to $K^0 - \overline{K}^0$ mixing comes from the box diagrams as shown in Fig. 1. We should notice that there are also contributions from double penguin diagrams and disconnected diagrams. These diagrams will start to contribute if we include higher order QCD effects. It is customary when discussing $K^0 - \overline{K}^0$ mixing to separate the results into the top contribution, the charm contribution and the charm-top contribution. These different contributions come from different choices of two quark propagators in the loop inside the box diagram.

To evaluate the mixing process, we can construct a low energy effective filed theory and encode the short distance physics into the Wilson coefficients in the effective theory. First the W bosons and the top quarks are integrated out. The top component is then described by a local $\Delta S = 2$ effective Hamiltonian. The charm component is described by bilocal operators, *i.e.*, the products of two $\Delta S = 1$ effective Hamiltonians. The charm-top component is more complicated due to the presence of penguin operators. The effective Hamiltonian of the charm-top component contains both bilocal operators and local counterterms. Fig. 2 gives some examples of the contributions from the bilocal operators. Next, these effective Hamiltonians are renormalized at the charm quark scale. Finally, we integrate out the charm quark and renormalize the effective Hamiltonian at a low scale μ. The effective low energy Hamiltonian can be written as:

$$\mathcal{H}_{eff}^{\Delta S=2}(\mu) = \frac{G_F^2}{16\pi^2} M_W^2 \left(\lambda_c^2 \eta_{cc} S_0(x_c) + \lambda_t^2 \lambda_t^2 \eta_{tt} S_0(x_t) + 2\lambda_c \lambda_t \eta_{ct} S_0(x_c, x_t)\right) b(\mu) Q(\mu) + \text{h.c.} \tag{2.21}$$

where $\lambda_i = V_{is}^* V_{id}$ and $x_i = m_i^2/m_W^2$. The basic electroweak loop contributions without QCD

correction are given by the function S_0:

$$S_0(x_c) \approx x_c, \tag{2.22}$$

$$S_0(x_t) = \frac{4x_t - 11x_t^2 + x_t^3}{4(1-x_t)^2} - \frac{3x_t^3 \ln x_t}{2(1-x_t)^3}, \tag{2.23}$$

$$S_0(x_t, x_c) = x_c \left(\ln \frac{x_t}{x_c} - \frac{3x_t}{4(1-x_t)} - \frac{3x_t^2 \ln x_t}{4(1-x_t)^2} \right). \tag{2.24}$$

The first approximation holds when $x_c \ll 1$. Short distance QCD corrections are given by factors η_{cc}, η_{tt}, η_{ct} and $b(\mu)$ in Eq. 2.21. The $\mathcal{H}_{eff}^{\Delta S=2}$ consist of a single local $\Delta S = 2$ four quark operator:

$$Q = (\bar{s}d)_{V-A}(\bar{s}d)_{V-A}. \tag{2.25}$$

The parameter $b(\mu)$ cancels the scale and scheme dependence of $Q(\mu)$. The product $b(\mu)Q(\mu)$ is related to the scale and scheme invariant bag parameter $\hat{B}_K$ by

$$\hat{B}_K = \frac{3}{8} \frac{\langle \overline{K}^0 | b(\mu) Q(\mu) | K^0 \rangle}{F_K^2 m_K^2}, \tag{2.26}$$

where F_K is the kaon decay constant.

Substituting Eq. 2.21 into Eq. 2.10 and 2.20, the mass difference is then given by:

$$\Delta M_K = \frac{G_F^2}{6\pi^2} F_K^2 \hat{B}_K m_K M_W^2 \left(\lambda_c^2 \eta_{cc} S_0(x_c) + \lambda_t^2 \eta_{tt} S_0(x_t) + 2\lambda_c \lambda_t \eta_{ct} S_0(x_c, x_t) \right). \tag{2.27}$$

The indirect CP violation parameter writes

$$|\epsilon_K| = \kappa_\epsilon C_\epsilon \hat{B}^K \mathrm{Im}\lambda_t \{ [\eta_{cc} S_0(x_c) - \eta_{ct} S_0(x_c, x_t)] - \mathrm{Re}\lambda_t \eta_{tt} S_0(x_t) \}, \tag{2.28}$$

where the correction factor κ_ϵ take into account the effects due to $\phi_\epsilon \neq 45^\circ$ and $\xi \neq 0$. The factor C_ϵ is given by

$$C_\epsilon = \frac{G_F^2 F_K^2 m_K M_W^2}{6\sqrt{2}\pi^2 \Delta M_K}. \tag{2.29}$$

The evaluation of the short distance QCD correction factors η_i is the most important part of the perturbation theory calculation. The calculation of η_{tt} is straightforward. We are left with a single operator Q after we integrate out the top quark and the W boson. Then

we renormalize this operator at a low scale μ and we will obtain η_{tt}. The calculation of η_{cc} is more complicated. We need to deal with the product of two $\Delta S = 1$ operators. At the charm quark threshold, we need to match this bi-local operator with the local operator Q. Top-charm contribution η_{ct} is the most difficult one due to the presence of penguin operators.

The top contribution is known at NLO:

$$\eta_{tt} = 0.5765(65) \tag{2.30}$$

The result is very precise since the perturbation theory should work well at the scale of top quark mass. Higher order calculation for η_{tt} is unnecessary. The charm-top contribution has been evaluated to NNLO:

$$\eta_{ct} = 0.496(47). \tag{2.31}$$

ΔM_K receives most of the contribution from the momenta at the scale of charm quark mass. Unfortunately, the factor η_{cc} has the largest uncertainty among all the three factors. The reason is the slow convergence of the perturbation series at the scale of charm quark mass. We can compare the NLO result with the NNLO result. At NLO,

$$\eta_{cc}^{\rm NLO} = 1.38 \pm 0.52_{\mu_c} \pm 0.07_{\mu_W} \pm 0.02_{\alpha_s}. \tag{2.32}$$

We can see that the largest uncertainty comes from the dependence on the charm threshold μ_c. If perturbation theory works well, we would expect a smaller dependence on μ_c at NNLO. However, the size of uncertainty is similar in the NNLO result:

$$\eta_{cc}^{\rm NNLO} = 1.86 \pm 0.53_{\mu_c} \pm 0.07_{\mu_W} \pm 0.06_{\alpha_s}. \tag{2.33}$$

Also, there is a large positive shift of 36% in the central value. The NNLO result suggest that the perturbation series converges very badly. Brod *et.al.* suggest we can use the size of NNLO correction as the theoretical uncertainty, which leads to a 36% total uncertainty:

$$\eta_{cc} = 1.87 \pm 0.76. \tag{2.34}$$

Combining all these ingredients together gives:

$$\Delta M_K = 3.1(1.2) \times 10^{-12}\text{MeV} \tag{2.35}$$

$$|\epsilon_K| = 1.81(28) \times 10^{-3}. \tag{2.36}$$

These results can be compared with the precisely measured experimental values:

$$\Delta M_K^{\text{exp}} = 3.483(6) \times 10^{-12}\text{MeV} \tag{2.37}$$

$$|\epsilon_K|^{\text{exp}} = 2.228(11) \times 10^{-3}. \tag{2.38}$$

The uncetanity of the standard model predictions are too large so that these comparisons are not very conclusive.

It is obvious that the perturbation theory calculation of ΔM_K is not satisfactory. First, the perturbation series at the scale of charm quark mass converges poorly. A higher order perturbation theory calculation will be extremely complicated if not impossible. Also, a next-to-next-to-next-leading-order (NNNLO) calculation seems meaningless at this point because the series may not converge. To give a more reliable prediction for ΔM_K, we need to treat the charm quark non-perturbatively. Another limitation of the perturbation theory calculation is that the long distance contributions are completely neglected. Such contributions are schematically described in Fig. 3. The separation between the two $\Delta S = 1$ weak Hamiltonian can be as large as $1/m_\pi$. A local $\Delta S = 2$ effective Hamiltonian can't capture these effects. At present, there are no available results with controlled error because the long distance contributions are highly non-perturbative. However, an estimation by Donoghue *et.al.* [9] suggest that the long distance contributions may be sizable.

The calculation of ϵ_K is under much better control, because the largest contribution involves momenta at the scale of the top quark mass, where perturbation theory should be reliable. However, the discrepancy between the NNLO result and the experimental value is about 1.5σ, making a more precise calculation of ϵ_K extremely interesting. The largest source

of uncertainty is the CKM matrix element V_{cb}. The second largest one is η_{cc}, which contribute about 8% uncertainty to ϵ_K. A more precise calculation will require non-perturbative method. In addition the long distance contribution to ϵ_K is estimated to be -3.6% by Buras *et al.* [10], again suggesting the need for a reliable, non-perturbative method.

Chapter 3

K^0-$\overline{K}$ mixing from Lattice QCD

In this chapter we will discuss the details about evaluating K^0-$\overline{K}^0$ mixing process with lattice QCD. In Section 3.1, we discuss the second order weak amplitude and the method to extract ΔM_K from such amplitude. In Section 3.2, we give all the Wick contractions in the lattice calculation. Section 3.3 will discuss the calculation of the Wilson coefficients. In Section 3.4, we will discuss the renormalization of the $\Delta S = 1$ four-quark operators. In Section 3.5, we give a theoretical discussion on the subtraction of short distance effects on lattice. In Section 3.6, we discuss the finite volume effects in the ΔM_K calculation.

3.1 Second Order Weak Amplitude

Lattice QCD has been used to calculate K^0-$\overline{K}^0$ mixing for a long time. However, all the previous calculation use a local $\Delta S = 2$ effective Hamiltonian defined in Eq. 3.1. The purpose of these calculations is to evaluate the $\hat{B}_K$ parameter. As we have discussed at the end of Section. 2.2, these calculations suffer from non-perturbative effects at the scale of charm quark mass and uncontrolled long distance effects. In order to resolve these difficulties, we need to evaluate the second order weak process directly using lattice QCD. The typical inverse lattice spacing is a few GeVs. So we can't directly put the W boson and the top

quark on the lattice. These heavy degree of freedoms are integrated out, leaving us with a product of two $\Delta S = 1$ effective weak Hamiltonians. The effective Hamiltonian in this process is given by:

$$H_W^{\Delta S=1} = \frac{G_F}{\sqrt{2}} \sum_{q,q'=u,c} V_{qd} V_{q's}^* (C_1 Q_1^{qq'} + C_2 Q_2^{qq'}) \tag{3.1}$$

where q and q' are each on of the u and c quarks, V_{qd} and $V_{q's}$ are the CKM matrxi elements, C_1 and C_2 are the Wilson coefficients and Q_i are the current-current operators, which are defined as:

$$Q_1^{qq\prime} = (\bar{s}_i q'_j)_{V-A} (\bar{q}_j d_i)_{V-A} \tag{3.2}$$

$$Q_2^{qq\prime} = (\bar{s}_i q'_i)_{V-A} (\bar{q}_j d_j)_{V-A} \,, \tag{3.3}$$

where i, j are color indices and the spinor indices are contracted within each pair of brackets.

In order to evaluate the K^0-$\overline{K}^0$ process, the most natural thoughts will be calculating the four point correlators:

$$G(t_f, t_2, t_1, t_i) = \langle 0 | T \left\{ \overline{K}^0(t_f) H_W(t_2) H_W(t_1) \overline{K}^0(t_i) \right\} \rangle, \tag{3.4}$$

where T is the usual time ordering operator. Here the initial K^0 states is generated by kaon source $\overline{K}^0(t_i)$ at time t_i and the final $\overline{K}^0$ state is destroyed by the anti-kaon sink $\overline{K}^0(t_f)$. The two $\Delta S = 1$ effective Hamiltonian acts at the time t_1 and t_2. Assuming that the time separations $t_f - t_k$ and $t_k - t_i$ for $k = 1$ and 2 are sufficiently large that kaon interpolating operator will project onto the K^0 and $\overline{K}^0$ initial and final states and after inserting a complete set of energy eigenstates, we find:

$$G(t_f, t_2, t_1, t_i) = N_K^2 e^{-M_K(t_f - t_i)} \sum_n \langle \overline{K}^0 | H_W | n \rangle \langle n | H_W | K^0 \rangle e^{-(E_n - M_K)|t_2 - t_1|}, \tag{3.5}$$

where N_k is the normalization factor for the kaon interpolation operator and $|n\rangle$ are the eigenstates of the QCD Hamiltonian. If we fix the time t_i and t_f, then this correlator depends only on the time separation between the two Hamiltonians $|t_2 - t_1|$. We will refer to

$G(t_f, t_2, t_1, t_i)$ as the unintegrated correlator. The unintegrated correlator receives contributions from all possible intermediate states. The terms in this sum over intermediate states show exponentially decreasing or increasing behavior with increasing separation $|t_2 - t_1|$ depending on whether E_n lies above or below M_K.

In order to get a lattice approximation to the mass difference defined in Eq. 2.10, we can start from the unintegrated correlator and integrate it over a time interval. There are a few possible choices of the integration time interval. We will discuss our choice of method first and discuss the alternatives later. We choose an integration time interval $[t_a, t_b]$ and obtain:

$$\mathcal{A} = \frac{1}{2} \sum_{t_2=t_a}^{t_b} \sum_{t_1=t_a}^{t_b} \langle 0 | T \left\{ \overline{K^0}(t_f) H_W(t_2) H_W(t_1) \overline{K^0}(t_i) \right\} | 0 \rangle, \tag{3.6}$$

where $t_a - t_i$ and $t_f - t_b$ should be sufficiently large that the kaon interpolating operators to project onto kaon states. We will refer to this correlator as the integrated correlator. The integrated correlator is represented schematically in Fig. 4. After inserting a sum over intermediate states and summing explicitly over t_1 and t_2 in the interval $[t_a, t_b]$ one obtains:

$$\begin{aligned} \mathcal{A} =& N_K^2 e^{-M_K(t_f - t_i)} \Bigg\{ \sum_{n \neq n_0} \frac{\langle \overline{K}^0 | H_W | n \rangle \langle n | H_W | K^0 \rangle}{M_K - E_n} \left(-T + \frac{e^{(M_K - E_n)T} - 1}{M_K - E_n} \right) \\ &+ \frac{1}{2} \langle \overline{K}^0 | H_W | n_0 \rangle \langle n_0 | H_W | K^0 \rangle T^2 \Bigg\}. \end{aligned} \tag{3.7}$$

Here $T = t_b - t_a + 1$ and the sum includes all possible intermediate states except a possible state $|n_0\rangle$ which is degenerate with the kaon, $E_{n_0} = M_K$. Such a state will exist only if the spatial volume of the lattice is adjusted to some specific value. The contribution from such a degenerate state appears separately as the final term on the right hand side of this equation. In order to obtain Eq. 3.7, we neglect all the $O((E_n a)^2)$ terms.

The coefficient of the term which is proportional to T in Eq. 3.7 gives the finite-volume approximation to ΔM_K up to some normalization factors:

$$\Delta M_K^{\mathrm{FV}} = 2 \sum_{n \neq n_0} \frac{\langle \overline{K}^0 | H_W | n \rangle \langle n | H_W | K^0 \rangle}{M_K - E_n}. \tag{3.8}$$

The other terms in Eq. 3.7 can be classified into four categories according to their dependence on T:

i) The term independent of T within the large parentheses. This constant does not affect our determination of the mass difference from $\mathcal{A}$.

ii) Terms exponentially decreasing as T increases coming from states $|n\rangle$ with $E_n > M_K$. These terms are negligible for sufficiently large T.

iii) Terms exponentially increasing as T increases coming from states $|n\rangle$ with $E_n < M_K$. These will be the largest contributions when T is large and must be removed as discussed in the paragraph below.

iv) The final term proportional to T^2 coming from states degenerate with the kaon. As discussed below, this term must be identified and removed in order to relate the finite- and infinite-volume expressions for ΔM_K following the method of Ref. [11].

The exponentially growing terms, introduced in item iii) above, pose a significant challenge. Fortunately, the two leading terms corresponding to the vacuum and single pion states can be computed separately and removed. The matrix elements $\langle \pi^0 | H_W | K^0 \rangle$ and $\langle 0 | H_W | K^0 \rangle$ can be obtained from the three-point and two-point correlation functions:

$$\langle 0 | \pi^0(t_\pi) H_W(t_O) K^{0\dagger}(t_K) | 0 \rangle = N_\pi N_K e^{-m_\pi (t_\pi - t_O)} e^{-m_K (t_O - t_K)} \langle \pi^0 | H_W | K^0 \rangle \tag{3.9}$$

$$\langle 0 | H_W(t_O) K^{0\dagger}(t_K) | 0 \rangle = N_K e^{-m_K (t_O - t_K)} \langle 0 | H_W | K^0 \rangle. \tag{3.10}$$

Here we assume that the time separations $t_\pi - t_O$ and $t_O - t_K$ are sufficiently large that only the pion and kaon states contribute. The matrix element can be obtained from the ratio of the correlation functions in Eqs. 3.9 and 3.10 with the corresponding two-point correlators of the operators K^0 and π^0 to remove the t_O dependence. This will be discussed in more details in Section. 4.3. Then we can use these matrix elements to remove the single-pion and vaccum exponential growing terms in Eq. 3.7.

A second approach to remove these two unwanted exponentially growing terms exploits the lattice Ward identities to add to H_W terms proportional to the scalar and pseudo-scalar densities, $\bar{s}d$ and $\bar{s}\gamma^5 d$. The new Hamiltonian is given by:

$$H'_W = H_W + c_1\bar{s}d + c_2\bar{s}\gamma^5 d. \tag{3.11}$$

The coefficients c_1 and c_2 are chosen to eliminate the single-pion and vacuum intermediate states:

$$\langle\pi^0|H_W + c_1\bar{s}d|K^0\rangle = 0, \tag{3.12}$$

$$\langle 0|H_W + c_2\bar{s}\gamma^5 d|K^0\rangle = 0. \tag{3.13}$$

The coefficient is then given by:

$$c_\pi = -\frac{\langle\pi^0|H_W|K^0\rangle}{\langle\pi^0|\bar{s}d|K^0\rangle}, \tag{3.14}$$

$$c_{\text{vac}} = -\frac{\langle 0|H_W|K^0\rangle}{\langle 0|\bar{s}\gamma^5 d|K^0\rangle}. \tag{3.15}$$

Since these two densities can be written as the divergence of the vector and axial currents respectively, they cannot contribute to the mass difference given in 3.8. This approach is similar to the subtraction in the previous paragraph, but instead of removing only the exponentially growing term in Eq. 3.7, such an addition will remove all single pion and vacuum contributions from that equation, including their appearance in the sum over intermediate states $|n\rangle$.

A third approach is multi-parameter fitting. We can add these exponential terms into the fitting ansatz. In principal, we can get the mass difference from a multi-parameter fit. However, this approach can be very noisy, especially when the exponentially increasing term is significantly larger than the linear term.

Two-pion states with energies below M_K may also exist and, if present, must be explicitly identified and removed. For the kinematics studies in this thesis, the only π-π state with

an energy possibly below M_K is the state with two pions at rest. In our calculation on 16^3 lattice with a 412 MeV pion mass, we study the contribution of this state as the kaon mass varied. However, we are not able to identify a clear two-pion signal. In our calculation on 24^3 with a 330 MeV pion mass and a 575 MeV kaon mass, there is no π-π state lighter than the kaon. Thus we don't need to worry about it. However, in a future calculation in physical kinematics, the two pion states will pose a significant challenge. More studies are needed to solve this problem.

So far all our discussions are based on Eq. 3.6. However, there are alternative choices for the integrated correlators. For example, a simpler alternative integrates the two weak operators over the full time interval $[t_i, t_f]$ between the kaon source and sink instead of the the restricted interval $[t_a, t_b]$ used here:

$$\mathcal{A}' = \frac{1}{2} \sum_{t_2=t_i}^{t_f} \sum_{t_1=t_i}^{t_f} \langle 0|T \left\{ \overline{K^0}(t_f) H_W(t_2) H_W(t_1) \overline{K^0}(t_i) \right\} |0\rangle. \tag{3.16}$$

After inserting a sum over intermediate states, the integrated correlator will contain the term of interest, $N_K^2 \Delta M_K (t_f - t_i) e^{-M_K(t_f - t_i)}/2$. The mass difference can be obtained by varying $t_f - t_i$, which is the separation of the two kaons. However, this method has two disadvantages when compared to the method which we use. The first is the need to vary the location of the source and sink positions of the kaons if the dependence on $t_t - t_i$ is to be identified. For the method which we use we are able to work with fixed t_f and t_i and simply vary the interval $[t_a, t_b]$ over which the weak operator insertions are integrated. Having fixed kaon source and sink locations largely reduced the number of quark propagators which must be evaluated in the calculation presented here. Thus the computing cost of this method will be much larger than the method we choose.

A second, far more serious difficulty with the expression in Eq. 3.16 arises from the analogue of the exponentially increasing terms given in Eq. 3.7. In that previous case the coefficient of an exponentially increasing term coming from a QCD energy eigenstate $|n\rangle$ with energy E_n lower than M_K is a simple matrix element of H_W between that state and a

physical kaon state, a quantity easily determined in a separate lattice calculation. However, for the correlator defined in Eq. 3.16 these unwanted terms come with coefficients that are very difficult to determine and hence cannot be easily removed. More specifically, in the correlators we will have terms like:

$$\sum_{n'',n'\neq n} \langle 0|K_\alpha|n''\rangle \frac{\langle n''|H_W|n\rangle}{E_n - E_{n''}} e^{-E_n(t_f-t_i)} \frac{\langle n|H_W|n'\rangle}{E_n - E_{n'}} \langle n'|K_\alpha|0\rangle. \tag{3.17}$$

Here $|n'\rangle$ and $|n^{''}\rangle$ are possible excited kaon states. A term with energy $E_n < M_K$ which must be removed has a complicated coefficient given by a sum over matrix elements of H_W between that state $|n\rangle$ and a series of excited states $|n'\rangle$, a combination apparently inaccessible to a lattice QCD calculation. Thus, a separate determination of the terms to be subtracted appears very difficult. Note this second difficulty only arises when there exist states of lower energy than that of the state being studied, in our case the kaon. All of these unwanted terms with $E_n > m_K$ will not contribute for sufficiently large $t_f - t_i$.

Finally, a third alternative that we can examine integrates the product of the two weak operators $H_W(t_2)H_W(t_1)$ over the entire time interval $[0, T_{\rm tot}]$:

$$\mathcal{A}'' = \frac{1}{2}\sum_{t_2=0}^{T_{\rm tot}} \sum_{t_1=0}^{T_{\rm tot}} \langle 0|T\left\{\overline{K^0}(t_f)H_W(t_2)H_W(t_1)\overline{K^0}(t_i)\right\}|0\rangle. \tag{3.18}$$

Again the mass difference is obtained by varying $t_f - t_i$. $\mathcal{A}''$ will also suffer from excited kaon states just like $\mathcal{A}'$. Further more, there will be severe around-the-world effects in $\mathcal{A}^{''}$. This problem may be solved by evaluating the propagator for both periodic and anti-periodic boundary conditions in time direction. However, this will require more computational cost.

3.2 Diagrams Needed for the Lattice Calculation

The four-point correlators Eq. 3.4 are given by combinations of Wick contractions on lattice. After inserting Eq. 3.1 into Eq. 3.4, we will get all the possible contractions. To simplify the

problem, we can apply the Fierz transformation to the Q_1 operators in Eq. 3.1:

$$Q_1^{qq'} = (\bar{s}_i q'_j)_{V-A}(\bar{q}_j d_i)_{V-A} = (\bar{s}_i d_i)_{V-A}(\bar{q}_j q'_j)_{V-A}. \tag{3.19}$$

The Fierz transformation is exact on lattice. After the transformation, the operator Q_1 becomes explicitly color unmixed. This will guarantee the color and spin will flow in the same way in all the diagrams. We list all the possible contractions contributing to the four-point correlators in Figs. 5-8. There are in total 16 diagrams which are labeled by circled numbers and we categorize them into four types according to their topology. There are six quark propagators in each diagram. Four of these propagators are connected to the kaon wall sources while two propagators connect one of the weak operators to the other or each weak operator to itself. We call these two quark propagators internal propagators. In a four flavor theory, the flavor of the internal quark propagators can be either up or charm. We therefore have four different combinations for each diagram: uu, cc, uc and cu. We use these labels in a subscript to denote the flavor of the two internal quark propagators. For example, the first diagram with two internal up quark propagators is represented by $①_{uu}$, and the GIM cancellation occurs in the combination:

$$①_{\text{GIM}} = ①_{uu} + ①_{cc} - ①_{uc} - ①_{cu}. \tag{3.20}$$

Because of the arrangement of quark flavors and spin contractions in the operators $Q_1^{qq'}$ and $Q_2^{qq'}$ the spin indices on quark fields which carry the same charge are always contracted with an interposed $\gamma_\mu(1-\gamma_5)$ spin matrix. Therefore, the pattern of spin contractions need not be represented in Figs. 5-8. Instead, the separation of each four-quark vertex into two pairs of two quark vertices shown in those figures indicates the pattern of color contractions. Thus, when two quark lines carrying the same charge are joined in those figures that arrangement of spin and color contractions is the same and the operator $Q_1^{qq'}$ appears at that vertex. If lines with different charge are joined, it is the operator $Q_2^{qq'}$ that appears.

All the correlation functions are given by combinations of these contractions. For example,

$$\langle \overline{K^0}(t_f) Q_1^{uu}(t_2) Q_1^{uu}(t_1) \overline{K^0}(t_i) \rangle = ①_{uu} - ⑤_{uu} + ⑨_{uu} - ⑬_{uu}, \tag{3.21}$$

where the contractions identified by circled numbers do not carry the minus sign coming from the number of fermion loops. Instead these minus signs appear explicitly in Eq. 3.21. Since our definition of the kaon interpolation operators is $K^0 = i(\bar{d}\gamma_5 s)$, there will be a minus sign, $i^2 = -1$, coming from two kaon sources. This minus sign is also not included in the contractions.

There are two other possible operator combinations in this calculation:

$$\langle \overline{K^0}(t_f) Q_2^{uu}(t_2) Q_2^{uu}(t_1) \overline{K^0}(t_i) \rangle = ④_{uu} - ⑧_{uu} + ⑫_{uu} - ⑯_{uu}, \tag{3.22}$$

$$\begin{aligned} \langle \overline{K^0}(t_f) (Q_1^{uu}(t_2) Q_2^{uu}(t_1) + Q_2^{uu}(t_2) Q_1^{uu}(t_1)) \overline{K^0}(t_i) \rangle &= -②_{uu} - ③_{uu} + ⑥_{uu} + ⑦_{uu}, \\ &\quad - ⑩_{uu} - ⑪_{uu} + ⑭_{uu} + ⑮_{uu}. \end{aligned} \tag{3.23}$$

After GIM cancellation, these become:

$$\langle \overline{K^0}(t_f) Q_{11}^{\mathrm{GIM}}(t_2, t_1) \overline{K^0}(t_i) \rangle = ①_{\mathrm{GIM}} - ⑤_{\mathrm{GIM}} + ⑨_{\mathrm{GIM}} - ⑬_{\mathrm{GIM}}, \tag{3.24}$$

$$\langle \overline{K^0}(t_f) Q_{22}^{\mathrm{GIM}}(t_2, t_1) \overline{K^0}(t_i) \rangle = ④_{\mathrm{GIM}} - ⑧_{\mathrm{GIM}} + ⑫_{\mathrm{GIM}} - ⑯_{\mathrm{GIM}}, \tag{3.25}$$

$$\begin{aligned} \langle \overline{K^0}(t_f) \left(Q_{12}^{\mathrm{GIM}}(t_2, t_1) + Q_{21}^{\mathrm{GIM}}(t_2, t_1) \right) \overline{K^0}(t_i) \rangle &= -②_{\mathrm{GIM}} - ③_{\mathrm{GIM}} + ⑥_{\mathrm{GIM}} + ⑦_{\mathrm{GIM}} \\ &\quad - ⑩_{\mathrm{GIM}} - ⑪_{\mathrm{GIM}} + ⑭_{\mathrm{GIM}} + ⑮_{\mathrm{GIM}}. \end{aligned} \tag{3.26}$$

Here the subscript "GIM" under the circles indicates the same combination of internal quark line flavors as is given in Eq. 3.20. The four operator products $Q_{ij}^{\mathrm{GIM}}(t_2, t_1)$ appearing on the left-hand side of Eq. 3.26 are each the appropriate sum of all four combinations of intermediate charm and up quarks:

$$\begin{aligned} Q_{ij}^{\mathrm{GIM}}(t_2, t_1) = Q_i^{uu}(t_2) Q_j^{uu}(t_1) + Q_i^{cc}(t_2) Q_j^{cc}(t_1) \\ - Q_i^{uc}(t_2) Q_j^{cu}(t_1) - Q_i^{cu}(t_2) Q_j^{uc}(t_1) \qquad i, j = 1, 2. \end{aligned} \tag{3.27}$$

As discussed in Sec. 3.1, we can add the scalar density $\bar{s}d$ and the pseudoscalar density $\bar{s}\gamma_5 d$ to the effective weak Hamiltonian. There will be 10 more diagrams after adding these operators. We categorize these new diagrams into two types according their topology. These two types are similar with type 3 and type 4 diagrams respectively. We will name these two new types as type 3-scalar and type 4-scalar. We list all the diagrams in Fig. 9-10.

We also need to evaluate the three-point correlators in order to remove the exponentially increasing terms in the four-point correlators. The diagrams contributing to $K^0 \to \pi^0$ process are listed in Fig. 11. These diagrams are labeled by boxed numbers. The correlators are given by combinations of contractions:

$$\langle 0|\pi^0(t_\pi)Q_1(t)K^{0\dagger}(t_K)|0\rangle = \frac{(\ -\boxed{1}+\boxed{3}\ )}{\sqrt{2}}, \tag{3.28}$$

$$\langle 0|\pi^0(t_\pi)Q_2(t)K^{0\dagger}(t_K)|0\rangle = \frac{(\ \boxed{2}-\boxed{4}\ )}{\sqrt{2}}, \tag{3.29}$$

$$\langle 0|\pi^0(t_\pi)\bar{s}d(t)K^{0\dagger}(t_K)|0\rangle = -\frac{\boxed{5}}{\sqrt{2}}. \tag{3.30}$$

The diagrams contributing to the $K^0 \to 0$ process are listed in Fig. 12. The correlators are given by combinations of contractions:

$$\langle 0|Q_1(t)K^{0\dagger}(t_K)|0\rangle = i\ \boxed{6}, \tag{3.31}$$

$$\langle 0|Q_2(t)K^{0\dagger}(t_K)|0\rangle = -i\ \boxed{7}, \tag{3.32}$$

$$\langle 0|\bar{s}d(t)K^{0\dagger}(t_K)|0\rangle = -i\ \boxed{8}. \tag{3.33}$$

3.3 Wilson Coefficients

The Wilson coefficients in Eq. 3.1 summarize the effect of short-distance contributions. We calculate them using renormalization group improved perturbation theory to the next-to-leading order by following exactly the method and techniques of Ref. [13]. The initial values, renormalization-group evolution and anomalous-dimension matrices, and quark threshold matching matrix are all described in Section IV/B in Ref. [13]. The calculated Wilson

coefficients at the renormalization energy scale $\mu = 2.15$ GeV with 4 flavors are given in the first two columns in Tab. 2. All the standard model parameters used in this calculation are taken from PDG 2010 [4] and are also listed in Tab. 1.

Notice that to obtain the results in Tab. 2, we used the same strategy as in [13]: keep only the leading order terms and discarding the other terms. For example, if we have an expression $(1 + a\alpha_s)(1 + b\alpha_s)$, we should not do the calculation by numerically multiplying the two factors in the two parentheses. Instead, the expression is expanded, and only the term $1 + (a + b)\alpha_s$ is kept and calculated. The initial condition, evolution operator matrix and matching matrix are all functions of α_s. We need to keep the analytical form, expand the expressions, and keep only the $O(1)$ and $O(\alpha_s)$ term at the end. By following this method, we were able to exactly reproduce the results in Ref. [13]. This is the advantage of explicitly dropping all higher order terms: others can recognize exactly what kinds of terms are dropped so the results can be easily reproduced exactly.

3.4 Non-perturbative Renormalization

In order to get ΔM_K in physical units, we need to combine our lattice results with the Wilson coefficients which describe the short-distance physics. However, the Wilson coefficients are calculated in the $\overline{\text{MS}}$ scheme. So we need to convert the operators renormalized in $\overline{\text{MS}}$ scheme into the lattice operators. We will discuss the details of this procedure in this section.

We will consider only the current-current operators defined in Eq. (3.3) which enter the

present calculation. In particular, we are only interested in the operators:

$$\begin{aligned}
\widetilde{Q}_1 &= (\bar{s}_i u_j)_{V-A}(\bar{u}_j d_i)_{V-A} - (\bar{s}_i c_j)_{V-A}(\bar{c}_j d_i)_{V-A}, \\
\widetilde{Q}_2 &= (\bar{s}_i u_i)_{V-A}(\bar{u}_j d_j)_{V-A} - (\bar{s}_i c_i)_{V-A}(\bar{c}_j d_j)_{V-A}, \\
Q_1^{cu} &= (\bar{s}_i u_j)_{V-A}(\bar{c}_j d_i)_{V-A}, \\
Q_2^{cu} &= (\bar{s}_i u_i)_{V-A}(\bar{c}_j d_j)_{V-A}, \\
Q_1^{uc} &= (\bar{s}_i c_j)_{V-A}(\bar{u}_j d_i)_{V-A}, \\
Q_2^{uc} &= (\bar{s}_i c_i)_{V-A}(\bar{u}_j d_j)_{V-A}.
\end{aligned} \tag{3.34}$$

These six operators can be categorized into three groups according to their different flavor structure. Operator mixing will take place within each group. The discussion of operator mixing is simplified if we define a second, equivalent basis:

$$\begin{aligned}
Q_+^X &= Q_1^X + Q_2^X, \\
Q_-^X &= Q_1^X - Q_2^X,
\end{aligned} \tag{3.35}$$

where the label X takes on the three values '$\sim$', cu, uc appearing in Eq. (3.34). Thus, we have three groups of operators $\widetilde{Q}_\pm$, $Q_\pm^{cu}$ and $Q_\pm^{uc}$. The advantage of this basis is that Q_+ belongs to the (84,1) irreducible representation of $SU(4)_L \times SU(4)_R$, while Q_- belongs to the (20,1) representation [14]. Since the renormalization will be carried out in the $SU(4)_L \times SU(4)_R$ symmetric limit of vanishing u, d, s and c quark masses, the operators Q_+ and Q_- will not mix with each other or any other dimension 6 operator. Finally $SU(4)_L \times SU(4)_R$ symmetry requires that the renormalization factors for all operators in the same representation will be identical.

Although the basis in Eq. (3.35) is favored theoretically, we choose to use the basis in Eq. (3.34) for our actual calculation since it is those operators whose matrix elements are obtained from the explicit contractions which we evaluate. The effects of the Wilson

coefficients and all operator renormalization and mixing can then be summarized by:

$$\begin{aligned} H_W &= \frac{G_F}{\sqrt{2}} \sum_{q,q'=u,c} V_{qd} V_{q's}^* \sum_{i,j,k=1,2} C_i^{\overline{\rm MS}}(\mu)(1+\Delta r^{\rm RI\to\overline{MS}})_{ij}(Z^{\rm lat\to RI})_{jk} Q_k^{qq',\rm lat}(\mu) \\ &= \frac{G_F}{\sqrt{2}} \sum_{q,q'=u,c} V_{qd} V_{q's}^* \sum_{i=1,2} C_i^{\rm lat}(\mu) Q_i^{qq',\rm lat}(\mu). \end{aligned} \tag{3.36}$$

From this formula we can see that the $\overline{\rm MS}$ operators are converted to the lattice operators in two steps. First the $\overline{\rm MS}$ operators are converted to the regularization independent (RI) Rome-Southampton scheme perturbatively. We use formula provided by Lehner and Sturm which extend their earlier, 2+1 flavor results [15] for the matching matrix $\Delta r^{\rm RI\to\overline{MS}}$ to the four-flavor case being studied here. Their 2×2 matching matrix is given by:

$$\Delta r = \frac{\alpha_s(\mu)}{4\pi} \begin{pmatrix} -4\ln(2) & -8+12\ln(2) \\ -8+12\ln(2) & -4\ln(2) \end{pmatrix}. \tag{3.37}$$

Here $\alpha_s(\mu)$ is calculated using the two-loop formula given by equation (3.19) in Ref. [13]. For $\mu = 2.15$ GeV, we obtain $\alpha_s = 0.2974$.

In the next step, the operators in the RI scheme are related to the lattice operators non-perturbatively following the non-perturbative renormalization (NPR) method developed in Ref. [16] but using non-exceptional momenta [17] at a scale $\mu = 2.15$ GeV. Specifically, we use the RI/SMOM(γ_μ,$\not{q}$) scheme [15]. Here the first γ_μ means that the projectors are constructed from γ matrices. The second $\not{q}$ identifies the wave function renormalization scheme. The renormalization condition is imposed on the amputated four-quark off-shell Green's funtcion. For our convenience, we will use operator basis Q_i^{cu} defined in Eq. 3.34. The renormalization matrices for $\widetilde{Q}_i$ and Q_i^{uc} are identical to that for Q_i^{cu}. The following Green's function:

$$\Gamma_i(p_1,p_2) = \frac{1}{V}\sum_x \langle Q_i^{cu}(x) s(p_1)\bar{u}(p_2) c(p_1)\bar{d}(p_2)\rangle, \tag{3.38}$$

where the quark fields are Fourier transformed and gauge is fixed to Landau gauge. The momemta satisfy the conditon:

$$p_1^2 = p_2^2 = (p_1-p_2)^2, \tag{3.39}$$

which is the non-exceptional momentum scheme. The Green's function defined in Eq. 3.38 is given by products of momentum source propagators:

$$G_1(p_1,p_2)^{abcd}_{\alpha\beta\gamma\delta} = \frac{1}{V}\sum_{x,\mu} e^{-2i(p_2-p_1)x} \left\{ \left(\gamma^5 S^\dagger(x,p_1)\gamma^5\right)^{aa'} \gamma^\mu(1-\gamma^5) S(x,p_2)^{c'b} \right\}^{\alpha\beta} \left\{ \left(\gamma^5 S^\dagger(x,p_1)\gamma^5\right)^{cc'} \gamma^\mu(1-\gamma^5) S(x,p_2)^{a'd} \right\}^{\gamma\delta}, \tag{3.40}$$

$$G_2(p_1,p_2)^{abcd}_{\alpha\beta\gamma\delta} = \frac{1}{V}\sum_{x,\mu} e^{-2i(p_2-p_1)x} \left\{ \left(\gamma^5 S^\dagger(x,p_1)\gamma^5\right)^{aa'} \gamma^\mu(1-\gamma^5) S(x,p_2)^{a'b} \right\}^{\alpha\beta} \left\{ \left(\gamma^5 S^\dagger(x,p_1)\gamma^5\right)^{cc'} \gamma^\mu(1-\gamma^5) S(x,p_2)^{c'd} \right\}^{\gamma\delta}. \tag{3.41}$$

Here $S(x,p)$ is a propagator with a momentum source and a sink position x, the contraction of color indices are given explicitly and the contraction of spin indices should follow the rule of matrix multiplication inside each large curly bracket. We don't specify the flavor of the quark propagators in this formula since the renormalization condition is defined in the chiral limit. The dependence of the results on the quark masses is usually extremely small. So we use the same light quark mass for all flavors of quarks in our calculation and do not perform a extrapolation to chiral limit.

Then we use the full quark propagator to amputate these Green's functions:

$$\Gamma_i(p_1,p_2)^{abcd}_{\alpha\beta\gamma\delta} = G_i(p_1,p_2)^{a'b'c'd'}_{\alpha'\beta'\gamma'\delta'} \left(\langle\gamma^5 S^\dagger(p_1)\gamma^5\rangle^{-1}\right)^{a'a'}_{\alpha'\alpha} \left(\langle S(p_2)\rangle^{-1}\right)^{b'b}_{\beta'\beta} \left(\langle\gamma^5 S^\dagger(p_1)\gamma^5\rangle^{-1}\right)^{c'c}_{\gamma'\gamma} \left(\langle S^\dagger(p_1)\rangle^{-1}\right)^{d'd}_{\delta'\delta}. \tag{3.42}$$

Here the quark propagator by $S(p) = \sum_x S(x,p)e^{-ipx}$. In order to define the renormalization condition, we introduce the following spin-color projectors:

$$(P_1)^{abcd}_{\alpha\beta\gamma\delta} = [(1-\gamma^5)\gamma^\mu]_{\alpha\beta}[(1-\gamma^5)\gamma^\mu]_{\gamma\delta}\delta^{ad}\delta^{cb}, \tag{3.43}$$

$$(P_2)^{abcd}_{\alpha\beta\gamma\delta} = [(1-\gamma^5)\gamma^\mu]_{\alpha\beta}[(1-\gamma^5)\gamma^\mu]_{\gamma\delta}\delta^{ab}\delta^{cd}. \tag{3.44}$$

The operator mixing matrix is given by:

$$M_{ij} = \Gamma_i P_j. \tag{3.45}$$

Due to the Firez symmetry:

$$Q_1^{cu} = (\bar{s}_i u_j)_{V-A}(\bar{c}_j d_i)_{V-A} = (\bar{s}_i d_i)_{V-A}(\bar{c}_j u_j)_{V-A}, \tag{3.46}$$

we can easily show that $M_{11} = M_{22}$ and $M_{12} = M_{21}$. So we only need to calculate one amputated Green's function G_1 or G_2. We choose to calculate G_2 since it is explicitly color unmixed. The renormalization condition is defined as:

$$\frac{Z^{\text{lat}\to\text{RI}}}{Z_q^2} M = F, \tag{3.47}$$

where Z_q is the quark wave renormalization factor and F is the operator mixing matrix in tree level. In our case, F is simply given by:

$$F = \begin{pmatrix} 1 & 1/3 \\ 1/3 & 1 \end{pmatrix}. \tag{3.48}$$

Finally the operator renormalization matrix is given by:

$$Z^{\text{lat}\to\text{RI}} = Z_q^2 F M^{-1}. \tag{3.49}$$

For all the mass differences presented in this work, we use the operator mixing matrix M calculated from the 16^3 lattice configurations. We use 20 configurations and the valence quark mass $m^{\text{val}} = 0.01$. We take the value $Z_q^{\not q} = 0.8016(3)$ from [18]. Combining all the ingredients we obtain the final coefficients $C_{i=1,2}^{\text{lat}}$ that must be applied to the bare lattice operators to construct the complete $\Delta S = 1$ effective weak Hamiltonian given in Eq. (3.36). The results for these coefficients and the ingredient from which they are constructed are given in Tab. 2. Note the diagonal character of the renormalization for the operator basis $Q_\pm^X$ can be seen from the structure of the 2×2 matrices given in this table, with equal diagonal and equal off-diagonal elements in our Q_i^X, $i = 1, 2$ basis.

3.5 Short distance correction

The product of operators appearing in Eq. 3.4 accurately describes the second order weak effects when the corresponding Hamiltonian densities $H(x_i)_{i=1,2}$ are separated by a few lattice spacings ($|x_2 - x_1| \gg a$). However, as $|x_2 - x_1| \to 0$ the behavior is unphysical, being dominated by lattice artifacts. For the diagrams like Fig.4, we expect a quadratic divergence by naive power counting. This divergence will be cutoff by the inverse lattice spacing rather than a physical W boson mass. Fortunately, the GIM mechanism will remove this divergence completely. One might expect that the GIM cancellation would reduce the quadratic divergence to a milder logarithmic divergence leaving an unphysical, short distance artifact of the form $\ln(\mathrm{m_c a})$ reflecting a physical $\ln(\mathrm{m_c/M_W})$ short distance contribution, inaccessible to a lattice calculation. However, because of the $V - A$ structure of the weak vertices in the standard model, there will be no logarithmic divergence after the GIM cancellation. The quark propagator connecting the two weak vertices is given by:

$$\frac{\not{p} - m_u}{\not{p}^2 + m_u^2} - \frac{\not{p} - m_c}{\not{p}^2 + m_c^2}. \tag{3.50}$$

This propagator will scale like $1/p^2$ for large momenta. However, if we consider the $V - A$ structure of the weak vertices, we will get:

$$\gamma^\mu(1-\gamma^5)\left(\frac{\not{p} - m_u}{\not{p}^2 + m_u^2} - \frac{\not{p} - m_c}{\not{p}^2 + m_c^2}\right)\gamma^\nu(1-\gamma^5). \tag{3.51}$$

Then the mass term in the numerator will become zero. The propagator will become:

$$\gamma^\mu(1-\gamma^5)\frac{\not{p}(m_c^2 - m_u^2)}{(\not{p}^2 + m_u^2)(\not{p}^2 + m_c^2)}\gamma^\nu(1-\gamma^5). \tag{3.52}$$

We can see that a factor of $m_c^2 - m_u^2$ is introduced for each of the two internal quark lines, reducing the overall degree of divergence by four units. Thus, the GIM cancellation is complete, leaving only convergent integrals in a theory built from the effective four-quark operator H_W, with all "short distance" contributions coming from distances on the order

of $1/m_c$. Thus, if potential lattice artifacts associated with the large value of $m_c a$ can be neglected, then the lattice calculation will capture all important aspects of ΔM_K.

Although there is no divergence in the lattice calculation of ΔM_K, this is not always true for generic calculation of a second order weak quantity. For example, there will be a logarithmic divergence in the lattice calculation of ϵ_K. This divergence must be subtracted before the lattice calculation becomes well defined. After this subtraction, we can add back the correct short distance behavior, which can be calculated using perturbation theory in the continuum. This correction can be summarized by:

$$\mathcal{A} - \mathcal{A}_{\rm SD}^{\rm lat} + \mathcal{A}_{\rm SD}^{\rm cont}, \tag{3.53}$$

where the second term is the short distance part of $\mathcal{A}$ in the lattice calculation and the third term gives the correct short distance behavior in the continuum. Similar to the perturbation theory calculation, we can use a local $\Delta S = 2$ operator to describe the short distance behavior:

$$O_{LL} = (\bar{s}d)_{V-A}(\bar{s}d)_{V-A}. \tag{3.54}$$

Then the short distance part of $\mathcal{A}$ is given by:

$$\mathcal{A}_{\rm SD}^{lat} = \langle \overline{K}^0(t_f) \sum_{t=t_a}^{t_b} c^{\rm lat}(\mu^2) O_{LL}(t) \overline{K}^0(t_i) \rangle, \tag{3.55}$$

where the scale μ is the scale at which we define the short distance part of the second order weak process, $c^{\rm lat}(\mu^2)$ is a Wilson coefficient which can be determined by using non-perturbative Rome-Southhampton method. The correct short distance behavior is given by:

$$\mathcal{A}_{\rm SD}^{lat} = \langle \overline{K}^0(t_f) \sum_{t=t_a}^{t_b} c^{\rm cont}(\mu^2) O_{LL}(t) \overline{K}^0(t_i) \rangle, \tag{3.56}$$

where the $c^{\rm lat}(\mu^2)$ can be evaluated using perturbation theory in continuum.

The complexity of this correction comes from the evaluation of $c^{\rm lat}(\mu^2)$. This can be done

by evaluating the four quark, off-shell Greens function:

$$\Gamma_{\alpha\beta\gamma\delta}(p_i) = \langle s_\alpha(p_1)\bar{d}_\beta(p_2)\int d^4x_1 \int d^4x_2 H_W(x_1)H_W(x_2)s_\gamma(p_3)\bar{d}_\delta(p_4)\rangle. \tag{3.57}$$

Here the quark fields are Fourier transformed and gauge is fixed to Landau gauge. The momentum of the quarks should satisfy $p_i^2 = \mu^2$. Also, the p_i should be chosen appropriately so that all the internal momenta contributing to $\Gamma(p_i)$ will be at the scale μ. The value for μ should satisfy:

$$\Lambda_{\rm QCD} \ll \mu \ll \frac{\pi}{a}. \tag{3.58}$$

Then we evaluate second Greens function for the local O_{LL} operator:

$$\Gamma^{\rm SD}_{\alpha\beta\gamma\delta}(p_i) = \langle s_\alpha(p_1)\bar{d}_\beta(p_2)\int d^4x O_{LL}(x)s_\gamma(p_3)\bar{d}_\delta(p_4)\rangle. \tag{3.59}$$

This Green's function describes the short distance component of $\Gamma(p_i)$. For a large enough μ, $\Gamma(p_i)$ will be dominated by short distance effect. Thus, the Wilson coefficients $c_s^{\rm lat}(\mu^2)$ can be determined by:

$$\left(\Gamma^{\rm amp}_{\alpha\beta\gamma\delta}(p_i) - c_s^{\rm lat}(\mu^2)\Gamma^{\rm amp,SD}_{\alpha\beta\gamma\delta}(p_i)\right) P_{\alpha\beta\gamma\delta} = 0, \tag{3.60}$$

where "amp" means that the Green functions should be amputated, $P_{\alpha\beta\gamma\delta}$ is a spin and color projector

$$P_{\alpha\beta\delta\gamma} = \left((1-\gamma^5)\gamma^\mu\right)_{\alpha\beta}\left((1-\gamma^5)\gamma^\mu\right)_{\gamma\delta} \tag{3.61}$$

3.6 Finite volume corrections

In this section, we will discuss the finite volume effects in the mass difference calculation. Finite volume effects exist in all lattice calculations. If there are no multi-particle states evolved, the finite volume effects will usually decrease exponentially as the lattice size L increases. We usually neglect such effects in the practical calculation. However, in the mass difference calculation, there will be lots of multi-particle intermediate states contributing to

the mass difference defined in Eq. 3.8. These states may be distorted in a finite volume and will introduce errors which are suppressed only by power of L. The finite volume effects have been studied in detail in Ref. [11]. The starting point is Luscher's relation [19] between finite-volume two-particle energy, $E = 2\sqrt{m_\pi^2 + k^2}$ and the two-particle scattering phase shift $\delta(E)$:

$$\phi(\frac{kL}{2\pi}) + \delta(E) = n\pi, \tag{3.62}$$

where n is an integer and the known function $\phi(q)$ is defined in Ref. [19]. This relation connects the π-π energy in finite volume with the phase shift in the infinite volume. We want to use this condition to relate ΔM_K computed in finite volume and infinite volume. We can choose a specially tuned lattice volume so that there is an π-π state $|n_0\rangle$ with energy E_{n_0} nearly degenerate with m_K. The $K_S \leftrightarrow \pi\pi$ system can be described by the degenerate perturbation theory. For simplicity we will assume CP symmetry so that π-π will only couple to the K_S state. Following second order degenerate perturbation theory, we can obtain the energies of K_S and the two pion state $|n_0\rangle$ as the eigenvalues of the 2×2 matrix:

$$\begin{pmatrix} m_K + \sum_{n \neq n_0} \frac{|\langle n|H_W|K_S\rangle_L|^2}{m_K - E_n} & \langle K_S|H_W|n_0\rangle_L \\ \langle n_0|H_W|K_S\rangle_L & E_{n_0} + \sum_{n \neq K_S} \frac{|\langle n|H_W|n_0\rangle_L|^2}{E_{n_0} - E_n} \end{pmatrix}. \tag{3.63}$$

The subscript L here means the matrix elements are evaluated in finite volume. Finite and infinite volume quantities can then be related by requiring that the eigenvalue of the 2×2 matrix in Eq. 3.63 solve Eq. 3.62. The infinite volume phase shift $\delta(E)$ is the sum of that arising from the strong interaction, $\delta_0(E)$, a resonant contribution from the K_S pole and the second-order Born terms:

$$\delta(E) = \delta_0(E) + \arctan\left(\frac{\Gamma(E)/2}{m_K + \Delta m_{K_s} - E}\right) - \pi \sum_{\beta \neq K_S} \frac{|\langle \beta|H_W|\pi\pi(E)\rangle|^2}{E - E_\beta}. \tag{3.64}$$

Here $\Gamma(E)$ is proportional to the square of the K_S-$\pi\pi$ vertex which becomes the K_S width when evaluated at $E = m_K$:

$$\Gamma(E) = 2\pi|\langle \pi\pi(E)|H_W|K_S\rangle|^2, \tag{3.65}$$

where for the two-pion state we choose the normalization $\langle\pi\pi(E)|\pi\pi(E')\rangle = \delta(E-E')$.

The easiest case to examine is that in which the energy difference $E_{n_0} - m_K$ is very small compared with m_K but large compared to ΔM_K or Γ, so that E_{n_0} and m_K are not exactly degenerate. With this choice, the π-π energy eigenvalue determined from Eq. 3.63 differs from E_{n_0} by the usual second order perturbation theory expression:

$$E_{\pi\pi} = E_{n_0} + \frac{|\langle K_S|H_W|n_0\rangle_L|^2}{E_{n_0} - M_K} + \sum_{n\neq K_S} \frac{|\langle n|H_W|n_0\rangle_L|^2}{E_{n_0} - E_n}. \tag{3.66}$$

Combining this expression with Eq. 3.62 and Eq. 3.64 and equating all terms of second order in H_W yeilds the relation:

$$\frac{\partial(\phi+\delta_0)}{\partial E}\left\{\frac{|\langle K_S|H_W|n_0\rangle_L|^2}{E_{n_0} - M_K} + \sum_{n\neq K_S} \frac{|\langle n|H_W|n_0\rangle_L|^2}{E_{n_0} - E_n}\right\} = \frac{\Gamma(E_{n_0})/2}{E_{n_0} - m_K} + \pi \sum_{\beta\neq K_S} \frac{|\langle\beta|H_W|\pi\pi(E)\rangle|^2}{E - E_\beta}. \tag{3.67}$$

This expression has two useful consequences. First we can equate the residues of the kaon poles, $E_{n_0} = m_K$ on the left- and right-hand sides. This gives us the original Lellouch-Luscher relation. Second we can subtract the pole terms and equate the remaining parts of Eq. 3.67 evaluated at $E_{n_0} = m_K$. This second result will be used below to remove the second-order Born terms.

Now we are ready to the examine a more complicated case in which $E_{n_0} = m_K$. The finite volume energies are still given by the eigenvalues of Eq. 3.63:

$$E_\pm = m_K \pm \langle n_0|H_W|K_S\rangle + \frac{1}{2}\left\{\sum_{n\neq n_0} \frac{|\langle n|H_W|K_S\rangle_L|^2}{m_K - E_n} + \sum_{n\neq K_S} \frac{|\langle n|H_W|n_0\rangle_L|^2}{m_K - E_n}\right\}, \tag{3.68}$$

but now the eigenstates are equal mixtures of the degenerate K and finite volume π-π state $|n_0\rangle$. We again require that these finite-volume energies satisfy Eq. 3.62. Expanding to second order of H_W, we get the relation between the finite and infinite volume expressions for the K_S mass shift:

$$\Delta m_{K_S} = \sum_{n\neq n_0} \frac{|\langle n|H_W|K_S\rangle_L|^2}{m_K - E_n} - \frac{\partial^2(\phi+\delta_0)/\partial E^2}{2\partial(\phi+\delta_0)/\partial E}|\langle n_0|H_W|K_S\rangle_L|^2 - \frac{\partial|\langle n_0|H_W|K_S\rangle_L|^2}{\partial E}, \tag{3.69}$$

where Eq. 3.67, evaluated at $E_{n_0} = m_K$ with the pole term subtracted has been used to eliminate the second-order Born terms.

Finally, the infinite-volume $K_L - K_S$ mass difference can be obtained as follows. First observe that the K_L second order mass shift is given by a formula similar to Eq. 3.69 in which K_L replaces K_S and all but the first term on the right-hand side are omitted since K_L does not couple to two pions in our approximation of CP symmetry. Subtracting this new formula from Eq. 3.69 gives:

$$\Delta m_K = \Delta m_K^{\mathrm{FV}} - \frac{\partial^2(\phi+\delta_0)/\partial E^2}{2\partial(\phi+\delta_0)/\partial E}|\langle n_0|H_W|K_S\rangle_L|^2 - \frac{\partial|\langle n_0|H_W|K_S\rangle_L|^2}{\partial E} \tag{3.70}$$

Here Δm_K^{FV} is defined in Eq. 3.8. The correction terms in this formula are all accessible in a lattice calculation. However, in our numerical work, we are not able to identify a clear signal from two pion intermediate states. So we are not able to study their finite volume effects in this work. In a future calculation at physical kinematics, the finite volume effects may be important and should be furthur studied.

Chapter 4

Measurement methods

In this chapter, we will discuss the techniques used in the lattice calculation of ΔM_K. In Section. 4.1, we will discuss the different sources of the quark propagators. In Section. 4.2, we will discuss the low-mode-deflation method based on the EigCG and Lanczos algorithms. Section 4.3 discuss the two-point and three-point correlators in our calculation. In Section. 4.4, we explain the strategies used to calculate the different type of diagram contributing to the four-point correlators. In Section. 4.5, we briefly summarize the data analysis techniques.

4.1 Propagator sources

The contractions shown in Chapter. 3 can be written as products of quark propagators. We need to compute all the needed quark propagators in order to evaluate these contractions. A quark propagator from the source point y to the sink point x is given by:

$$S(x, y) = D^{-1}(x, y), \tag{4.1}$$

where $S(x, y)$ is a 12×12 spin-color matrix, D is the Dirac matrix on lattice. For a resonantly large lattice, it is impossible to invert the Dirac matrix since the dimension of the matrix is extremely high. So we are not able to compute $S(x, y)$ for all source and sink locations.

However, in most cases, it is sufficient to calculate the quark propagator for some specific choices of source. A general source $b(x)$ can be any function of space-time. The propagator for this source is given by:

$$S(x) = \sum_y S(x,y)b(y). \tag{4.2}$$

This propagator can be obtained by solving the following linear system:

$$\sum_y D(x,y)S(y) = b(x). \tag{4.3}$$

A important property of $S(x,y)$ is γ^5-hermiticity. In our domain wall fermion formalism, the dirac matrix satisfies $D^\dagger = \gamma^5 D \gamma^5$. Hence the propagator should satisfy:

$$S(y,x) = \gamma^5 S(x,y)^\dagger \gamma^5. \tag{4.4}$$

This relation can tell us $S(y,x)$ immediately after $S(x,y)$ has been computed.

The choice of sources depends on the specific problem. The point source $b(x)$ is 1 at the source point and a specific spin and color and 0 anywhere else. The propagator given by Eq. 4.1 is a point source propagator from source location y. We will use point sources for the two internal propagators in the type 1 and type 2 contractions given in Section. 3.2.

To better overlap with the ground state, we use wall source for the kaon and pion interpolating field. The wall source at time t with spin α and color a is defined as:

$$b(\vec{y}, t_y) = \begin{cases} \chi_{a\alpha}, & t_y = t \\ 0, & t_y \neq t_y, \end{cases} \tag{4.5}$$

where $\chi_{a\alpha}$ is a 12-component vector with 1 at spin α and color a and 0 at anywhere else. The corresponding wall source propagator is given by:

$$S^W(\vec{x}, t_x; t_y) = \sum_{\vec{y}} S(\vec{x}, t_x; \vec{y}, t_y). \tag{4.6}$$

Sometimes we will need a wall-source-wall-sink propagator. It is obtained by summing over the sink position:

$$S^{WW}(t_x;t_y)=\sum_{\vec{x},\vec{y}}S(\vec{x},t_x;\vec{y},t_y). \tag{4.7}$$

In the type 3 and type 4 diagrams given in Section. 3.2, we have some fermion loops in the contractions. So we need to calculate $S(x,x)$ for all possible x. Naively, we can calculate a point source propagators at each possible source location. However, this is impossible in practice since there are too many source locations. The solution is to use random wall source propagators to give a unbiased estimation for the fermion loops. The random wall source at time t and spin α and color a is given by:

$$b(\vec{y},t_y)=\begin{cases}\chi_{a\alpha}\eta(\vec{y},t_y), & t_y=t\\ 0, & t_y\neq t,\end{cases} \tag{4.8}$$

where $\chi_{a\alpha}$ is a 12-component vector with 1 at spin α and color a and 0 at anywhere else, $\eta_{\vec{y},t_y}$ are random numbers satisfying:

$$\langle\eta^\dagger(\vec{x},t_x)\eta(\vec{y},t_y)\rangle=\delta(\vec{x}-\vec{y})\delta(t_x-t_y), \tag{4.9}$$

where $\langle\cdots\rangle$ means average for many hits of random numbers. The random wall source propagators is given by:

$$S^{RW}(\vec{x},t_x;t_y)=\sum_{\vec{y}}S(\vec{x},t_x;\vec{y},t_y)\eta(\vec{y},t_y). \tag{4.10}$$

Finally the fermion loops are given by:

$$S(\vec{x},t_x;\vec{x},t_x)=\sum_{\eta}S^{RW}(\vec{x},t_x;t_x)\eta^\dagger(\vec{x},t_x) \tag{4.11}$$

In our operator renormalization calculation, we use a volume momentum source. This source is defined as $b(y)=e^{ipy}$. The momentum source propagator is then given by:

$$S(x;p)=\sum_{y}S(x,y)e^{ipy}. \tag{4.12}$$

The momentum-source-momentum-sink propagator is defind as:

$$S(p) = \sum_{x,y} S(x,y) e^{-ip(x-y)}. \tag{4.13}$$

4.2 Low Mode Deflation

Solving for the quark propagators is the most expensive part of our calculation. We use the conjugate gradient (CG) methods to solve the linear system:

$$Dx = b \tag{4.14}$$

The conjugate gradient method can't be applied directly here since it requires D to be Hermitian. So the equation we actually solve is the normal equation:

$$D^\dagger D x = D^\dagger b = b' \tag{4.15}$$

The convergence of the CG algorithm is controlled by the condition number $\kappa(D^\dagger D)$, which is given by the ratio between the largest and the smallest eigenvalues of $D^\dagger D$. The CG algorithm converges faster for a smaller κ. $\kappa(D^\dagger D)$ for the light quark is larger than the value for the heavy quark. So a light quark CG solver needs more iteration steps to converge. A straightforward method to reduce κ is the low-mode-deflation method. If we know lowest n eigenvalue and eigenvectors of $D^\dagger D$, we can define a projection operator:

$$P = \sum_{i=1}^{n} |\lambda_i\rangle\langle\lambda_i| \tag{4.16}$$

where λ_i is the i-th smallest eigenvalue and $|\lambda_i\rangle$ is the corresponding eigenvector. Since P commutes with $D^\dagger D$, the normal equation 4.15 can be split into decoupled equations:

$$(D^\dagger D)_\parallel x_\parallel = b_\parallel, \quad x_\parallel = Px, \quad b_\parallel = Pb \tag{4.17}$$

$$(D^\dagger D)_\perp x_\perp = b_\perp, \quad x_\perp = Px, \quad b_\perp = Pb \tag{4.18}$$

where $(D^\dagger D)_\parallel = PD^\dagger DP$ and $(D^\dagger D)_\perp = (1-P)D^\dagger D(1-P)$ are referred to as the little operator and the deflated operator respectively.

The solution to the little system 4.17 is simply given by:

$$x_\parallel = \sum_{i=1}^{n} \frac{|\lambda_i\rangle\langle\lambda_i|b\rangle}{\lambda_i}. \tag{4.19}$$

The deflated system 4.18 can be solved using the CG algorithm. The new condition number

$$\kappa(D^\dagger D)_\perp = \frac{\lambda_1}{\lambda_{n+1}}\kappa(D^\dagger D) \tag{4.20}$$

is reduced and one expects the solver to be accelerated.

The deflation method usually works well for the light quark CG solver and works poorly when the quark mass is heavy. This can be explained by the ratio λ_n/λ_1. For light quark, this ratio is usually large. When the quark mass becomes heavy, the spectrum of the low eigenvalues is bounded from below by the large quark mass. Thus this ratio will become smaller for a heavier quark mass. In our lattice calculation, we only use low mode deflation only for the up and down quark CG solver. For the strange and charm quark propagator, we will use the plain CG algorithm.

The calculation of the low-lying eigenvectors can be expensive. However, in our calculation, we need to solve Eq. 4.15 for many different sources. We can use the same eigenvectors for CG solver with the same quark mass. Hence the time spent on calculating eigenvectors is amortized and is only a small part of the total time.

There are two different method to calculate the low modes: the EigCG algorithm and the implicit restarted Lanczos (IRL) algorithm. The details of the EigCG algorithm can be found in Ref. [1]. The EigCG algorithm builds an explicitly restarted Lanczos algorithm on top of the conjugate gradient method. The idea is to reuse the search directions generated by the conjugate gradient algorithm to compute the necessary Lanczos vectors. It does not require additional matrix-vector multiplications. So the increase in the time cost is small. Since the Lanczos algorithm is numerically unstable, it also restarts itself when a certain

number of eigenvectors obtained. A certain number of new eigenvectors will be calculated in each EigCG solver. At the beginning of a EigCG solver, we can use the already calculated eigenvectors to do deflation, thus accelerating the solver. A new solver will not only collect more eigenvectors but also will improve the precision of the already calculated eigenvectors. Gradually, the solver will become faster and faster. Finally we stop collecting eigenvectors and simply do deflation to speed up the solver.

In Ref. [2], Qi observed a clear slowing-down point (around $res \approx 10^{-6}$) on the convergence curve for the sped-up solver. This was due to the inaccuracy of the eigenvectors we obtained from EigCG solvers. Qi's strategy was to do multiple projections by restarting the CG algorithm using the residual of previous inversion as the new right hand side. Qi found that a single restart point at $res \approx 10^{-5}$ can solve the slowing-down problem if our target precision is 10^{-8}. With this restart technique, in our 16^3 lattice calculation with a 412 MeV pion mass, we obtain a factor of 6 speed up using 100 low-lying eigenvectors calculated by the EigCG algorithm.

The IRL algorithm is a standard textbook algorithm which can calculate the eigenvectors of hermitian matrix to machine precision. The details of this algorithm are discussed in Ref. [3]. This algorithm can give more precise eigenvectors which can provide better deflation effects. However, unlike EigCG, the IRL algorithm needs many extra Dirac matrix multiplication and is more expensive. A Chebyshev polynomial is used to change the spectrum of the Dirac matrix and thereby to accelerate the IRL algorithm. In our 24^3 lattice calculation with a 330 MeV pion mass, the number of propagators we need to calculate is much greater than were required for the 16^3 calculation. To get a better deflation effect, we use the IRL algorithm to calculate 300 exact eigenvectors and get a factor of 8 speed up.

4.3 Two-point and three-point correlators

Two-point correlator are usually calculated to extract meson masses. The correlator is given by:

$$C(t_1, t_2) = \langle O(t_2) O^\dagger(t_1) \rangle, \tag{4.21}$$

where O is the interpolating field for the meson. After inserting a set of QCD energy eigenstates, one obtains:

$$\begin{aligned} C(t_1, t_2) &= \sum_n \langle 0|O|n\rangle e^{-E_n(t_2-t_1)} \langle n|O^\dagger|0\rangle \\ &\overset{t_2 \gg t_1}{=} N_m^2 e^{-E_m(t_2-t_1)}, \end{aligned} \tag{4.22}$$

where $|m\rangle$ is the lowest energy eigenstate which can couple to O and $N_m = |\langle 0|O|m\rangle|$ is the normalization factor of the interpolating field. We can obtain the meson mass and the normalization factor of the interpolating field from a simple exponential fit. There are many choices of the meson interpolating field. We use a wall source in our calculation. For example, the wall source kaon interpolating filed can be written as:

$$O_{K^0}(t) = i \sum_{\vec{x},\vec{y}} s(\vec{x}, t) \gamma^5 \bar{d}(\vec{y}, t). \tag{4.23}$$

The wall source is not explicitly gauge invariant. This can be solved by fixing the gauge field at time t to Coulomb gauge. A Coulomb gauge fixed wall source has better overlap with the ground state of kaon than a plain point source. We find that a Coulomb gauge fixed wall source works well in both our 16^3 and 24^3 calculations. However, in a larger volume, a Coulomb gauge fixed wall source kaon can be much larger than the size of a physical kaon. So it may be better to use a more localized source in future calculation on larger lattices.

We also use a two-point correlator to extract the matrix element $\langle 0|O|K^0\rangle$. The correlator can be written as:

$$C(t_1, t_2) = \langle O(t_2) K^\dagger(t_1) \rangle. \tag{4.24}$$

We have listed the needed diagrams in Fig. 12. It is convenient to define the ratio:

$$\begin{aligned} R(t_1,t_2) =& \frac{\langle O(t_2)K^\dagger(t_1)\rangle}{\langle K(t_2)K^\dagger(t_2)\rangle} \\ =& \frac{1}{N_K}\langle 0|O|K^0\rangle, \end{aligned} \tag{4.25}$$

where N_K is the kaon normalization factor and the second equality is valid for sufficiently large $t_2 - t_1$. The plateau value for $R(t_1, t_2)$ will give us $\langle 0|O|K^0\rangle$.

The three-point correlator is used to compute the matrix element $\langle \pi^0|O|K^0\rangle$. The correlator can be written as:

$$C(t_1, t_o, t_2) = \langle \pi^0(t_2)O(t_o)K(t_1)^\dagger\rangle. \tag{4.26}$$

We have listed the possible diagrams in Fig. 11. We can define the ratio:

$$\begin{aligned} R(t_1,t_O,t_2) =& \frac{\langle \pi^0(t_2)O(t_o)K(t_1)^\dagger\rangle}{\langle \pi^0(t_2)\pi(t_o)^\dagger\rangle\langle K(t_o)K(t_1)^\dagger\rangle} \\ =& \frac{\langle \pi^0|O|K^0\rangle}{N_K N_\pi}, \end{aligned} \tag{4.27}$$

where N_K is the kaon normalization factor, N_π is the pion normalization factor and the second equality is valid for sufficiently large $t_2 - t_1$. Again $\langle \pi^0|O|K^0\rangle$ is given by the plateau in $R(t_1, t_O, t_2)$ as a function of t_O for sufficiently large $|t_2 - t_1|$.

4.4 Evaluation of four-point correlators

The four-point correlators can be written as combinations of the contractions listed in Fig. 5-8. Each contraction is composed of six quark propagators. In order to make our discussion more concrete, we will discuss some example here. The contraction ① in Fig. 5 is written as:

$$\begin{aligned} ①_{uu} =& \mathrm{Tr}(\gamma^\mu(1-\gamma^5)L^W(x,t_i)\gamma^5\gamma^5 S^{W\dagger}(x,t_i)\gamma^5)\mathrm{Tr}(\gamma^\nu(1-\gamma^5)L(y,x)\gamma^\mu(1-\gamma^5)L(x,y)) \\ & \mathrm{Tr}(\gamma^\nu(1-\gamma^5)L^W(y,t_f)\gamma^5\gamma^5 S^{W\dagger}(y,t_f)\gamma^5). \end{aligned} \tag{4.28}$$

The two kaon wall sources are located at t_i and t_f. The two weak Hamiltonian densities are located at x and y. $L(x,y)$ is a point source propagator from y to x. $L^W(x,t)$ is a light quark propagator with Coulomb gauge fixed wall source at time t and sink at space time position x. $S^W(x,t)$ is similar but for strange quark. We also use the γ^5-hermiticity for the strange quark wall source propagators. For simplicity we only give the contraction for the case that both internal quarks are up, the other cases are very similar.

Another useful example is contraction ⑨:

$$\begin{aligned}⑨_{uu} =&\mathrm{Tr}(\gamma^\mu(1-\gamma^5)L^W(x,t_i)\gamma^5\gamma^5 S^{W\dagger}(x,t_i)\gamma^5\gamma^\nu(1-\gamma^5)L^W(y,t_f)\gamma^5\gamma^5 S^{W\dagger}(y,t_f)\gamma^5)\\ &\mathrm{Tr}(\gamma^\mu(1-\gamma^5)L^{RW}(x,t_x)\eta^\dagger(x))\mathrm{Tr}(\gamma^\nu(1-\gamma^5)L^{RW}(y,t_y)\eta^\dagger(y)).\end{aligned} \tag{4.29}$$

Here $L^{RW}(x,t)$ is a light quark propagator with random wall source at time t and sink at space-time position x. The fermion loop $L(x,x)$ is replaced with $L^{RW}(x,t_x)\eta^\dagger(x)$ this expression. All the contractions can be written down similarly. We give all the expression in Appendix. A.

We use a different strategy to calculate type 1,2 and type 3,4 diagrams. This is due to the different topologies of these contractions. We will discuss the evaluation of type 1,2 diagram first. For given values of t_x and t_y, each of the two effective operators should be integrated over the whole spatial volume. However, there is no easy way to do this because of two difficulties. First, we are not able to compute all of the light-quark propagators connecting the two operators. It is impractical to use point source propagators since there will be V point sources on each time slice. In simpler cases, this difficulty can be avoided by the use of a stochastic source distributed over the time slice. However, an attempt to use this technique in the present case failed to give a signal that could be recognized above the noise. Even if this first difficulty of generating the multitude of needed point source propagators could be overcome, we would still face a second difficulty: the number of operations needed to calculate all the contractions would be $O(V^2)$, where V is the space-time volume of the lattice. This also would be too time consuming. Thus, we sum the location of only one of

the two operators over the spatial volume and, relying on the translational symmetry of the other ingredients in the calculation, fix the spatial location of other operator at the origin $(0, 0, 0)$. For each of the contractions in our calculation, these two weak operators enter in distinct ways and we average the two cases where one operator is fixed at the origin and the other integrated over the spatial volume to improve the statistics.

For type 3,4 contractions, the situation is very different. There are no propagators connecting the two weak Hamiltonian densities. We can use random wall source propagators to evaluate the fermion loops. Also, the number of operations needed to calculate the contractions is $O(V)$ instead of $O(V^2)$. The reason is that we can always write the contractions as the product of two independent parts. Each part will only need $O(V)$ time to compute. For type 3 diagrams in Fig. 7, the contractions can be split into an upper half and a lower half. For type 4 diagrams in Fig. 8, we can split the contractions into a left half and a right half. The subtraction diagrams shown in Fig. 9-10 can be calculated similarly.

4.5 Data Analysis

After we have calculated all the correlation functions like $C(t)$ for the meson from N configurations, we usually need to fit the data to some specific function form $f(t, \theta)$ as in Eq. 4.22. This is accomplished by minimizing the correlated χ^2,

$$\chi^2 = \sum_{t,t'} [f(t,\theta) - \overline{C(t)}] V^{-1}(t,t') [f(t',\theta) - \overline{C(t')}], \tag{4.30}$$

where $V(t, t')$ is the covariance matrix for C(t),

$$V(t,t') = \frac{1}{N(N-1)} \sum_{i=1}^{N} (C^i(t) - \overline{C(t)})(C^i(t') - \overline{C(t')}). \tag{4.31}$$

If the functional form fits the data well, we will get a χ^2 per degree of freedom ($\chi^2/d.o.f$) around 1. A large χ^2/d.o.f usually suggests a poor fit. The most common reason is that the fitting model does not describe the data well. For example, we will usually find a large

χ^2/d.o.f if we include data from the case in which the meson source and the meson sink are too close. There will be contamination from excited states and we will be unable to fit all the data with a single exponential term. To improve the quality of fit, we can either drop these data points or add another exponential term into the fitting formula. Another reason for large $\chi^2/d.o.f$ is that the covariance matrix is so singular that the fitting does not work properly. In that case, we can turn to an uncorrelated fit. The χ^2 minimized in uncorrelated fit is given by:

$$\chi^2 = \sum_t [\frac{f(t,\theta) - \overline{C(t)}}{\sigma(t)}]^2, \tag{4.32}$$

where $\sigma^2(t) = \sum_{i=1}^{N}(C^i(t) - \overline{C(t)})^2/N$. We can see that the uncorrelated fit uses only the diagonal terms in the correlation matrix. Hence the correlation between difference data points are not taken into account. The χ^2 values in the uncorrelated fit are usually very small and not very meaningful.

If the data are strongly correlated, it is usually better to use the correlated fit which may give a more accurate result and a clear meaning for the χ^2. In some cases, we will get significantly smaller error bar with the correlated fit than the uncorrelated fit. However, for simplicity, we can just use an uncorrelated fit for simple exponential or cosh fits for the mesons since it usually gives equally good results.

The correlated or uncorrelated fit gives us the central value only. We usually use the jackknife method to compute the error on the fitting parameters. Let $C_i(t)$ to stand for the correlation function we calculated from the configuration number i, and we have a sample of N configurations $S = \{C_i(t), i = 1 \ldots N\}$. Based on the correlated fit or uncorrelated fit from the previous discussion, we can compute the parameter θ. Let us label it as $\theta = g(S)$.

The jackknife method is :

1. Obtain N samples from S by leaving one out a time: S_{-i}, for i=1, ..., N.

2. For each such sample, fit the parameter θ and determine the value $\theta_i = g(S_{-i})$.

3. Compute the average of θ_i: $\overline{\theta}$. The square of the error on θ is estimated as

$$\sigma_J^2 = \frac{N-1}{N}\sum_{i=1}^{N}(\theta_i - \overline{\theta})^2. \tag{4.33}$$

Chapter 5

Results from the $16^3 \times 32 \times 16$ Lattice

In this chapter, we will discuss the results from the $16^3 \times 64 \times 16$ lattice ensemble. Section. 5.1 gives the details of the simulation. In Section. 5.2, we discuss the short distance effects and the GIM cancellation. In Section. 5.3, we discuss the long distance effects for each definite parity channels separately. Section. 5.4 gives the mass difference results in physical units. In Section. 5.5, we compare our lattice calculation with the NLO perturbation theory calculation.

5.1 Simulation Details

The calculation is performed on a lattice ensemble generated with the Iwasaki gauge action and 2+1 flavors of domain wall fermions at a coupling $\beta = 2.13$. The space-time volume is $16^3 \times 32$ and the inverse lattice spacing $a^{-1} = 1.729(28)$GeV. The fifth-dimensional extent is $L_s = 16$ and the residual mass is $m_{\rm res} = 0.00308(4)$ in lattice units. The sea light and strange quark masses are $m_l = 0.01$ and $m_s = 0.032$ respectively, corresponding to a pion mass $M_\pi = 421\,\mathrm{MeV}$ and a kaon mass $M_K = 563\,\mathrm{MeV}$. We use 800 configurations, each separated by 10 time units. This ensemble is described in greater detail in Ref. [18] and is also similar to the earlier ensembles described and analyzed in Ref. [20], except that the

current ensemble has a more physical value for the sea quark mass and was generated with a better RHMC algorithm.

We will use Fig. 4 to explain the set up of this calculation. Two Coulomb gauge-fixed kaon sources are located at time slices $t_i = 0$ and $t_f = 27$ respectively. The two effective weak operators $H_W(t_i)_{i=1,2}$ are introduced in the interval $4 \leq t_1, t_2 \leq 23$. We calculate the four-point function defined in Eq. 3.4 for all possible choices of t_1 and t_2. Note that the diagram given in Fig. 4 is only one type of possible contraction. The details of the contractions have already been discussed in Section. 3.2. In a unitary calculation, we need to include all types of diagrams. However, we only include type 1 and type 2 contractions in this pilot calculation. The evaluation of these two types of contractions is discussed in Section. 4.4. Neglecting type 3 and type 4 diagrams, reduces Eq. 3.26 to:

$$\langle \overline{K^0}(t_f) Q_{11}^{\mathrm{GIM}}(t_2, t_1) \overline{K^0}(t_i) \rangle = ①_{\mathrm{GIM}} - ⑤_{\mathrm{GIM}}, \tag{5.1}$$

$$\langle \overline{K^0}(t_f) Q_{22}^{\mathrm{GIM}}(t_2, t_1) \overline{K^0}(t_i) \rangle = ④_{\mathrm{GIM}} - ⑧_{\mathrm{GIM}}, \tag{5.2}$$

$$\langle \overline{K^0}(t_f) \left(Q_{12}^{\mathrm{GIM}}(t_2, t_1) + Q_{21}^{\mathrm{GIM}}(t_2, t_1) \right) \overline{K^0}(t_i) \rangle = -②_{\mathrm{GIM}} - ③_{\mathrm{GIM}} + ⑥_{\mathrm{GIM}} + ⑦_{\mathrm{GIM}}. \tag{5.3}$$

We neglect type 3 and type 4 diagrams in this calculation for two reasons. The first reason is practical. We would need to compute an additional stochastic wall source for each time slice to evaluate the new loop graphs which appear in the type 3 and 4 contractions. This would approximately double the computation time. More importantly, type 4 diagrams are disconnected diagrams which are extremely noisy and would require a far larger statistical sample than is being used here [18]. The second reason is phenomenological. There is some empirical evidence suggesting that the contribution from type 3 and type 4 diagrams may be small. For example, disconnected graphs similar to those of type 4 are often small when contributing to other processes where they are said to be "Zweig suppressed" [21, 22, 23]. The omission of such diagrams is also consistent with the results of the recent study of $\Delta I = 1/2$ $K \to \pi\pi$ decays [18] in which the contribution of disconnected diagrams was found to be zero within rather large errors. Of course in a complete calculation these diagrams must be

calculated explicitly after which the precision of the Zweig suppression will be known.

As discussed in Sec. 3.1, we need to calculate the matrix elements $\langle\pi^0|H_W|K^0\rangle$ and $\langle 0|H_W|K^0\rangle$ in order to remove the exponentially growing terms in the second order correlator. In this non-unitary calculation, there will be no contribution from vacuum intermediate state. Also the definition of the π^0 intermediate state must be reconsidered. In a unitary theory, $\bar{u}u$, $\bar{d}d$ and $\bar{s}s$ will mix with each other through disconnected diagrams. Then the resulting energy eigenstates are π^0, η and η', where π^0 is defined as $i(\bar{u}\gamma_5 u - \bar{d}\gamma_5 d)/\sqrt{2}$. However, in our non-unitary calculation, all disconnected diagrams are neglected and correlators of the operators $i(\bar{u}\gamma_5 u \pm \bar{d}\gamma_5 d)$ will reveal independent but symmetrical "states" with the same mass. Since only up quarks can appear in our intermediate state, we must use the interpolating operator $i\bar{u}\gamma_5 u$ to create our π^0 state and can neglect the effects of the symmetrical state created by $\bar{d}\gamma_5 d$. Thus, in our calculation of $\langle\pi^0|H_W|K^0\rangle$, we use $\pi^0 = i\bar{u}\gamma_5 u$ (with no $1/\sqrt{2}$ factor) and only include contraction $\boxed{1}$ and $\boxed{2}$ shown Fig. 11.

We use periodic boundary conditions in the spatial directions for the Dirac operator when computing the propagators. In the temporal direction, we calculate propagators for both periodic and anti-periodic boundary conditions and take their average for the propagator that we use. This effectively doubles the temporal extent of the lattice and suppresses around-the-world effects to a negligible level. (This approach is equivalent to working on a lattice of size $16^3 \times 64$ with gauge fields invariant under a translation of 32 sites in the time direction.) The most expensive part of this simulation is solving for the light quark propagators. There are 2 wall source light quark propagators and 20 point source light quark propagators, one on each time slice between $t_a = 4$ and $t_b = 23$. So in total we need to calculate $(20+2)\times 2 = 44$ propagators, where the factor of two comes from our two choices of temporal boundary conditions. Further each propagator requires 12 Dirac operator inversions, one for each spin and color. This large number of light-quark Dirac operator inversions makes this calculation a good candidate for the use of the EigCG technique [1, 2]. We collect the lowest 100

eigenvectors and use them to accelerate the light-quark Dirac operator inversions. The overhead associated with collecting these low modes is amortized over many inversions and the number of conjugate gradient iterations is reduced by a factor of 6. The parameters for the EigCG algorithm is given in Tab. 5. For the strange and charm quark propagators, we use the plain CG solver. We use a 10^{-8} stopping condition for all the CG and EigCG solvers.

The time separation between the kaon wall source and the $\Delta S = 1$ weak operators should be large enough to project onto kaon states. In the set up of this calculation, the two operators can be located at any time slice between $[4, 23]$. So the time separation between the kaon source or sink and either effective weak operator is guaranteed to be equal or larger than 4. In Fig. 13 we give a sample kaon effective mass plot for $m_l = 0.01$ and $m_s = 0.032$. This plot suggests that the effects of excited kaon states will be negligible when the separation between source and sink is 5 or larger. We therefore use the restricted range $[5, 22]$ for t_k in the following analysis, discarding the results when either operator is at the location $t_k = 4$ or 23 for $k = 1$ and 2.

In order to reduce short distance effects to a level which can be accurately controlled using lattice methods, we introduce a valence charm quark into our calculation. We will use six different charm quark masses in order to investigate the resulting GIM cancellation. These masses are given in Tab. 3, where we use the mass renormalization factor $Z_m^{\overline{MS}}(2$ Gev)=1.498 [24]. When we discuss the long distance effects in Sec. 5.3, we choose a 863 MeV valence charm quark mass and several different valence strange quark masses. The strange quark masses and corresponding kaon masses are given in Tab. 4. The up and down quark masses are kept at their unitary value, equal to the 0.01 mass of the sea quark.

5.2 Short Distance Contribution

In this section, we discuss the short distance contribution to our calculation of ΔM_K in detail. We begin by discussing results without a charm quark and their dependence on a

short distance, position-space cutoff. We then introduce a charm quark and examine the resulting GIM cancellation.

All the results presented in this section are for integrated correlators composed of the operator combination $Q_1 \cdot Q_1$, *i.e.* both four quark operators are Q_1 operators. (This case is presented for illustration since it is for this combination of operators that we have data which includes a short distance, position-space cutoff.) The results are the average of 600 configurations separated by 10 time units, with valence quark masses $m_l = 0.01$ and $m_s = 0.032$. The resulting pion and kaon masses are $m_\pi = 0.2431(8)$ and $m_K = 0.3252(7)$ respectively. The π^0 state is the only intermediate state lying below the kaon mass for these kinematics.

Quadratic divergence at short distance

In Eq. 3.7, we can see that the integrated correlator depends only on the separation between t_a and t_b which we defined earlier as $T = t_b - t_a + 1$, the number of discrete times lying in the interval $[t_a, t_b]$. For a given value of T, all $(t_a, t_b = t_a + T - 1)$ pairs which lie in the range $[5, 22]$ are possible choices of this integration interval. We calculate all of them and use the averaged result after normalization as the final definition of integrated correlator:

$$\overline{\mathcal{A}}(T; t_i, t_f) = \frac{1}{19 - T} \frac{e^{M_K(t_f - t_i)}}{N_K^2} \sum_{t_a=5}^{23-T} \mathcal{A}(t_a, t_b = t_a + T - 1; t_i, t_f). \tag{5.4}$$

In the top panel of Fig. 14, we plot the integrated correlator as a function of the integration time interval T. Here the valence charm quark is not included, so there is no GIM cancellation. There are two curves in this plot: the red squares correspond to the integrated correlator defined in Eq. 5.4, the blue diamonds represent the results after the exponentially growing π^0 term is removed. The π^0 contribution to the integrated correlator can be determined using Eq. 3.7, where the $\langle \pi^0 | H_W | K^0 \rangle$ matrix element is determined from a three point correlator calculation. Note that only the exponentially growing π^0 term and a constant term coming from the π^0 are removed; the π^0 contribution to the term proportional to

T is retained as required by Eq. 3.7. The top plot suggests that the exponentially growing π^0 term is only a small part of the result. This can be explained as follows. The integrated correlator receives contributions from all possible intermediate states. The short distance part, which comes from heavy intermediate states, is expected to be power divergent. The π^0 contribution, which is long distance physics, contains no such divergence and is small compared to the divergent short distance part even though it is exponentially growing with T.

To investigate the divergent character of short distance part in detail, we introduce an artificial position-space cutoff radius R. When we perform the double integration, we require the space-time separation between the positions of the two operators to be larger than or equal to this cutoff radius:

$$\sqrt{(t_2 - t_1)^2 + (\vec{x}_2 - \vec{x}_1)^2} \geq R \tag{5.5}$$

The bottom plot in Fig. 14 presents the result with a cutoff radius of 5. Comparing this plot with the left plot, we can see that the amplitude of the integrated correlator is reduced by a factor of approximately 10 and the exponentially growing π^0 term is now a very important part of the result which significantly changes the behavior of the correlator at long distance. All these observations suggest that the short distance contribution is substantially reduced after we impose the cutoff. We can also plot the mass difference ΔM_K as a function of this cutoff radius R. The mass difference on a finite lattice is defined in Eq. 3.8. However, we consider only the operator Q_1 here, so we define:

$$\Delta M_K^{11} = 2 \sum_{n \neq n_0} \frac{\langle \overline{K}^0 | Q_1^{uu} | n \rangle \langle n | Q_1^{uu} | K^0 \rangle}{M_K - E_n}, \tag{5.6}$$

where the superscript 11 means both operators are Q_1. This quantity is given by the slope of the coefficient of linear term in Eq. 3.7 when T is sufficiently large that the exponentially falling terms can be neglected. We choose to fit the slope of the integrated correlator in the range $9 \leq T \leq 18$. In Fig. 15 we show the dependence of ΔM_K^{11} on the cutoff radius R. The

blue curve is a naive uncorrelated two parameter fit:

$$\Delta M_K^{11}(R) = \frac{b}{R^2} + c\,, \tag{5.7}$$

where b and c are constants. The fitting result shows a convincing, power divergent short distance contribution.

Valence charm quark and GIM cancellation

The short distance contribution in a lattice calculation is necessarily unphysical, principally determined by the lattice cutoff. To control these short distance effects, we introduce a valence charm quark. The resulting GIM mechanism will then substantially reduce the short distance contribution. The implementation of the GIM cancellation in this calculation is quite straightforward. We simply replace the two internal up quark propagators in the contractions with the appropriate difference between up quark and charm quark propagators. We use six different valence charm quark masses which are given in Tab. 3. In Fig. 16 we plot the integrated $Q_1 \cdot Q_1$ correlator after GIM cancellation with a 863 MeV valence charm quark mass. We can compare this plot with those in Fig. 14. The behavior of the integrated correlator after GIM cancellation is quite similar to the result after introducing the artificial position-space cutoff. The GIM cancellation reduces the amplitude by approximately a factor of 10. Thus, as expected, the short distance contribution is substantially reduced by the GIM mechanism.

In Fig. 17, we plot the mass difference for different valence charm masses. The definition of the mass difference ΔM_K^{11} is similar to that given in Eq. 5.6, but the GIM cancellation is now included. The mass difference is obtained from the slope in T of the integrated correlator using the fitting range $T \in [9, 18]$. The values of ΔM_K^{11} are listed in the Tab. 6. The plot shows that the mass difference increases as the charm quark mass increases. This is expected since the cancellation between the up and charm quark propagators will be more complete for a lighter charm quark.

5.3 Long Distance Contribution

In this section we will examine the long distance contribution to our calculation of ΔM_K in detail. As we have discussed in Sec. 3.1, the intermediate states lying below the kaon mass will contribute terms which grow exponentially as the time interval T, over which the bi-local, second order weak interaction operators are integrated, is increased. These terms do not contribute to the physical mass difference ΔM_K and must be identified and removed. For physical quark masses such states include the vacuum, π^0, π-π and three π states. There is no vacuum state contribution in this work and for our kinematics the kaon mass is below the three-pion threshold. Thus, in the present calculation we are most interested in the π^0 and π-π intermediate states. The different parity of these two states allows us to study their contributions separately. Each left-left, $\Delta S = 1$ four quark operator can be separated into parity conserving and violating parts:

$$LL = (VV + AA) - (VA + AV). \tag{5.8}$$

The product of the two left-left operators can then be written as the sum of four terms:

$$\begin{aligned} LL \otimes LL = (VV + AA) \otimes (VV + AA) + (VA + AV) \otimes (VA + AV) \\ - (VV + AA) \otimes (VA + AV) - (VA + AV) \otimes (VV + AA). \end{aligned} \tag{5.9}$$

The third and fourth terms of Eq. 5.9 change the parity and hence cannot contribute to the matrix element between K^0 and $\overline{K^0}$ states. In the first term on the right-hand side of Eq. 5.9 both operators are parity conserving, which implies that the intermediate state must have odd parity. In the second term, both operators are parity violating, so the intermediate states have even parity. We can distinguish these two contributions and investigate the π^0 (parity odd) and π-π (parity even) intermediate states separately.

The integrated correlator receives contributions from both short and long distances. Therefore, in this section we examine the unintegrated correlators in Eq. 3.4, where we can explicitly study the case of large time separation between the two $\Delta S = 1$ operators.

The results presented in this section are for an average of 800 configurations separated by 10 time units, with a valence light quark mass $m_l = 0.01$ which corresponds to a pion mass $m_\pi = 0.2431(8)$ and eight valence strange quark masses whose values together with the corresponding kaon masses are given in Tab. 4.

Parity-odd channel

For this case, corresponding to the contribution of the first term in Eq. 5.9, both operators are parity conserving which implies that all intermediate states have odd parity. As can be seen from Eq. 3.5, in the limit of large time separation $|t_2 - t_1|$ the contribution from heavier states will decrease exponentially and only the lightest states will survive. For the parity-odd case this lightest state is the π^0 so that the unintegrated correlator becomes:

$$G(t_f, t_2, t_1, t_i) = N_K^2 e^{-M_K(t_f - t_i)} \langle \overline{K^0} | H_W | \pi^0 \rangle \langle \pi^0 | H_W | K^0 \rangle e^{-(M_\pi - M_K)|t_2 - t_1|}. \tag{5.10}$$

The unintegrated correlator only depends on the time separation $T_H = t_2 - t_1$ at given t_i and t_f. For a given value of T_H, all $(t_1, t_1 + T_H)$ pairs in the range $[5, 22]$ are possible choices. We compute all of them, take their average and remove the normalization factor N_K^2. The result is the unintegrated correlator $\overline{G}(T_H; t_f, t_i)$:

$$\overline{G}(T_H; t_f, t_i) = \frac{1}{t_f - t_i - 9 - T_H} \frac{e^{M_K(t_f - t_i)}}{N_K^2} \sum_{t_1 = t_i + 5}^{t_f - 5 - T_H} G(t_f, t_2 = t_1 + T, t_1, t_i)\,, \tag{5.11}$$

where we have adopted the order $t_2 > t_1$ and imposed the restriction $t_1 \geq t_i + 5$ and $t_f - 5 \geq t_2$.

We also compute the three point correlator needed to extract the matrix element $\langle \pi^0 | Q_i | K^0 \rangle$. We can then compare our lattice result for the unintegrated correlator given in Eq. 5.11 for large T_H with the contribution of a single π^0 shown in Eq. 5.10. The single-pion matrix elements are given in Tab. 7 for the set of 8 kaon masses. As we have explained in Sec. 5.1, we use $\pi^0 = i\bar{u}\gamma_5 u$ and only compute the first two diagrams shown in Fig. 11.

In Figs. 18-20, we plot the unintegrated correlators and resulting effective masses for the kaon mass $M_K = 0.4848(8)$. The three figures correspond to the different operator combinations: $Q_1 \cdot Q_1$, $Q_1 \cdot Q_2$ and $Q_2 \cdot Q_2$, respectively. In the plots of the unintegrated correlators we show both original results and the results after the subtraction of the π^0 contribution. This subtraction is done using the numerical results in Tab. 7. Since only the π^0 term should be present for large time separations, we expect that the results after subtraction should be consistent with zero for large T_H. In the effective mass plots, we calculate the effective mass $M_X - M_K$ from the unintegrated correlators, here M_X is the mass of the intermediate state. For this parity conserving case, the lightest state is the pion. The "exact" $M_\pi - M_K$ mass obtained from two point correlator calculation is shown in the plots as a blue horizontal line which agrees well with the computed effective mass. Although all three figures show the expected behavior, we find that the statistical errors seen for the different operator combinations are quite different. The operator combination $Q_1 \cdot Q_1$ has the smallest errors while $Q_2 \cdot Q_2$ has the largest.

In Fig. 21, we plot the intermediate state masses obtained from unintegrated correlators at eight different kaon masses for the $Q_1 \cdot Q_1$ case. The mass $M_X - M_K$ is obtained from a two parameter exponential fit and compared with the difference of M_K and M_π obtained directly from the two point correlators. The intermediate state mass agrees very well with the single pion mass for all choices of kaon mass.

Parity even channel

In this section, we examine the case where parity violating operators appear at both vertices. This requires that the intermediate states have even parity. The long distance behavior is expected to be dominated by the two-pion intermediate state, which is the lightest parity-even state.

In Figs. 22-24, we present the unintegrated correlators and the effective mass calculated

from type 2 diagrams for the three different products of parity violating operators evaluated at a kaon mass $M_K = 0.4848(4)$. This kaon mass is very close to the energy of two pions at rest, so we expect to get a plateau at large time separation T_H. However, our results are extremely noisy at long distance and we are not able to identify such a plateau. This large noise can be explained as follows. Although the signal should come from two-pion intermediate states, we will also have noise, whose size can be estimated from the square of the Green's functions being studied. In this squared Green's function the source and sink are composed of the product of two parity-violating operators and two kaon sources and sinks. Such a Green's function will receive a contribution from a two-pion intermediate state. The noise will fall with increasing separation $|t_2 - t_1|$ between the weak operators as the square root of this Green's function, implying that this noise will behave as $e^{-|t_2-t_1|m_\pi}$, dominating the two-pion signal which falls more rapidly as $e^{-|t_2-t_1|2m_\pi}$. Thus, the signal to noise ratio will fall exponentially for large time separation. The situation here is very similar to what is found for disconnected diagrams. This argument is consistent with our observation that most of the noise comes from type 1 diagrams, because the topology of type 2 diagrams does not allow a single-pion contribution to their noise. This argument is confirmed by plotting the results from type 2 contractions only.

If we analyze the type 2 diagrams alone, and fit the resulting intermediate state masses the results agree with the two-pion mass very well, as seen in the lower right panel of Fig. 25. However, the effective mass obtained from type 2 diagrams alone are not physical. We do not expect the effective mass shown in this last plot to be either the $I = 0$ or $I = 2$ finite volume $\pi - \pi$ energy. We view the agreement with $2m_\pi$ as coincidental.

5.4 The $K_L - K_S$ Mass Difference from the 16^3 Lattice

We now combine the lattice correlators with the Wilson coefficients given in Tab. 2 and determine the mass difference ΔM_K in physical units. The mass difference ΔM_K can be

obtained by fitting the integrated correlator in the limit that the integration region $[t_a, t_b]$ becomes large. We fit the dependence of the integrated correlator on $T = t_b - t_a + 1$ to a linear function over the range $9 \leq t_b - t_a \leq 18$. In Figs. 26-32, we show the computed values for the integrated correlator as a function of T and the corresponding effective slope plots for each of the three operator products $Q_1 \cdot Q_1$, $Q_1 \cdot Q_2$ and $Q_2 \cdot Q_2$. The three straight lines correspond to linear fits to the data points in the range $[9, 18]$. The χ^2/d.o.f given in the figure suggests that these fits describe the data well. Another method to check the quality of these fits is to plot the effective slope in analogy to the effective mass plots used when determining a mass from a correlation function. The effective slope at a given time T is calculated using the difference between the data points at $T - 1$ and $T + 1$. The horizontal lines with error bands give our final fitting results. The plateau at large T for these effective slope plots again suggests that these linear fits describe the data well. In Tab. 8, we give the lattice integrated correlators for $\Delta M_K = 563$ MeV. The fitting results for various kaon masses are given in Tab. 9. The lattice mass differences given in these table have a common factor 10^{-2} which is not shown. The errors given in the table are statistical only.

Although we have data for eight different kaon masses, we present results for only the seven kaon masses ranging from 563 MeV to 1162 MeV. We do not give results for the lightest kaon because it is degenerate with the pion while the standard formula for ΔM_K, which we are using, assumes that the K^0 and $\overline{K^0}$ are the only coupled, single-particle, degenerate states. While listed for completeness, the three heavier kaon masses of 918, 993 and 1162 MeV are more massive than the threshold two-pion intermediate state and will therefore contain an unknown, exponentially growing contamination which we have been unable to identify and remove.

Given our pion mass of 421 MeV, the two-pion intermediate state will be close to degenerate with the kaon for the $M_K = 839$ MeV case. Were we to follow the prescription proposed in Ref. [11] to control finite volume effects, we should choose this degenerate case and then

remove completely the contribution of the degenerate, two-pion intermediate state, which should appear in the integrated correlator with the time dependence $(t_b - t_a)^2$. However, as explained earlier, we are not able to identify the two-pion intermediate state within errors. This implies that the approximately on-shell, two-pion intermediate state contributes only a small part to the mass difference in our calculation and should have a small effect, at least on the results for 563 MeV $\leq M_K \leq$ 834 MeV.

5.5 Comparison with NLO Perturbative Calculation

A direct comparison between our results and the experimental value of ΔM_K has limited value because our kaon and pion masses are far from physical and we have not included all diagrams. However, we can learn something about the degree to which the present perturbative calculations describe ΔM_K for our unphysical kinematics by comparing our result with that obtained perturbatively by evaluating the perturbative formula at the kaon and pion masses used in our present calculation. While there are now results for ΔM_K computed at NNLO given in Ref. [8], complete expressions for the results are not given in that brief letter. Therefore, we choose to compare with the NLO result of Herrlich and Nierste [25] for which complete information is available in published form. Since the full results at NLO and NNLO orders differ by 36% at the physical point, the agreement with our result should be only approximate and this use of the NLO result adequate for our purpose. This comparison with NLO perturbation theory may also lessen the significance of our omission of disconnected diagrams, which do not appear at NLO. We will compare this NLO result, evaluated at our kinematics, with our lattice calculation carried out using 600 configurations at the unitary quark masses $m_l = 0.01$ and $m_s = 0.032$ ($M_\pi = 421$ MeV and $M_K = 563$ MeV) for a series of valence charm quark masses.

The mass difference in the perturbative calculation is given by:

$$\Delta M_K = \frac{G_F^2}{6\pi^2} f_K^2 \hat{B}_K M_K (\lambda - \frac{\lambda^3}{2})^2 \eta_1(\mu_c, m_c(\mu_c)) m_c^2(\mu_c), \tag{5.12}$$

which can be obtained, for example, from Eq. (12.1) in Ref. [13]. Here λ is the sine of the Cabibbo angle, one of the four Wolfenstein parameters entering the CKM matrix, μ_c is the scale at which the four-flavor theory is matched to that with three flavors and the kaon decay constant f_K is defined using conventions which make its physical value equal to 155 MeV. The two non-perturbative parameters, the kaon decay constant f_K and the kaon bag parameter $\hat{B}_K$, evaluated in the renormalization group invariant (RGI) scheme, can also be computed for the unphysical values of m_l and m_s listed above. For the present calculation we find it convenient to directly compute the matrix element of the left-left operator:

$$\langle O_{LL} \rangle = \langle \bar{K}^0 | (\bar{s}d)_{V-A} (\bar{s}d)_{V-A} | K^0 \rangle, \tag{5.13}$$

obtaining the value 0.00462(5) for $m_l = 0.01$ and $m_s = 0.032$. Here we use non-relativistic normalization for the kaon states: $\langle K(\vec{p}) | K(\vec{p}\,') \rangle = \delta^3(\vec{p} - \vec{p}\,')$. This lattice result can be converted to the RGI scheme by multiplying by the factor:

$$Z_{VV+AA}^{\mathrm{RGI}} = Z_{B_K}^{RGI} Z_A^2, \tag{5.14}$$

where $Z_{BK}^{\mathrm{RGI}} = 1.27$ and $Z_A = 0.7161$ are taken from Ref. [26].

The expression for the mass difference then becomes:

$$\Delta M_K = \frac{G_F^2}{8\pi^2} Z_{B_K}^{\mathrm{RGI}} Z_A^2 \langle O_{LL} \rangle (\lambda - \frac{\lambda^3}{2})^2 \eta_1 \left(\mu_c, m_c(\mu_c)\right) m_c^2(\mu_c). \tag{5.15}$$

Here the factors $Z_{B_K}^{\mathrm{RGI}} Z_A^2 \langle O_{LL} \rangle$ are lattice quantities determined for the kinematics studied here while η_1 is determined from the NLO perturbation theory calculation of Ref. [25], summarized in Ref. [13]. Specifically, Eq. (5.15) corresponds to the term in Eq. (12.1) of Ref. [13] containing η_1. Note, the two right-most factors in Eq. (12.1) do not appear in our Eq. (5.15) since they have been incorporated in $\hat{B}_K$, changing it to the RGI scheme. We

evaluate η_1 using Eq. (12.31) of Ref. [13]. We now compare this perturbative result with our non-perturbative, lattice calculation of the same box topology and for the same quark masses.

In our lattice calculation, we determine ΔM_K for a series of charm quark masses. We can exploit this mass dependence to attempt to separate the complete lattice result into short and long distance parts as follows. The dominant contribution to ΔM_K is proportional to the CKM matrix element product $|V_{cd}V_{cs}^*|^2$ and for large m_c grows as m_c^2 as is suggested by the perturbative result in Eq. 5.15. As is also implied by that equation, additional factors of $\ln(m_c^2)$ will appear in higher order perturbation theory. If ΔM_K is examined for $m_c \gg \Lambda_{\rm QCD}$, in addition to such $m_c^2 \ln^n(m_c^2)$ terms, we should also expect a constant piece, coming from long distance effects in which the charm quark mass plays a negligible role, with the remaining mass dependence behaving as $1/m_c^2$ for large m_c. As explained in the discussion of the GIM subtraction in Sec 3.5, the charm quark mass enters only as m_c^2 which implies there are no terms behaving as m_c or $1/m_c$. Note, the non-zero density of Dirac eigenvalues, $\rho(\lambda)$ at zero eigenvalue $\lambda = 0$ would induce a non-perturbative, chiral symmetry breaking m_c term in the limit of small m_c, but has no effect on the large m_c limit being considered here. This limit is determined only by the large λ behavior of ρ.

We use this large m_c expansion to parameterize the dependence of ΔM_K on m_c by adopting the ansatz:

$$\Delta M_K(m_c) = a + b\, m_c^2 + c\, m_c^2 \ln(m_c), \tag{5.16}$$

where we drop the possible $1/m_c^2$ term. The quadratic plus quadratic times logarithmic form of the terms with coefficients b and c can be found in the NLO perturbative expansion Eq. (5.15) if we use a fixed value of μ_c as m_c varies. Thus, the constants b and c are determined by short distance physics, arising from length scales of order $1/m_c$ and should be accessible to a perturbation theory calculation. In contrast the a term involves non-perturbative phenomena and long distances. The perturbative calculation also contains a

long distance part which contributes to the constant a. However, this is suppressed by a factor of $(m_{ud}/m_c)^2$ which is at most 0.5% for our lightest charm quark mass.

In Fig. 33 we plot our results for ΔM_K as a function of the charm quark mass as well as the result from the fit to the ansatz given in Eq. (5.16). The upper solid curve shows the entire fitting function given in Eq. (5.16) while the lower solid curve has the non-perturbative terms proportional to a removed. A comparison of these two solid curves in Fig. 33 suggests that for unphysically massive $M_\pi = 421$ MeV and $M_K = 563$ MeV and a charm quark mass of 1.2 GeV, approximately 50% of ΔM_K comes from long-distance effects.

Chapter 6

Results from the $24^3 \times 64 \times 16$ Lattice

In this chapter, we will give the results from the $24^3 \times 64 \times 16$ lattice ensemble. Section. 6.1 describes the details of the simulation. Section. 6.2 gives the mass difference calculated on this lattice. In Section. 6.3, we discuss the contribution to the mass difference from the different diagrams. In Section. 6.4, we discuss the effects of the scalar and pseudo-scalar operators in our calculation.

6.1 Simulation Details

The 24^3 lattice ensemble is similar to the 16^3 lattice ensemble described in the last chapter except for a larger volume and a lighter pion mass. The sea light and strange quark masses are $m_l = 0.005$ and $m_s = 0.04$ respectively, corresponding to a 330 MeV pion mass and a 575 MeV kaon mass. A valence charm quark mass $m_c = 0.363$ is used to implement GIM cancellation, corresponding to $m_c^{\overline{\rm MS}}(2{\rm GeV}) = 949$. We use 800 configurations, each separated by 10 time units.

We will refer to Fig. 4 to explain the setup of this calculation. All the diagrams are included in this unitary calculation. For the kaon sources, we continue to use Coulomb-gauge-fixed wall sources. However, unlike the 16^3 calculation, the locations of the two kaons

are not fixed. All the diagrams are averaged over all time translations to increase statistics. We also vary the separation between the two kaons in this calculation. This separation should be sufficient large so that the interaction range is large enough to extract the mass difference. Since we are not using periodic plus anti-periodic boundary condition here, this separation should not be too large so that we don't need to worry about the around-the-world effects. In our calculation, we vary this separation from 24 lattice units to 31 lattice units. The separation between the kaon sources and the effective Hamiltonian densities should be sufficiently large to suppress the contribution from the excited kaons. We choose this separation to be equal or larger than 6 lattice units in our calculation.

The evaluation of the four point correlators are explained in Section. 4.4. For type 1 and type 2 diagrams, 64 propagators are computed using a point source on each of the 64 time slices. For type 3 and type 4 diagrams, we use 64 random wall source propagators to construct the quark loops. In order to reduce the noise coming from the random numbers, we use 6 sets of random sources for each time slice, color and spin. For the light quark propagators, which make up the most expensive part of this calculation, we calculate the lowest 300 eigenvectors of the Dirac operator using the Lanczos method and get a factor of 8 speed up by using mix precision CG solver with low mode deflation. The parameters for the Lanczos algorithm are given in Tab. 11. For the strange and charm quarks, we use the mix precision CG solver. We use a 10^{-5} stopping condition for the inner single precision CG solvers and use a 10^{-8} stopping condition for the outer CG solvers.

6.2 The $K_L - K_S$ Mass Difference from the 24^3 Lattice

In this unitary calculation, the integrated correlators will receive exponentially increasing contributions from both the π^0 and vacuum intermediate states. We evaluate the matrix element $\langle \pi^0 | H_W | K^0 \rangle$ and subtract the single pion exponentially increasing term explicitly from Eq.3.7. For the vacuum state, we add a pseudo-scalar density $\bar{s}\gamma^5 d$ to the weak Hamil-

tonian to eliminate the matrix element $\langle 0|H_W + c_2\bar{s}\gamma^5 d|K^0\rangle$. The details of this subtraction has been discussed in Section. 3.1. We also perform a subtraction for the exponentially decreasing term from η state. We have a 624 MeV η meson mass in our calculation, which is slightly heavier than the 575 MeV kaon mass. The exponential term from η state will decreasing slowly and may contaminate the linear behavior at large T. So we subtract this term by using the $\langle\eta|H_W|K^0\rangle$ matrix element calculated from three point correlators. This has a less than 10% effect on the final result. The $(\pi\pi)_{I=0}$ state energy is 629 MeV for our kinematics [27]. The exponential term from $(\pi\pi)_{I=0}$ state will also be decreasing slowly. We are not able to subtract this term since we don't have the date for the $K \to (\pi\pi)_{I=0}$ matrix element. This is a source of the systematic errors in our calculation. However, the χ^2/d.o.f values for our linear fits suggest this term gives small contribution.

In Fig. 34, we show the the integrated correlators and the corresponding effective mass plots for the case in which the kaon sources are separated by 31 lattice units. The three curves correspond to the three different operator combinations: $Q_1 \cdot Q_1$, $Q_1 \cdot Q_2$ and $Q_2 \cdot Q_2$. The numbers are bare lattice results without any Wilson coefficients or renormalization factors. All the exponentially increasing terms have been removed from the correlators, so we expect a linear behavior for sufficiently large T. When T becomes too large, the errors explode as should be expected since the disconnected diagrams have an exponentially decreasing signal-to-noise ratio. The straight lines correspond to linear fits to the data points in the range $[7, 20]$. The χ^2/d.o.f given in the figure suggest that these fits describe the data well. The effective slope is calculated using a method different from the method in Section. 5.4. At a given time T, the effective slope is calculated using a correlated linear fit to three data points at $T-1$, T and $T+1$. For each operator combination we get good plateaus starting from $T = 7$. In Tab. 12, we give the lattice integrated correlators without any Wilson coefficients for the three operator products. The fitting results for the mass differences are given in Tab. 13. We give both the bare lattice numbers without Wilson coefficients and the physical

values with Wilson coefficients.

We have the results for the integrated correlators for two kaon separation $\Delta_{\rm KK}$ from 24 lattice units to 31 lattice units. In Tab. 14, we give the fitting results from difference choices of $\Delta_{\rm KK}$. We choose a minimum fitting time $T_{\rm min} = 7$ for all the fits. The fitting results from different choices of $\Delta_{\rm KK}$ have similar central values and error-bars. The error-weighted-average given in the last row suggest that combining these results dose not improve the errors. Hence we conclude that these results are highly correlated and we will not benefit from a combined analysis. So we will focus on the $\Delta_{\rm KK} = 31$ case in the following analysis.

We have also tried different fitting ranges to make sure that our results do not depend sensitively on these choices. We varied two parameters: the lower limit on the linear fitting range $T_{\rm min}$ and the minimum separation between the kaon sources and weak Hamiltonians $\Delta_{\rm min}$. We first fixed $\Delta_K = 6$ and varied $T_{\rm min}$ from 7 to 9. The results are given in Table. 15. While the central value of the fitting results are quite stable, the errors are sensitive to the choice of $T_{\rm min}$, which is caused by the exponentially decreasing signal-to-noise ratio of disconnected diagrams. We vary $\Delta_{\rm min}$ to check whether or not there are contaminations from excited kaon states. We give the results in Table. 16 for fixed $T_{\rm min} = 7$ but with $\Delta_{\rm min}$ varying from 6 to 8. Both the central values and the errors are very stable, suggesting that a separation of 6 is large enough to suppress excited kaon states.

6.3 Contributions from Different Diagrams

In our 16^3 calculation, only the first two types of diagrams were included in the calculation. We can now determine the accuracy of this approximation by calculating the contribution from each type of diagrams separately. The subtraction of π^0 exponentially increasing term will become more complicated if we analyze different types of diagrams separately. We perform a subtraction for each type of diagrams separately following a procedure similar to our previous work [**?**]. For the vacuum state, we don't have such complication since the

vacuum state contribution only comes from type 4 diagrams. Adding a $\bar{s}\gamma^5 d$ term will remove the exponentially increasing term for type 4 diagrams and will not cause any other problem. We perform a subtraction for the slow decaying term from η state in the full analysis. In a separate analysis, we are not able to separate the η contribution into different types of diagram in a reasonable way. So we don't perform any subtraction for η state here.

In Fig. 35-38, we give the integrated correlators and the corresponding diagrams for the four different types of diagrams. For type 1 and type 4 diagrams, we obtain good linear behavior at the large T range and the χ^2/d.o.f are also close to 1. Although the integrated correlators for type 2 diagrams appears to be very linear, the χ^2/d.o.f values are very large. For type 3 diagrams, the effective slope plots suggest that there are some non-linear behavior at large T. This unexpected behavior can be caused by excited kaon states. We can increase $\Delta_{\min}$ to suppress the contribution from excited kaon states. In Fig. 39-42, we give the results for the choice of $\Delta_{\min} = 8$. The effective slopes from the type 3 diagrams do become better. The large χ^2/d.o.f for type 2 diagrams improve but are still not perfect. We should keep in mind that the results from each individual type of diagram are not physical and we are not guaranteed to obtain a linear behavior for each type of diagram separately.

In Tab. 17, we give ΔM_K from different types of diagrams. We choose $\Delta_{\min} = 8$ for all the fittings. If we focus on the final results for ΔM_K, the contribution from type 1 and type 3 diagrams are relatively small. There is a large cancellation between the disconnected (type 4) diagrams and other type of diagrams. The mass difference is reduced by almost a factor of two after the inclusion of the disconnected diagrams. This is a surprisingly large failure of the"OZI suppression" [21, 22, 23], naively expected for these disconnected diagrams. Although this result may be surprising, we should always keep in mind that the mass difference from individual types of diagrams is not a physical quantity and we should not put too much emphasis on these results.

6.4 Effects of Scalar and Pseudoscalar Operators

In this calculation, we remove the π^0 exponentially increasing term explicitly and add a pseudoscalar density to the Hamiltonian to eliminate the vacuum states. We can also add a scalar density to the Hamiltonian to eliminate the π^0 state so that we do not need to perform the subtraction. In Section. 3.1, we claim that adding these operators will not change the final results. We will examine the effects of the scalar and the pseudo-scalar operators numerically in this section.

We can rewrite the Hamiltonian in Eq. 3.11 as:

$$H'_W = H_W + c_s c_\pi \bar{s}d + c_p c_{\rm vac} \bar{s}\gamma^5 d, \tag{6.1}$$

where c_π and $c_{\rm vac}$ are defined in Eq. 3.15 and c_s and c_p are dimensionless numbers. In Sec. 6.2, we choose $c_s = 0$ and $c_p = 1$. If we choose other value for c_s and c_p, the resulting exponentially increasing terms can be determined and subtracted by using the matrix elements $\langle \pi^0 | H'_W | K^0 \rangle$ and $\langle 0 | H'_W | K^0 \rangle$. In the next few paragraphs, we will vary c_s and c_p and determine whether or not the mass difference will depend on the choice of these values.

We will first vary the value of c_s while the value of c_p is fixed to 1. Generally, the mass differences can have both linear and quadratic dependence on c_s. The linear dependence comes from the operator products $Q_i \cdot \bar{s}d$ and the quadratic dependence comes from the operator product $\bar{s}d \cdot \bar{s}d$. In Fig. 43, we give the integrated correlator for these operator products. We expect these plots have a zero slope at large T range since $\bar{s}d$ operator should not contribute to the mass difference. The results are within our expectation.

In Figure. 47, the mass differences are plotted for c_s value from -1 to 1. The plots for the integrated correlators and the corresponding effective slope plots for $c_s = 1$ and $c_s = -1$ are given in Figs. 44-45. We can also tune the value of c_s to eliminate the η intermediate state. The corresponding results are given in Fig.46. In Tab. 18, we give the mass differences for various choices of c_s. We conclude that the mass differences are independent of the choices

of the value of c_s.

Next we will vary the value of c_p while the value of c_s is fixed to 0. The first difficulty here is the subtraction of the exponentially increasing term coming from the vacuum state when $c_p \neq 1$. This term is two orders of magnitude larger than the linear term if we choose $c_p = 0$. For the subtraction of the π^0 exponentially increasing term, we use the formula given in Eq. 3.7. However, this formula uses a second order integration instead of a second order sum. The more precise formula for the vacuum exponential term is given by:

$$N_K^2 e^{-M_K(t_f-t_i)} \langle \overline{K}^0 | H_W | 0 \rangle \langle 0 | H_W | K^0 \rangle \frac{e^{M_K T} - 1}{4\mathrm{Sinh}^2(M_K/2)}. \tag{6.2}$$

Comparing with Eq. 3.7, the M_K^2 in the denominator is changed to $4\mathrm{sinh}^2(M_K/2)$, which is a $O(a^2)$ correction. For $M_K = 0.3325$, there is a 1% differnce between these two formula. Since we are trying to subtract a term which is two orders of magnitude larger than the signal, this is a necessary correction here.

In the first plot in Fig. 48, we give the integrated correlator for $c_p = 0$, $c_s = 1$ and $\Delta_{\min} = 6$. We can see that the integrated correlators at large T are not linear. A possible cause of this behavior is a contribution from excited kaon states. Our signal is supposed to behave as $e^{-M_K(t_f-t_i)}$. The exponentially increasing vacuum term receives its largest contribution when the vacuum travels the entire interaction range T and behaves like $e^{-M_K(t_f-t_i-T)}$. This is $e^{m_K T}$ larger than the signal and needs to be removed. For the $c_p = 1$ case, the vacuum state will not couple to kaon ground state and the contribution from the excited kaon state K^* behaves as $e^{-M_{K^*}(t_f-t_i-T)}$. We have repeated the calculation for different values of $t_f - t_i - T$ and the results are consistent. So the excited kaon state is likely not important if $c_p = 1$. However, if we consider $c_p = 0$, we can have a kaon on onside and an excited kaon on the other side. The excited kaon contribution will behave like $e^{-M_K(t_f-t_i-T)/2} e^{-M_{K^*}(t_f-t_i-T)/2}$, which for large T grows exponentially larger than the signal which behaves as $e^{-m_K(t_f-t_i)}$ and may obscure the expected linear large T behavior.

In the second plot in Fig. 48, we increase $\Delta_{\min}$ from 6 to 10. Although the integrated

correlator appears to be better, the results are actually not changed much if we just compare the integrated correlator in the time interval $[7, 12]$. This is confirmed by the fitting results given in Tab. 19. The results we get for two $\Delta_{\min}$ values are consistent and are quite different from the mass difference from $c_p = 1$ data. This difference is likely due to some uncontrolled systematic effects introduced when we try to remove the exponentially increasing term from the $c_p = 0$ data. In Tab. 20, we can see that the mass difference comes from the vacuum state when is 10 times larger than the final mass difference. This large contribution is mostly non-physical and coming from the mixing between the Hamiltonian and the $\bar{s}\gamma^5 d$ operator. If we can control every aspect of this calculation, this large contribution will be canceled by other unphysical contributions and the final results should be consistent with $c_p = 1$ results. However, it seems that we can only remove the vacuum contribution to 10% precision and hence the mass difference for $c_p = 0$ case is completely different from the $c_p = 1$ results. In Fig. 49, we give the mass difference for different values of c_p while c_s is fixed to 1. The mass difference will change rapidly when c_p deviates from 1. ΔM_{11}'s dependence on c_p is mild while ΔM_{22}'s dependence is the strongest among the three operator products. This is consistent with the fact that the contribution to ΔM_{22} from the vacuum state is 30 times larger than the vacuum contribution to ΔM_{11}. So we conclude that there are some uncontrolled systematic effect from the subtraction of vacuum states and we are not able to obtain a reliable value for ΔM_K if c_p is too far away from 1. If we choose $c_p = 1$, there will be no contribution from vacuum states and we will not have this trouble.

Chapter 7

Conclusions

In the previous two chapters, we showed detailed results for ΔM_K from a $16^3 \times 32 \times 16$ lattice ensemble and a $24^3 \times 64 \times 16$ lattice ensemble. These calculations are all done with a heavy pion mass and with a quenched charm quark on a relatively coarse lattice. We also neglect the type 3 and 4 diagrams in the 16^3 lattice calculation. However, as a pioneering calculation on this difficult problem with four-point correlators and disconnected diagrams, our results have shown the success for application of the lattice QCD as a first-principal method to compute a second-order weak processes.

In the 24^3 full calculation, we compute ΔM_K for a case of unphysical kinematics with pion, kaon and charmed quark masses of 330, 575 and 949 MeV respectively, each quite different from their physical values of 135, 495 and 1100 MeV. Our results is:

$$\Delta M_K = 3.19(41)(96) \times 10^{-12} MeV. \tag{7.1}$$

Here the first error is statistical and the second an estimate of largest systematic error, the discretization error which results from including a 949 MeV charm quark in a calculation using an inverse lattice spacing $1/a = 1.73$ GeV. This 30% estimate for the discretization error can be obtained either by simple power counting, $(m_c a)^2 = 30\%$, or from the failure of the calculated energy of the η_c meson to satisfy the relativistic dispersion relation. We find

$(E^2(p) - m^2)/p^2 = 0.740(3)$ instead of 1.0 when evaluated at $p = 2\pi/L$. Our result agrees well with the experimental value of $3.483(6) \times 10^{-12}$ MeV. However, since we are not using physical kinematics, this agreement could be easily fortuitous.

One further potential difficulty which should be discussed associated with our inclusion of a heavy charm quark is a phenomena particular to the domain wall fermion formalism. When the input lattice quark mass, m_f, becomes large, about 0.4 or larger in lattice units, the eigenvalues of the physical, 4-dimensional modes become as large as the smallest eigenvalues of unphysical, 5-dimensional modes [28, 29]. These 5-dimensional modes could add a further, unphysical contribution to the GIM cancellation. We have examined the propagator for the η_c state and determined that our value of $m_f = 0.363$ lies slightly below the threshold at which there is significant coupling to these bulk, 5-dimensional states.

This is illustrated in Fig. 50, where we show $f_{q\bar{q}}$ as a function of m_f. The quanity f_{qq} is the product of the pseudoscalar decay constant and the normalization factor for the gauge fixed wall source. Here one sees a monotonic increase in $f_{q\bar{q}}$ as m_f increases up to a value of $m_f \approx 0.4$, above which $f_{q\bar{q}}$ turns over and decreases with increasing m_f, suggesting the appearance of propagating 5-dimensional modes, which necessarily have a substantially reduced coupling to operators with support on the 4-dimensional $s = 0$ and $s = L_s - 1$ walls. A significant contribution from such unphysical, 5-dimension modes would add further lattice artifacts to the GIM cancellation that controls the largest component of ΔM_K. One of the benefits of carrying out future calculations at smaller lattice spacing will be the further reduction of the possibility of such contributions.

The disconnected diagrams contribute a large amount to our result for ΔM_K. This is a surprising result and suggests the failure of "OZI suppression" in the $K_L - K_S$ mass difference. The OZI rule argues that the disconnect diagrams are connected by hard gluons and hence suppressed by a small gauge field coupling constant. However, there is no strict theoretical basis for this rule. We view our result as strong evidence that the OZI rule is not

solid and may be unreliable in some cases.

To perform a calculation with physical kinematics and controlled systematic errors, two difficulties must be overcome. First, we need to perform the calculation on a four-flavor lattice ensemble with two or more, smaller lattice spacings. This would remove the difficult-to-estimate error associated with quenching the charm quark and allow the $O(m_c^2 a^2)$ discretization errors to be removed. Second, we must perform a finite volume correction associated with $\pi - \pi$ re-scattering which will be needed for physical kinematics, when the two-pion threshold lies below the kaon mass. In this case, ΔM_K in infinite volume contains the principal part of the integral over the two-pion relative momentum, which can be substantially different from a finite-volume momentum sum. A generalization of the Lellouch-Luscher method has been devised to correct this potentially large finite volume effect [11] and a more general method has been proposed in Ref. [30].

Similar techniques can be used to determine the long distance contribution to ϵ_K. However, the calculation of ϵ_K involves two additional complexities described in Appendix A of Ref. [31]. First, we must introduce new QCD penguin operators representing top quark effects. Second, an overall, logarithmic divergence must be removed from the lattice calculation using non-perturbative methods. In summary, a full calculation of ΔM_K and ϵ_K, including their long distance contributions, should be accessible to lattice QCD with controlled systematic errors within a few years, substantially increasing the importance of these quantities in the search for new phenomena beyond the standard model.

Appendix A

K^0-$\overline{K}^0$ Mixing Contractions

As we discussed in Section 3.2, there are 16 contractions in total that are shown in Figure 5-8 and we label them with numbers from 1 to 16. We have explicitly explained a few examples already. In this appendix, we write down the specific expressions for all the diagrams. For simplicity, we only give the contractions for the case that both internal quarks are up quarks, the other cases are very similar. It is convenient to define the following quantities:

$$\begin{aligned}
\text{LW1X} =& \gamma^\mu(1-\gamma^5)L^W(x,t_i) \\
\text{SW1X} =& \gamma^5\gamma^5 S^{W\dagger}(x,t_i)\gamma^5 \\
\text{SW1Y} =& \gamma^5\gamma^5 S^{W\dagger}(y,t_i)\gamma^5 \\
\text{LW2Y} =& \gamma^\nu(1-\gamma^5)L^W(y,t_f) \\
\text{SW2X} =& \gamma^5\gamma^5 S^{W\dagger}(x,t_f) \\
\text{SW2Y} =& \gamma^5\gamma^5 S^{W\dagger}(y,t_f) \\
\text{LXY} =& \gamma^\mu(1-\gamma^5)L(x,y) \\
\text{LYX} =& \gamma^\nu(1-\gamma^5)L(y,x) \\
\text{LOOP1} =& L^{\text{RW}}(x,t_x)\eta^\dagger(x) \\
\text{LOOP2} =& L^{\text{RW}}(y,t_y)\eta^\dagger(y)
\end{aligned}$$

In these expressions, $L(x,y)$ is a point source propagator from y to x. $L^W(x,t)$ is a light quark propagator with Coulomb gauge fixed wall source at time t and sink at space-time position x. $S^W(x,t)$ is similar but for strange quark. $L^{\mathrm{RW}}(x,t)$ is a light quark propagator with Coulomb gauge fixed wall source at time t and sink at space-time position x. With these definitions, the contractions ①-⑯ can be written down easily.

① =Tr(LW1X · SW1X)Tr(LXY · LYX)Tr(LW2Y · SW2Y)

② =Tr(LW1X · SW1X · LXY · LW2Y · SW2Y · LYX)

③ =Tr(LW1X · SW1X · LXY · LYX)Tr(LW2Y · SW2Y)

④ =Tr(LW1X · SW1X)Tr(LYX · LXY · LW2Y · SW2Y)

⑤ =Tr(LW1X · SW2Y · LW2Y · SW2X)Tr(LXY · LYX)

⑥ =Tr(LW1X · SW2Y · LYX)Tr(LXY · LW2Y · SW2X)

⑦ =Tr(LW1X · SW2Y · LW2Y · SW2X · LXY · LYX)

⑧ =Tr(LW1X · SW2Y · LYX · LXY · LW2Y · SW2X)

⑨ =Tr(LW1X · SW2Y · LW2Y · SW2X)Tr(LOOP1)Tr(LOOP2)

⑩ =Tr(LW1X · SW2Y · LOOP2 · LW2Y · SW2X · LOOP1)

⑪ =Tr(LW1X · SW2Y · LW2Y · SW2X · LOOP1)Tr(LOOP2)

⑫ =Tr(LW1X · SW2Y · LOOP2 · LW2Y · SW2X)Tr(LOOP1)

⑬ =Tr(LW1X · SW1X)Tr(LOOP1)Tr(LW2Y · SW2Y)Tr(LOOP2)

⑭ =Tr(LW1X · SW1X · LOOP1)Tr(LW2Y · SW2Y · LOOP2)

⑮ =Tr(LW1X · SW1X · LOOP1)Tr(LW2Y · SW2Y)Tr(LOOP2)

⑯ =Tr(LW1X · SW1X)Tr(LOOP1)Tr(LW2Y · SW2Y · LOOP2)

Bibliography

[1] A. Stathopoulos and K. Orginos, SIAM J.Sci.Comput. **32**, 439 (2010), 0707.0131.

[2] Q. Liu (2011), 1110.2143.

[3] R. Arthur (2012), Ph.D. thesis, University of Edinburgh, unpublished.

[4] K. Nakamura *et al.* (Particle Data Group), J.Phys.G **G37**, 075021 (2010).

[5] R. Mohapatra, J. S. Rao, and R. Marshak, Phys.Rev. **171**, 1502 (1968).

[6] S. Glashow, J. Iliopoulos, and L. Maiani, Phys.Rev. **D2**, 1285 (1970).

[7] M. Gaillard and B. W. Lee, Phys.Rev. **D10**, 897 (1974).

[8] J. Brod and M. Gorbahn, Phys.Rev.Lett. **108**, 121801 (2012), 1108.2036.

[9] J. F. Donoghue, E. Golowich, and B. R. Holstein, Phys.Lett. **B135**, 481 (1984).

[10] A. J. Buras, D. Guadagnoli, and G. Isidori, Phys. Lett. **B688**, 309 (2010), 1002.3612.

[11] N. H. Christ (RBC and UKQCD Collaborations), PoS **LATTICE2010**, 300 (2010).

[12] L. Lellouch and M. Luscher, Commun. Math. Phys. **219**, 31 (2001), hep-lat/0003023.

[13] G. Buchalla, A. J. Buras, and M. E. Lautenbacher, Rev.Mod.Phys. **68**, 1125 (1996), hep-ph/9512380.

[14] L. Giusti, P. Hernandez, M. Laine, P. Weisz, and H. Wittig, JHEP **0411**, 016 (2004), hep-lat/0407007.

[15] C. Lehner and C. Sturm, Phys.Rev. **D84**, 014001 (2011), 1104.4948.

[16] T. Blum *et al.* (RBC Collaboration), Phys.Rev. **D68**, 114506 (2003), hep-lat/0110075.

[17] C. Sturm, Y. Aoki, N. Christ, T. Izubuchi, C. Sachrajda, *et al.*, Phys.Rev. **D80**, 014501 (2009), 0901.2599.

[18] T. Blum, P. Boyle, N. Christ, N. Garron, E. Goode, *et al.*, Phys.Rev. **D84**, 114503 (2011), 40 pages, 12 figures, 1106.2714.

[19] M. Luscher, Nucl.Phys. **B354**, 531 (1991).

[20] C. Allton *et al.* (RBC and UKQCD), Phys. Rev. **D76**, 014504 (2007), hep-lat/0701013.

[21] G. Zweig (1964).

[22] S. Okubo, Phys.Lett. **5**, 165 (1963).

[23] J. Iizuka, Prog.Theor.Phys.Suppl. **37**, 21 (1966).

[24] Y. Aoki *et al.* (RBC Collaboration, UKQCD Collaboration), Phys.Rev. **D83**, 074508 (2011), 1011.0892.

[25] S. Herrlich and U. Nierste, Nucl. Phys. **B419**, 292 (1994), hep-ph/9310311.

[26] C. Allton *et al.* (RBC-UKQCD Collaboration), Phys.Rev. **D78**, 114509 (2008), 0804.0473.

[27] Q. Liu (2012), Ph.D. thesis, Columbia University, unpublished.

[28] N. H. Christ and G. Liu, Nucl. Phys. Proc. Suppl. **129**, 272 (2004).

[29] G. Liu (2003), Ph.D. thesis, Columbia University, unpublished.

[30] N. H. Christ, G. Martinelli, and C. T. Sachrajda, PoS **LATTICE2013**, 399 (2014), 1401.1362.

[31] N. Christ, T. Izubuchi, C. Sachrajda, A. Soni, and J. Yu (2012), 1212.5931.

Tables

Table 1: The values of the parameters used in the calculation of Wilson coefficients.

$\alpha_s(m_Z)$	0.1184
m_Z	91.1876 GeV
m_W	80.399 GeV
$m_b(m_b)$	4.19 GeV
$m_c(m_c)$	1.27 GeV

Table 2: The Wilson coefficients, the RI $\rightarrow$ $\overline{\text{MS}}$ matching matrix, the non-perturbative lat $\rightarrow$ RI operator renormalization matrix and their final product, all at a scale $\mu = 2.15$ GeV shown in columns one through four respectively.

$C_1^{\overline{\text{MS}}}$	$C_2^{\overline{\text{MS}}}$	$\Delta r_{11} = \Delta r_{22}$	$\Delta r_{12} = \Delta r_{21}$	$Z_{11} = Z_{22}$	$Z_{12} = Z_{21}$	C_1^{lat}	C_2^{lat}
-0.2967	1.1385	-6.562×10^{-2}	7.521×10^{-3}	0.5916	-0.05901	-0.2216	0.6439

Table 3: Valence charm quark masses used to implement the GIM cancellation on the 16^3 lattice. The upper row gives the bare masses in lattice units. The lower row contains the $\overline{\text{MS}}$ masses at a scale of 2 GeV.

m_c	0.132	0.165	0.198	0.231	0.264	0.330
m_c (MeV)	350	435	521	606	692	863

Table 4: Valence strange quark mass (upper row) and kaon mass (lower row) on the 16^3 lattice, both in lattice units.

m_s	0.01	0.032	0.06	0.075	0.09	0.11	0.13	0.18
M_K	0.2431(8)	0.3252(7)	0.4087(7)	0.4480(7)	0.4848(8)	0.5307(8)	0.5738(8)	0.6721(10)

Table 5: The values of the parameters used in the EigCG algorithm on the 16^3 lattice. The meaning of these parameters are given in Ref. [1, 2].

nev	8
m	24
max_def_len	100
max_eig_cut	1
restart	10^{-5}

Table 6: The mass difference ΔM_K^{11}, defined in Eq. (5.6) after GIM cancellation, evaluated for different charm quark masses. These results were obtained on the 16^3 lattice, use 600 configurations and a kaon mass of 563 MeV and are the matrix elements of bare lattice operators without Wilson coefficients or renormalization factors.

m_c (MeV)	350	435	521	606	692	863
$\Delta M_K^{11,GIM}$	0.0452(13)	0.0481(14)	0.0511(15)	0.0542(15)	0.0575(16)	0.0647(18)

Table 7: Results for single-pion matrix elements, $\langle \pi^0 | Q_i | K^0 \rangle$, at various kaon masses on the 16^3 lattice. We use $\pi^0 = i\bar{u}\gamma_5 u$ and only include the first two diagrams in Fig. 11.

M_K	$\langle \pi^0 \vert Q_1^{uu} \vert K^0 \rangle$	$\langle \pi^0 \vert Q_2^{uu} \vert K^0 \rangle$
0.2431(8)	0.02107(29)	-0.00779(26)
0.3252(7)	0.02729(30)	-0.00954(23)
0.4087(7)	0.03300(33)	-0.01067(22)
0.4480(7)	0.03550(35)	-0.01103(22)
0.4848(8)	0.03773(36)	-0.01128(22)
0.5307(8)	0.04037(39)	-0.01149(22)
0.5738(8)	0.04271(42)	-0.01160(23)
0.6721(10)	0.04753(49)	-0.01156(25)

Table 8: Results for the integrated correlators for $M_k = 563$ MeV and $m_c = 863$ MeV on the 16^3 lattice. The quantities in columns two through four are the simple lattice integrated correlators of the operator products $Q_i^{qq'} Q_j^{q'q}$ for each $i, j = 1, 2$, summed over the four values of $q, q' = u, c$, without Wilson coefficients or renormalization factors and have been scaled to remove a factor 10^{-2}.

T	$Q_1 \cdot Q_1$	$Q_1 \cdot Q_2$	$Q_2 \cdot Q_2$
1	-1.422(26)	0.662(23)	-0.1641(68)
2	-3.604(66)	1.450(51)	-0.474(20)
3	-6.23(12)	2.401(87)	-0.937(39)
4	-9.13(17)	3.48(13)	-1.514(66)
5	-12.18(24)	4.67(18)	-2.172(97)
6	-15.31(30)	5.93(23)	-2.89(12)
7	-18.49(36)	7.22(29)	-3.63(16)
8	-21.69(43)	8.54(35)	-4.40(20)
9	-24.88(49)	9.84(41)	-5.18(24)
10	-28.07(56)	11.14(49)	-5.97(28)
11	-31.24(62)	12.39(58)	-6.74(32)
12	-34.42(69)	13.62(68)	-7.49(36)
13	-37.61(75)	14.88(79)	-8.24(41)
14	-40.79(82)	16.11(91)	-8.96(46)
15	-43.95(90)	17.3(10)	-9.65(51)
16	-47.07(99)	18.5(12)	-10.32(57)
17	-50.2(11)	19.6(14)	-10.97(64)
18	-53.3(12)	20.7(16)	-11.63(70)

Table 9: The contribution of the three operator products evaluated here to the mass difference ΔM_K for the seven different choices of the kaon mass listed in the first column in MeV. The quantities in columns two through four are the simple lattice matrix elements of the operator products $Q_i^{qq'}Q_j^{q'q}$ for each $i, j = 1, 2$, summed over the four values of $q, q' = u, c$, without Wilson coefficients or renormalization factors and have been scaled to remove a factor 10^{-2}. These results are obtained from a fitting range [9,18]. The final column gives the complete contribution to ΔM_K, expressed in physical units. The results for the three largest values of the kaon mass are contaminated by an unknown, exponentially growing two-pion contribution which we have been unable to identify and subtract but are given here for completeness. These results come from the 16^3 lattice and use 800 configurations and a charm quark mass of 863 MeV.

M_K (MeV)	ΔM_K^{11}	ΔM_K^{12}	ΔM_K^{22}	ΔM_K ($\times 10^{-12}$ MeV)
563	6.42(15)	-2.77(16)	1.56(9)	6.58(30)
707	8.94(23)	-3.16(27)	2.26(14)	8.85(48)
775	10.65(29)	-3.49(35)	2.67(18)	10.32(62)
834	12.55(37)	-3.84(46)	3.11(24)	11.89(81)
918	15.36(50)	-4.34(66)	3.75(34)	14.20(115)
993	18.51(69)	-4.91(93)	4.49(48)	16.83(164)
1162	28.23(154)	-6.97(220)	6.99(112)	25.58(382)

Table 10: The quantity ΔM_K for various charm quark masses and $M_K = 563$ MeV on the 16^3 lattice. Here the charm quark mass is given in the $\overline{\text{MS}}$ scheme at a scale $\mu = 2$ GeV. The third and fourth columns give the lattice results and NLO perturbation result respectively. For the perturbative result, the matching between four and three flavors is done at $\mu_c = m_c(m_c)$. The second column contains the values of $(m_c a)^2$ as an indication of the size of finite lattice spacing errors which may corrupt the comparison between the lattice and NLO perturbative results.

m_c (MeV)	$(m_c a)^2$	ΔM_K (10^{-15} GeV)	ΔM_K^{NLO} (10^{-15} GeV)
350	0.04	4.76(27)	3.24
435	0.06	5.06(29)	2.82
521	0.09	5.36(31)	2.63
606	0.12	5.66(32)	2.56
692	0.16	5.96(33)	2.56
863	0.25	6.58(35)	2.68
1086	0.39	7.37(38)	2.99
1449	0.70	8.61(41)	3.67

Table 11: The values of the parameters used in the Lanczos algorithm on the 24^3 lattice. The meaning of these parameters are given in Ref. [3].

Mass	0.005
Stop_rsd	10^{-9}
QR_rsd	10^{-14}
Operator	$D^\dagger D$
Precondition	1
K	305
M	340
$N_{\rm eig}$	300
$n_{\rm poly}$	100
α	5.5
β	0.12
μ	0

Table 12: Results for the integrated correlators on the 24^3 lattice. The kaon sources are separated by 31 lattice units. The quantities in columns two through four are the simple lattice integrated correlators of the operator products $Q_i^{qq'}Q_j^{q'q}$ for each $i, j = 1, 2$, summed over the four values of $q, q' = u, c$, without Wilson coefficients or renormalization factors and have been scaled to remove a factor 10^{-4}.

T	$Q_1 \cdot Q_1$	$Q_1 \cdot Q_2$	$Q_2 \cdot Q_2$
1	0.14(20)	-0.220(93)	-0.188(5)
2	-0.44(40)	-0.50(19)	-0.496(14)
3	-1.37(60)	-0.77(28)	-0.921(27)
4	-2.51(80)	-0.99(38)	-1.438(46)
5	-3.73(99)	-1.18(49)	-2.017(71)
6	-5.0(12)	-1.34(60)	-2.64(10)
7	-6.3(14)	-1.48(72)	-3.28(15)
8	-7.6(16)	-1.66(86)	-3.92(21)
9	-8.8(18)	-1.9(10)	-4.57(29)
10	-10.1(20)	-2.1(12)	-5.21(40)
11	-11.3(23)	-2.4(15)	-5.85(54)
12	-12.5(25)	-2.8(18)	-6.46(73)
13	-13.5(28)	-3.5(22)	-6.99(99)
14	-14.4(31)	-4.4(28)	-7.43(135)
15	-14.9(35)	-5.7(35)	-7.7(18)
16	-15.4(40)	-6.7(46)	-8.2(25)
17	-15.1(47)	-8.0(61)	-8.7(35)
18	-14.5(56)	-8.4(81)	-9.5(47)
19	-14.1(69)	-7.4(107)	-10.7(65)
20	-13.8(87)	-5.1(143)	-13.3(90)

Table 13: Results for the mass difference for each of the three operator products on the 24^3 lattice. The kaon sources are separated by 31 lattice units. These results are obtained from a fitting range [7,20]. The quantities in the second row are bare lattice numbers without any Wilson coefficients and have been scaled to remove a factor of 10^{-4}. The third row gives the corresponding physical values with Wilson coefficients in units of 10^{-15} GeV.

	$Q_1 \cdot Q_1$	$Q_1 \cdot Q_2$	$Q_2 \cdot Q_2$	ΔM_K
Lattice value	2.71(42)	0.24(25)	1.278(88)	
Physical value	0.68(10)	-0.18(18)	2.69(19)	3.19(41)

Table 14: Results for the mass difference from each of the three operator products for different choices of Δ_{KK}, which is the separation between the two kaon sources. In the last row, we give the error-weighted-average (EWR) of these fitting results. These results are obtained on the 24^3 lattice. All the masses are in units of 10^{-15} GeV.

Δ_{KK}	$Q_1 \cdot Q_1$	$Q_1 \cdot Q_2$	$Q_2 \cdot Q_2$	ΔM_K
24	0.68(10)	-0.21(17)	2.59(17)	3.06(38)
25	0.67(10)	-0.24(17)	2.63(17)	3.06(39)
26	0.68(10)	-0.19(17)	2.63(17)	3.11(39)
27	0.69(10)	-0.12(17)	2.75(18)	3.32(39)
28	0.68(10)	-0.17(17)	2.62(18)	3.13(40)
29	0.68(10)	-0.18(18)	2.62(18)	3.12(40)
30	0.68(10)	-0.24(18)	2.62(18)	3.06(40)
31	0.68(10)	-0.18(18)	2.69(19)	3.19(41)
EWR	0.68(10)	-0.19(17)	2.64(17)	3.13(38)

Table 15: Results for the mass difference from each of the three operator products for different choices of $T_{\min}$, which is the minimum fitting time. We fix $\Delta_{\min} = 6$, which is the minimum separation between the kaon sources and the weak Hamiltonians. These results are obtained on the 24^3 lattice. All the masses are in units of 10^{-15} GeV.

$\Delta_{\min}$	$T_{\min}$	$Q_1 \cdot Q_1$	$Q_1 \cdot Q_2$	$Q_2 \cdot Q_2$	ΔM_K
	7	0.68(10)	-0.18(18)	2.69(19)	3.19(41)
6	8	0.68(10)	-0.11(20)	2.85(24)	3.42(48)
	9	0.68(11)	-0.18(25)	2.69(34)	3.18(63)

Table 16: Fitting results for the mass difference from each of the three operator products for different choices of $\Delta_{\min}$, which is the minimum separation between the kaon sources and the weak Hamiltonians. The minimal fitting time $T_{\min}$ is 7. These results are obtained on the 24^3 lattice. All the masses are in units of 10^{-15} GeV.

$T_{\min}$	$\Delta_{\min}$	$Q_1 \cdot Q_1$	$Q_1 \cdot Q_2$	$Q_2 \cdot Q_2$	ΔM_K
	6	0.68(10)	-0.18(18)	2.69(19)	3.19(41)
7	7	0.68(10)	-0.20(18)	2.64(19)	3.13(41)
	8	0.67(10)	-0.19(18)	2.61(19)	3.09(41)

Table 17: Comparison of mass difference from different types of diagrams on the 24^3 lattice. We choose $\Delta_{\min} = 8$, which is the minimum separation between the kaon sources and the weak Hamiltonians. These results are obtained on the 24^3 lattice. All the numbers here are in units of 10^{-15} GeV.

Diagrams	Fitting range	$Q_1 \cdot Q_1$	$Q_1 \cdot Q_2$	$Q_2 \cdot Q_2$	ΔM_K
Type 1	[9,16]	1.330(7)	0.075(32)	-0.978(58)	0.427(91)
Type 2	[12,16]	0.164(3)	1.600(24)	4.807(50)	6.572(75)
Type 3	[7,16]	0.009(3)	0.187(17)	0.619(32)	0.816(47)
Type 4	[7,16]	-0.541(21)	-1.77(10)	-1.33(18)	-3.64(29)

Table 18: Comparison of mass difference for different values of c_s. The value of c_p is chosen to be 1. c_s and c_p are defined in Eq. 6.1. The value of $c_s = -1.14(26)$ will eliminate the η intermediate state. We choose $\Delta_{\min} = 6$, which is the minimum separation between the kaon sources and the weak Hamiltonians. The minimal fitting time $T_{\min}$ is 7. These results are obtained on the 24^3 lattice. All the numbers here are in units of 10^{-15} GeV.

c_s	$Q_1 \cdot Q_1$	$Q_1 \cdot Q_2$	$Q_2 \cdot Q_2$	ΔM_K
-1	0.68(12)	-0.20(21)	2.63(27)	3.12(40)
0	0.68(10)	-0.18(18)	2.69(19)	3.19(41)
1	0.672(88)	-0.15(28)	2.69(25)	3.21(59)
-1.14(26)	0.661(58)	-0.14(19)	2.70(19)	3.22(40)

Table 19: The mass difference for $c_s = 1$ and $c_p = 0$. c_s and c_p are defined in Eq. 6.1. $\Delta_{\min}$ is the minimum separation between the kaon sources and the weak Hamiltonians. These results are obtained on the 24^3 lattice. All the numbers here are in units of 10^{-15} GeV.

$\Delta_{\min}$	Fitting range	ΔM_{11}	ΔM_{12}	ΔM_{22}	ΔM_K
6	[7,12]	0.56(10)	-1.35(19)	-0.79(23)	-1.57(44)
10	[7,12]	0.55(11)	-1.43(22)	-0.78(32)	-1.66(55)

Table 20: Contribution to the mass differences from the π^0 and the vacuum states. We do not add $\bar{s}d$ or $\bar{s}\gamma_5 d$ operators to the weak Hamiltonian while evaluating these values. These results are obtained on the 24^3 lattice. All the numbers are in units of 10^{-15} GeV.

State	ΔM_{11}	ΔM_{12}	ΔM_{22}	ΔM_K
π^0	0.00609(15)	-0.1258(13)	0.6498(41)	0.5301(47)
η	0.330(99)	0.07(14)	0.04(14)	0.40(23)
vacuum	-0.863(40)	-9.630(31)	-26.866(78)	-37.36(11)

Figures

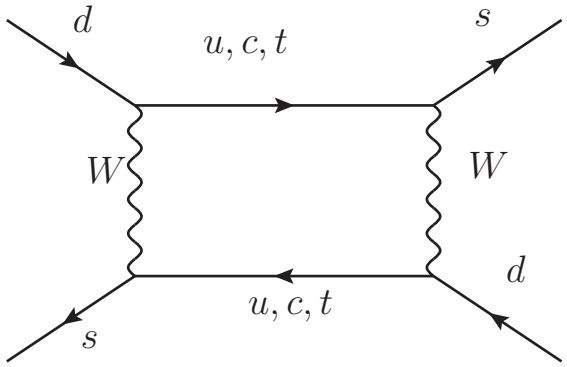


Figure 1: $K^0 - \overline{K}^0$ mixing in the standard model through the box diagram.

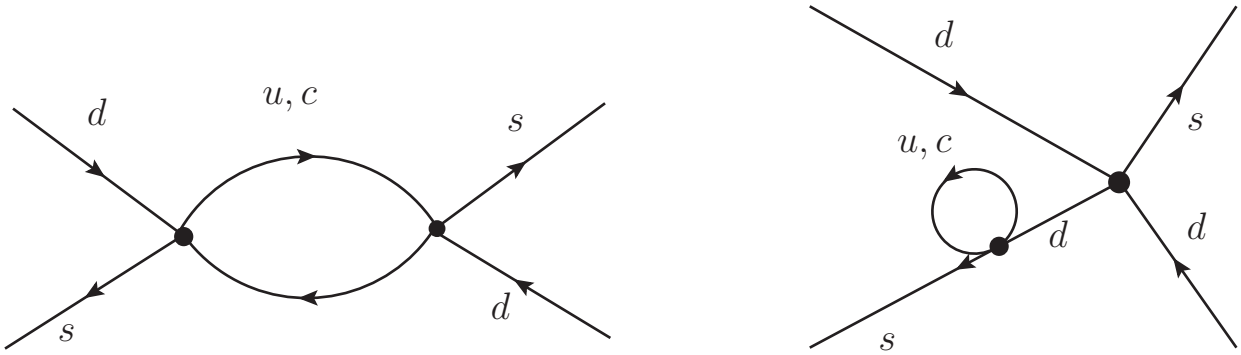
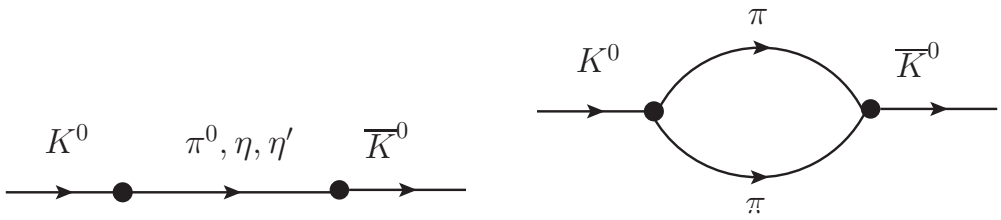


Figure 2: Examples of diagrams contributing to the $K^0 - \overline{K}^0$ mixing in the standard model. The W bosons have been integrated out and the weak interaction is described by two local four-quark $\Delta S = 1$ operators.



(a). Single meson channel (b). Two pion channel

Figure 3: Example of long distance contributions to the K^0-$\overline{K}^0$ mixing process.

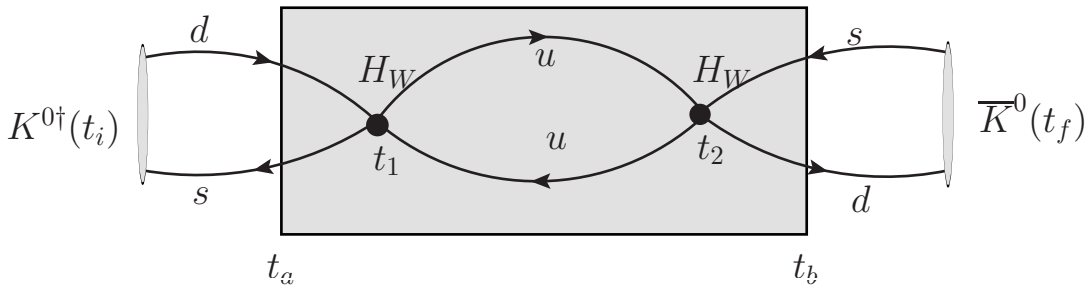


Figure 4: One type of diagram contributing to $\mathcal{A}$ in Eq. 3.6. Here t_2 and t_1 are integrated over the time interval $[t_a, t_b]$, represented by the shaded region.

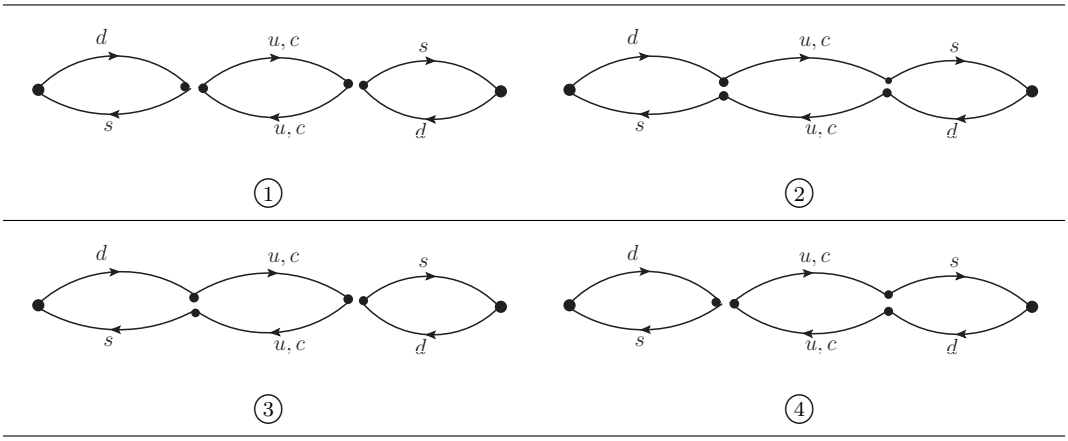


Figure 5: Diagrams for type 1 contractions. The two two-quark vertices associated with the kaon sources correspond to a spinor product including a γ_5 matrix. Each of the four two-quark vertices associated with four quark operators correspond to a contraction of color indices. The spinor products, which include the matrix $\gamma_\mu(1-\gamma_5)$, connect incoming and outgoing quark lines which carry the same electric charge. Vertices where the quark lines are joined in this fashion then have the color and spin contracted in the same pattern and correspond to the operator Q_1. Where the quark lines and corresponding color contractions for quarks with different electric charges are joined, the operator Q_2 appears.

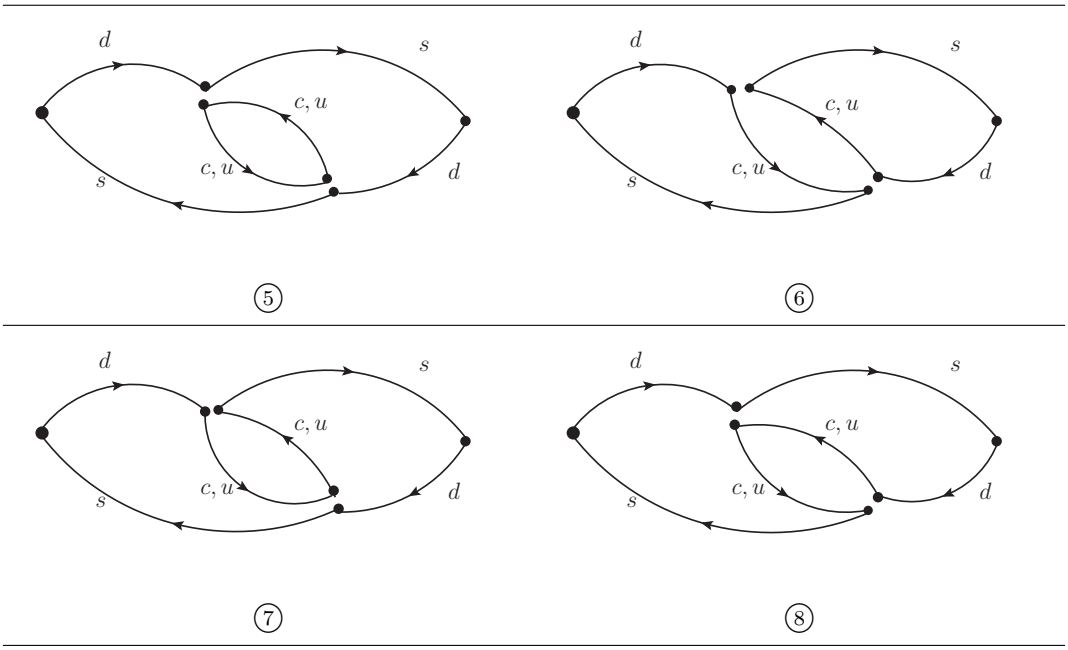


Figure 6: Diagrams for type 2 contractions. The conventions used here are the same as those explained in the caption to Fig. 5.

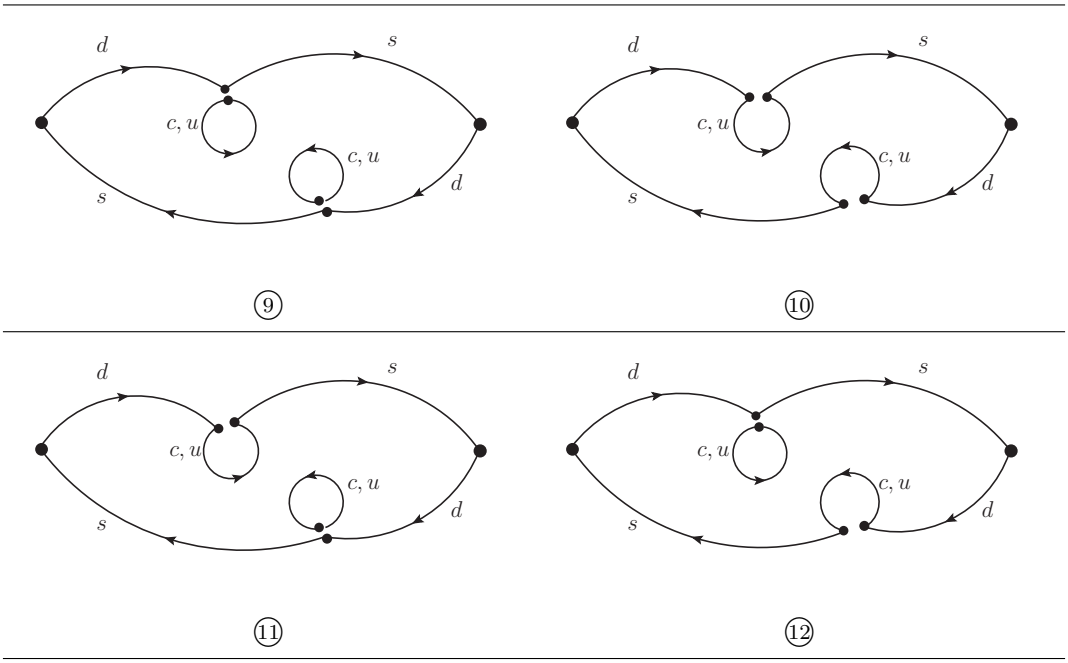


Figure 7: Diagrams for type 3 contractions. The conventions used here are the same as those explained in the caption to Fig. 5.

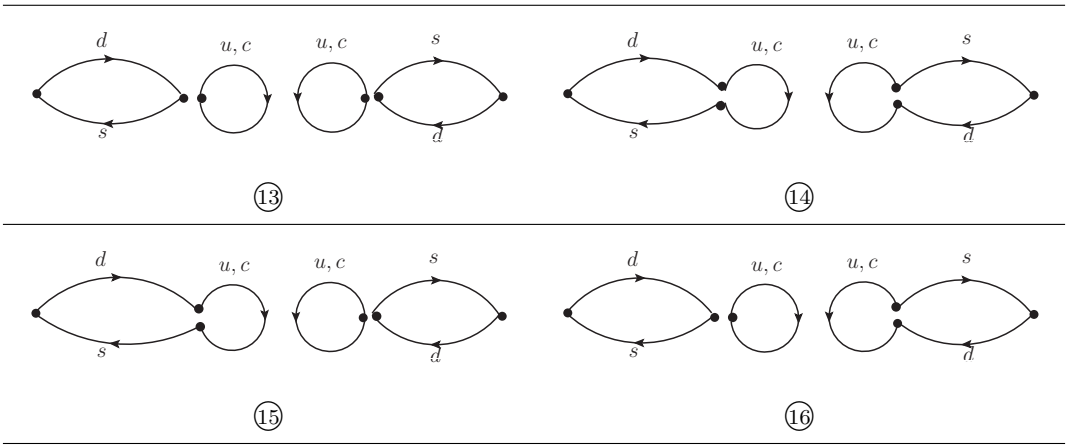


Figure 8: Diagrams for type 3 contractions. The conventions used here are the same as those explained in the caption to Fig. 5.

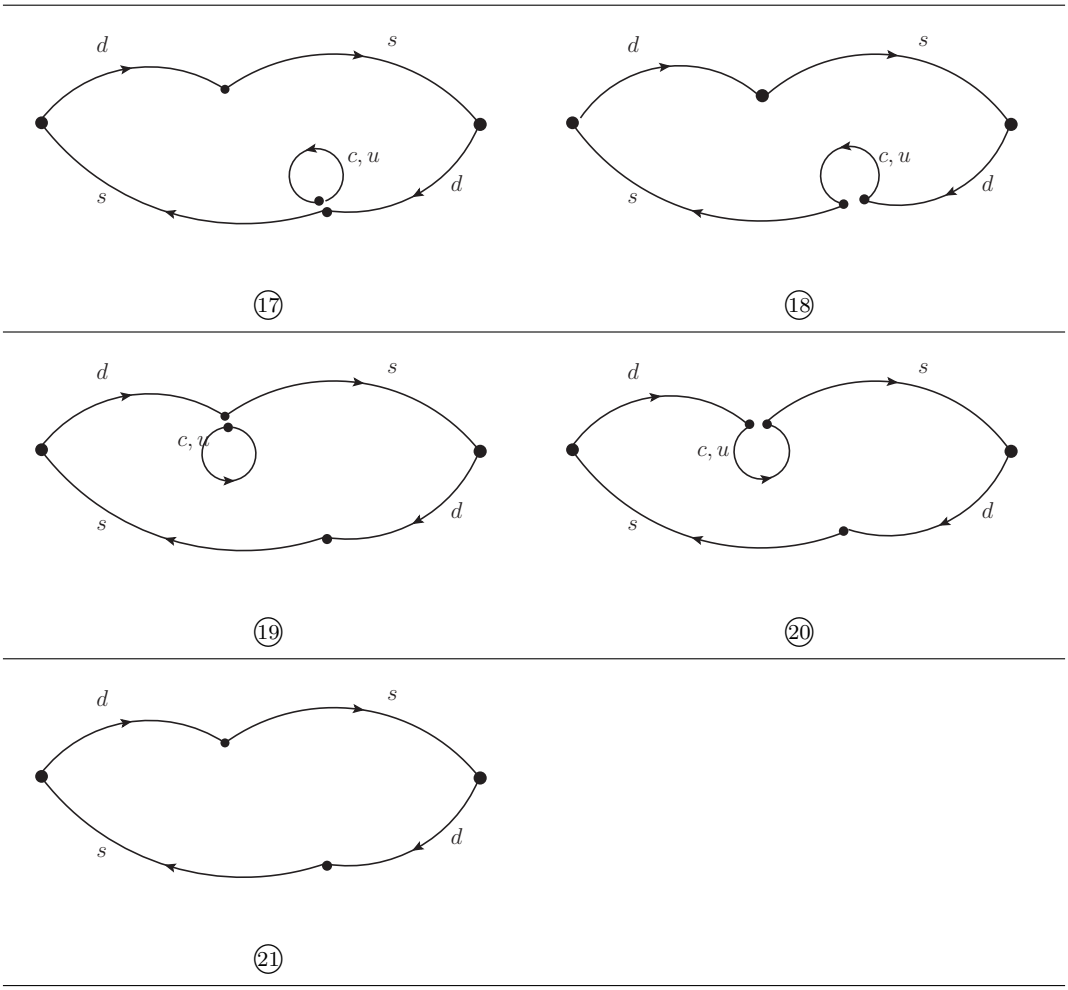


Figure 9: Diagrams for type 3-scalar contractions. The conventions used here are the same as those explained in the caption to Fig. 5.

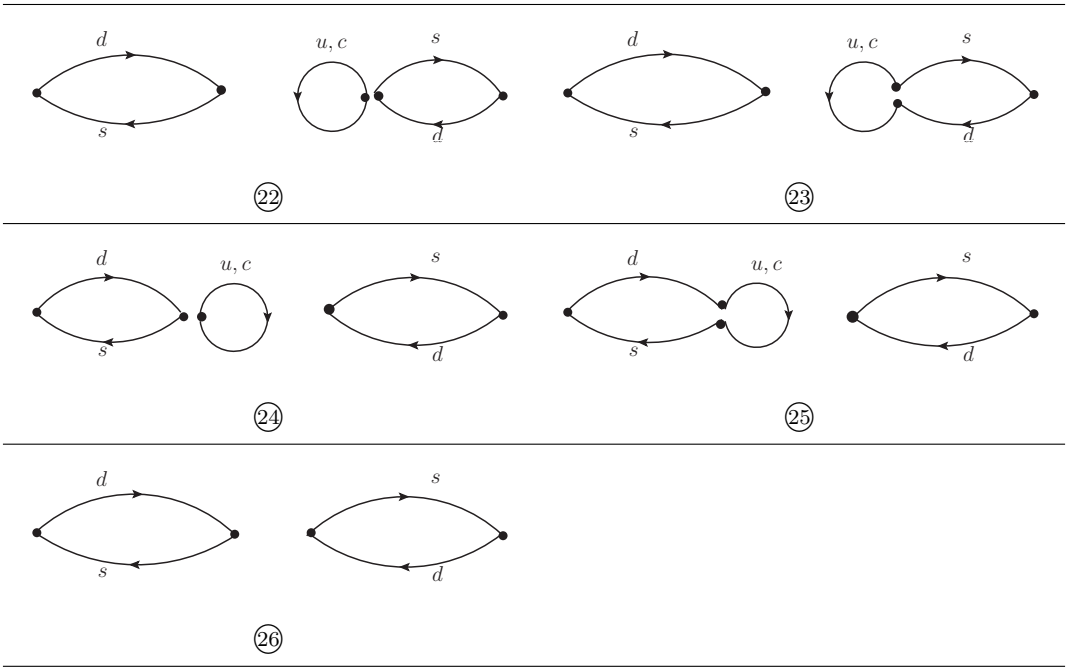


Figure 10: Diagrams for type 4-scalar contractions. The conventions used here are the same as those explained in the caption to Fig. 5.

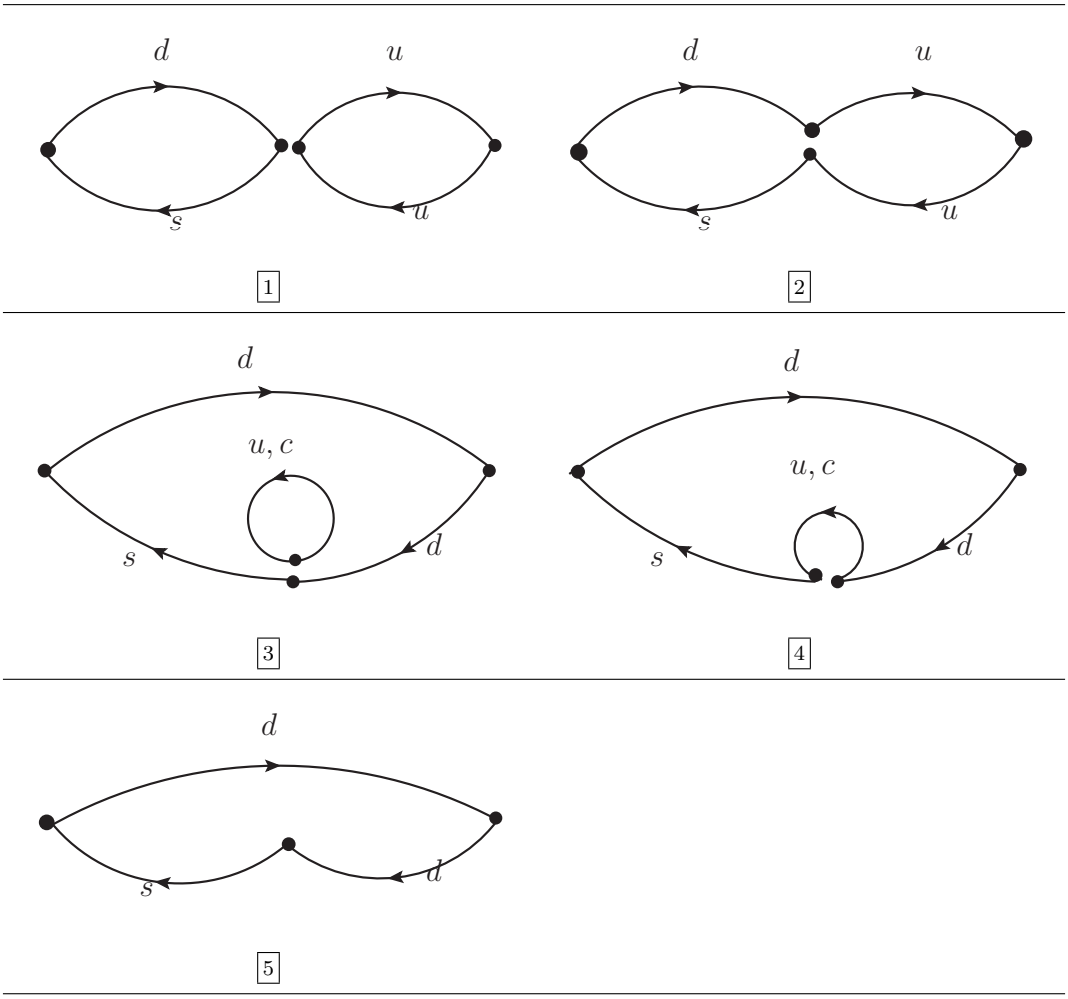


Figure 11: Diagrams contributing to the $\langle 0|K^0(t_K)O(t_O)\pi^0(t_\pi|0\rangle$ correlator. There will be a γ^5 insertion for the kaon and pion sources. Each of the two quark vertices associated with the four quark operators include a $\gamma^\mu(1-\gamma^5)$ insertion. The operator in the last diagram is the scalar operator $\bar{s}d$.

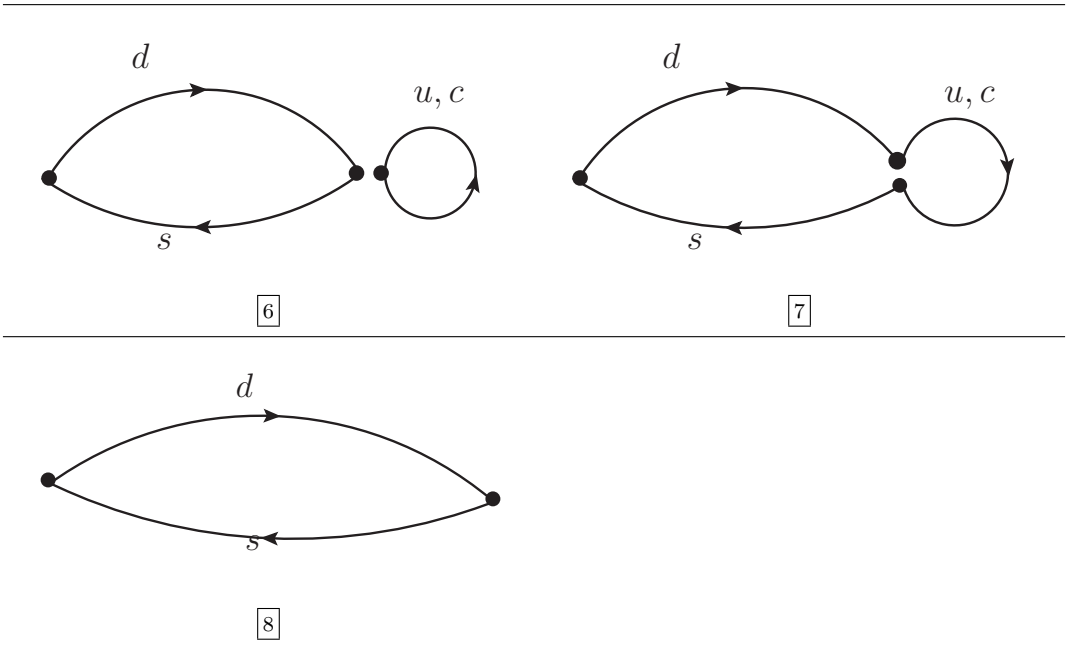


Figure 12: Diagrams contributing to the $\langle 0|K^0(t_K)O(t_O)|0\rangle$ correlator. There will be a γ^5 insertion for the kaon and pion sources. Each of the two quark vertices associated with the four quark operators include a $\gamma^\mu(1-\gamma^5)$ insertion. The operator represented by the vertex on the right in the last diagram is the pseudoscalar operator $\bar{s}\gamma^5 d$.

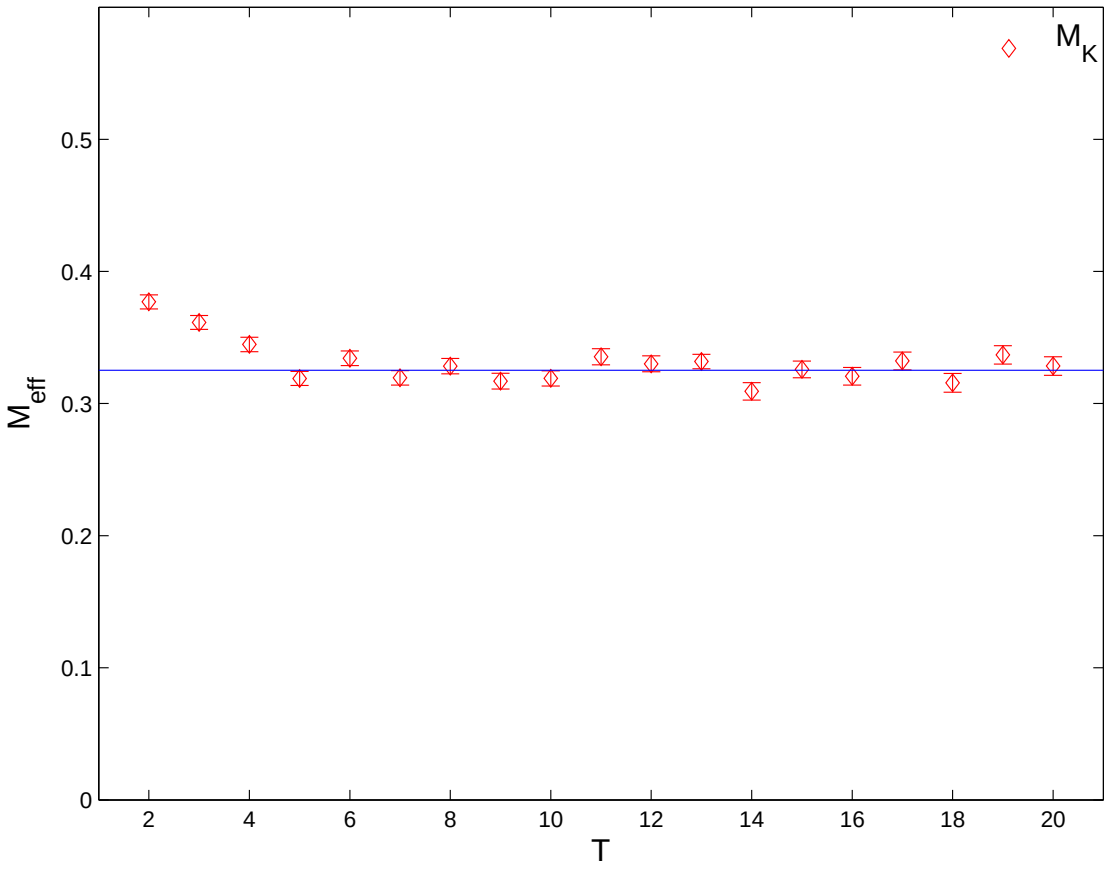
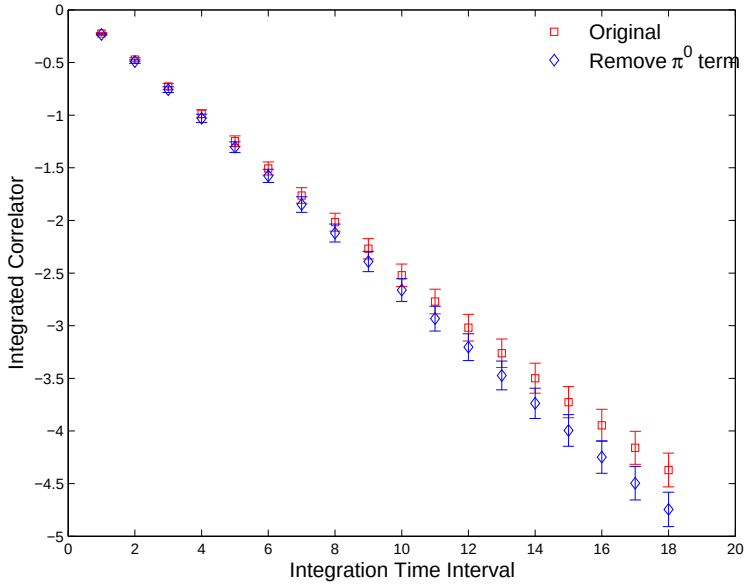
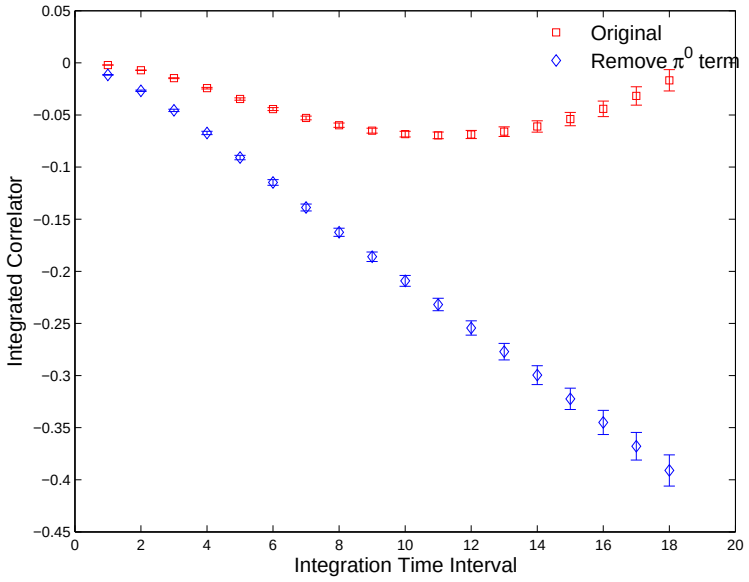


Figure 13: A plot of the kaon effective mass found from the two point correlator between a wall source and a wall-sink using $m_l = 0.01$ and $m_s = 0.032$ on the 16^3 lattice. The blue line shows the result of a hyperbolic cosine fit to the correlator in the time interval [6, 26].



(a)



(b)

Figure 14: The integrated correlator as a function of integration time interval T on the 16^3 lattice. (a) The original result without any artificial position-space cutoff; (b) The result with a cutoff radius of 5. The red squares and blue diamonds are the results before and after the subtraction of the exponentially growing π^0 term, respectively. For both plots we include only the operator combination $Q_1 \cdot Q_1$.

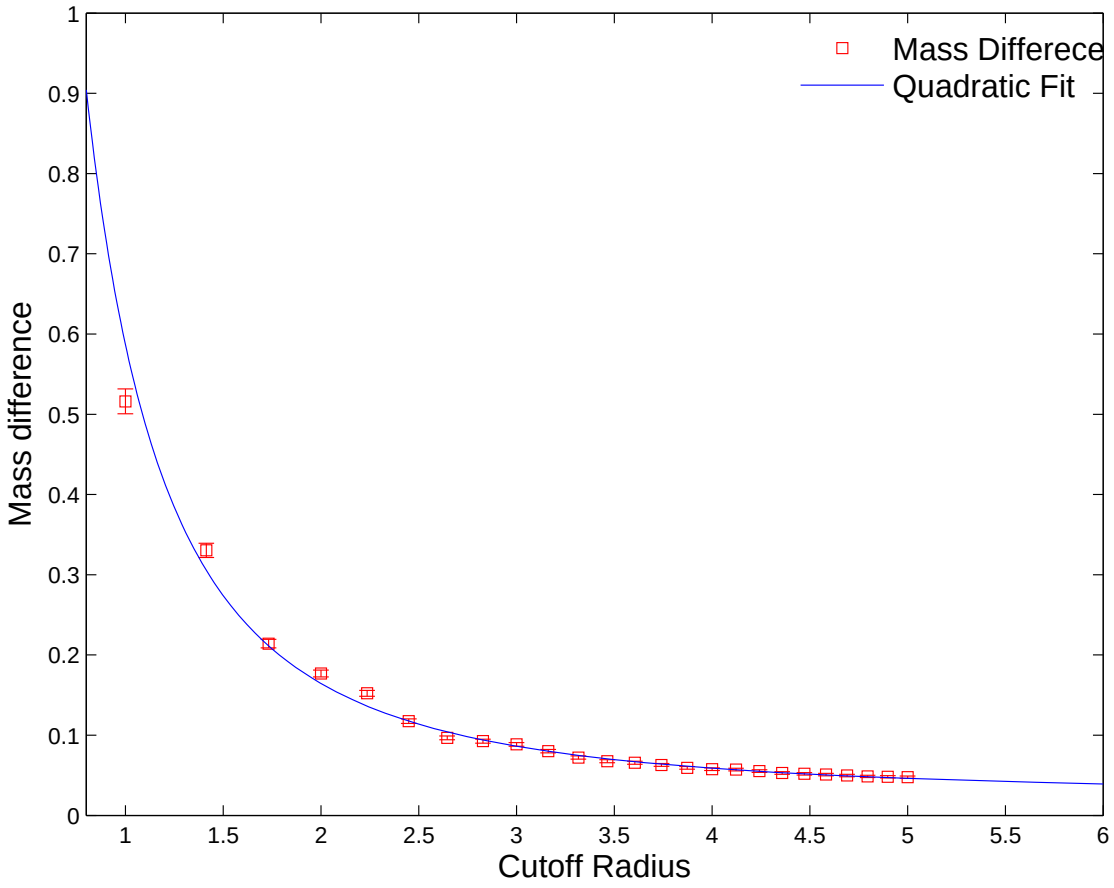


Figure 15: The mass difference ΔM_K^{11} defined in Eq. (5.6) for different values of the cutoff radius R on the 16^3 lattice. The blue curve is the two parameter fit to a $1/R^2$ behavior as defined in Eq. (5.7)

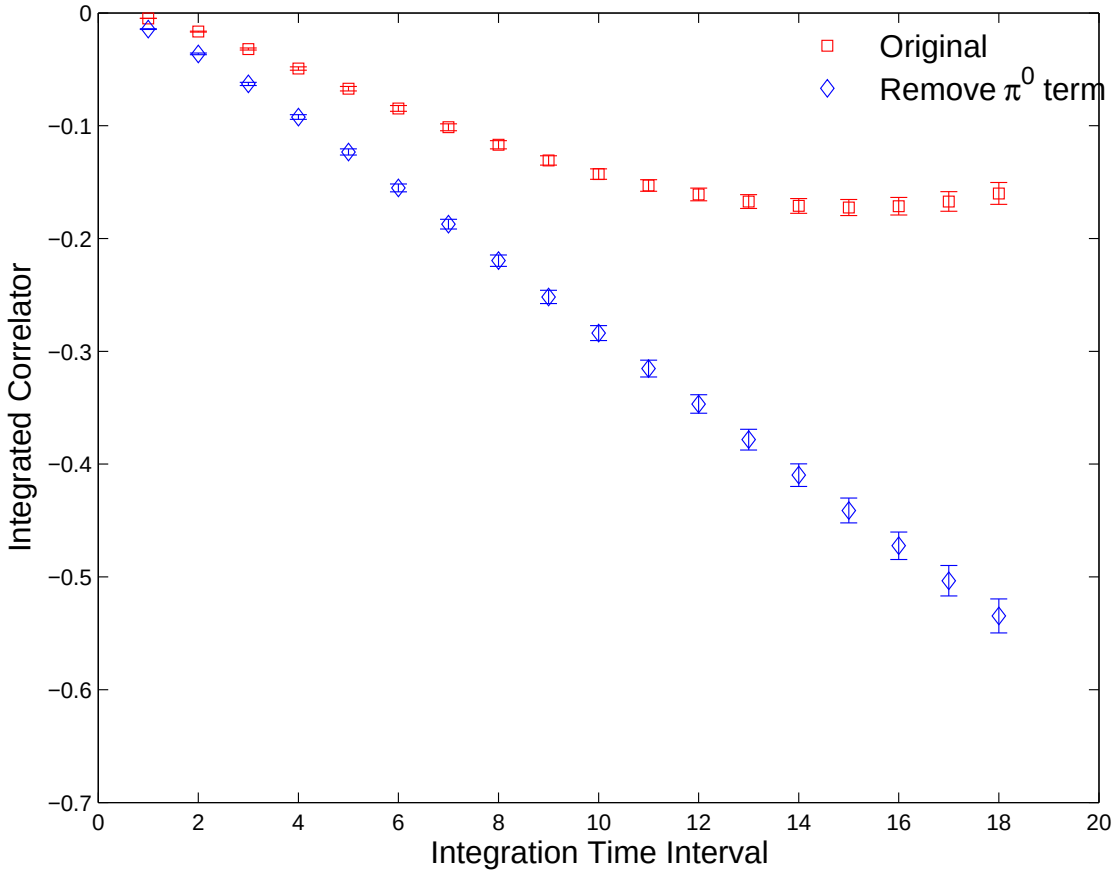


Figure 16: The integrated correlator after GIM cancellation with a 0.863 GeV valence charm quark on the 16^3 lattice. The red squares and blue diamonds are the results before and after the subtraction of the exponentially increasing π^0 term respectively. We include only the $Q_1 \cdot Q_1$ operator combination in this plot.

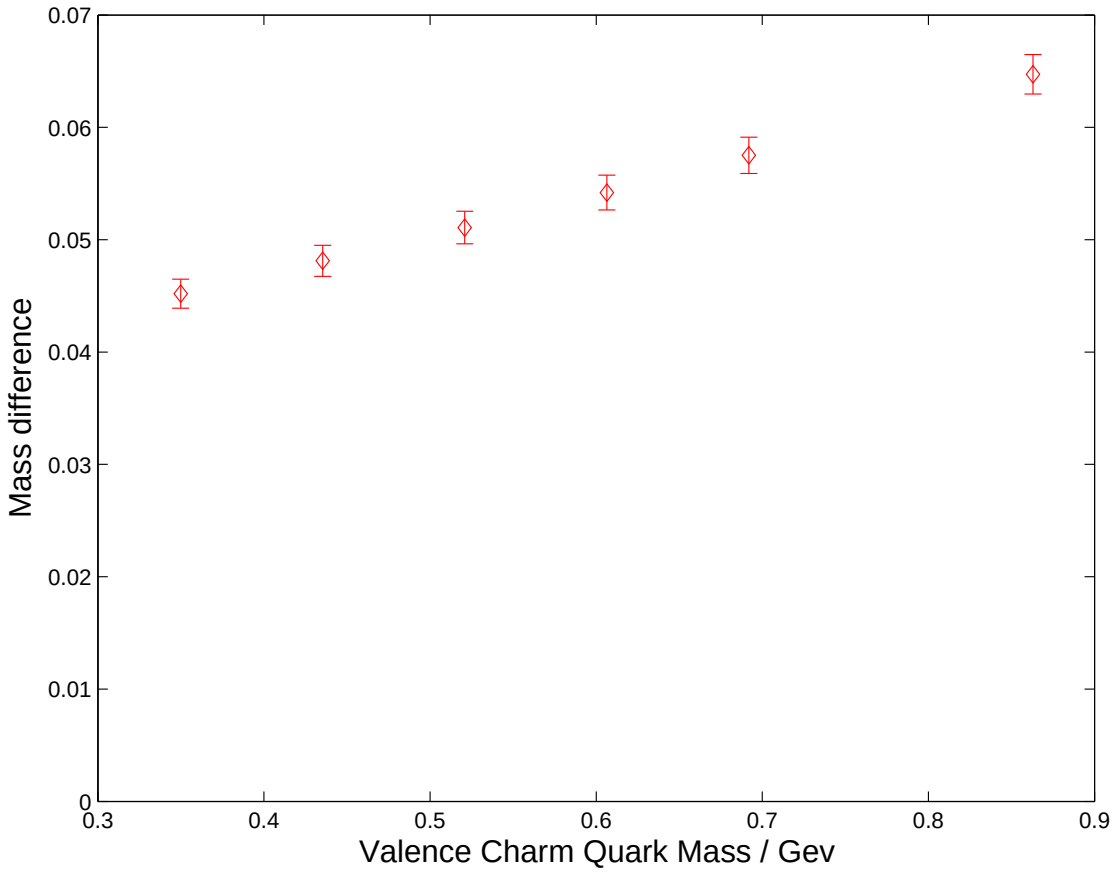


Figure 17: The mass difference ΔM_K^{11}, defined in Eq. (5.6) after GIM cancellation as a function of the valence charm quark mass. These results are obtained on the 16^3 lattice.

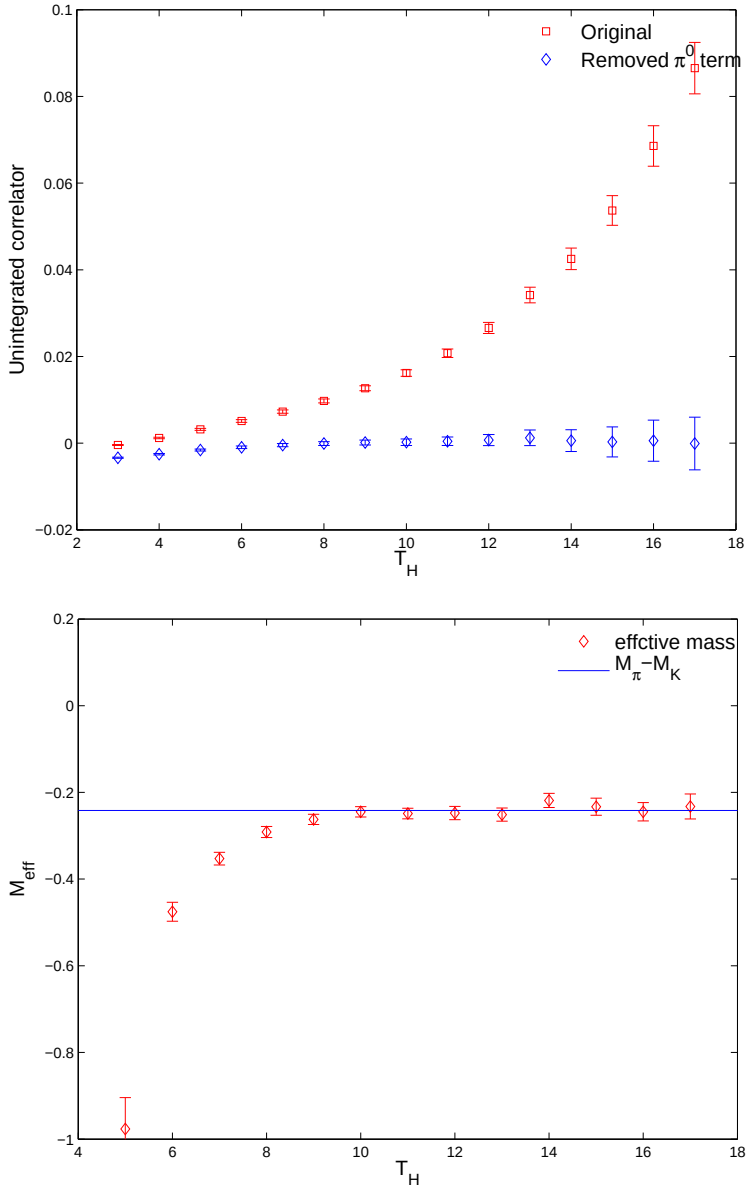


Figure 18: A plot of the unintegrated correlator $\overline{G}$ and resulting effective mass for the combination of operators $Q_1 \cdot Q_1$ and a kaon mass $M_K = 0.4848(8)$ on the 16^3 lattice. Only the product of the parity even components of the two operators is included. In the left-hand plot, the red diamonds and blue squares show the result before and after subtraction of the π^0 term. In the right-hand plot, the red diamonds are effective masses obtained from the integrated correlator. The blue horizontal line shows the "exact" value of $M_\pi - M_K$ obtained from the two point correlator calculation.

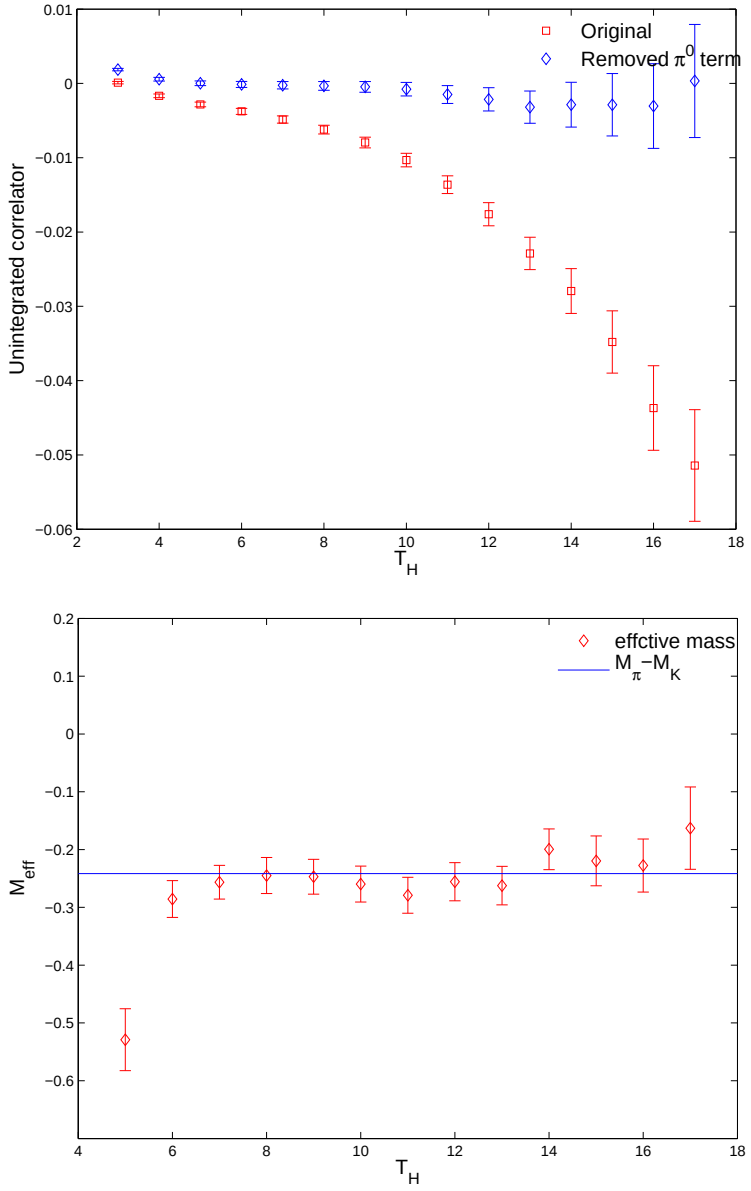


Figure 19: Plots of the unintegrated correlator $\overline{G}$ and corresponding effective mass for the operator combination $Q_1 \cdot Q_2$ at a kaon mass $M_K = 0.4848(8)$ on the 16^3 lattice. Only the product of the parity even components of the two operators is included.

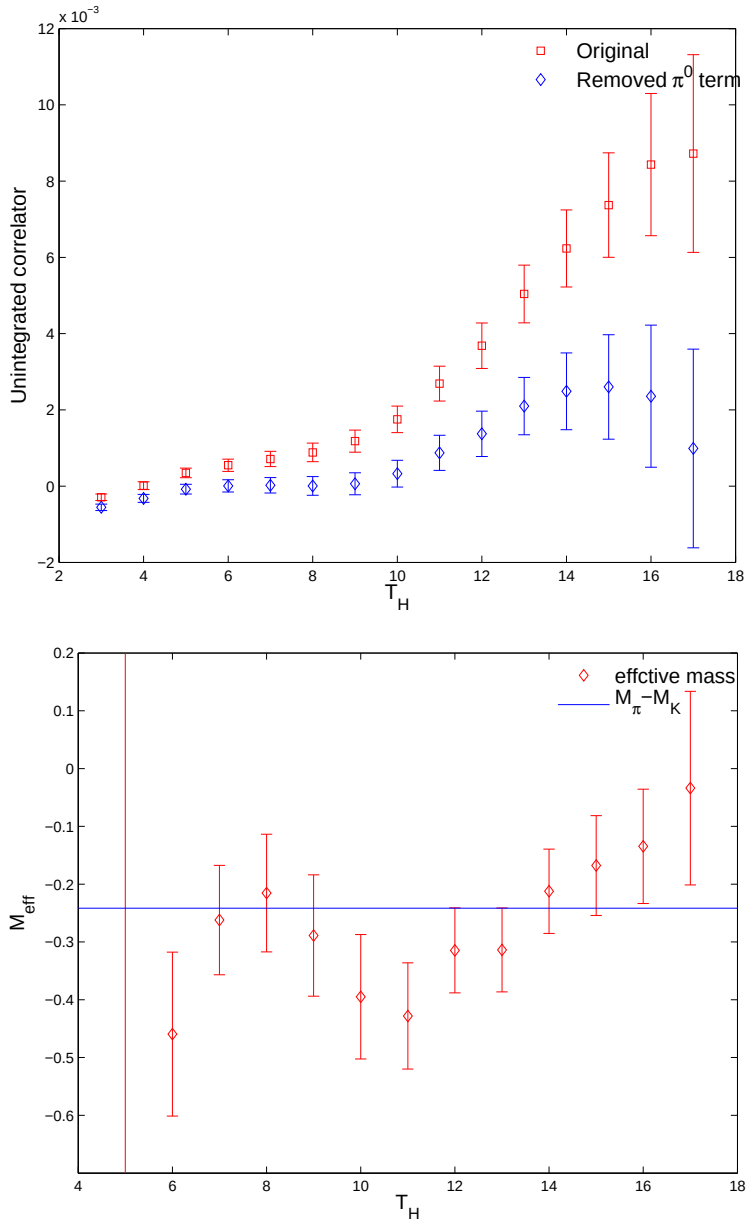


Figure 20: Plots of the unintegrated correlator and corresponding effective mass for the operator combination $Q_2 \cdot Q_2$ at a kaon mass $M_K = 0.4848(8)$ on the 16^3 lattice. Only the product of the parity even components of the two operators is included.

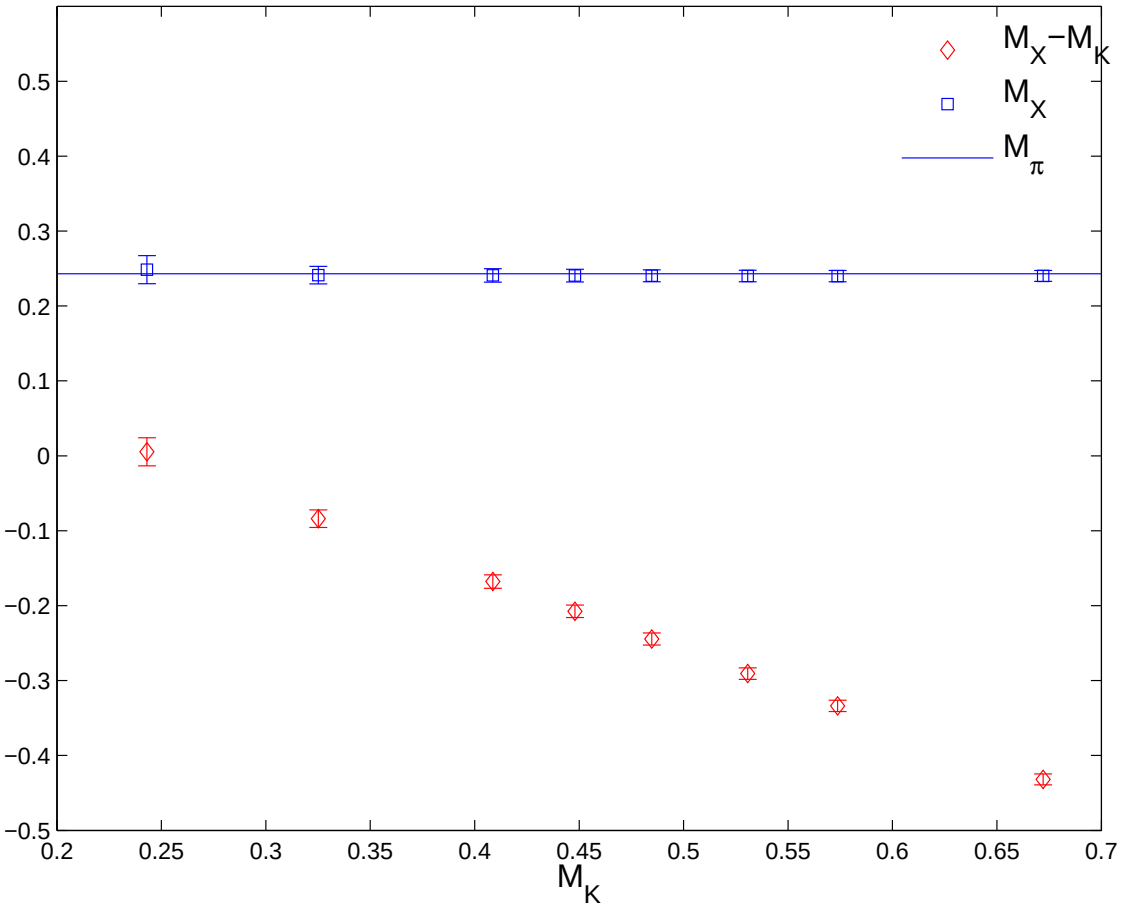


Figure 21: Intermediate state masses determined for all eight kaon masses from the unintegrated correlators of the parity even portion of the operators $Q_1 \cdot Q_1$. The red diamonds are the fitting results and should correspond to the difference $M_X - M_K$. The blue squares are obtained from the results for $M_X - M_K$ by adding the result for M_K obtained from the two-point kaon correlators. The blue horizontal line is the "exact" pion mass given by the two point function calculation. These results are obtained on the 16^3 lattice.

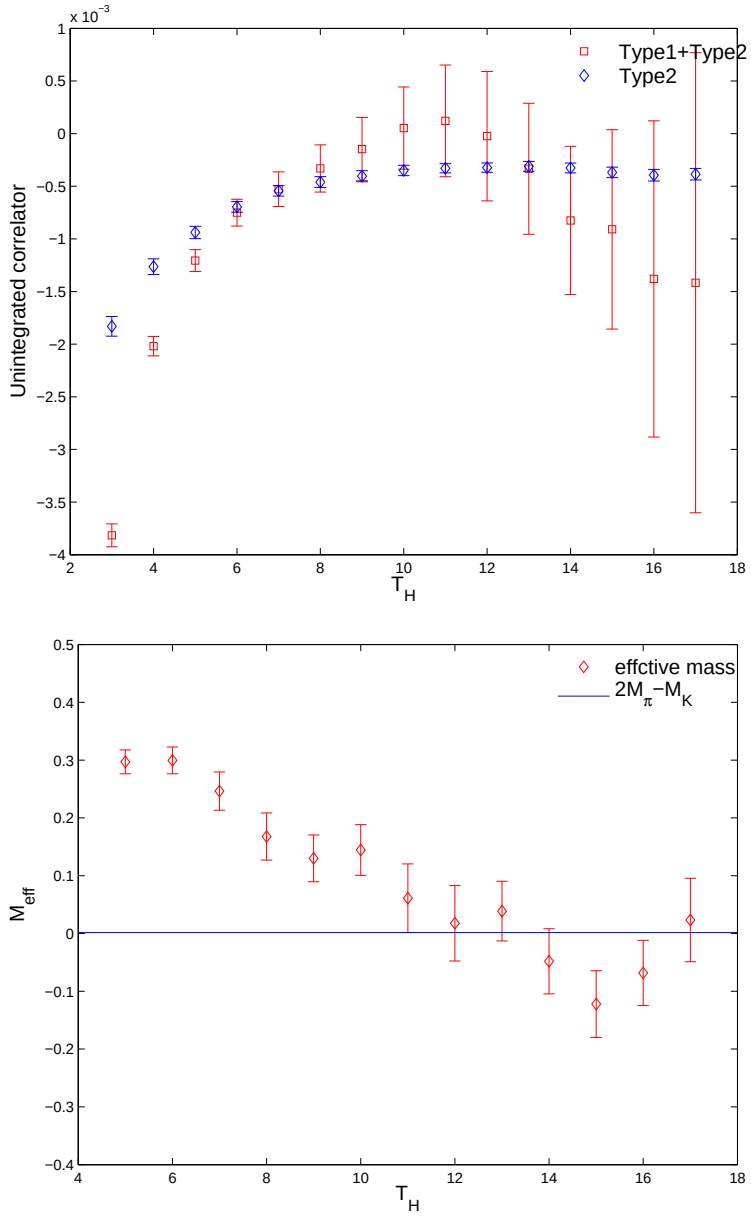


Figure 22: A plot of the unintegrated correlator $\overline{G}$ and resulting effective mass for the combination of operators $Q_1 \cdot Q_1$ and a kaon mass $M_K = 0.4848(8)$ on the 16^3 lattice. Only the product of the parity odd components of the two operators is included. In the left-hand plot, the red diamonds and blue squares show the full results and the results from the type 2 diagrams only. In the right-hand plot, the red diamonds are effective masses obtained from the type 2 diagrams. The blue horizontal line shows the "exact" value of $2M_\pi - M_K$ obtained from the two point correlator calculation.

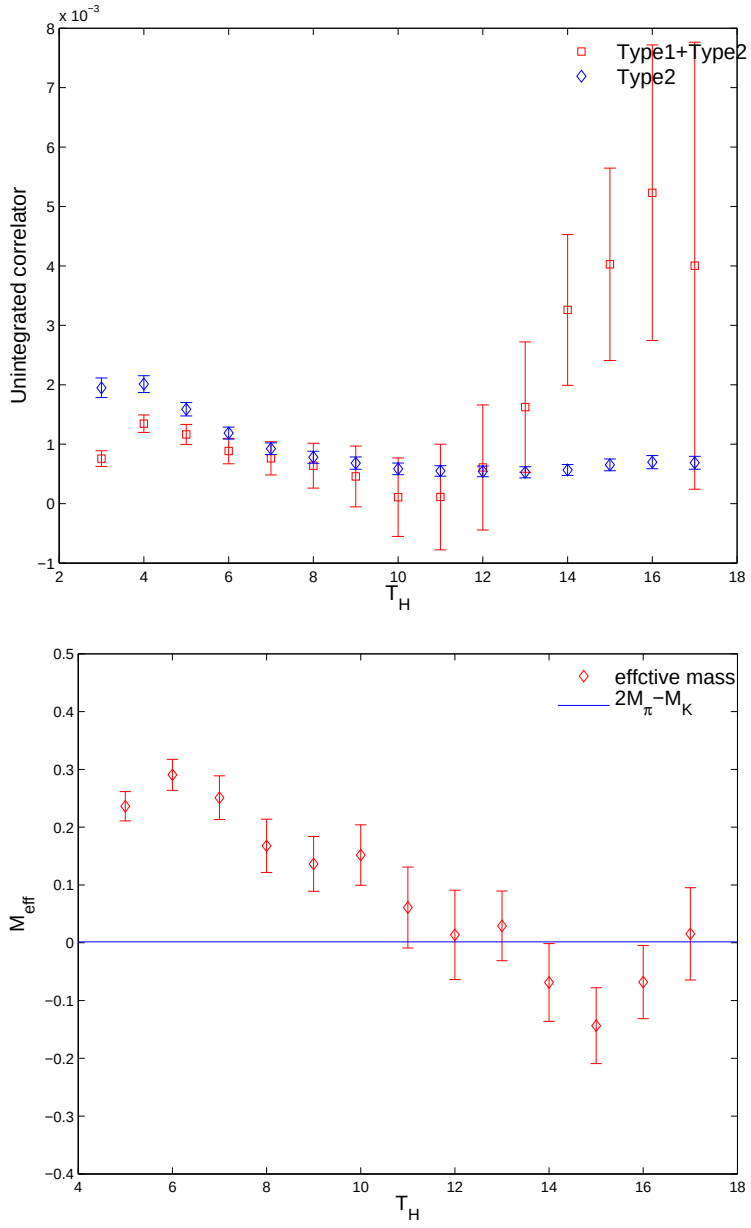


Figure 23: Plots of the unintegrated correlator $\overline{G}$ and corresponding effective mass for the operator combination $Q_1 \cdot Q_2$ at a kaon mass $M_K = 0.4848(8)$ on the 16^3 lattice. Only the product of the parity odd components of the two operators is included. The effective mass is obtained from the type 2 diagrams alone.

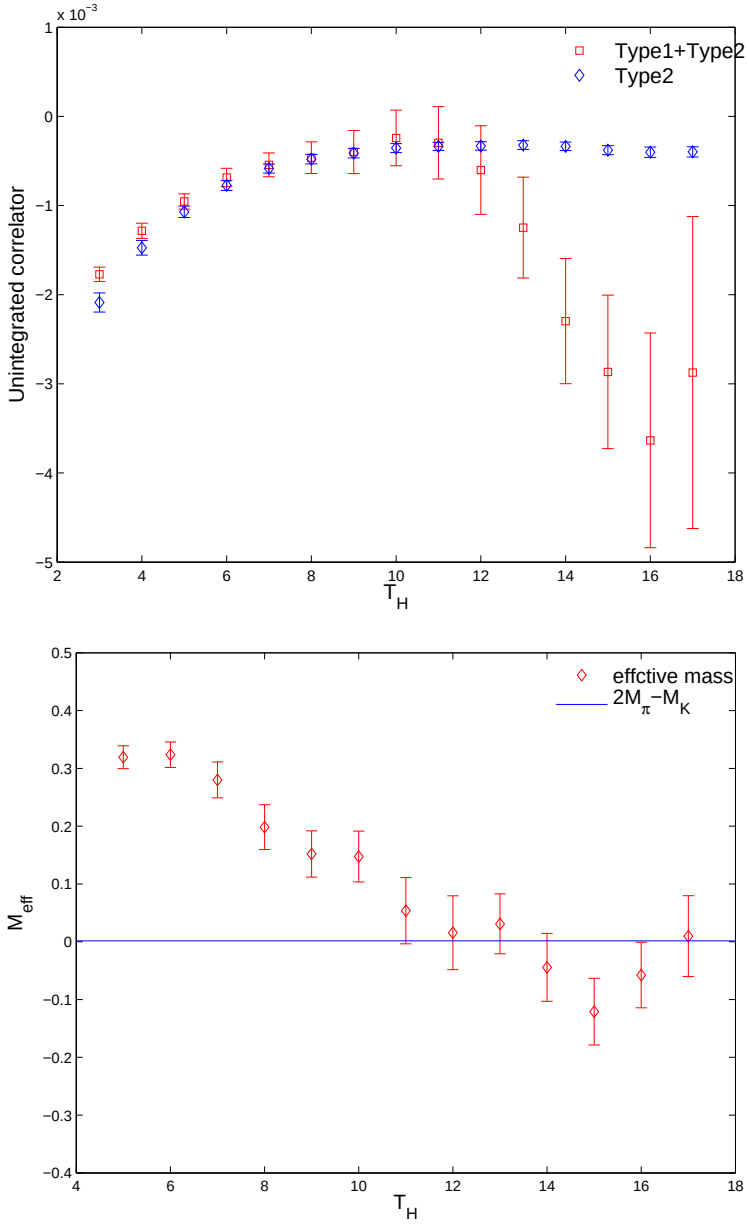


Figure 24: Plots of the unintegrated correlator and corresponding effective mass for the operator combination $Q_2 \cdot Q_2$ at a kaon mass $M_K = 0.4848(8)$ on the 16^3 lattice. Only the product of the parity odd components of the two operators is included. The effective mass is obtained from the type 2 diagrams alone.

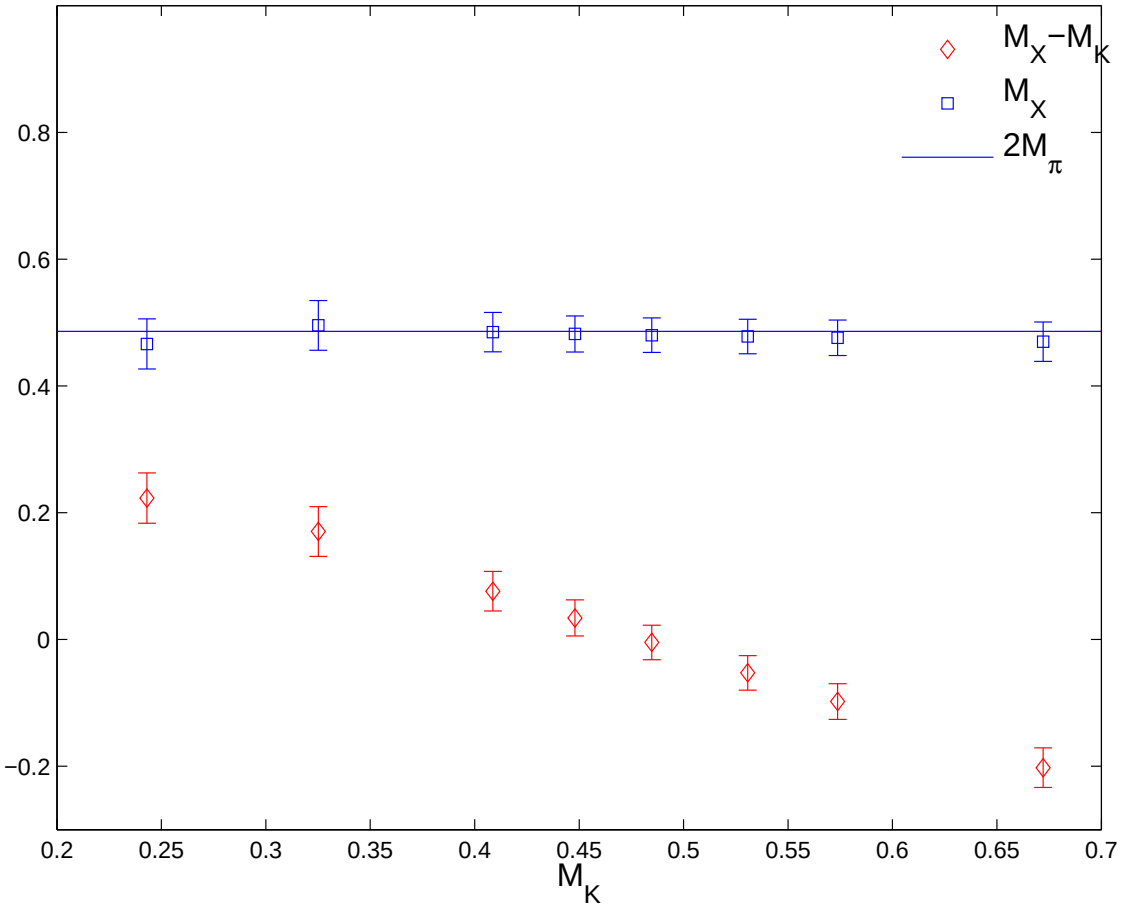


Figure 25: Intermediate state masses determined for all eight kaon masses from the unintegrated correlators of the parity odd portion of the operators $Q_1 \cdot Q_1$. The results shown in this last plot are obtained from fitting the type 2 diagrams alone. Because the type 2 diagrams shown in the last plot are only a subset of those needed for a physical calculation, we do not expect the effective mass shown in this last plot to be either the $I = 0$ or $I = 2$ finite volume $\pi - \pi$ energy. We view the agreement with $2m_\pi$ as coincidental. These results are obtained on the 16^3 lattice.

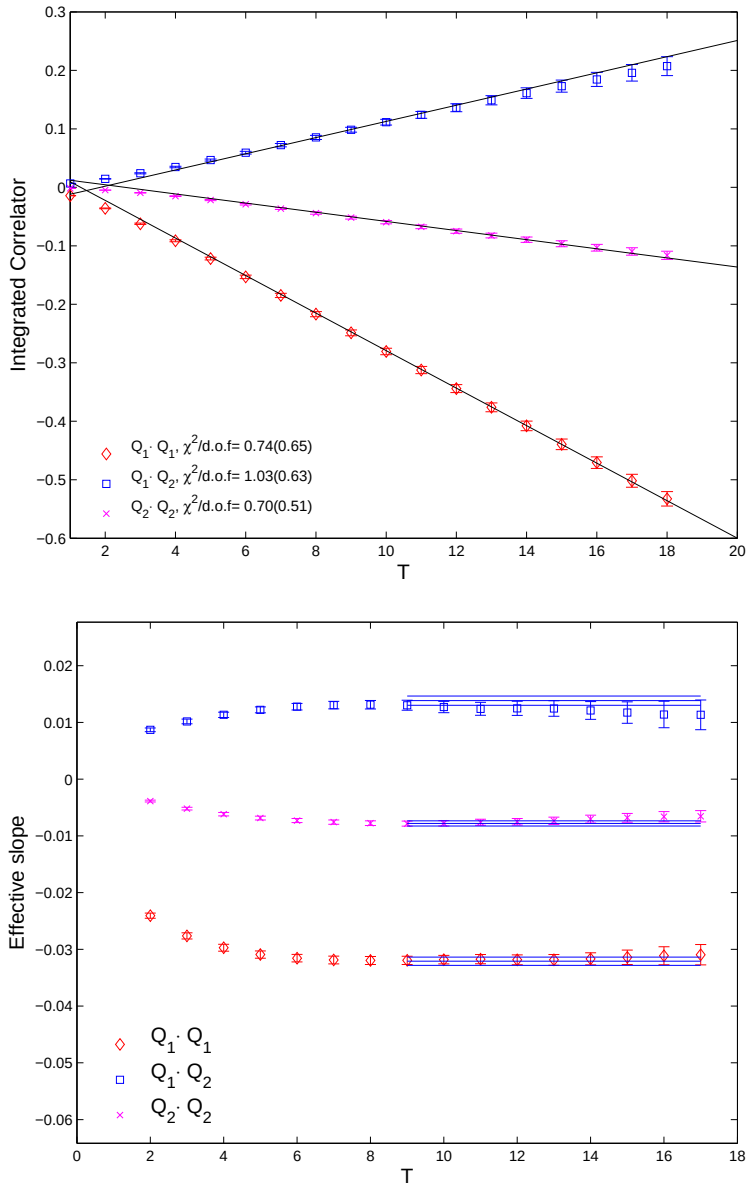


Figure 26: Lattice results for the integrated correlator given in Eq. 3.7 and the corresponding effective slope plots for the three operator products $Q_1 \cdot Q_1$, $Q_1 \cdot Q_2$ and $Q_2 \cdot Q_2$ in the case $M_K = 563$ MeV on the 16^3 lattice. The three lines in the upper panel give the linear fits to the data in the time interval [9,18] used to extract the corresponding values given in Tab. 9. The horizontal lines with error bands in the lower panel show slopes from the same three linear fits with errors. (Note the slope of the integrated correlator as a function of the time T given in Eq. (3.7) must be multiplied by -2 to obtain the corresponding contribution to ΔM_K.)

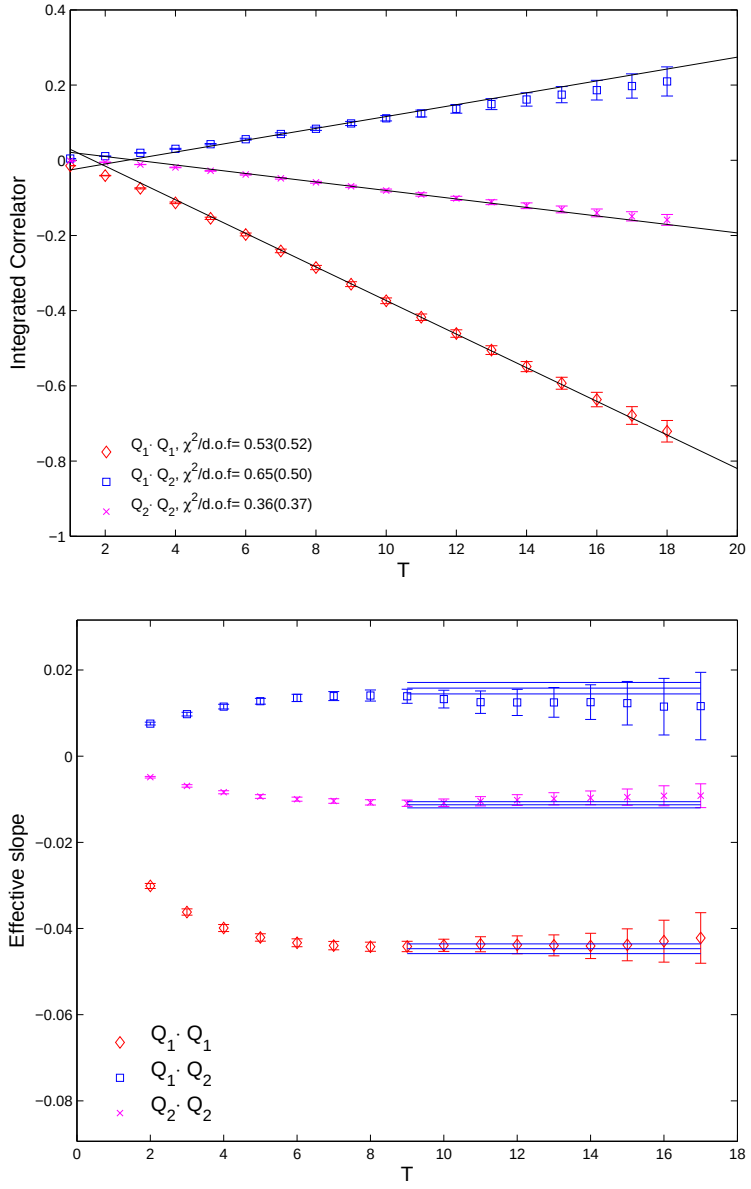


Figure 27: Lattice results for the integrated correlator given in Eq. 3.7 and the corresponding effective slope plots for the three operator products $Q_1 \cdot Q_1$, $Q_1 \cdot Q_2$ and $Q_2 \cdot Q_2$ in the case $M_K = 707$ MeV on the 16^3 lattice. The three lines in the upper panel give the linear fits to the data in the time interval [9,18] used to extract the corresponding values given in Tab. 9. The horizontal lines with error bands in the lower panel show slopes from the same three linear fits with errors. (Note the slope of the integrated correlator as a function of the time T given in Eq. (3.7) must be multiplied by -2 to obtain the corresponding contribution to ΔM_K.)

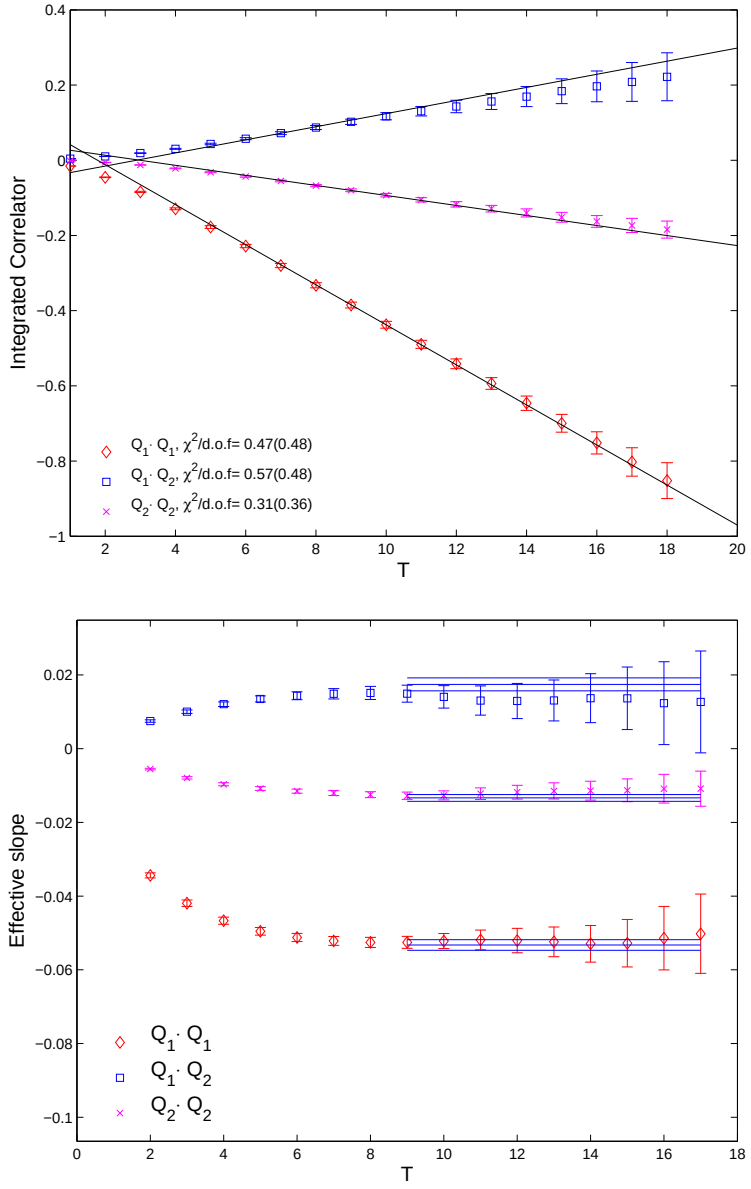


Figure 28: Lattice results for the integrated correlator given in Eq. 3.7 and the corresponding effective slope plots for the three operator products $Q_1 \cdot Q_1$, $Q_1 \cdot Q_2$ and $Q_2 \cdot Q_2$ in the case $M_K = 775$ MeV on the 16^3 lattice. The three lines in the upper panel give the linear fits to the data in the time interval [9,18] used to extract the corresponding values given in Tab. 9. The horizontal lines with error bands in the lower panel show slopes from the same three linear fits with errors. (Note the slope of the integrated correlator as a function of the time T given in Eq. (3.7) must be multiplied by -2 to obtain the corresponding contribution to ΔM_K.)

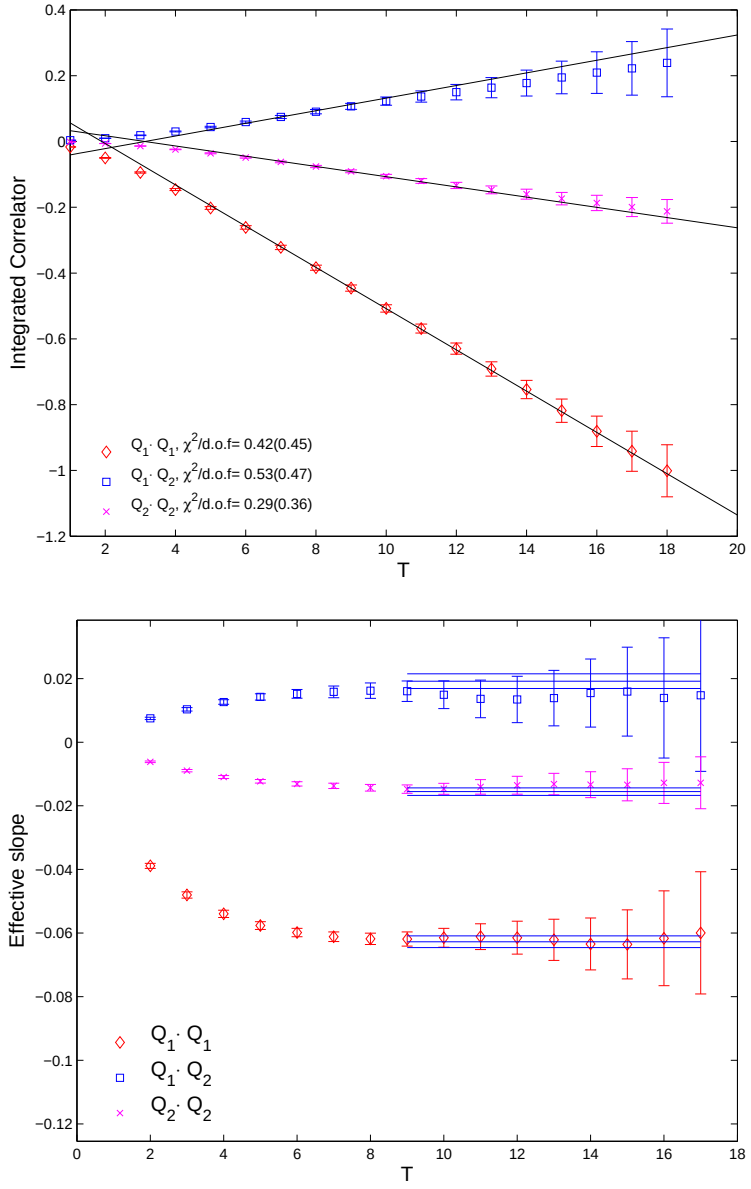


Figure 29: Lattice results for the integrated correlator given in Eq. 3.7 and the corresponding effective slope plots for the three operator products $Q_1 \cdot Q_1$, $Q_1 \cdot Q_2$ and $Q_2 \cdot Q_2$ in the case $M_K = 834$ MeV on the 16^3 lattice. The three lines in the upper panel give the linear fits to the data in the time interval [9,18] used to extract the corresponding values given in Tab. 9. The horizontal lines with error bands in the lower panel show slopes from the same three linear fits with errors. (Note the slope of the integrated correlator as a function of the time T given in Eq. (3.7) must be multiplied by -2 to obtain the corresponding contribution to ΔM_K.)

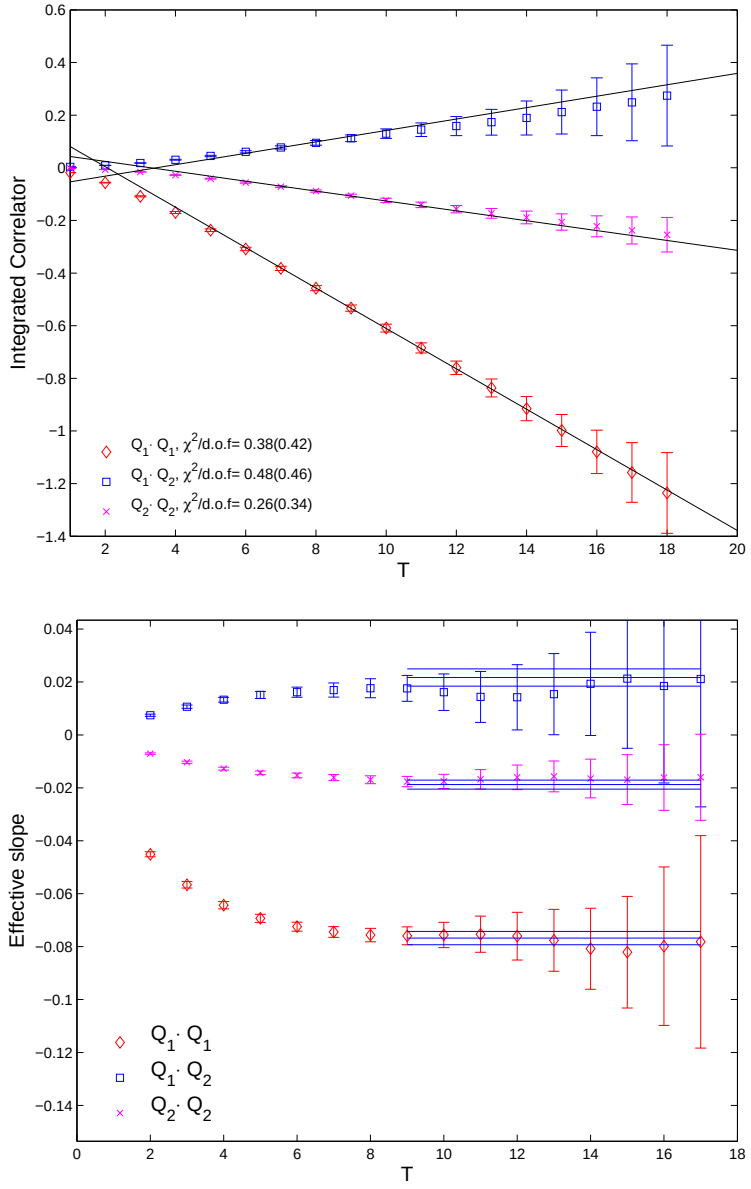


Figure 30: Lattice results for the integrated correlator given in Eq. 3.7 and the corresponding effective slope plots for the three operator products $Q_1 \cdot Q_1$, $Q_1 \cdot Q_2$ and $Q_2 \cdot Q_2$ in the case $M_K = 918$ MeV on the 16^3 lattice. The three lines in the upper panel give the linear fits to the data in the time interval [9,18] used to extract the corresponding values given in Tab. 9. The horizontal lines with error bands in the lower panel show slopes from the same three linear fits with errors. (Note the slope of the integrated correlator as a function of the time T given in Eq. (3.7) must be multiplied by -2 to obtain the corresponding contribution to ΔM_K.)

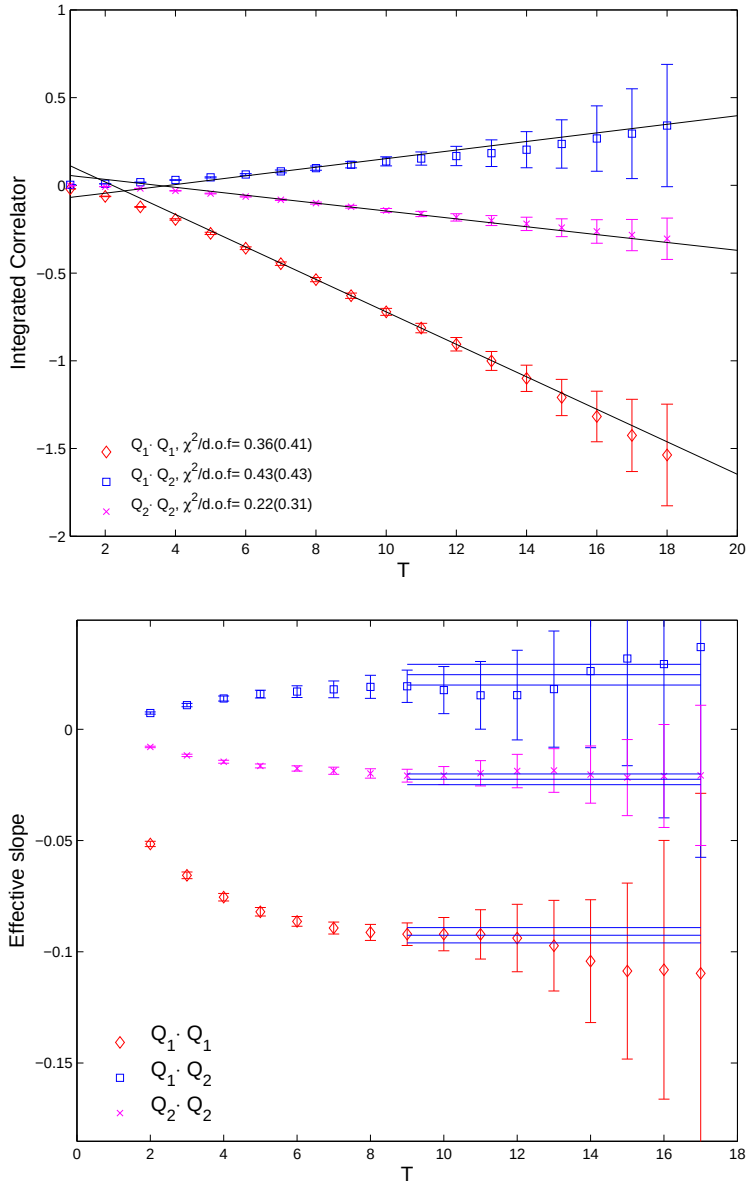


Figure 31: Lattice results for the integrated correlator given in Eq. 3.7 and the corresponding effective slope plots for the three operator products $Q_1 \cdot Q_1$, $Q_1 \cdot Q_2$ and $Q_2 \cdot Q_2$ in the case $M_K = 993$ MeV on the 16^3 lattice. The three lines in the upper panel give the linear fits to the data in the time interval [9,18] used to extract the corresponding values given in Tab. 9. The horizontal lines with error bands in the lower panel show slopes from the same three linear fits with errors. (Note the slope of the integrated correlator as a function of the time T given in Eq. (3.7) must be multiplied by -2 to obtain the corresponding contribution to ΔM_K.)

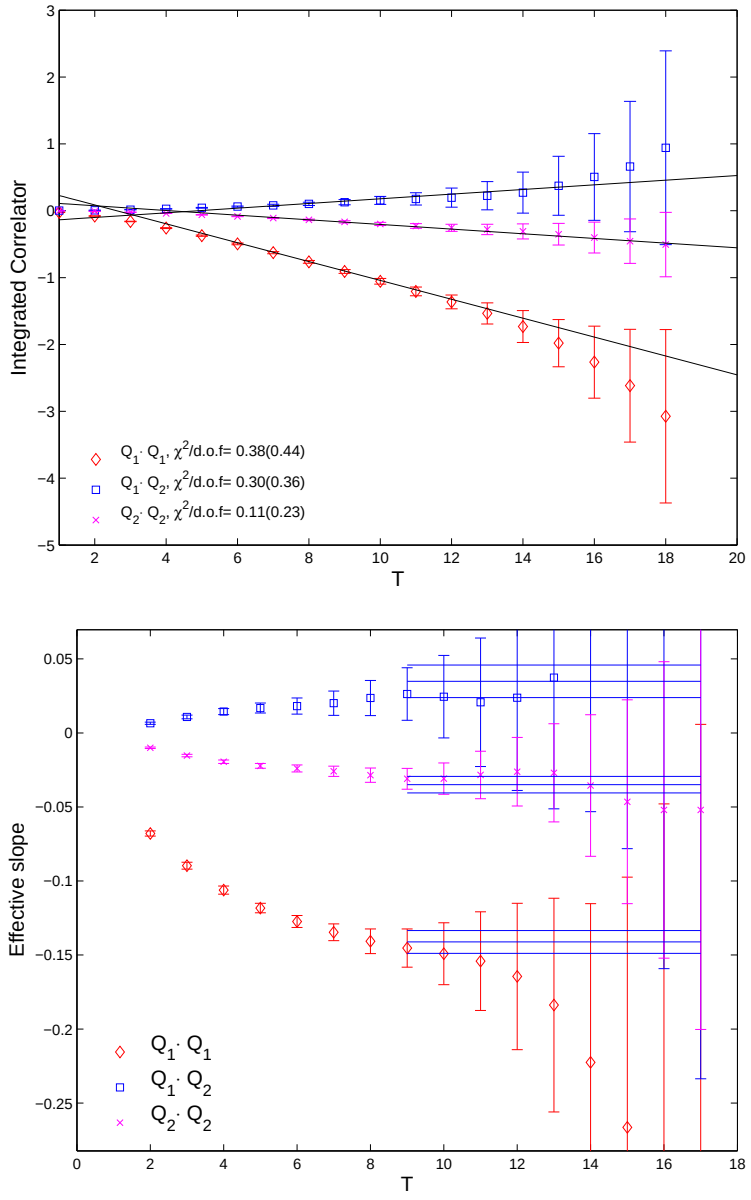


Figure 32: Lattice results for the integrated correlator given in Eq. 3.7 and the corresponding effective slope plots for the three operator products $Q_1 \cdot Q_1$, $Q_1 \cdot Q_2$ and $Q_2 \cdot Q_2$ in the case $M_K = 1162$ MeV on the 16^3 lattice. The three lines in the upper panel give the linear fits to the data in the time interval [9,18] used to extract the corresponding values given in Tab. 9. The horizontal lines with error bands in the lower panel show slopes from the same three linear fits with errors. (Note the slope of the integrated correlator as a function of the time T given in Eq. (3.7) must be multiplied by -2 to obtain the corresponding contribution to ΔM_K.)

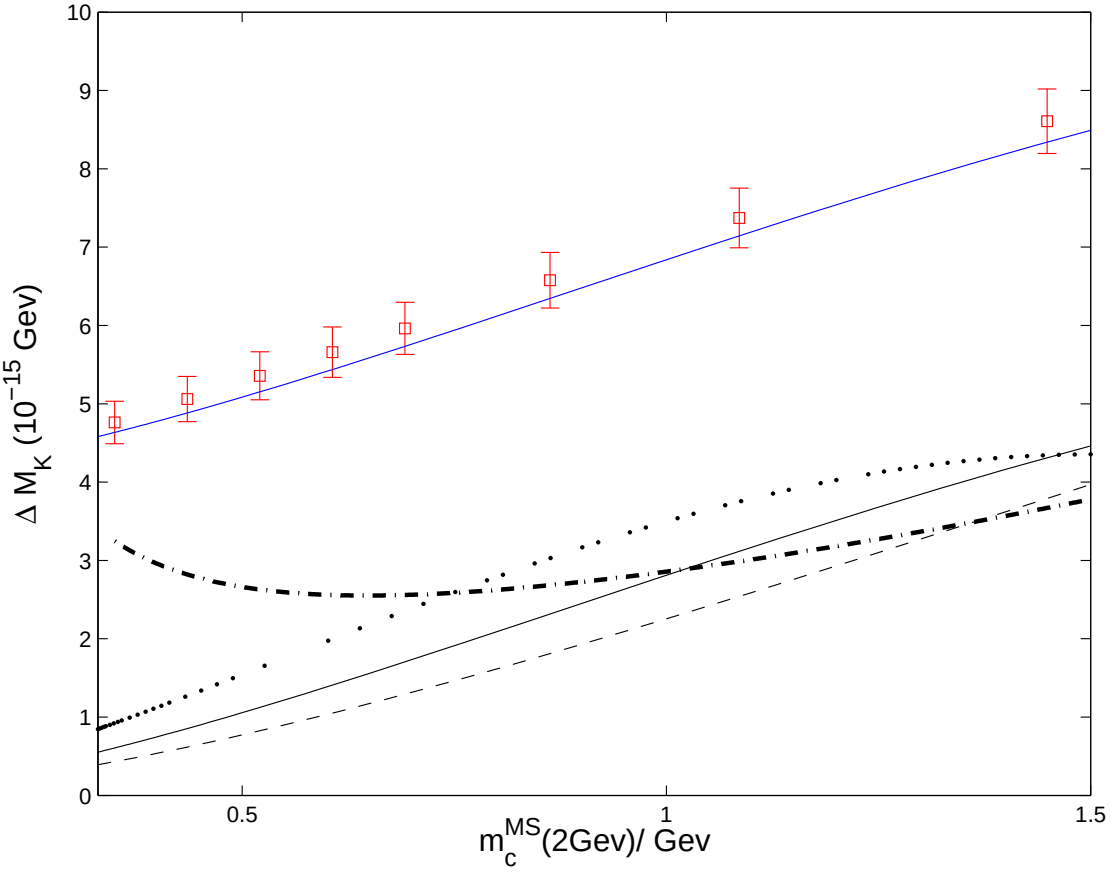


Figure 33: The lattice results for ΔM_K plotted as a function of m_c for the single kaon mass $M_K = 563$ on the 16^3 lattice. Here the charm quark mass is defined in the $\overline{\text{MS}}$ scheme at a scale $\mu = 2$ GeV. The top solid curve is the result of a correlated fit to the ansatz given in Eq. (5.16). The same result but with the long distance constant a omitted gives the lowest, solid curve. The dotted and dashed lines give the perturbative result for the choices of matching scale $\mu_c = 1$ and 1.5 GeV respectively. Finally the dash-dot curve corresponds to the choice $\mu = m_c$.

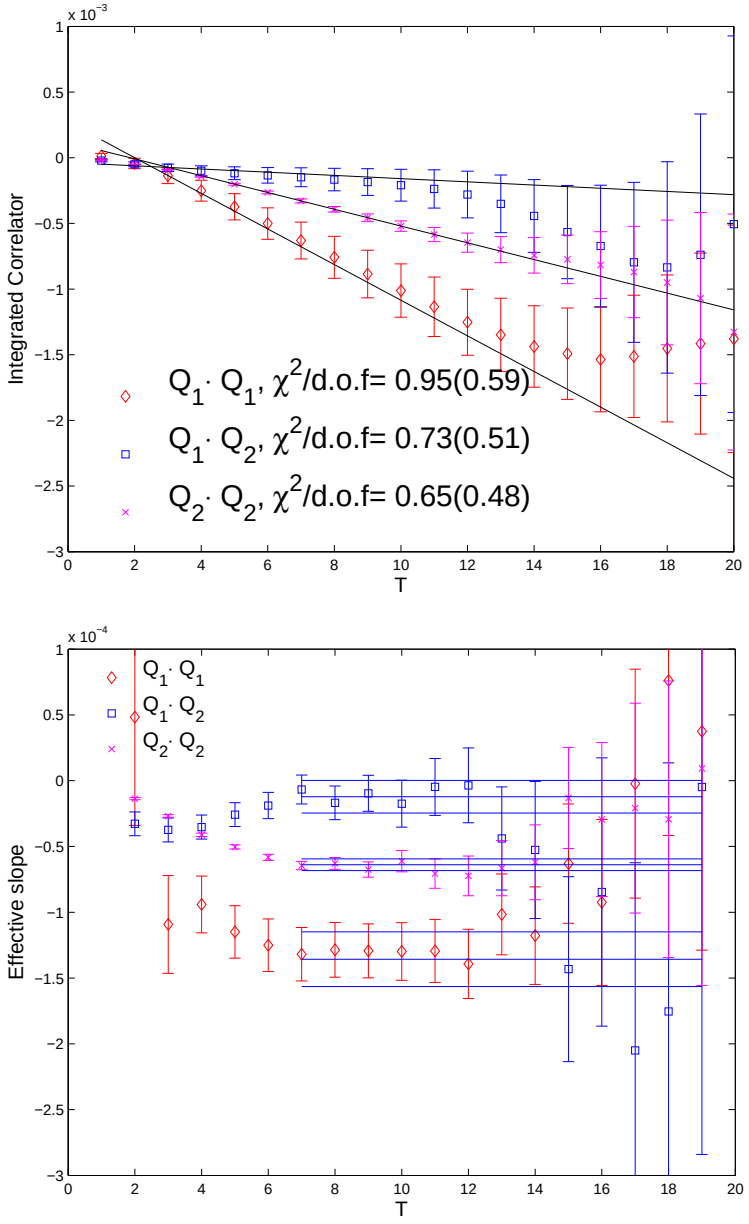


Figure 34: Integrated correlators and the corresponding effective slope plots for the three products of operators $Q_1 \cdot Q_1$, $Q_1 \cdot Q_2$ and $Q_2 \cdot Q_2$. We choose $\Delta_{\min} = 6$, which is the minimum separation between kaon sources and the effective Hamiltonian densities. The three lines in the upper panel give the linear fits to the data in the time interval $[7, 20]$. The horizontal lines with error bands in the lower panel show slopes from the same three linear fits with their errors. These results are obtained on the 24^3 lattice.

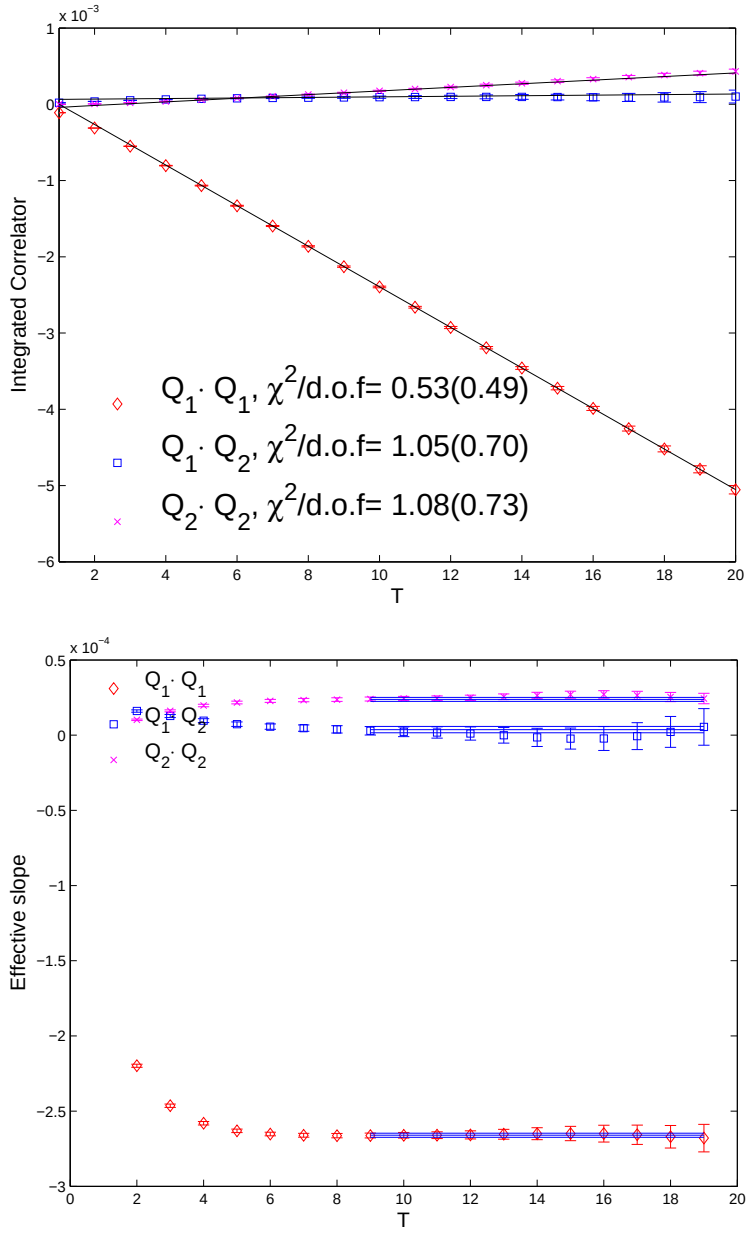


Figure 35: Integrated correlators and the corresponding effective slope plots from type 1 diagrams. We choose $\Delta_{\min} = 6$, which is the minimum separation between kaon sources and the effective Hamiltonian densities. The three lines in the upper panel give the linear fits to the data in the time interval $[9, 20]$. The horizontal lines with error bands in the lower panel show slopes from the same three linear fits with their errors. These results are obtained on the 24^3 lattice.

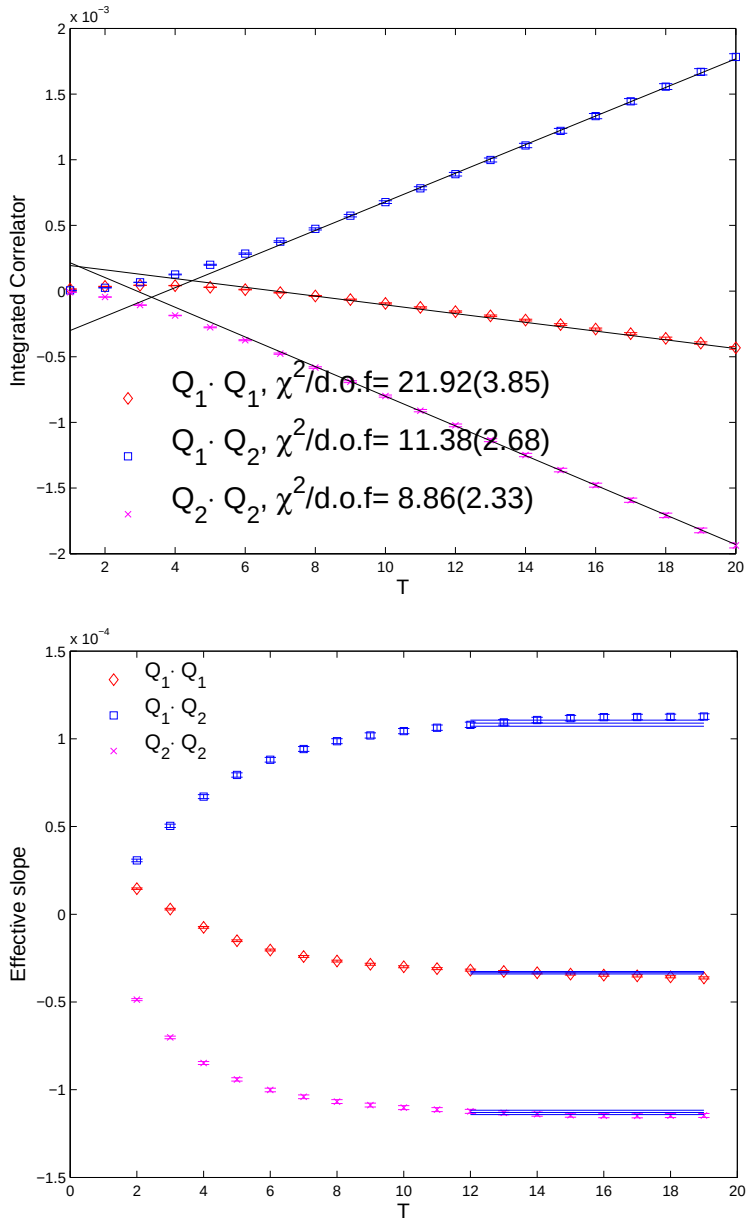


Figure 36: Integrated correlators and the corresponding effective slope plots from type 2 diagrams. We choose $\Delta_{\min} = 6$, which is the minimum separation between kaon sources and the effective Hamiltonian densities. The three lines in the upper panel give the linear fits to the data in the time interval $[12, 20]$. The horizontal lines with error bands in the lower panel show slopes from the same three linear fits with their errors. These results are obtained on the 24^3 lattice.

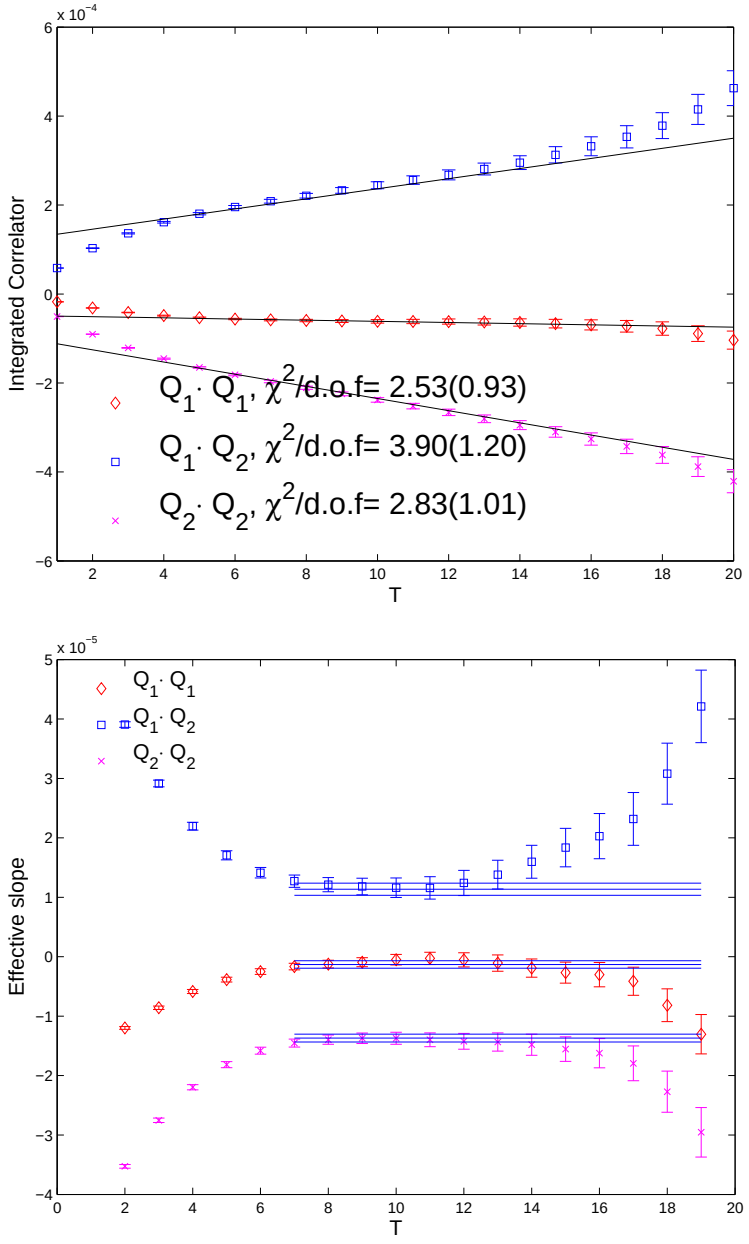


Figure 37: Integrated correlators and the corresponding effective slope plots from type 3 diagrams. We choose $\Delta_{\min} = 6$, which is the minimum separation between kaon sources and the effective Hamiltonian densities. The three lines in the upper panel give the linear fits to the data in the time interval $[7, 20]$. The horizontal lines with error bands in the lower panel show slopes from the same three linear fits with their errors. These results are obtained on the 24^3 lattice.

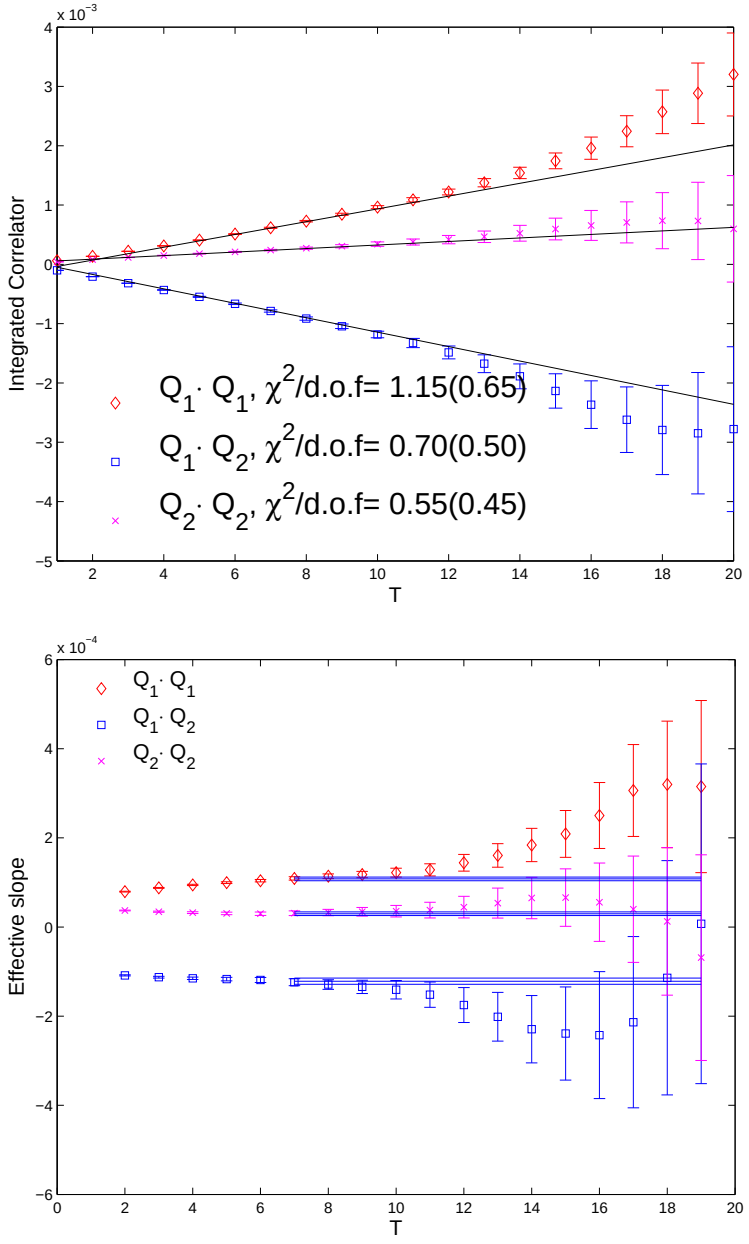


Figure 38: Integrated correlators and the corresponding effective slope plots from type 4 diagrams. We choose $\Delta_{\min} = 6$, which is the minimum separation between kaon sources and the effective Hamiltonian densities. The three lines in the upper panel give the linear fits to the data in the time interval $[7, 20]$. The horizontal lines with error bands in the lower panel show slopes from the same three linear fits with their errors. These results are obtained on the 24^3 lattice.

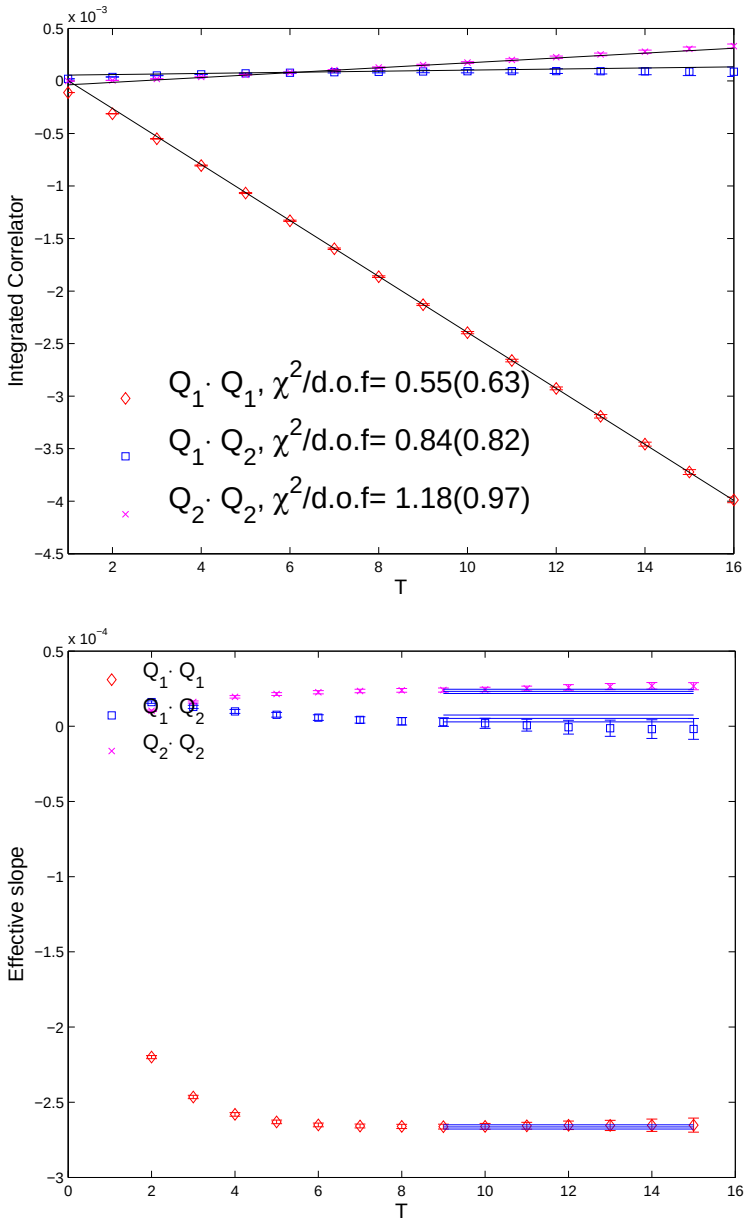


Figure 39: Integrated correlators and the corresponding effective slope plots from type 1 diagrams. We choose $\Delta_{\min} = 8$, which is the minimum separation between kaon sources and the effective Hamiltonian densities. The three lines in the upper panel give the linear fits to the data in the time interval $[9, 16]$. The horizontal lines with error bands in the lower panel show slopes from the same three linear fits with their errors. These results are obtained on the 24^3 lattice.

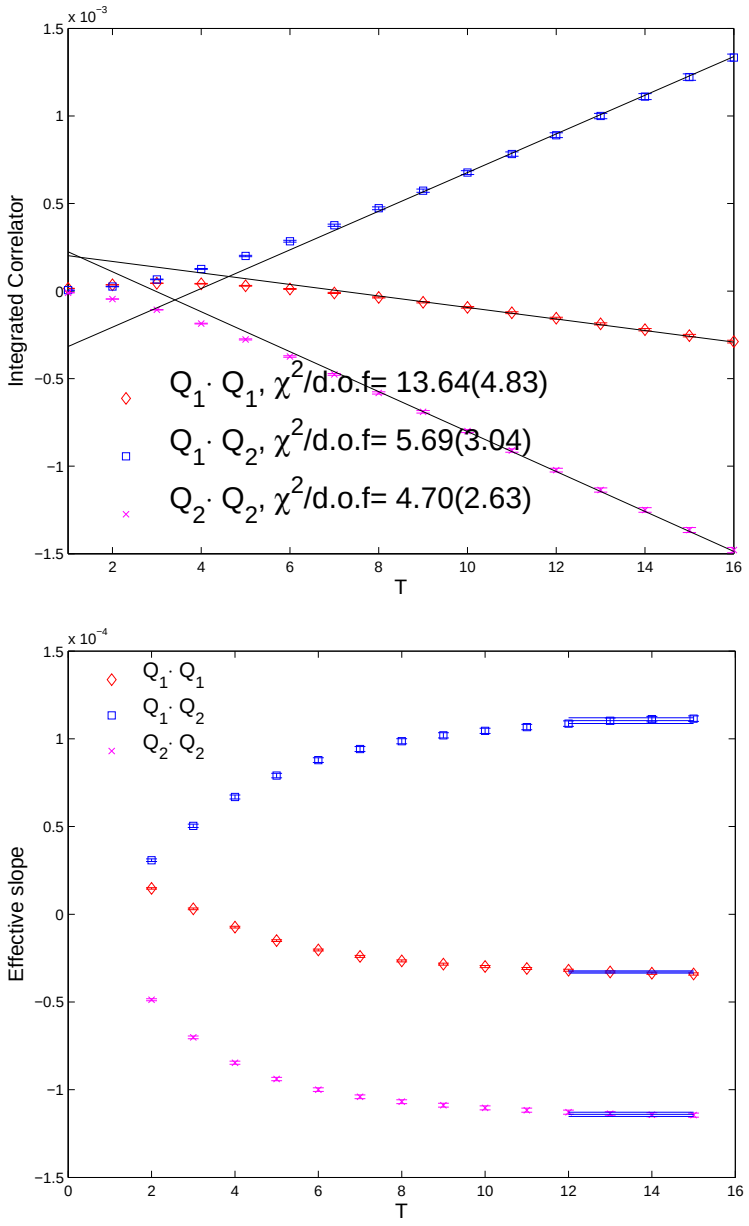


Figure 40: Integrated correlators and the corresponding effective slope plots from type 2 diagrams. We choose $\Delta_{\min} = 8$, which is the minimum separation between kaon sources and the effective Hamiltonian densities. The three lines in the upper panel give the linear fits to the data in the time interval $[12, 16]$. The horizontal lines with error bands in the lower panel show slopes from the same three linear fits with their errors. These results are obtained on the 24^3 lattice.

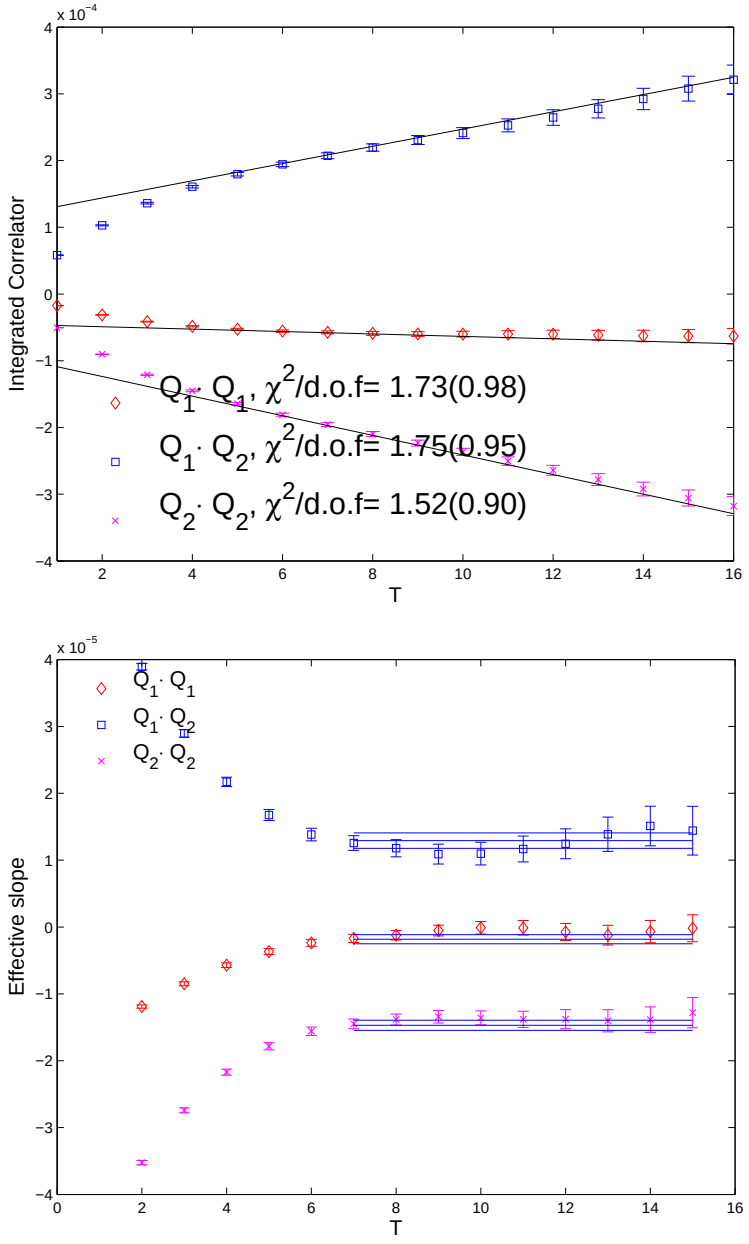


Figure 41: Integrated correlators and the corresponding effective slope plots from type 3 diagrams. We choose $\Delta_{\min} = 8$, which is the minimum separation between kaon sources and the effective Hamiltonian densities. The three lines in the upper panel give the linear fits to the data in the time interval $[7, 16]$. The horizontal lines with error bands in the lower panel show slopes from the same three linear fits with their errors. These results are obtained on the 24^3 lattice.

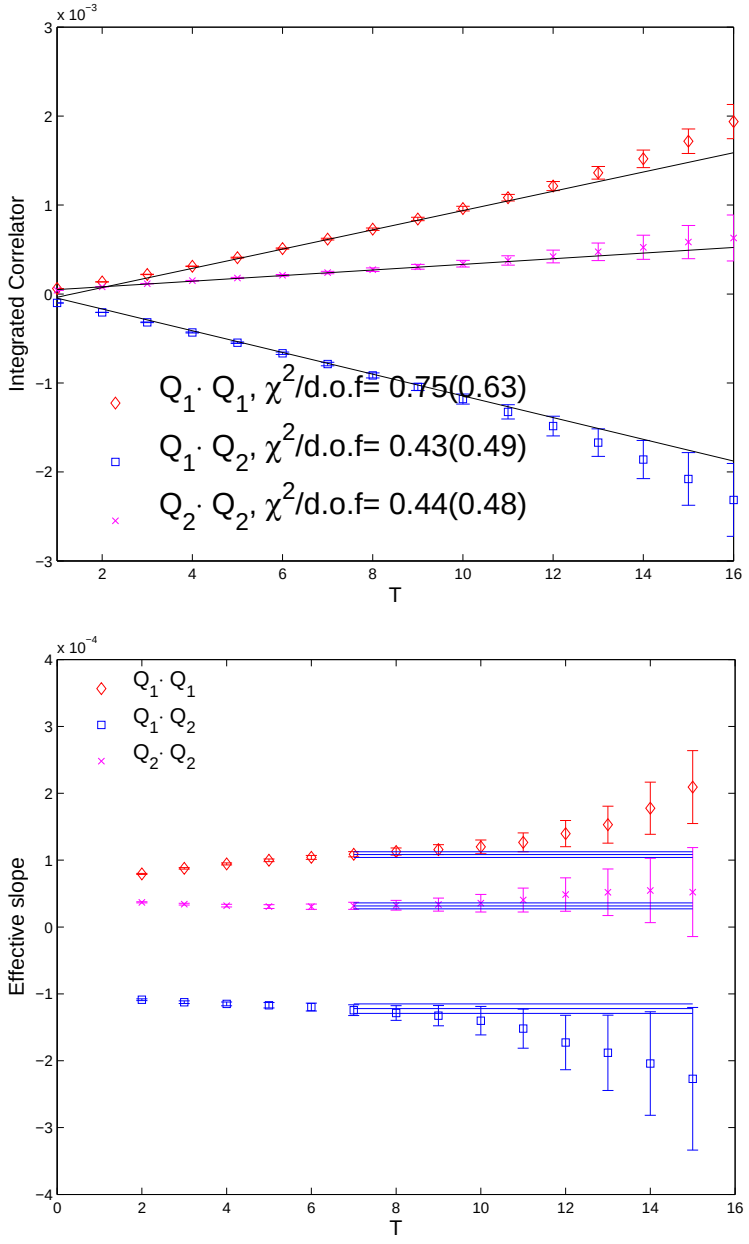


Figure 42: Integrated correlators and the corresponding effective slope plots from type 4 diagrams. We choose $\Delta_{\min} = 8$, which is the minimum separation between kaon sources and the effective Hamiltonian densities. The three lines in the upper panel give the linear fits to the data in the time interval $[7, 16]$. The horizontal lines with error bands in the lower panel show slopes from the same three linear fits with their errors. These results are obtained on the 24^3 lattice.

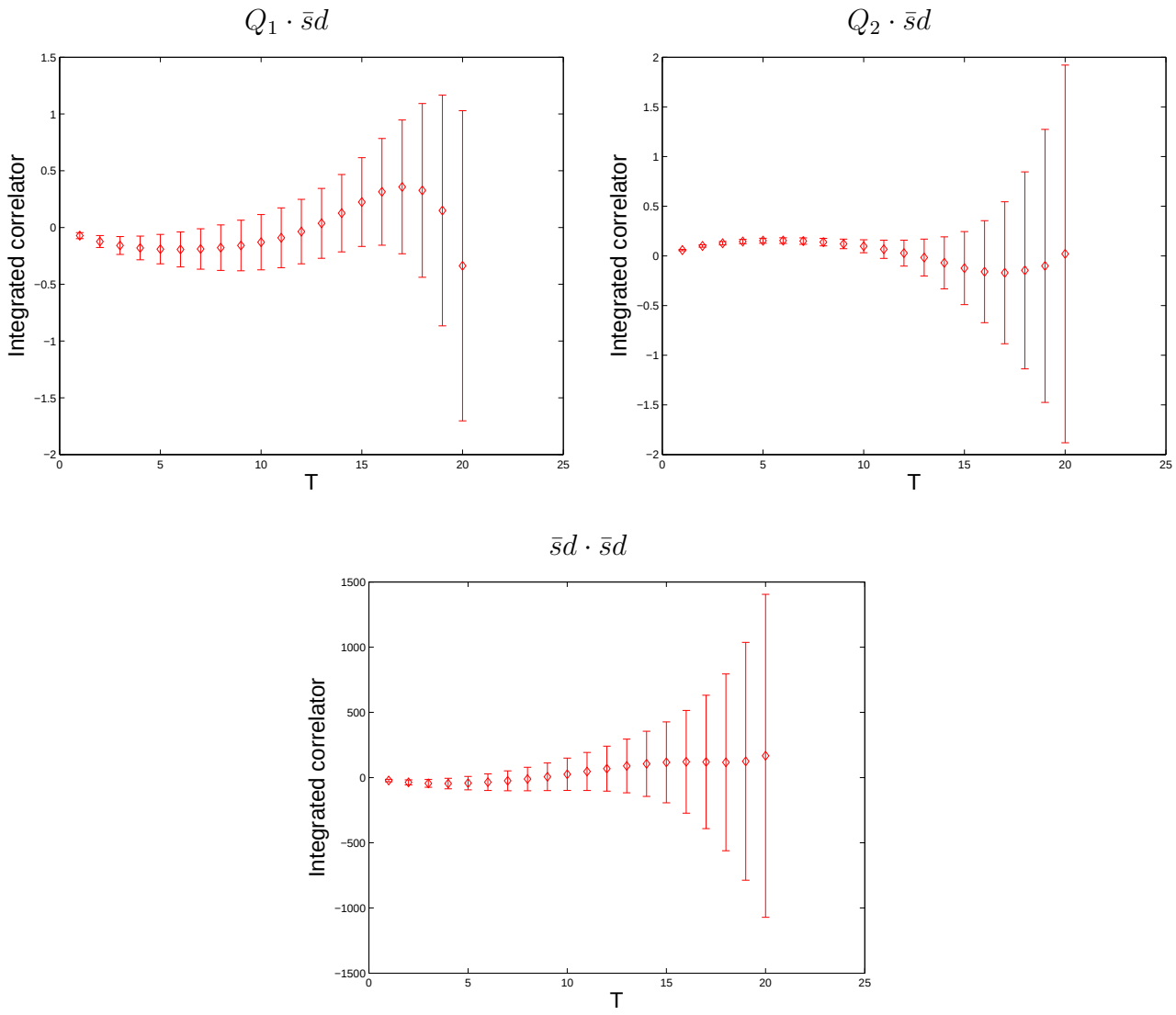


Figure 43: Integrated correlators for the operator products $(Q_i \cdot \bar{s}d)$ and $(\bar{s}d \cdot \bar{s}d)$. We choose $\Delta_{\min} = 6$, which is the minimum separation between the kaon sources and the weak Hamiltonians. These results are obtained from the 24^3 lattice.

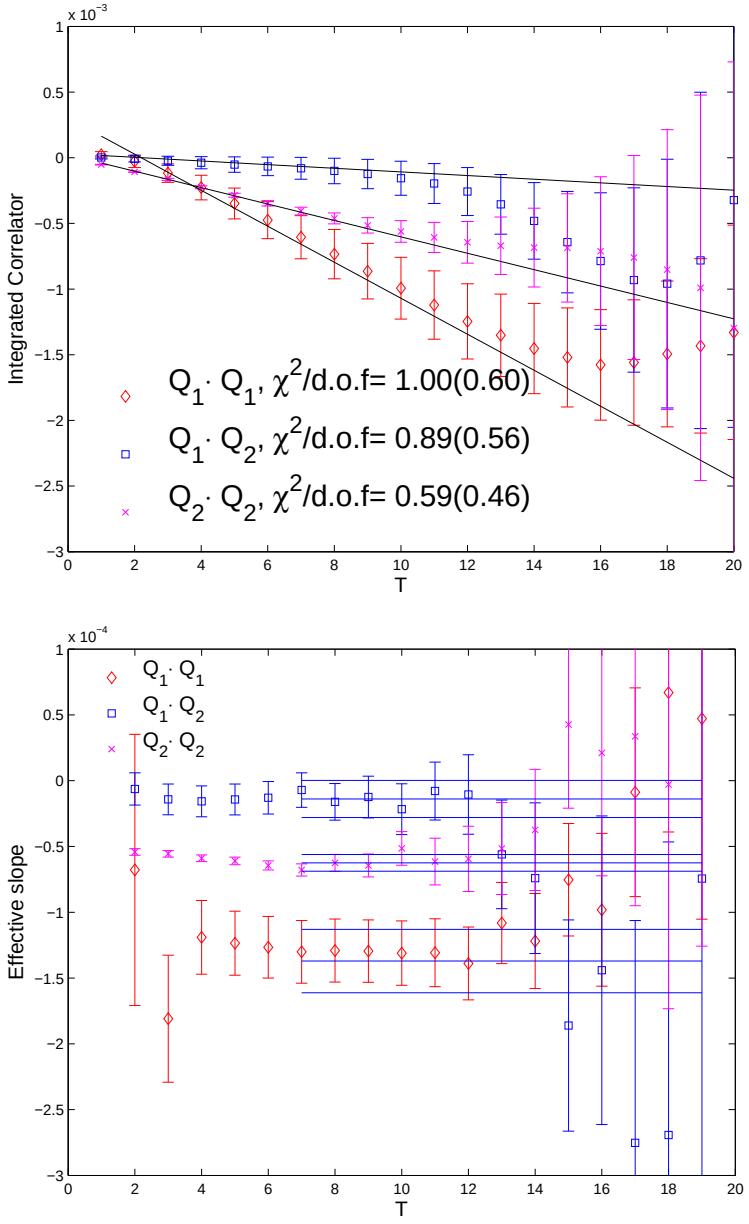


Figure 44: Integrated correlators and the corresponding effective slope plots for $c_p = 1$ and $c_s = -1$. c_s and c_p are defined in Eq.6.1. We choose $\Delta_{\min} = 6$, which is the minimum separation between kaon sources and the effective Hamiltonian densities. The three lines in the upper panel give the linear fits to the data in the time interval $[7, 20]$. The horizontal lines with error bands in the lower panel show slopes from the same three linear fits with their errors. These results are obtained on the 24^3 lattice.

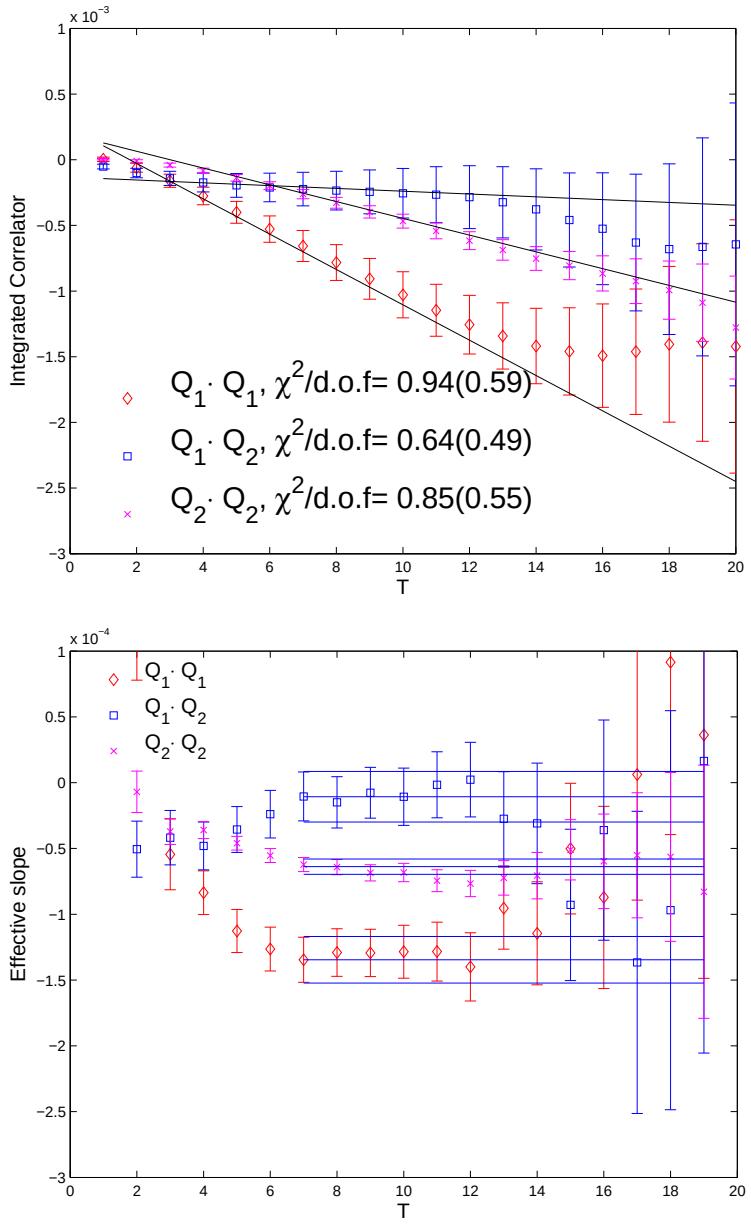


Figure 45: Integrated correlators and the corresponding effective slope plots for $c_p = 1$ and $c_s = 1$. c_s and c_p are defined in Eq.6.1. We choose $\Delta_{\min} = 6$, which is the minimum separation between kaon sources and the effective Hamiltonian densities. The three lines in the upper panel give the linear fits to the data in the time interval $[7, 20]$. The horizontal lines with error bands in the lower panel show slopes from the same three linear fits with their errors. These results are obtained on the 24^3 lattice.

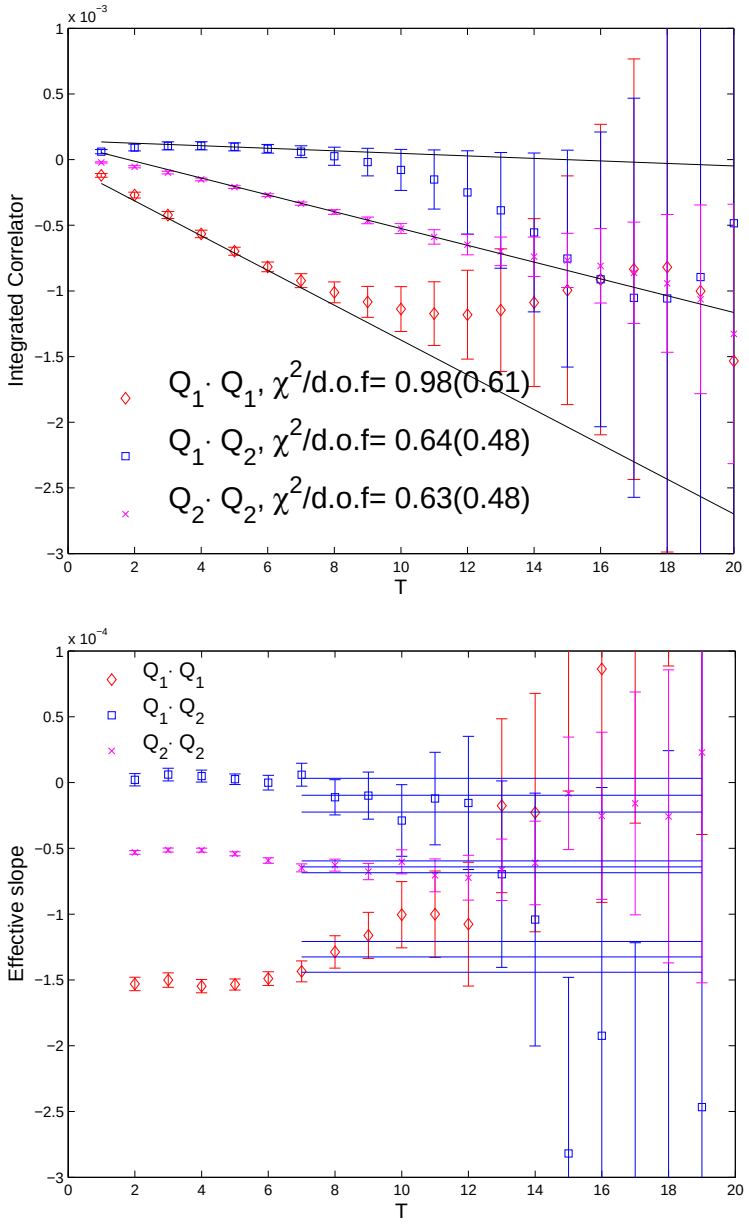


Figure 46: Integrated correlators and the corresponding effective slope plots for $c_p = 1$ and $c_s = -1.14(26)$. c_s and c_p are defined in Eq.6.1. This value of c_s will eliminate the η intermediate state. We choose $\Delta_{\min} = 6$, which is the minimum separation between kaon sources and the effective Hamiltonian densities. The three lines in the upper panel give the linear fits to the data in the time interval $[7, 20]$. The horizontal lines with error bands in the lower panel show slopes from the same three linear fits with their errors. These results are obtained on the 24^3 lattice.

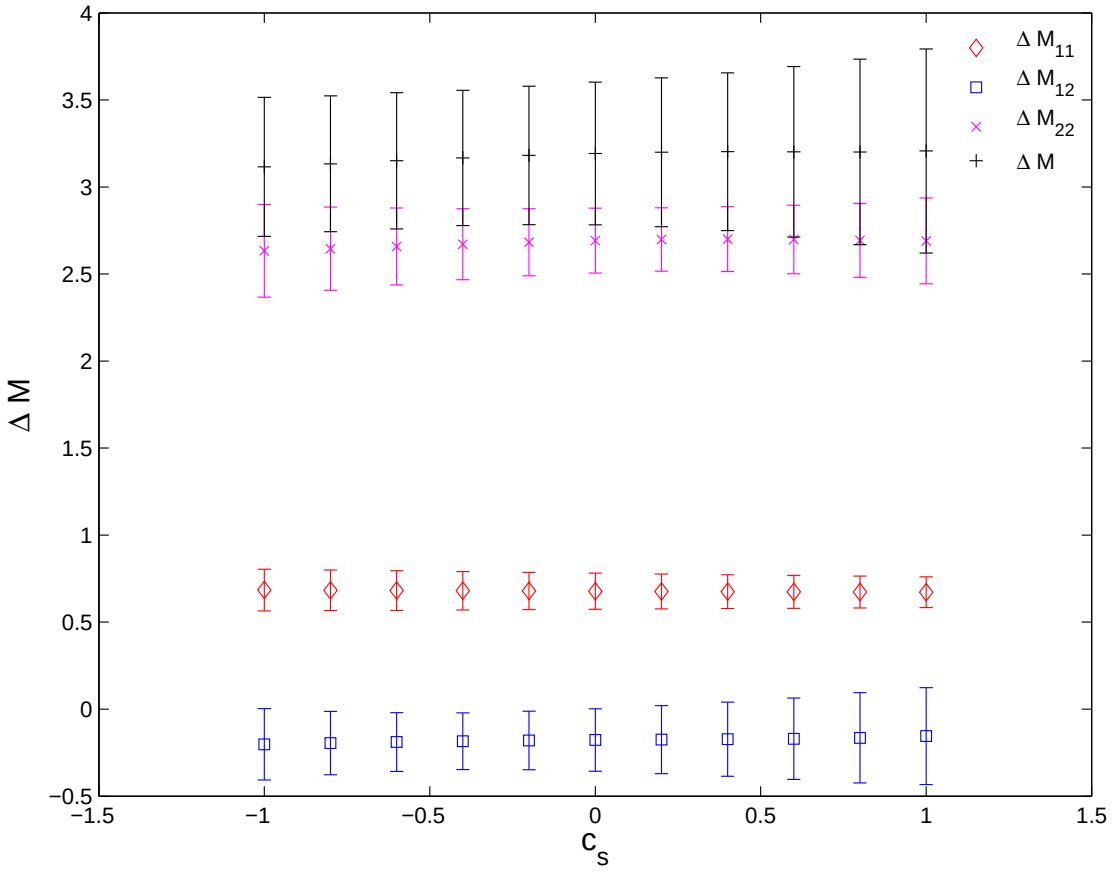
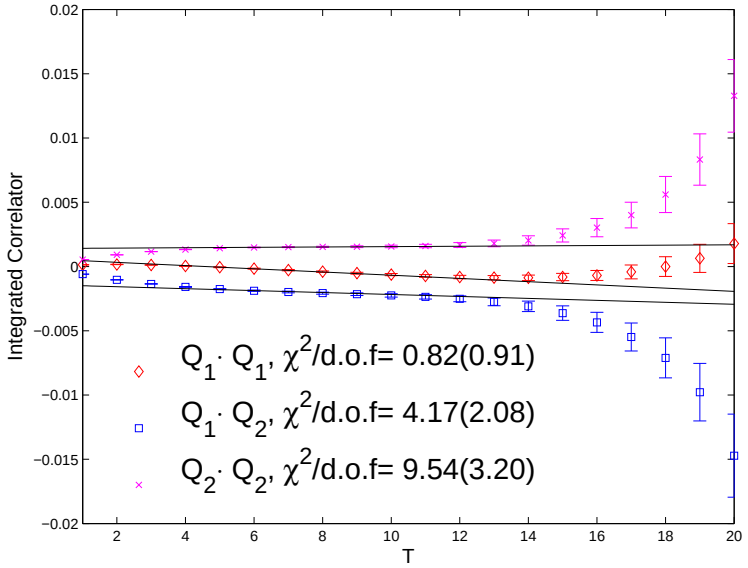
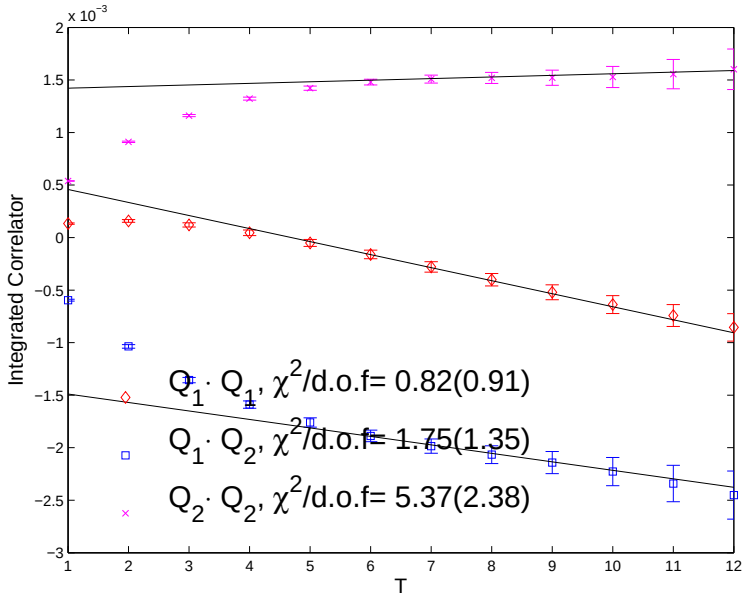


Figure 47: Mass difference for different choice of c_s while fixing $c_p = 1$. We choose $\Delta_{\min} = 6$, which is the minimum separation between the kaon sources and the weak Hamiltonians. The minimal fitting time $T_{\min}$ is 7. Results are given for different operator products and for their sum. These results are obtained on the 24^3 lattice.



$c_p = 0,\ c_s = 1,\ \Delta_{\min} = 6$



$c_p = 0,\ c_s = 1,\ \Delta_{\min} = 10$

Figure 48: Integrated correlator for $c_p = 0$ and $c_s = 0$. We give the results for both $\Delta_{\min} = 6$ and $\Delta_{\min} = 10$. $\Delta_{\min}$ is the minimum separation between the kaon sources and the weak Hamiltonians. All the standard exponentially increasing terms have been removed. The straight lines give the linear fits to the data in the time interval [7,12]. These results are obtained on the 24^3 lattice.

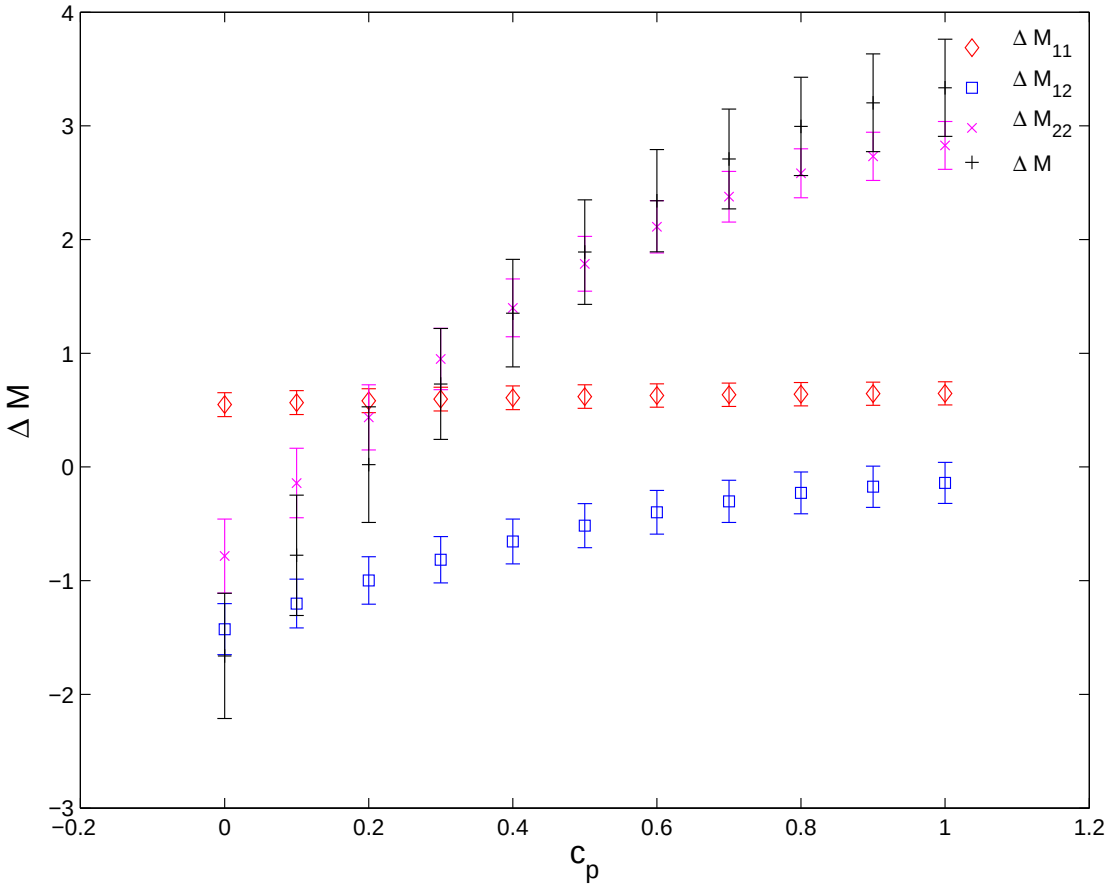


Figure 49: Mass difference for different choice of c_p while fixing $c_s = 0$. We choose $\Delta_{\min} = 10$, which is the minimum separation between the kaon sources and the weak Hamiltonians. The minimal fitting time $T_{\min}$ is 7. Results are given for different operator products and for their sum. These results are obtained on the 24^3 lattice.

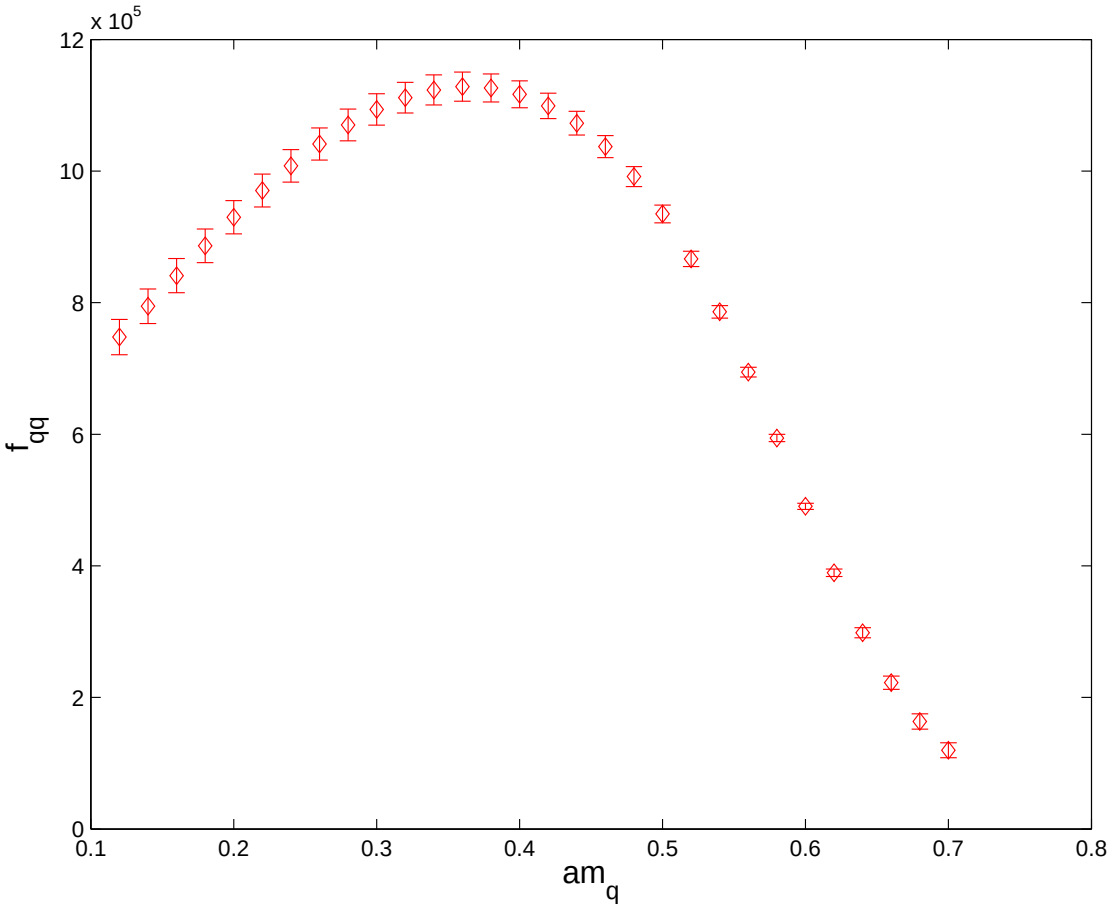


Figure 50: $f_{q\bar{q}}$ as a function of quark mass m_q on the 24^3 lattice. The quantity $f_{q\bar{q}}$ is the product of the pseudoscalar decay constant and the normalizaiton factor of the gauge fixed wall source. (The author thanks Ziyuan Bai for providing this figure.)