Explorations beyond dilaton chiral perturbation theory in the eight-flavor SU(3) gauge theory

Maarten Golterman[1] and Yigal Shamir[2]

[1]*Department of Physics and Astronomy, San Francisco State University,*
San Francisco, California 94132, USA
[2]*Raymond and Beverly Sackler School of Physics and Astronomy,*
Tel Aviv University, 69978 Tel Aviv, Israel

We continue our study of spectroscopy data for the SU(3) gauge theory with eight fundamental fermions, motivated by the effective field theory framework of dilaton chiral perturbation theory (dChPT). At leading order dChPT predicts a constant mass anomalous dimension γ_m, consistent with the assumed proximity of an infrared fixed point. For the relatively large fermion masses simulated by the LatKMI Collaboration, the influence of the infrared fixed point diminishes, and our fits suggest that γ_m starts running. Since a complete higher-order analysis is not feasible with presently available data, we adopt a more phenomenological approach. We propose a partial extension to higher orders, which incorporates the running of γ_m into the tree-level Lagrangian. We find that this extension successfully describes the full fermion-mass range of the LatKMI data, including the pion taste splittings which arise from using staggered fermions in the lattice simulations. We also investigate a more general class of dilaton potentials proposed in the literature, using both the LSD and LatKMI datasets, concluding that these data favor the form predicted by dChPT.

I. INTRODUCTION

Lattice simulations of the SU(3) gauge theory with eight Dirac fermions in the fundamental representation have revealed the existence of a flavor-singlet scalar particle, which, at the fermion masses explored in these simulations, is approximately degenerate with the pions—the Nambu-Goldstone bosons associated with chiral symmetry breaking [1–4]. A similar light scalar has also been found in the SU(3) gauge theory with two sextet fermions [5–9], or with four light and six [10] or eight [11] heavy fundamental fermions.[1]

The existence of a light flavor-singlet scalar particle roughly degenerate with the pions means that, besides the pions, any effective field theory (EFT) description of the low-energy behavior has to include a field that represents this scalar particle. Here, our starting point is dilaton chiral perturbation theory (dChPT), an EFT in which the lightness of the scalar particle is assumed to arise from approximate scale invariance of the underlying theory in the infrared

[17–21].[2] Increasing the number of (massless) fermionic degrees of freedom will eventually take the theory into the conformal window, where the non-Abelian gauge theory is still asymptotically free, but develops an infrared fixed point (IRFP). The idea is that, with eight flavors, the SU(3) gauge theory is still outside the conformal window, but close enough to the conformal sill—the number of flavors where the IRFP first develops—that the breaking of scale invariance in the infrared is governed by the proximity of the IRFP. The key assumption is then that the distance to the conformal sill can be treated as a small parameter, in which a systematic power counting can be developed. The scalar particle, which we will refer to as the dilaton, is interpreted as a pseudo–Nambu-Goldstone boson (pNGB) for the approximate scale symmetry [17]. The mass of the dilaton is controlled by this small parameter, just as the fermion mass leads to a parametrically small pion mass. Since the fermion mass breaks scale invariance too, the dilaton mass will also depend on the fermion mass.

In a previous paper [33] we applied leading-order (LO) dChPT to numerical data for the eight-flavor SU(3) gauge theory produced in lattice simulations by the LSD Collaboration [3]. We showed that, over the fermion mass

---

[1]For reviews of lattice work, see Refs. [12–16].

[2]For early work, and for other low-energy approaches, see Refs. [22–32].

range in these simulations, LO dChPT successfully describes the pNGB sector of the theory, including the dilaton. In Ref. [3] staggered fermions were used, which exhibit taste splittings—a lattice artifact mass splitting of the pion multiplet caused by a partial breaking of the flavor symmetry group in the staggered fermion formulation.[3] We showed that dChPT explains the pattern of taste splittings in the pion sector observed in Ref. [3] as a function of the fermion mass. The vacuum expectation value of the dilaton field depends on the fermion mass already in LO, leading to a fermion-mass dependence of pNGB decay constants and masses that is qualitatively different from QCD. This includes the taste splittings, which are also qualitatively different from the pattern seen in QCD with staggered fermions.

Given this success, our goal in this paper is to investigate whether dChPT can also be applied to the other major lattice study of the eight-flavor SU(3) gauge theory, by the LatKMI Collaboration [4].[4] This study also used staggered fermions and presented extensive spectroscopy data for the pNGB sector, including taste splittings. The KMI simulations were done at larger fermion masses than those of LSD. Even if dChPT is the correct EFT, the question arises whether one can fit the KMI data using LO dChPT, or, alternatively, whether higher orders in the EFT expansion would be needed. Indeed, unlike for the LSD data [3], we found that LO dChPT does not quantitatively describe the KMI data over the full fermion mass range, as will be discussed in detail in this paper.

For the LSD data, we found that as the fermion mass is varied, hadron masses and decay constants respond with an approximate hyperscaling behavior [20]. As the fermion mass increases, the theory is drawn further away from the influence of the IRFP at the nearby conformal sill. Once the fermion mass becomes large enough, we expect that the running of the coupling will become noticeable, and thus also the running of the mass anomalous dimension γ_m.[5] In dChPT, at leading order, the mass anomalous dimension is constant, $\gamma_m = \gamma_*$, where γ_* is the mass anomalous dimension at the nearby IRFP. dChPT allows for a nonconstant γ_m, but the power counting underlying dChPT accommodates corrections to a constant γ_m only through higher orders. In order to systematically compare dChPT with the KMI data, we would thus have to consider dChPT to next-to-leading order (NLO) or beyond. However, the relatively large number of additional parameters that would be needed already at NLO, and limitations of the presently available lattice data, to be discussed below, prevent us from attempting a complete NLO fit.

Instead, we will take a more phenomenological approach, based on the following observation. The salient difference between the KMI and LSD data appears to be that a constant γ_m cannot account for the full range of (larger) fermion

masses explored in the KMI data. We will thus extend LO dChPT by only including higher-order effects that are directly related to γ_m; we will refer to this extension as γ-dChPT. This makes our approach not systematic, since most NLO and higher-order effects are left out. Strictly speaking, γ-dChPT should thus be viewed as a model approach.

In order for LO dChPT to accommodate a varying γ_m, we will modify the mass-dependent part of the potential, as described in detail in Sec. II. This raises the question of what happens if one also considers a generalization of the dilaton part of the potential. A class of potentials depending on a new parameter Δ, generalizing the dilaton potential of dChPT, has been proposed before [24,25,28,32], and we will refer to this different extension of LO dChPT as Δ-dChPT. One recovers LO dChPT, including its dilaton potential, by taking $\Delta \to 4$. It is interesting to also confront Δ-dChPT with the data. We will revisit the analysis of the LSD data using Δ-dChPT by Ref. [32], and extend this investigation to the KMI data. Despite claims in the literature [32], Δ-dChPT takes us outside the systematic power counting of dChPT, and should thus be considered as a more phenomenological approach to the low-energy behavior of the $N_f = 8$ theory.

This paper is organized as follows. In Sec. II we introduce γ-dChPT, in which LO dChPT is extended to accommodate a varying γ_m. In Sec. III we first present our evidence that γ_m, as well as other LO parameters, are changing over the KMI mass range in a fit to LO dChPT. We then apply γ-dChPT to the pNGB sector of the KMI data. We find that a rather simple model for a varying γ_m provides good fits of the KMI data, including taste splittings. In Sec. IV we consider the generalized class of dilaton potentials, reviewing the application of Δ-dChPT to the LSD data, and applying it to the KMI data. Combining these results provides some evidence that the dilaton potential of LO dChPT is preferred by the data, i.e., that the preferred value in Δ-dChPT is close to $\Delta = 4$. Finally, Sec. V contains our conclusions. In Appendix A we elaborate on the choice of a mass-independent scale setting prescription. In Appendix B we investigate the claim of Ref. [32] that Δ-dChPT admits a systematic power counting for any value of Δ and show that this claim is incorrect.

II. DILATON CHPT AND γ_m

In Sec. II A, we begin with a summary of LO dChPT. This is the EFT that was applied to the LSD data in Ref. [33]. In Sec. II B we revisit the physics of hyperscaling and its manifestation in LO dChPT. This leads us in Sec. II C to introduce γ-dChPT, where we generalize the low-energy Lagrangian to accommodate a nonconstant mass anomalous dimension. We emphasize that this extension takes us outside the strict EFT framework. In Sec. II D we present the hadronic quantities to be fit to the KMI data of Ref. [4] in the rest of this paper.

---

[3]For reviews, see Refs. [34,35].
[4]We will often shorten "LatKMI" to just "KMI."
[5]For an early study of γ_m in the $N_f = 8$ theory, see Ref. [36].

A. dChPT at lowest order

The Euclidean LO Lagrangian for dChPT is given by

$$\mathcal{L} = \frac{1}{2} f_\tau^2 e^{2\tau} \partial_\mu \tau \partial_\mu \tau + \frac{1}{4} f_\pi^2 e^{2\tau} \text{tr}(\partial_\mu \Sigma^\dagger \partial_\mu \Sigma)$$
$$+ \mathcal{L}_m(\tau, \Sigma) + \mathcal{L}_d(\tau). \tag{2.1}$$

The potential terms are

$$\mathcal{L}_d(\tau) = f_\tau^2 B_\tau e^{4\tau}(c_0 + c_1 \tau), \tag{2.2a}$$

$$\mathcal{L}_m(\tau, \Sigma) = -\frac{1}{2} f_\pi^2 B_\pi m e^{(3-\gamma_*)\tau} \text{tr}(\Sigma + \Sigma^\dagger). \tag{2.2b}$$

Here Σ is the usual nonlinear field describing the pion multiplet, while τ is the dilaton effective field. $\mathcal{L}$ depends on the low-energy constants (LECs) f_τ, f_π, B_π, B_τ, γ_*, c_0, and c_1.

We define the theory in the Veneziano limit [37], in which $N \equiv N_c \propto N_f$ is taken to infinity keeping the ratio $n_f = N_f/N_c$ fixed, with N_f the number of fundamental-representation flavors and N_c the number of colors. The power counting is [17]

$$p^2 \sim m \sim n_f - n_f^* \sim 1/N. \tag{2.3}$$

The relation $p^2 \sim m$ defines the power counting of ordinary ChPT.[6] The small parameter controlling the hard breaking of scale invariance is $n_f - n_f^*$, where n_f^* is the limiting value of n_f for the theory at the conformal sill: the boundary between the regime where the massless theory undergoes chiral symmetry breaking, and the regime where this theory is conformal in the infrared, i.e., where the gauge coupling g runs into an infrared fixed point g_*.

Invoking the proximity of the sill of the conformal window, we assume that the β function is small at the chiral symmetry breaking scale, and that the corresponding value of g is close to g_*. We can then expand the mass anomalous dimension $\gamma(g)$ in powers of $n_f - n_f^*$ around $\gamma_* = \gamma(g_*)$, the mass anomalous dimension at the infrared fixed point at the conformal sill. For a detailed discussion of the construction of the LO Lagrangian, and the underlying power counting, see Refs. [17,20].

In the dilaton potential (2.2), c_0 is $\mathcal{O}(1)$, while c_1 is proportional to the small expansion parameter $n_f - n_f^*$.[7] For $m = 0$, we shift the τ field to $\tau + v_0$, with $v_0 = \langle \tau \rangle|_{m=0}$ (before the shift). After the shift, the dilaton expectation value $v(m) = \langle \tau \rangle$ vanishes in the massless theory. Defining

---

[6]The dimensionful quantities, p^2 and m, are measured in units of the dynamically generated infrared scale of the massless theory.

[7]For a few more details about the power counting, see Appendix B.

$$\hat{f}_{\pi,\tau} = e^{v_0} f_{\pi,\tau}, \tag{2.4a}$$

$$\hat{B}_\tau = e^{2v_0} B_\tau, \tag{2.4b}$$

$$\hat{B}_\pi = e^{(1-\gamma_*)v_0} B_\pi, \tag{2.4c}$$

the Lagrangian becomes

$$\mathcal{L} = \frac{1}{2} \hat{f}_\tau^2 e^{2\tau} \partial_\mu \tau \partial_\mu \tau + \frac{1}{4} \hat{f}_\pi^2 e^{2\tau} \text{tr}(\partial_\mu \Sigma^\dagger \partial_\mu \Sigma)$$
$$+ \mathcal{L}_m(\tau, \Sigma) + \mathcal{L}_d(\tau), \tag{2.5}$$

with

$$\mathcal{L}_d(\tau) = \hat{f}_\tau^2 \hat{B}_\tau e^{4\tau} V_d(\tau), \tag{2.6a}$$

$$V_d(\tau) = c_1\left(\tau - \frac{1}{4}\right), \tag{2.6b}$$

$$\mathcal{L}_m(\tau, \Sigma) = -\frac{1}{2} \hat{f}_\pi^2 \hat{B}_\pi m e^{(3-\gamma_*)\tau} \text{tr}(\Sigma + \Sigma^\dagger). \tag{2.6c}$$

The shift sets $c_0 = -c_1/4$, and now the whole LO Lagrangian is $\mathcal{O}(p^2)$ in the power counting (2.3). We will assume $c_1 > 0$, so that the potential $\mathcal{L}_d + \mathcal{L}_m$ is bounded from below.

Assuming $m \geq 0$, the potential is minimized by $\Sigma = 1$. The dilaton expectation value $v = v(m)$ solves the saddle-point equation

$$\frac{(3-\gamma_*)m}{4c_1 \mathcal{M}} = v e^{(1+\gamma_*)v}, \qquad \mathcal{M} = \frac{\hat{f}_\tau^2 \hat{B}_\tau}{\hat{f}_\pi^2 \hat{B}_\pi N_f}. \tag{2.7}$$

The solution is positive and monotonically increasing with m. The spectroscopy data we considered in Ref. [33] can then be expressed as functions of m,

$$\frac{M_\pi^2}{F_\pi^2} = \frac{1}{d_1} v(m) \equiv h(m), \tag{2.8a}$$

$$F_\pi = \hat{f}_\pi e^{v(m)} \tag{2.8b}$$

$$= \left(\frac{d_0 m}{h(m)}\right)^{\frac{1}{1+\gamma_*}}, \tag{2.8c}$$

$$\frac{M_\tau^2}{F_\tau^2} = d_3(1 + (1+\gamma_*)d_1 h(m)). \tag{2.8d}$$

Explicitly,

$$h(m) = \frac{1}{(1+\gamma_*)d_1} W_0\left(\frac{(1+\gamma_*)d_1}{d_2}m\right), \tag{2.9}$$

where W_0 is the Lambert W function. The parameters $d_{0,1,2,3}$ are defined in terms of the LECs of the tree-level Lagrangian,

$$d_0 = \frac{2\hat{B}_\pi}{\hat{f}_\pi^{1-\gamma_*}}, \quad d_1 = \frac{(3-\gamma_*)\hat{f}_\pi^2}{8\hat{B}_\pi c_1 \mathcal{M}}, \quad d_2 = \frac{\hat{f}_\pi^2}{2\hat{B}_\pi}, \quad d_3 = \frac{4c_1 \hat{B}_\pi}{\hat{f}_\pi^2}.$$

$$(2.10)$$

In Ref. [33] we applied LO dChPT, as summarized above, to the LSD data [3]. The key assumptions underlying this analysis were as follows: (a) the $N_f = 8$, $N_c = 3$ theory undergoes chiral symmetry breaking; (b) for the LSD mass range, the β function is small enough that the dChPT power counting is applicable. The results of our analysis corroborated these assumptions.

B. Hyperscaling

Consider momentarily a mass-deformed infrared conformal theory. We can probe the theory over a range of scales where g is so close to the infrared fixed-point g_* that all effects of its running can be neglected. The breaking of scale invariance is then driven entirely by the input bare fermion mass m_0. Under these circumstances, any hadronic mass M follows a simple hyperscaling law,

$$\frac{M}{\Lambda_{\rm UV}} \sim \left(\frac{m_0}{\Lambda_{\rm UV}}\right)^{\frac{1}{1+\gamma_*}}. \qquad (2.11)$$

Here $\Lambda_{\rm UV}$ is an ultraviolet scale for which the approximation $\gamma_m(\mu) = \gamma_*$ is valid for any $\mu \le \Lambda_{\rm UV}$, and $m_0 = m(\Lambda_{\rm UV})$, where $m(\mu)$ is the running renormalized mass. Hyperscaling is based on the following simple observations:

(i) The renormalized mass, $m = m(\mu)$, runs as dictated by its anomalous dimension. By contrast, the renormalized coupling has attained its fixed-point value g_* (up to negligible corrections); hence also the mass anomalous dimension has a fixed value $\gamma_* = \gamma_m(g_*)$.

(ii) No physical scale is generated dynamically in the massless theory. When the fermion mass is nonzero, the induced physical scale M is set by the condition $M \sim m(M)$.

Indeed, starting from the solution for $m(\mu)$ for a constant mass anomalous dimension,

$$\frac{m(\mu)}{m_0} = \left(\frac{\mu}{\Lambda_{\rm UV}}\right)^{-\gamma_*}, \qquad (2.12)$$

the hyperscaling law (2.11) immediately follows by postulating that the typical hadron mass M satisfies $M \sim m(\mu)$ for $\mu = M$. For any $\gamma_* > 0$, the existence of the physical scale M is guaranteed if $m_0 \ll \Lambda_{\rm UV}$. Starting from $m(\mu) = m_0 \ll \mu$ at $\mu = \Lambda_{\rm UV}$, $m(\mu)$ keeps increasing as μ is decreased, until eventually the equality $M = m(M)$ is reached.

Returning to dChPT, in Ref. [33] we found that the LSD data are in the "large-mass" regime [20], where

$$|n_f - n_f^*| \sim c_1 \ll \frac{m_0}{\mathcal{M}}, \qquad (2.13)$$

for all (bare) masses. As follows from the previous subsection,[8] in LO dChPT, c_1 encodes the magnitude of the β function at the chiral symmetry breaking scale. The large-mass regime is thus an approximate hyperscaling regime, where the input fermion mass dominates the breaking of scale invariance. Indeed, in Ref. [20] we showed that the leading mass dependence predicted by LO dChPT in the large-mass regime is the hyperscaling relation (2.11), for all hadronic masses and decay constants. We also calculated corrections to this relation, which are present in dChPT already at LO, because the β function at the chiral symmetry breaking scale, hence c_1, is (by assumption) parametrically small, but not vanishingly small as in a mass-deformed infrared conformal theory. Moreover, we showed that as long as

$$|n_f - n_f^*| \log\left(\frac{m_0}{|n_f - n_f^*|\mathcal{M}}\right) \ll 1, \qquad (2.14)$$

dChPT provide a systematic expansion, even though $m_0/\mathcal{M}$ can be large. By Eq. (2.7), $\mathcal{M}$ is constructed from LECs which can be defined in the chiral limit. It is a striking difference between ordinary ChPT and dChPT that, because of the nearby IRFP, in dChPT a systematic low-energy expansion exists even if the fermion mass is not small relative to the infrared scale of the massless theory, so long as inequality (2.14) holds.

The fermion mass range explored in the KMI data is higher than in the LSD data. The comparison can be made, for example, in units of t_0 (see Fig. 5 of Ref. [2]). We will return to the comparison between the LSD and KMI data, and its limitations, in Sec. III E below. As mentioned in the Introduction, when we increase the input fermion mass the influence of the IRFP diminishes. Eventually, we will reach energy scales where the running of the coupling picks up,[9] and, as a result, so does the running of the mass anomalous dimension. In the next subsection, guided by this consideration, we will develop a generalized notion of hyperscaling, which is founded on the same principles as above, except that the assumption of a constant mass anomalous dimension is relaxed. This will lead to the framework of γ-dChPT, where LO dChPT is extended to accommodate a varying mass anomalous dimension. We stress that the power counting of dChPT allows for corrections to a constant γ_m, but only via higher-order terms in the expansion in $n_f - n_f^*$. In seeking an extension of LO dChPT that accommodates a varying γ_m we are thus asking for a partial resummation of these higher-order terms, under the assumption that these are the dominant higher-order corrections.

---

[8]See also Appendix B.
[9]At extremely high-energy scales perturbation theory will eventually take over, and the β function will tend to zero as dictated by asymptotic freedom.

We conclude this subsection with a technical comment. The hyperscaling law (2.11) can be rewritten as

$$\frac{m_0}{M} \sim \left(\frac{m_0}{\Lambda_{\text{UV}}}\right)^{\frac{\gamma_*}{1+\gamma_*}}. \tag{2.15}$$

It follows that the fermion mass m_0 is always much smaller than any hadronic mass M (as long as $m_0 \ll \Lambda_{\text{UV}}$), and the same is true for the decay constants F_π and F_τ. Moreover, in Ref. [20] we showed that this conclusion extends to $n_f < n_f^*$, below the conformal window, and that it applies also to the masses of the pNGBs, M_π and M_τ. We will assume that the ratio m_0/M remains small also when the simple hyperscaling relations, Eqs. (2.11) and (2.15), are generalized to account for the running of γ_m. Indeed, for the LSD data, m_0/M_π ranges between 0.015 and 0.04, while for the KMI data it ranges between 0.07 and 0.17. Since $m_0/M_\pi \ll 1$, this allows us to use a *mass-independent* renormalization scheme.[10] As we will see below, this greatly simplifies our considerations.

C. Varying γ_m and γ-dChPT

We will now proceed to develop the extension of LO dChPT allowing for a scale-dependent γ_m. The renormalization group (RG) equation governing the dependence of the renormalized mass m on the renormalization scale μ is closely related to the behavior of the renormalized mass under scale transformations. In order to relate the two, we first review how a scale is introduced into the bare theory; we will do this using dimensional regularization. For more details, we refer to Ref. [19]. We regulate the action of the microscopic theory as

$$S = \int d^d x \mu_0^{d-4} \mathcal{L}(x), \tag{2.16}$$

where $\mathcal{L}$ is the bare Lagrangian and d is the number of dimensions. With the factor μ_0^{d-4}, the bare action S is invariant under scale transformations if we promote the bare parameters μ_0 and m_0 to spurions. The scale transformation rules are

$$m_0 \to \lambda m_0, \tag{2.17a}$$

$$\mu_0 \to \lambda \mu_0, \tag{2.17b}$$

$$A_\mu(x) \to \lambda A_\mu(\lambda x), \tag{2.17c}$$

$$\psi(x) \to \lambda^{3/2} \psi(\lambda x), \tag{2.17d}$$

where A_μ is the bare gauge field and ψ the bare fermion field.

The function γ_m, defined by the RG equation

$$\frac{\mu}{m}\frac{dm}{d\mu} = -\gamma_m, \tag{2.18}$$

describes the response of the renormalized mass m to a change of the renormalization scale μ. In a mass-independent scheme, all renormalization factors depend on the scales μ and μ_0 only through their ratio, μ/μ_0. Hence,

$$\gamma_m = \gamma_m(g(\mu/\mu_0)), \tag{2.19}$$

where $g = g(\mu/\mu_0)$ is the running coupling. From now on, we will write $\gamma_m(\mu/\mu_0)$ for $\gamma_m(g(\mu/\mu_0))$, with a slight abuse of notation. We choose μ not to transform under scale transformations: the transformation (2.17) describes a rescaling of all the dimensionful bare quantities relative to a fixed renormalization scale.

Once γ_m is known we can express $m(\mu)$, the renormalized mass at an arbitrary renormalization scale μ, in terms of the bare mass, $m_0 = m(\mu_0)$, by integrating Eq. (2.18) between μ_0 and μ. Introducing the formal solutions

$$E^\pm(\mu/\mu_0) = e^{\pm \int_0^{\log \mu/\mu_0} dt \gamma_m(e^t)} \tag{2.20}$$

of the RG equations

$$\mu \frac{dE^\pm}{d\mu} = \pm \gamma_m(\mu/\mu_0) E^\pm, \tag{2.21}$$

one has

$$m(\mu) = E^-(\mu/\mu_0) m_0. \tag{2.22}$$

Using Eq. (2.17) for the dependence of the bare parameters m_0 and μ_0 on the scale transformation parameter λ, it follows that an infinitesimal scale transformation of the renormalized mass is governed by the differential equation [19]

$$\frac{\partial m(\lambda; \mu)}{\partial \log \lambda} = \left(1 + \gamma_m\left(\frac{\mu}{\lambda \mu_0}\right)\right) m(\lambda; \mu), \tag{2.23}$$

which is solved by

$$m(\lambda; \mu) = \lambda E^-(\mu/(\lambda \mu_0)) m_0. \tag{2.24}$$

For constant $\gamma_m = \gamma_*$, Eq. (2.20) simplifies to

$$E^\pm(\mu/\mu_0) = \left(\frac{\mu}{\mu_0}\right)^{\pm \gamma_*}, \tag{2.25}$$

and hence

---

[10]The β and γ functions in a mass-dependent scheme can be expanded in powers of m_0/M, and the first term in this expansion yields a mass-independent scheme that is a good approximation if $m_0/M \ll 1$.

$$m(\lambda;\mu) = \lambda^{1+\gamma_*}m(\mu) = \lambda^{1+\gamma_*}\left(\frac{\mu_0}{\mu}\right)^{\gamma_*}m_0. \qquad (2.26)$$

The second equation explains the origin of the factor $\lambda^{1+\gamma_*}$. A factor λ comes from the transformation of m_0, Eq. (2.17), while the remaining factor λ^{γ_*} comes from the transformation of μ_0, Eq. (2.17b). With the transformation rules of the effective fields

$$\tau(x) \to \tau(\lambda x) + \log\lambda, \qquad (2.27a)$$

$$\Sigma(x) \to \Sigma(\lambda x), \qquad (2.27b)$$

it follows that $\mathcal{L}_m(x)$ in Eq. (2.2b) transforms into $\lambda^4\mathcal{L}_m(\lambda x)$, as required for the invariance of the action.

In order to accommodate a nonconstant γ_m, we replace $\mathcal{L}_m$ of Eq. (2.2b) by

$$\mathcal{L}_m = -\frac{1}{2}f_\pi^2 e^{3\tau}E^-(e^\tau f_\pi/\mu_0)B_\pi(\mu/\mu_0)m(\mu/\mu_0)\mathrm{tr}(\Sigma+\Sigma^\dagger). \qquad (2.28)$$

Let us derive the transformation properties of this Lagrangian. The combination $B_\pi(\mu/\mu_0)m(\mu/\mu_0)$ is by assumption RG invariant, and we can write $B_\pi(\mu/\mu_0)$ as

$$B_\pi(\mu/\mu_0) = B_\pi^{\mathrm{RG}}E^+(\mu/\mu_0). \qquad (2.29)$$

The new LEC, B_π^{RG}, is both RG invariant and scale invariant, also by assumption. Hence $B_\pi(\mu/\mu_0)m(\mu/\mu_0) = B_\pi^{\mathrm{RG}}m_0$, and using Eq. (2.17a) it follows that under a scale transformation

$$\frac{\partial}{\partial\log\lambda}B_\pi(\mu/(\lambda\mu_0))m(\lambda;\mu) = +B_\pi(\mu/(\lambda\mu_0))m(\lambda;\mu). \quad (2.30)$$

The factor $E^-(e^\tau f_\pi/\mu_0)$ in Eq. (2.28) is invariant under a scale transformation by construction, because the combination $e^\tau f_\pi/\mu_0$ is.[11] Noting that the scaling dimension of Σ is zero, and taking the contribution from the factor $e^{3\tau}$ into account, we obtain

$$\frac{\partial}{\partial\log\lambda}\mathcal{L}_m|_{\lambda=1} = 4\mathcal{L}_m + x_\mu\frac{\partial}{\partial x_\mu}\mathcal{L}_m = \frac{\partial}{\partial x_\mu}(x_\mu\mathcal{L}_m), \quad (2.31)$$

which establishes the invariance of the action. This conclusion is valid for any choice of the function γ_m.

The Lagrangian for dChPT with a varying γ_m function is given by Eq. (2.1), now with $\mathcal{L}_m$ given by Eq. (2.28). The theory is invariant under the scale transformation of the

effective fields, Eq. (2.27), combined with the spurion transformation rules[12]

$$m(\mu) \to m(\lambda;\mu), \qquad (2.32a)$$

$$\mu_0 \to \lambda\mu_0, \qquad (2.32b)$$

$$c_0 \to c_0 - \log\lambda, \qquad (2.32c)$$

$$c_1 \to c_1. \qquad (2.32d)$$

The transformation rule (2.32c) is needed to ensure the invariance of (the spacetime integral of) $\mathcal{L}_d$ in Eq. (2.2a).[13] As usual, once the spurions m, μ_0, and c_0 are set equal to their fixed values, this breaks the scale symmetry explicitly.

We may again shift the τ field, as we did in Sec. II A, such that after the shift it has a vanishing expectation value for $m = 0$. The LECs $f_{\pi,\tau}$ and B_τ are redefined as in Eq. (2.4), but now $\hat{B}_\pi$ is defined as

$$\hat{B}_\pi(\mu/\mu_0) = \hat{B}_\pi^{\mathrm{RG}}E^+(\mu/\mu_0), \qquad \hat{B}_\pi^{\mathrm{RG}} = e^{v_0}B_\pi^{\mathrm{RG}}, \qquad (2.33)$$

so that $\hat{B}_\pi = e^{v_0}B_\pi$. The Lagrangian after the shift is again given by Eq. (2.5), but now with

$$\mathcal{L}_m = -\frac{1}{2}\hat{f}_\pi^2 e^{3\tau}E^-(e^\tau\hat{f}_\pi/\mu_0)\hat{B}_\pi(\mu/\mu_0)m(\mu/\mu_0)\mathrm{tr}(\Sigma+\Sigma^\dagger), \qquad (2.34)$$

instead of Eq. (2.6c). Note that, instead of being a function of $e^\tau f_\pi/\mu_0$, now E^- is a function of $e^\tau\hat{f}_\pi/\mu_0$.

Let us now reconsider the trace anomaly. We first apply the scale transformation only to the effective fields, setting the spurions equal to their fixed values. In this case,[14]

$$\frac{\partial}{\partial\log\lambda} = \frac{\partial\tau}{\partial\log\lambda}\frac{\partial}{\partial\tau} = \frac{\partial}{\partial\tau}, \qquad (2.35)$$

and we obtain the contribution of $\mathcal{L}_m$ to $\partial_\mu S_\mu$, the divergence of the dilatation current S_μ (see Appendix D of Ref. [17]),

$$\left(\frac{\partial}{\partial\tau} - 4\right)\mathcal{L}_m = -(1+\gamma_m(e^\tau\hat{f}_\pi/\mu_0))\mathcal{L}_m$$

$$= -(1+\gamma_m(e^\tau\hat{f}_\pi/\mu_0))m\bar{\psi}\psi(\mathrm{EFT}). \quad (2.36)$$

In the last step we identified $\mathcal{L}_m$ with the EFT representation of $m\bar{\psi}\psi$ in the underlying theory. This reproduces, in the EFT, the contributions from the fermions to the trace anomaly [38]. Recall that we have defined γ_m to be a

---

[11]Being μ independent, $E^-(e^\tau f_\pi/\mu_0)$ is trivially RG invariant.

[12]In Ref. [17] we introduced a spacetime dependent spurion field $\chi(x)$ for the renormalized mass, but for our present purposes, a spacetime independent spurion for m is sufficient.

[13]The transformation rules of c_0 and c_1 get modified at higher orders. For a detailed discussion of $\mathcal{L}_d$, see Refs. [17,20].

[14]We omit the contribution from the scale dependence of the spacetime coordinates [compare Eq. (2.31)].

function of μ/μ_0; cf. Eq. (2.19). Replacing τ by $v(m)$, its vacuum expectation value at nonvanishing m, we see that Eq. (2.36) effectively identifies the renormalization scale μ with $F_\pi = e^{v(m)}\hat{f}_\pi$; cf. Eq. (2.8b). This reveals a key feature of our construction of γ-dChPT: γ_m is evaluated at a renormalization scale equal to the physical scale F_π, which, in turn, is a function of the input fermion mass. We comment that we chose the hadronic scale inside E^- in Eq. (2.28) to be f_π, but, to achieve the desired scaling behavior, we could equivalently choose f_τ, or, more generally, any other hadronic scale m_h that enters the dChPT Lagrangian (or generalization thereof) via the combination $e^\tau m_h$, such as the nucleon mass in the chiral limit.

We now specialize to specific choices for the function γ_m. First, for constant $\gamma_m = \gamma_*$,

$$\hat{B}_\pi(\mu/\mu_0)E^-(e^\tau\hat{f}_\pi/\mu_0) = \hat{B}_\pi^{\rm RG}\left(\frac{\mu}{\mu_0}\right)^{\gamma_*}e^{-\gamma_*\tau}\left(\frac{\mu_0}{\hat{f}_\pi}\right)^{\gamma_*}$$
$$= \hat{B}_\pi(\mu/\hat{f}_\pi)e^{-\gamma_*\tau}, \qquad (2.37)$$

and the Lagrangian $\mathcal{L}_m$ in Eq. (2.34) reduces to Eq. (2.6c).[15] This also implies $\hat{B}_\pi(\mu/\hat{f}_\pi) = e^{(1-\gamma_*)v_0}B_\pi(\mu/f_\pi)$, consistent with Eq. (2.4c).

We next introduce a new choice for γ_m that we will be using for the actual fits to the KMI data. With $t = \tau + \log(\hat{f}_\pi/\mu_0)$ we define

$$E^-(e^\tau\hat{f}_\pi/\mu_0) = E^-(e^t) = e^{-\tilde{F}(t)}, \qquad (2.38)$$

where

$$\tilde{F}(t) = \tilde{\gamma}_0 t - \frac{1}{2}\tilde{b}t^2 + \frac{1}{3}\tilde{c}t^3, \qquad (2.39)$$

a cubic polynomial in t. The variable t is invariant under scale transformations, and, consistent with our general discussion, $\tilde{\gamma}_0$, $\tilde{b}$, and $\tilde{c}$ are LECs that do not depend on μ or μ_0. Reexpressing t in terms of τ, we write

$$\tilde{F}(t) = \tilde{F}(\log(\hat{f}_\pi/\mu_0)) + F(\tau), \qquad (2.40)$$

$$F(\tau) = \gamma_0\tau - \frac{1}{2}b\tau^2 + \frac{1}{3}c\tau^3, \qquad (2.41)$$

which defines the coefficients of the cubic polynomial $F(\tau)$ in terms of those of $\tilde{F}(t)$, and $\log(\hat{f}_\pi/\mu_0)$. Substituting into Eq. (2.34), and absorbing $e^{-\tilde{F}(\log(\hat{f}_\pi/\mu_0))}$ into $\hat{B}_\pi$, the final form of the Lagrangian becomes

---

$$\mathcal{L}_m = -\frac{1}{2}\hat{f}_\pi^2\hat{B}_\pi me^{3\tau-F(\tau)}{\rm tr}(\Sigma + \Sigma^\dagger). \qquad (2.42)$$

We will use the acronym γ-dChPT for the Lagrangian defined by Eq. (2.1), with $\mathcal{L}_d$ given by Eq. (2.2a) and $\mathcal{L}_m$ by Eq. (2.42) for some general function $F(\tau)$. Of course, for the case of a linear $F(\tau)$, Eq. (2.42) reduces to Eq. (2.2b), and the Lagrangian is just LO dChPT.

As an EFT, dChPT is based on the power counting established in Refs. [17,20] and reviewed above. As in ordinary ChPT, loop corrections in dChPT can be included systematically; the power counting (2.3) dictates which terms occur at the NLO [17], at the next-to-next-to-leading order (NNLO), and so on. The same is true in the large-mass regime, where the power counting is controlled by Eq. (2.14). This raises the question of how much γ-dChPT deviates from the strict EFT framework of dChPT itself. If we rely on algebraic structure and symmetries only, this allows $E^-(e^\tau\hat{f}_\pi/\mu_0)$ in Eq. (2.34), or, equivalently, $F(\tau)$ in Eq. (2.42), to depend on an infinite number of parameters, reflecting the model nature of γ-dChPT. But if, on the other hand, we assume that $F(\tau)$ takes the form of Eq. (2.41), with

$$\gamma_0 \sim (n_f - n_f^*)^0 = 1, \quad b \sim n_f - n_f^*, \quad c \sim (n_f - n_f^*)^2, \qquad (2.43)$$

then the factor $e^{-F(\tau)}$ may be obtained via partial resummation of terms from all orders in the expansion in powers of $n_f - n_f^*$. It thus reflects a fairly modest departure from dChPT, in that we will be taking into account some higher-order analytic terms, resummed into $e^{-F(\tau)}$, while omitting other higher-order terms. In addition, we will not calculate any nonanalytic higher-order corrections when fitting γ-dChPT to data. We will reexamine the scenario of Eq. (2.43) after presenting our fits to the KMI data in Sec. III.

D. Hadronic quantities for varying γ_m

As in Sec. II A, we begin with the saddle-point equation. For $m \geq 0$ the potential is minimized by setting $\Sigma = 1$ in Eq. (2.34), and $v = v(m)$ is the solution of [compare Eq. (2.7)]

$$\frac{(3-\gamma_m)m}{4c_1\mathcal{M}} = ve^{v+F(v)}, \qquad (2.44)$$

where now

$$\gamma_m = F'(v). \qquad (2.45)$$

When $F(\tau)$ is linear in τ we reproduce the results of Sec. II A, whereas for $F(\tau)$ in Eq. (2.41) we have

$$\gamma_m = \gamma_0 - bv + cv^2. \tag{2.46}$$

Equation (2.44) can be rewritten as

$$m = \frac{d_2}{\tilde{d}_1} \frac{1}{3 - \gamma_m} v e^{v + F(v)}, \tag{2.47}$$

with

$$\tilde{d}_1 = \frac{\hat{f}_\pi^2}{8 \hat{B}_\pi c_1 \mathcal{M}}. \tag{2.48}$$

For a general function F, Eq. (2.47) cannot be explicitly inverted analytically. We will, in effect, solve it numerically for m as a function of v, as described in Sec. III. In terms of v, F_π is still given by Eq. (2.8b). The pion mass is now

$$M_\pi^2 = 2 \hat{B}_\pi m e^{v - F(v)}, \tag{2.49}$$

so that, using Eq. (2.47), the ratio M_π^2 / F_π^2 is given by

$$\frac{M_\pi^2}{F_\pi^2} = \frac{1}{\tilde{d}_1} \frac{v}{3 - \gamma_m}. \tag{2.50}$$

The three equations (2.47), (2.8b), and (2.50) contain six parameters, $\tilde{d}_1$, d_2, $\hat{f}_\pi$ and the three parameters inside F: γ_0, b, and c.

We will not fit M_τ to the KMI data, as the errors found in Ref. [4] are too large for such a fit to have statistical relevance. We will, however, fit the staggered taste splittings obtained in Ref. [4]. With M_{Γ_i} the masses of the taste-split pions corresponding to the tastes

$$\Gamma_i \in \{\Gamma_5, \Gamma_{\mu 5}, \Gamma_{\mu\nu}, \Gamma_\mu, \Gamma_I\}, \tag{2.51}$$

we will fit the differences[16]

$$\Delta(\Gamma_i) \equiv a^2 (M_{\Gamma_i}^2 - M_\pi^2), \tag{2.52}$$

according to [39,40]

$$\Delta(\Gamma_5) \equiv \Delta_P = 0, \tag{2.53a}$$

$$\Delta(\Gamma_{\mu 5}) \equiv \Delta_A = C_1 E(\gamma_1) + 3 C_3 E(\gamma_3) + C_4 E(\gamma_4) + 3 C_6 E(\gamma_6), \tag{2.53b}$$

$$\Delta(\Gamma_{\mu\nu}) \equiv \Delta_T = 2 C_3 E(\gamma_3) + 2 C_4 E(\gamma_4) + 4 C_6 E(\gamma_6), \tag{2.53c}$$

$$\Delta(\Gamma_\mu) \equiv \Delta_V = C_1 E(\gamma_1) + C_3 E(\gamma_3) + 3 C_4 E(\gamma_4) + 3 C_6 E(\gamma_6), \tag{2.53d}$$

$$\Delta(\Gamma_I) \equiv \Delta_S = 4 C_3 E(\gamma_3) + 4 C_4 E(\gamma_4). \tag{2.53e}$$

Here $C_{1,3,4,6}$ are LECs associated with the taste-breaking potential [40], and

$$E(\gamma_i) = e^{(4 - \gamma_i)v}. \tag{2.54}$$

Equation (2.54) assumes that γ_i, the anomalous dimensions of the taste-breaking four-fermion operators, are constant (see Ref. [33] for more details). A global fit of the data including all the taste splittings has eight new parameters, coming from Eq. (2.53), in addition to the six parameters of the basic fit. This is a large number of parameters, and, as we will see, some of them are not sufficiently constrained by the available data. Thus, we will not venture into an exploration of any scale dependence of the γ_i.

We end this section with a comment. While in LO dChPT the potential is bounded from below, in γ-dChPT with general $F(v)$ the potential can be unbounded from below.[17] Mathematically, this appears to be a problem, but we contend that it is physically irrelevant. Within the EFT framework, the potential can only be known for $\mathcal{O}(1)$ values of the fields. While the pion field is always $\mathcal{O}(1)$ because it is a compact field, this is not the case for τ. We thus need to restrict the EFT to $\mathcal{O}(1)$ values of τ "by hand." In practice, this means that after fits to the data, we need to check that indeed values of v predicted by the fits are $\mathcal{O}(1)$, and do not land in the large-field region. In all our fits with a varying γ_m indeed unphysical regions of the potential occur at very large values of v, but they are separated from the physical region by an exponentially large potential barrier. Consistently, our fits never explore the unphysical region of the potential.

III. FITS TO THE LATKMI DATA

In this section, we will present our fits to data reported in Ref. [4], obtained by the LatKMI Collaboration for the eight-flavor SU(3) gauge theory. We begin in Sec. III A with a discussion of these data and the policies we will follow when we use them. In Sec. III B, we present "window" fits. These are fits of M_π^2 / F_π^2 and aF_π to the predictions of LO dChPT, for successive quintets of fermion masses, from the five lightest masses to the five heaviest ones. Altogether, ten different fermion masses were simulated in Ref. [4], making six (overlapping) windows. The window fits test the constancy of the LO dChPT parameters. We find a systematic trend of change

---

[16]We note that $M_{\Gamma_5} = M_\pi$ is the mass of the Nambu-Goldstone pion.

[17]For polynomial $F(v)$, a necessary and sufficient condition that the potential will be bounded from below is that the highest power of v is even, and its coefficient is positive.

for all fit parameters, by much more than their errors allow, proving that the full KMI mass range cannot be fit to LO dChPT. Then, in Sec. III C we fit the data at all ten fermion masses simultaneously to γ-dChPT, the extension of LO dChPT with a varying γ_m constructed in Sec. II C, with the special choice of γ_m in Eq. (2.46). We find that this extension of dChPT successfully describes the KMI data-set. Data for taste-split pion masses is available for a more limited set of fermion masses, and we present our fits including the taste splittings in Sec. III D. We end with a discussion of the scale dependence of γ_m found in our fits in Sec. III E.

The simulations of Ref. [4] were all performed at the same bare coupling. Invoking a mass-independent scale setting prescription, this implies that all ensembles have a common lattice spacing a. We elaborate on the choice of a scale setting prescription in Appendix A.

We will be using lattice units in all our fits. This means taking $\mu = \mu_0 = 1/a$, and thus $m(\mu) = m(\mu_0) = m_0$.

A. The LatKMI data

The pion mass M_π and decay constant F_π were measured in Ref. [4] at ten bare-mass values

$$am_0 \in \{0.012, 0.015, 0.02, 0.03, 0.04, 0.05,$$

$$0.06, 0.07, 0.08, 0.1\}. \tag{3.1}$$

In Ref. [4] a great effort was made to also determine the dilaton mass M_τ. It was found that indeed a dilaton exists, roughly degenerate with the pions. M_τ was measured for only six fermion masses, leaving out $am_0 = 0.05, 0.07,$ 0.08, and 0.1. More seriously, the statistical errors of M_τ turn out to be too large to have any real impact on our fits. In the window fits to LO dChPT (next subsection), we found that when we include a fit of M_τ^2/F_π^2 to Eq. (2.8d) in our global fit, d_3 remains largely undetermined, while all other fit parameters do not change. The only noticeable change is a higher p value, as might be expected. We thus omit the dilaton mass from the fits discussed in this paper.

Other hadron masses were also determined, notably the vector meson mass aM_ρ and the nucleon mass aM_N.[18] For these hadrons, the prediction from LO dChPT is that the ratios M_ρ/F_π and M_N/F_π should be independent of am_0 [20]; this is also true if we extend LO dChPT to include a varying γ_m. Excluding the two largest fermion masses, $am_0 = 0.08$ and 0.1, we found that we can fit M_ρ/F_π to a constant, with a p value of 0.31. M_N was measured only for a subset of the fermion masses,

$$am_0 \in \{0.012, 0.015, 0.02, 0.03, 0.04, 0.06, 0.08\}, \tag{3.2}$$

which leaves out $am_0 = 0.05, 0.07,$ and 0.1. Keeping only the five lightest masses, we found that a fit of M_N/F_π to a

constant has a p value of 0.07. This suggests that for larger fermion masses, higher-order corrections in dChPT (other than a varying γ_m) would be needed to fit these ratios. In addition, discretization effects could be playing a bigger role (see below). We will thus focus in this paper on the pion sector, considering M_π^2/F_π^2 and aF_π in Secs. III B and III C, and adding taste splittings in Sec. III D.

Information on the systematic errors of aM_π and aF_π is incomplete. Mostly, they were measured on at least two different volumes, and we estimate the finite-volume error by taking the difference between the results at the largest two volumes. For $am_0 = 0.012$ only one volume is available. In this case we took the finite-volume errors to be the same as for $am_0 = 0.015$. The latter was simulated on the same volume as $am_0 = 0.012$, as well as on a somewhat smaller volume. We note that, since $am_0 = 0.012$ is the lightest fermion mass, this procedure may underestimate its finite-volume errors. A single volume was reported also for $am_0 = 0.08$ and 0.1. For these fermion masses, the two largest ones, $M_\pi L$ is very large, and finite-volume corrections should be very small. We thus took the finite-volume errors for these two masses to vanish. We added the statistical error and the finite-volume error of aM_π and aF_π in quadrature. These errors were propagated to the ratio M_π^2/F_π^2, and correlations between this ratio and aF_π were kept.[19]

As the simulations of Ref. [4] were done at a single bare coupling, no direct information is available on the lattice spacing dependence, and it is not possible to take the continuum limit. We are thus forced to ignore scaling violations in our fits, but it should be kept in mind that these affect our results in an unknown way. Generally speaking, M_ρ and M_N are larger than M_π, and are thus prone to larger discretization effects. Also, as an example, for $am_0 = 0.08$ Ref. [4] finds the central values $aM_\pi = 0.51$, $aM_\rho = 0.68$, and $aM_N = 1.02$; hence, at the largest fermion masses discretization effects could be significant for the pions as well. We will briefly mention evidence for scaling violations in the determination of the gradient flow scale t_0 in Sec. III E. The only other information on lattice spacing effects comes from pion taste splittings. The masses of taste-split pions, which were measured only on the seven ensembles with bare masses (3.2), will be considered in Sec. III D.

B. Window fits

We begin with fitting M_π^2/F_π^2 and aF_π to the predictions of LO dChPT, Eqs. (2.8a) and (2.8b). We consider sets of five successive fermion masses, taking first the lightest five masses from the set (3.1), then the second to the sixth masses, etc., for a total of six quintets. The results are

---

[18]The pions are too heavy for the ρ to decay.

[19]Correlations between aM_π and aF_π on each ensemble are not available. We note that, in Ref. [33], we found that these correlations are small in the LSD data.

TABLE I.   Fits of the KMI data to Eqs. (2.8a) and (2.8b), using selections of five successive fermion masses from the set (3.1). All parameter errors reported in this paper are Hessian.

	A	B	C	D	E	F
Range	0.012–0.04	0.015–0.05	0.02–0.06	0.03–0.07	0.04–0.08	0.05–0.1
χ^2/dof	9.37/6	9.85/6	4.81/6	4.38/6	4.56/6	3.83/6
p value	0.15	0.13	0.57	0.63	0.60	0.70
γ_*	0.608(8)	0.589(10)	0.543(10)	0.534(12)	0.527(8)	0.498(13)
$a\hat{f}_\pi$	0.0050(7)	0.0067(6)	0.0089(8)	0.010(2)	0.011(1)	0.011(1)
$\tilde{d}_1$	0.0716(44)	0.0629(28)	0.0545(23)	0.0512(56)	0.0500(28)	0.0484(28)
$-\log(a d_2)$	10.5(3)	10.0(2)	9.5(2)	9.2(4)	9.1(2)	9.0(2)

shown in Table I.[20] All the fits are good. However, the parameter values change with the partial mass range, more than allowed by their errors. In particular, the lowest mass range (fit 1A) and the highest mass range (fit 1F) do not overlap; hence their parameter errors are statistically independent. These fits are thus not consistent with each other. A simultaneous fit of LO dChPT to all ten masses has a p value of order 10^{-11}. Clearly, the whole KMI mass range cannot be fit to LO dChPT.

As dChPT admits a systematic expansion, the failure to describe a set of data at LO means that higher orders in the expansion are needed. However, already at LO, dChPT contains more parameters than ordinary ChPT. Depending on the observables being fitted, many more would be needed for an NLO fit. We believe that much better data are required for a meaningful NLO fit. As discussed in Sec. III A, the LSD and KMI datasets both contain only a single lattice spacing, leaving discretization errors as an uncontrolled source of systematic uncertainty. In addition, it may well be that more refined data, for additional bare masses and/or with smaller statistical errors, would be needed to determine all the parameters in the NLO fit.

C. Fits with a varying γ_m

Being unable to carry out a full NLO fit at present, we are left with the option of partially extending LO dChPT by exploring different "directions" in "higher-order parameter space." By its very nature, no such extension is fully systematic, and each extension should thus be considered a model. Our assumption is that our model, γ-dChPT, captures the relevant physics better than other extensions of LO dChPT.

As we have discussed in Sec. II B, the physical mechanism that underlies the behavior of the LSD data is hyperscaling. The KMI mass range is higher than the LSD one, which motivates us to consider a minimal modification of this physical picture. We assume that the KMI mass range is still governed by the same principles that produce hyperscaling in the LSD mass range, except

that, because of the diminishing influence of the IRFP, we now have to allow the mass anomalous dimension to vary. That consideration has led us to the framework of γ-dChPT, developed in Sec. II C.

In this subsection, we will thus consider fits of the KMI data to γ-dChPT. Specifically, we consider fits of M_π^2/F_π^2 and aF_π to Eqs. (2.50) and (2.8b), where γ_m is quadratic in v; cf. Eq. (2.46). We begin with a technical issue. The independent variable in these equations is v, which, in turn, can be determined in terms of am_0 using Eq. (2.47). However, unlike in LO dChPT discussed in Sec. II A, Eq. (2.47) cannot be analytically inverted.[21] Instead, in addition to the parameters defining the γ-dChPT Lagrangian, we introduce new parameters v_i, one per ensemble.[22] We fit the corresponding bare mass $am_{0,i}$ to Eq. (2.47), while simultaneously also fitting $(M_\pi^2/F_\pi^2)_i$ and $(aF_\pi)_i$, all as functions of the same parameter v_i. Artificially introducing a tiny error for $am_{0,i}$, the fit in effect solves Eq. (2.47) numerically for v_i in terms of $am_{0,i}$. Thus, for given values of the γ-dChPT parameters, v_i is equal to $v(am_{0,i})$ with numerical precision set by the "error" of the "data" $am_{0,i}$. We have varied the errors on $am_{0,i}$ between 10^{-6} and 10^{-7}, finding no discernible differences in the results of our fits. χ^2 values remain equal to four decimal places, whether one includes the "am_0 part" in the computation of χ^2 or not.

As in Ref. [33], we can calculate $(a\hat{B}_\pi)_i$ on each ensemble using Eq. (2.49) and our fit result for v_i. In all cases studied in this paper the so-obtained values of $(a\hat{B}_\pi)_i$ are equal within error. This confirms the self-consistency of our assumption that the lattice spacing a is independent of the fermion mass.

The results of our fits are shown in Table II. Fit 2A includes all ten ensembles, fit 2B leaves out the $am_0 = 0.1$

---

[20]We will label fits with a number for the table, and a letter for the fit in the table. For example, fit 1A refers to fit A in Table 1.

[21]In principle, the formal inverse function $m = m(v)$ may not be single valued. In practice, we found that v is monotonically increasing with m over the entire KMI mass range.

[22]The total number of parameters increases by the number of v_i parameters, i.e., by the number of ensembles included in the fit. The number of data increases by the same amount (the $am_{0,i}$), leaving the number of degrees of freedom unchanged.

TABLE II.   Fits of M_π^2/F_π^2 and aF_π to γ-dChPT, the extension of LO dChPT discussed in Sec. II C. The "omitted" row shows bare mass values from the set (3.1) which are not included in the fit, if any.

	A	B	C	D
Omitted	$\cdots$	0.1	0.1, 0.08	0.1, 0.08
χ^2/dof	20.7/14	11.5/12	10.0/10	14.8/11
p value	0.11	0.48	0.44	0.19
$\hat{f}_\pi$	0.0104(4)	0.0102(5)	0.0101(9)	0.0085(5)
$\tilde{d}_1$	0.0506(10)	0.0512(12)	0.0516(23)	0.0559(17)
$-\log(a d_2)$	10.1(2)	10.4(2)	10.6(3)	9.9(1)
γ_0	1.69(23)	2.11(31)	2.29(57)	0.85(5)
b	0.97(22)	1.38(31)	1.57(67)	0.12(2)
c	0.20(5)	0.30(8)	0.35(16)	$\cdots$

ensemble, and fit 2C leaves out both $am_0 = 0.1$ and 0.08. All the fits are good, but fits 2B and 2C are better than fit 2A. We also carried out fits setting $c = 0$, i.e., taking γ_m in Eq. (2.46) to be a linear function of v. Fits with $c = 0$ including all ten ensembles, or omitting the $am_0 = 0.1$ ensemble, have very low p values, 0.001 and 0.01, respectively. We do not show them in the table. However, if we omit both the $am_0 = 0.1$ and 0.08 ensembles, we obtain fit 2D, which is a good fit. The parameters $a\hat{f}_\pi$, $\tilde{d}_1$, and $\log(ad_2)$ are relatively stable between the fits

with c as a free parameter, and fit 2D, where $c = 0$. By contrast, the parameters defining the function γ_m change substantially: Fit 2D yields much smaller values for both γ_0 and b than the other fits of Table II.

The results of fits 2B and 2D are shown in Fig. 1. The black points are data that were included in the fits, whereas the magenta points were excluded. The lower left panel shows that if we simplify our *ansatz* for γ_m to be linear in v, then the $am_0 = 0.08$ and 0.1 ensembles must be excluded.

We have proposed in Sec. II C that the exponential factor $e^{-F(v)}$ may originate from a resummation of the dominant contributions from all orders in the expansion in $n_f - n_f^*$. According to the hypothesis (2.43), b is an NLO parameter, while c is an NNLO parameter. One way to test this scenario is to examine the effect of truncating the Taylor expansion of the exponential factor. The range of values we find for v in the fits to the KMI data is $1.5 \leq v \leq 2.5$. Considering first fit 2B, we can compare the numerical values of $\exp\left(\frac{1}{2}bv^2 - \frac{1}{3}cv^3\right)$, and its version truncated at NNLO, namely $1 + \frac{1}{2}bv^2 + \frac{1}{8}b^2v^4 - \frac{1}{3}cv^3$. When we vary v from 1.5 to 2.5, the exponential and its truncated version take values ranging from 3.4 to 15, respectively, 3.4 to 13. The differences (taking the correlations into account) are $-0.05(8)$ and $2(4)$, respectively, so that the exponential and truncated forms are consistent with each other. The situation is somewhat different for fit IID, where the smallness of both b and its relative error allows for a more
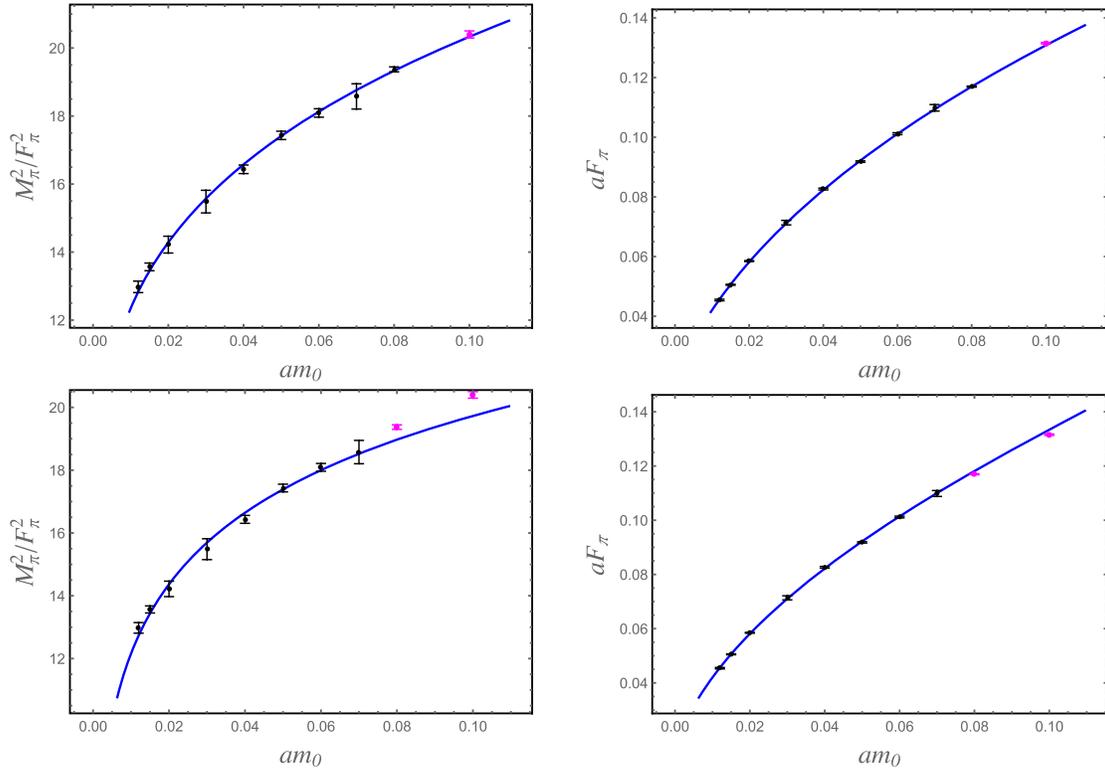

FIG. 1.   Upper panels: Fit results for M_π^2/F_π^2 (left panel) and aF_π (right panel) using fit 2B. Lower panels: similar, using fit 2D. Black points are fitted data, while magenta points were not included in the fit.

TABLE III. Fits of M_π^2/F_π^2, aF_π, and taste splittings to γ-dChPT. The "omitted" row shows bare mass values from the set (3.2) which are not included in the fit, if any. For description see text.

	A	B	C	D	E
Omitted	$\cdots$	$\cdots$	$\cdots$	$\cdots$	0.08
χ^2/dof	17.5/28	38.2/34	29.5/29	50.1/35	22.9/29
p value	0.94	0.28	0.44	0.05	0.78
$\hat{f}_\pi$	0.0102(5)	0.0105(4)	0.0095(4)	0.0099(4)	0.0085(6)
$\tilde{d}_1$	0.0512(12)	0.0503(10)	0.0526(12)	0.0514(10)	0.0562(20)
$-\log(ad_2)$	10.3(2)	10.2(2)	9.68(8)	9.65(7)	9.98(13)
γ_0	1.85(27)	1.81(26)	0.82(3)	0.86(2)	0.92(3)
b	1.12(28)	1.09(27)	0.121(17)	0.142(10)	0.154(17)
c	0.24(7)	0.23(7)	$\cdots$	$\cdots$	$\cdots$
$-\log C_1$	10(10)	$\cdots$	$-6(45)$	$\cdots$	$\cdots$
γ_1	4(6)	$\cdots$	14(29)	$\cdots$	$\cdots$
$-\log C_3$	11(4)	$\cdots$	8(8)	$\cdots$	$\cdots$
γ_3	3(2)	$\cdots$	5(5)	$\cdots$	$\cdots$
$-\log C_4$	12.3(3)	12.1(2)	12.6(3)	12.3(2)	12.5(3)
γ_4	1.29(10)	1.36(8)	1.26(10)	1.34(8)	1.42(11)
$-\log C_6$	18(6)	$\cdots$	18(4)	$\cdots$	$\cdots$
γ_6	0(2)	$\cdots$	0(2)	$\cdots$	$\cdots$

precise comparison. Varying again v from 1.5 to 2.5, $\exp\left(\frac{1}{2}bv^2\right)$ varies from 1.15 to 1.48, while the expansion to NLO, $1 + \frac{1}{2}bv^2$, varies from 1.14 to 1.39. The (correlated) differences are 0.010(4) and 0.09(3), respectively. Thus, while the behavior of both forms is qualitatively similar, the differences are statistically significant. Fits with the truncated version give results consistent with fits 2B and 2D, but with lower p values.

Without more data it is difficult to decide which fit in Table II is the preferred one. Clearly, unless the two heaviest masses are dropped, c must be kept in the fit. Given its (conjectural) role as an NNLO parameter, it is to be expected that eventually c will be needed to describe the data as the mass range is increased. Still, we cannot rule out that the main reason why fit 2D does not accommodate the two heaviest masses is large scaling violations at those mass values.

In all fits where the parameter c is present, it is always small compared to b, consistent with the conjectured hierarchy (2.43). However, in the same fits, one cannot say that b is small compared to γ_0. By contrast, in fit 2D, where $c = 0$, also b is clearly small compared to γ_0. The most appealing scenario thus appears to be the following. We exclude the two largest fermion mass values, because they require going to (at least) NNLO in the EFT expansion, and/or because they are afflicted by too large scaling violations. The remaining mass range may be amenable to an NLO dChPT fit,[23] for which fit 2D is our closest substitute.

D. Taste splittings

We now turn to fits which also include the taste splittings (2.52), i.e., fits of M_π^2/F_π^2, aF_π, and $\Delta_{A,T,V,S}$ to γ-dChPT, augmented by Eq. (2.53). Our fits are limited to the smaller ensemble set (3.2), where the taste-split pion masses were measured.

We show five different fits in Table III. Fit 3A includes all the parameters: the basic γ-dChPT parameters of Sec. III C, namely $a\hat{f}_\pi$, $\tilde{d}_1$, $\log(ad_2)$, γ_0, b, and c, as well as all eight taste-splitting parameters of Eq. (2.53). Data from all seven ensembles in the set (3.2) are included in the fit. The p value is very high. The results for the six basic γ-dChPT parameters are consistent with fit 2B.[24] As for the taste-splitting parameters, most of them, namely, $\gamma_{1,3,6}$ and $\log C_{1,3,6}$, are not well determined by the fit. We conclude that fit 3A gives an excellent description of the data, but the data are not precise enough to determine all parameters in the fit.

We next consider fits omitting poorly determined parameters. Among the taste-splitting parameters, only $\log C_4$ and γ_4 were determined with good precision. As for C_1, C_3, and C_6, if we take their errors seriously, using them as 1σ bounds, these parameters are "allowed" to be very small relative to C_4 (by factors $\sim 2 \times 10^3$, ~ 10, and $\sim 10^5$, respectively). Setting $C_1 = C_3 = C_6 = 0$, we obtain fit 3B. This is a good fit, even though its p value is much smaller than fit 3A, as one would expect. The results of fits 3A and 3B are in very good agreement. The dominance of the taste splittings generated by the $C_4 E(\gamma_4)$ term is consistent with the results we obtained for the LSD

---

[23]With the caveats discussed in Sec. III B.

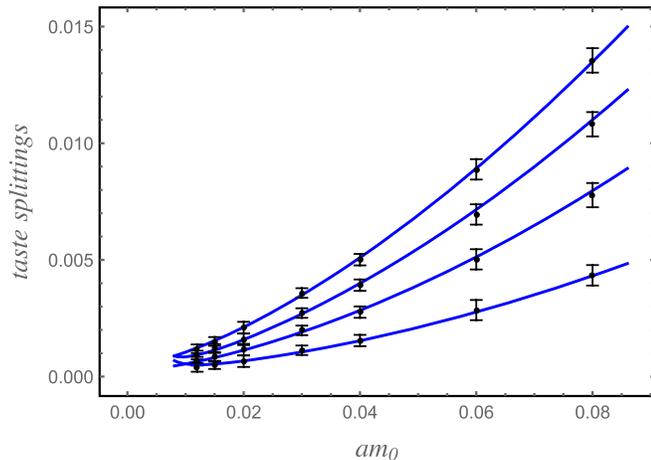[24]Note that the ensemble set (3.2) does not include $am_0 = 0.1$.

FIG. 2. Fit 3C of the taste splittings $\Delta_{A,T,V,S}$ of Eq. (2.53), as a function of am_0. From top to bottom: Δ_S, Δ_V, Δ_T, and Δ_A.

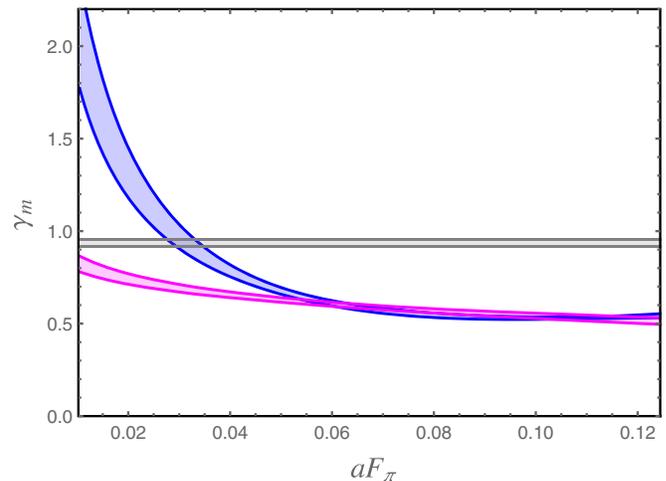

FIG. 3. The running mass anomalous dimension γ_m, obtained from fit 2B (blue band) and fit 2D (magenta band), plotted as a function of aF_π (see text). The gray horizontal band is $\gamma_* = 0.936 \pm 0.019$, from our fit to the LSD data [33]. The fitted KMI data have values of aF_π between 0.045 and 0.12.

data [33], as well as with the familiar taste splittings found in QCD.

In Sec. III C we saw that the parameter c can be omitted if the fermion masses $am_0 = 0.1$ and 0.08 are not included in the fit. While $am_0 = 0.08$ is present in the ensemble set (3.2), we also repeated fits 3A and 3B while setting $c = 0$, obtaining fits 3C and 3D, respectively. Finally, fit 3E is similar to fit 3D, except that the $am_0 = 0.08$ ensemble is not included. Fit IIIC, where we set $c = 0$ but keep all the taste-splitting parameters, is very good. Setting both $c = 0$ and $C_1 = C_3 = C_6 = 0$ leads to a relatively low p value in fit 3D. After dropping the $am_0 = 0.08$ ensemble, in fit 3E the p value is again very high.

Our results for $a\hat{f}_\pi$, $\tilde{d}_1$, and $\log(ad_2)$ are fairly consistent in all the fits reported in Tables II and III. The values of the parameters defining the function γ_m are consistent among the fits where $c \neq 0$: fits 2A, 2B, 2C, 3A, and 3B. In the fits with $c = 0$ the values of γ_0 and b are different, but again consistent across this group: fits 2D, 3C, 3D, and 3E. The values of the taste-splitting parameters $\log C_4$ and γ_4 are consistent in all the fits of Table III, while the (poorly determined) values of the remaining taste-splitting parameters are consistent between fits 3A and 3C.

In fit 3C, LO dChPT has been minimally extended (within the framework of γ-dChPT) to include an NLO correction to the function γ_m. This fit gives an excellent description of the ensemble set (3.2) with taste splittings included; the parameter c is not needed. We thus consider fit 3C to be the preferred fit from Table III. We plot the taste splittings of this fit in Fig. 2. A caveat is that, even though all the taste-split pion masses were measured in Ref. [4], the data are not precise enough to determine all taste-splitting parameters.[25] We recall that the QCD taste splittings are

essentially independent of the fermion mass [34,40].[26] By contrast, as for the LSD data [33], also in the KMI mass range the taste splittings vary with the fermion mass. This behavior can be successfully described in dChPT, where the scale dependence of the taste-breaking operators gives rise to mass dependent tree-level taste splittings, through the factors $E(\gamma_i)$ in Eq. (2.53).

E. Scale dependence of γ_m

The anomalous dimension function γ_m obtained from two of the fits of Table II is shown in Fig. 3. The blue band represents fit IIB, where $\gamma_m = F'(v)$ is quadratic in v [Eq. (2.46)], while the magenta band represents fit 2D, where γ_m is linear in v. With Eq. (2.8b), we take the argument of γ_m to be $v = \log(aF_\pi/a\hat{f}_\pi)$, and then we plot γ_m as a function of aF_π. The two γ_m functions agree well in most of the interval containing the fitted data, $0.045 \lesssim aF_\pi \lesssim 0.12$. The good agreement deteriorates toward the lower end of the interval, below which these functions diverge from each other. If we would overlay the (constant) results of each window fit from Sec. III B as a set of horizontal bands (each stretching over its corresponding range of aF_π), these bands would be consistent with the blue and magenta bands in that interval.[27]

Figure 3 also shows the value $\gamma_* = 0.936(19)$ obtained from our fits of the LSD data to LO dChPT [33], as a gray horizontal band. The LSD mass range is lower than the KMI range, and the (generalized) hyperscaling behavior we

---

[25]By contrast, the LSD data, which we fitted in Ref. [33], contains only $M_{\mu5}$ and $M_{\mu\nu}$ [3].

[26]Thanks to the dominance of C_4, the QCD taste splittings are also roughly equal to each other.

[27]We do not show window fits in Fig. 3 because the different bands become visually difficult to see.

have observed implies that the LSD range of F_π should also be lower than the corresponding KMI range, in physical units. Equivalently, the LSD values of aF_π, properly converted to KMI lattice units, should lie to the left of the KMI range of aF_π in Fig. 3.

Since the LSD data are successfully described by a constant $\gamma_m = \gamma_*$, we expect that also in the chiral limit γ_m will remain constant, at a value consistent with γ_*. The continuity of γ_m as a function of F_π thus requires that, as F_π is lowered from the KMI range into the LSD range, γ_m will rise to a value consistent with γ_*, and then stay roughly constant all the way to the chiral limit. It is intriguing that the strong dynamics of the $N_f = 8$ system might induce this behavior of γ_m.[28] Figure 3 shows that, when extrapolated below the KMI range, the quadratic γ_m of fit 2B overshoots γ_*, while the linear γ_m of fit 2D undershoots it. The desired behavior of γ_m over the combined KMI and LSD ranges cannot be described by simple *ansatzes* such as the ones we have used. One cannot rule out, however, that the combined LSD and KMI mass ranges could be described by including higher orders in dChPT systematically.

Clearly, an investigation of the combined LSD and KMI mass ranges would be extremely interesting. However, this is just not possible with the existing datasets. We already pointed out that the LSD and KMI datasets were each produced at a single lattice spacing. Moreover, the lattice actions used by LSD and by KMI differ in their details, and scaling violations can potentially differ significantly between the two lattice actions and axial currents. This means that the only way to reliably compare these results is by first taking the continuum limit separately for the LSD lattice action and for the KMI lattice action. The minimal requirement to make this possible is a second set of data at a different lattice spacing, for each lattice action.[29]

We have attempted a comparison of the LSD and KMI lattice scales, using $t_{0,\mathrm{ch}}$, the chiral-limit value of the gradient-flow scale t_0 [42], which we have determined for the LSD dataset in Ref. [33]. The comparison is deficient for several reasons. First, unlike in ordinary ChPT [43], dChPT does not predict the behavior of t_0 as a function of the fermion mass [33], so the best we can do is a phenomenological fit. Second, usually the gradient flow scale (or its chiral limit) is used to compare the lattice spacings of ensembles generated with different bare couplings, but with the same lattice action. By contrast, here we are comparing results obtained using two different lattice actions, hence the meaning of the comparison is less clear. Finally, there are also scaling violations in the lattice observables used to extract t_0, as well as in the

gradient-flow equation. KMI used two lattice definitions for t_0 which should agree in the continuum limit, but which consistently differ by some 15% over the entire KMI mass range; we do not have equivalent information about uncertainties associated with the LSD data. With all these caveats in mind, our findings suggest that the ratio $r = a(\mathrm{KMI})/a(\mathrm{LSD})$ is smaller than one. Using Eq. (3.1) together with Eq. (4.5) below, it follows that the KMI mass range is indeed higher than the LSD mass range, in agreement with the physical picture reflected in Fig. 5 of Ref. [2]. But, we are unable to turn this conclusion into a more quantitative statement.

We close this section with a comment. As discussed above, our experimentation with t_0 (and its chiral extrapolation) suggests that $r < 1$. Now, an alternative way to estimate r would be to take advantage of the fact that $\hat{f}_\pi$, the chiral-limit value of the pion decay constant, is a physical observable. Expecting $\sqrt{2}\hat{f}_\pi(\mathrm{LSD}) \approx \hat{f}_\pi(\mathrm{KMI})$ in physical units,[30] it follows that $a\hat{f}_\pi(\mathrm{KMI})/(\sqrt{2}a\hat{f}_\pi(\mathrm{LSD})) \approx r$. The reason why we only expect an approximate equality between $\sqrt{2}\hat{f}_\pi(\mathrm{LSD})$ and $\hat{f}_\pi(\mathrm{KMI})$ is the different scaling violations of the two lattice actions. In reality, using the value of $a\hat{f}_\pi(\mathrm{LSD})$ from Ref. [2], and taking $a\hat{f}_\pi(\mathrm{KMI}) \sim 0.01$, we find $a\hat{f}_\pi(\mathrm{KMI})/(\sqrt{2}a\hat{f}_\pi(\mathrm{LSD})) \sim 10$, in stark conflict with the estimate $r < 1$ obtained from the gradient flow scale. It is unlikely that scaling violations *per se* can account for this inconsistency. The problem must be related to the long extrapolation to the chiral limit inherent in the extraction of $a\hat{f}_\pi$. It does not necessarily imply that $(\gamma$-)dChPT cannot be trusted. The factor $e^{v(m)} = F_\pi/\hat{f}_\pi$ is very sensitive to m, which makes a long extrapolation to the chiral limit much more difficult than in the case of QCD. For at least one of the datasets our fit result for $a\hat{f}_\pi$ is likely to contain a large, and unaccounted for, source of systematic error. A comparison of the values of ad_2 obtained from the two datasets reveals a similar and, in fact, more severe problem, which presumably has a similar source, given that $d_2 = \hat{f}_\pi^2/(2\hat{B}_\pi)$. We comment that in order to compare $a\hat{B}_\pi$ between the LSD and KMI lattice scales we have to apply an RG transformation, but once again, it is hard to see how such a transformation would suffice to match the values of d_2 found in the two simulations.

IV. THE Δ CLASS OF DILATON POTENTIALS

So far, we have considered a model modification of the LO dChPT form of $\mathcal{L}_m$, based on the observation that the coupling of the underlying theory may start running at the physical scale determined by a growing fermion mass, thereby inducing a varying mass anomalous dimension as well. In this section we turn to a class of modifications to the dilaton-potential term $\mathcal{L}_d$. Alternate forms of the dilaton

---

[28]A γ_m function that saturates to a constant value at strong coupling was observed in the SU(2) theory with two adjoint Dirac fermions [41].

[29]To make sure that the same physical mass range is covered, one can, for example, monitor the values of some observable, such as a hadron mass or a decay constant, in units of $\sqrt{t_0}$.

[30]The factor of $\sqrt{2}$ is due to different normalization conventions.

potential were first applied to the LSD data in Ref. [28]. In Ref. [32] a class of dilaton potentials $\mathcal{L}_\Delta$ was proposed, defined by [compare Eq. (2.6)]

$$\mathcal{L}_\Delta(\tau) = \hat{f}_\tau^2 \hat{B}_\tau e^{4\tau} V_\Delta(\tau), \tag{4.1a}$$

$$V_\Delta(\tau) = \frac{c_1}{4-\Delta}\left(1 - \frac{4}{\Delta}e^{(\Delta-4)\tau}\right), \tag{4.1b}$$

where Δ is a new free parameter.[31] We have translated the notation of Ref. [32] to our notation. In the limit $\Delta \to 4$, the potential $\mathcal{L}_d$ of Eq. (2.6a) is recovered. For $\Delta = 2$, $\mathcal{L}_\Delta$ becomes the linear σ-model potential considered in Ref. [8]. We will refer to the low-energy Lagrangian with $\mathcal{L}_d$ replaced by $\mathcal{L}_\Delta$ as Δ-dChPT.

Applying Δ-dChPT to the LSD data, Ref. [32] concluded that these data appear to favor a value of Δ around 3.5, with a large uncertainty. Correlations in these data were not taken into account [32]. Moreover, correlations which occur because of the appearance of F_π in all three equations fitted in Ref. [32], as well as the appearance of M_π in two of them, apparently were not taken into account either. In Sec. IV A we begin by collecting the expressions needed to fit Δ-dChPT. In Sec. IV B we revisit the determination of Δ using the LSD data, taking all correlations into account. This analysis departs from the framework of LO dChPT (Sec. II A) only by replacing the dilaton potential $\mathcal{L}_d$ by $\mathcal{L}_\Delta$. At this stage the mass anomalous dimension is held fixed; cf. Eq. (2.2b). Then, in Sec. IV C, we explore fits of the KMI data to the Δ class of potentials. As in the previous section, we consider both fixed-γ_m fits to subsets of the KMI data, as well as fits with a varying γ_m to the entire KMI dataset. We summarize our findings in Sec. IV D.

Unlike the modification of $\mathcal{L}_m$ to accommodate a running γ_m, we are not aware of a concrete physical motivation to replace $\mathcal{L}_d$ by the more general form $\mathcal{L}_\Delta$. A closely related question is whether Δ-dChPT is the leading order in a systematic low-energy expansion for an arbitrary value of Δ.

The potential $\mathcal{L}_d$, Eq. (2.6), is based on the systematic power counting developed in Ref. [17]. Since $\mathcal{L}_d$ corresponds to the limit $\Delta \to 4$ in Eq. (4.1), it follows by continuity that there must exist a neighborhood of $\Delta = 4$ where the dChPT systematic expansion is still applicable. For arbitrary Δ, a power counting was proposed in Ref. [32]. We prove in Appendix B that the arguments given in Ref. [32] are not correct. Δ-dChPT, i.e., the low-energy Lagrangian consisting of Eq. (2.1) with $\mathcal{L}_d$ replaced by $\mathcal{L}_\Delta$, should thus be considered to be a model.

---

[31]$\mathcal{L}_\Delta(\tau)$ is bounded from below for any $-\infty < \Delta < \infty$.

A. Fitting data to $\mathcal{L}_\Delta$

For the case of a constant $\gamma_m = \gamma_*$, combining Eq. (4.1) with $\mathcal{L}_m$ of Eq. (2.2b), one finds the saddle-point equation relating v to m,

$$m = \frac{d_2}{d_1}\frac{1 - e^{(\Delta-4)v}}{4-\Delta}e^{(1+\gamma_*)v}. \tag{4.2}$$

It is then straightforward to derive the relations

$$\frac{M_\pi^2}{F_\pi^2} = \frac{1}{d_1}\frac{1 - e^{(\Delta-4)v}}{4-\Delta} \equiv h_\Delta(m), \tag{4.3a}$$

$$F_\pi = \hat{f}_\pi e^v \tag{4.3b}$$

$$= \left(\frac{d_0 m}{h_\Delta(m)}\right)^{\frac{1}{1+\gamma_*}}, \tag{4.3c}$$

$$\frac{M_\tau^2}{F_\pi^2} = d_3(1 + (\Delta + \gamma_* - 3)d_1 h_\Delta(m)), \tag{4.3d}$$

where we used the definitions (2.10).

In the case of a varying γ_m, Eq. (4.3b) is still applicable, while combining Eq. (4.1) with $\mathcal{L}_m$ of Eq. (2.42), Eqs. (4.2) and (4.3a) generalize to

$$m = \frac{d_2}{\tilde{d}_1}\frac{1 - e^{(\Delta-4)v}}{4-\Delta}\frac{e^{v+F(v)}}{3 - \gamma_m}, \tag{4.4a}$$

$$\frac{M_\pi^2}{F_\pi^2} = \frac{1}{\tilde{d}_1(3 - \gamma_m)}\frac{1 - e^{(\Delta-4)v}}{4-\Delta}, \tag{4.4b}$$

where γ_m is given in Eq. (2.45) and $\tilde{d}_1$ is defined in Eq. (2.48).

We now turn to fits of the LSD and KMI data, in order to explore to what extent they constrain the value of Δ. We emphasize again that this investigation is empirical, as no systematic power counting is available for this model for arbitrary values of Δ.

B. The LSD data

Data reported in Ref. [3] includes results at five different fermion masses,

$$am_i \in \{0.00125, 0.00222, 0.005, 0.0075, 0.00889\}. \tag{4.5}$$

All ensembles have the same bare coupling, and, in a mass-independent scheme, the same lattice spacing [33]. We fitted the LSD data to LO dChPT in Ref. [33]. Here, we repeat some of those fits replacing $\mathcal{L}_d$ by $\mathcal{L}_\Delta$, keeping Δ as a free parameter. Our results are shown in Table IV. These fits correspond to four fits presented in Ref. [33]: Fits 4A and 4B are to be compared to the fits shown in Table 1 of Ref. [33], while fits 4C and 4D are to be compared with the

TABLE IV. Fits of the LSD data to Δ-dChPT. Fits C and D include taste breaking, while fits A and B do not. The "omitted" row shows bare mass values from the set (4.5) which are not included in the fit, if any.

	A	B	C	D
Omitted	$\cdots$	0.00889	$\cdots$	0.00889
$\chi^2/$dof	8.72/9	2.50/6	15.18/13	5.52/8
p value	0.56	0.87	0.30	0.70
Δ	2.8(7)	3.5(7)	2.7(6)	3.5(7)
γ_*	0.935(19)	0.936(19)	0.933(19)	0.937(19)
$\log d_0$	1.94(6)	1.93(6)	1.94(6)	1.93(6)
d_1	0.042(20)	0.083(84)	0.037(15)	0.073(66)
$-\log(ad_2)$	11.6(9)	12.9(2.5)	11.3(7)	12.6(2.1)
d_3	17(9)	9(9)	20(8)	10(9)
$-\log C_1$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
γ_1	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$-\log C_3$	$\cdots$	$\cdots$	9.7(6)	10(2)
γ_3	$\cdots$	$\cdots$	2.0(1)	2.4(7)
$-\log C_4$	$\cdots$	$\cdots$	8.3(7)	10(2)
γ_4	$\cdots$	$\cdots$	1.96(6)	2.1(4)
$-\log C_6$	$\cdots$	$\cdots$	36(7)	17(11)
γ_6	$\cdots$	$\cdots$	$-11(4)$	0(3)

third column of Table 3 and the second column of Table 4 in Ref. [33].

As discussed in great detail in Ref. [33], it is not possible to fit all parameters in the taste-breaking sector with the available LSD data. Here we kept those taste-breaking parameters that gave rise to the best fits of Ref. [33]. Furthermore, in Ref. [33] we argued that four-ensemble fits, which exclude the ensemble with the largest fermion mass, are better behaved. While the five-ensemble fits reported in Table IV already have good p values, again we find that p values for the four-ensemble fits are significantly better.

Parameter values for γ_* and $\log d_0$ are in good agreement with the corresponding fits in Ref. [33]. The parameters d_1 and d_3 are very poorly determined by the fits; especially by those with four ensembles. This is no surprise, as d_1 and d_3 relate directly to the dilaton potential $\mathcal{L}_\Delta$, in which now a new parameter, Δ, has been introduced. The results for the taste-breaking parameters are in reasonable agreement with

Ref. [33] for the five-ensemble fit and in good agreement for the four-ensemble fit. By holding Δ fixed in the fit, we verified that in the limit $\Delta \to 4$ the results of Ref. [33] are reproduced.

The parameter Δ itself is reasonably well determined by each fit. However, there is a visible difference between the four-ensemble and five-ensemble fits. From the four-ensemble fits, we conclude that $\Delta = 3.5(7)$. This is consistent with the hypothesis that dChPT, which predicts $\Delta \to 4$, is the correct low-energy EFT. The linear σ-model value, $\Delta = 2$, is disfavored. By contrast, the values found in the five-ensemble fits average to 2.8(7). This is 1.7σ away from $\Delta \to 4$, and, in fact, between the two options, it slightly favors the linear σ-model value.

C. The KMI data

We next turn to fits of the KMI data, with $\mathcal{L}_\Delta$ replacing $\mathcal{L}_d$. We first consider again window fits similar to those of Table I, but now with Δ as an additional free parameter. The results are reported in Table V. The fits are reasonably consistent with $\Delta = 4$, while the other parameters are generally consistent between Tables V and I. As before, a constant γ_m is not sufficient to describe the KMI data over the full mass range. However, while γ_* varies with the mass range selected in the fit, Δ does not. If we compare the values of Δ between two of the fits in Table V, these values are always consistent within the smaller of the two errors (with the exception of the second fit, for which Δ has an anomalously small error). The first and last values, 3.8(5) and 4.0(6), coming from the lowest and highest mass ranges, are statistically independent, in agreement with $\Delta = 4$ and with each other.

As in Sec. III, our next step is to consider fits to all, or most, of the KMI data, with $\mathcal{L}_m$ of Eq. (2.42), and a varying γ_m as defined in Eq. (2.46). As before, this introduces two more parameters (b and c) into the fits, for a total of seven parameters. We will refer to this flavor of the low-energy Lagrangian as $\gamma\Delta$-dChPT.

In Table VI we show a scan in Δ: at each chosen value of Δ, we fit the other six parameters. The fit for $\Delta = 3.9999$ coincides with fit 2A, as one would expect. If we decrease Δ, we find that the p value rapidly decreases, dipping

TABLE V. Fits of the KMI data to Δ-dChPT (with a constant $\gamma_m = \gamma_*$), with selections of five successive fermion masses in Eq. (3.1), shown in the top row.

Range	0.012–0.04	0.015–0.05	0.02–0.06	0.03–0.07	0.04–0.08	0.05–0.1
$\chi^2/$dof	9.16/5	8.11/5	4.81/5	3.42/5	3.69/5	3.82/5
p value	0.10	0.15	0.44	0.64	0.59	0.57
Δ	3.8(5)	4.4(1)	4.0(5)	3.2(8)	3.3(8)	4.0(6)
γ_*	0.608(8)	0.590(10)	0.543(12)	0.535(12)	0.524(9)	0.498(13)
$a\hat{f}_\pi$	0.008(7)	0.0000(2)	0.009(10)	0.025(12)	0.027(13)	0.013(18)
$\tilde{d}_1$	0.047(35)	2(15)	0.056(59)	0.019(12)	0.019(12)	0.044(56)
$-\log(ad_2)$	9.7(1.3)	20(24)	9.5(1.7)	7.8(7)	7.7(8)	8.8(2.1)

TABLE VI.   Fits of M_π^2/F_π^2 and aF_π to $\gamma\Delta$-dChPT, for fixed values of Δ. All fits have 14 degrees of freedom. The fit with the asterisk may not have fully converged, and its χ^2 value is an upper bound to the true minimum.

Δ	χ^2	p value	$a\hat{f}_\pi$	$\tilde{d}_1$	$-\log(ad_2)$	γ_0	b	c
4.5	11.0*	0.69	0.00004(10)	2.2(2.9)	19(17)	0.2(8.0)	−0.3(2.1)	−0.03(12)
4.4	10.2	0.75	0.00159(31)	0.237(23)	15(1)	2.33(79)	0.73(41)	0.07(5)
4.3	12.0	0.61	0.00384(39)	0.123(6)	12.6(5)	2.04(49)	0.81(32)	0.11(5)
4.2	14.4	0.42	0.00613(42)	0.084(3)	11.4(3)	1.88(35)	0.87(28)	0.14(5)
4.1	17.3	0.24	0.00831(43)	0.063(2)	10.6(2)	1.78(28)	0.93(25)	0.17(5)
3.9999	20.7	0.11	0.01034(43)	0.051(1)	10.1(2)	1.69(23)	0.97(22)	0.20(5)
3.9	24.7	0.04	0.01222(42)	0.0423(7)	9.7(1)	1.62(20)	1.00(21)	0.23(5)
3.8	29.3	0.01	0.01396(41)	0.0363(5)	9.4(1)	1.56(17)	1.03(19)	0.26(5)

below 0.01 for $\Delta < 3.8$. We verified that the p value keeps decreasing down to $\Delta = 2$ (where the p value is of order 10^{-30}). If we increase Δ above 4, the p value increases until Δ reaches 4.5, where the p value appears to start decreasing again. However, we found that fits with $\Delta \geq 4.5$ become very difficult. This is reflected in the very large errors in the six fit parameters: for $\Delta = 4.5$, essentially all of them are not determined by the fit. We have repeated the fits of Table VI omitting the $am_0 = 0.1$ ensemble, or the $am_0 = 0.1$ and 0.08 ensembles, and we have also redone such fits setting $c = 0$ (as in fit 2D). The conclusions are always the same as for the fits shown in Table VI. The fit at $\Delta = 3.9999$ is consistent with the corresponding fit in Table II; values of Δ below roughly 3.8 are strongly disfavored; and the fit starts to deteriorate at $\Delta = 4.5$. If we attempt to include Δ as a parameter in the fit itself (instead of scanning over Δ) fits appear to be unstable.

Given the difficulty fitting the KMI data with the $\mathcal{L}_\Delta$ potential, we have not attempted to include taste splittings in the KMI case.

D. Discussion

Taking the fits of the LSD and KMI data together, it is clear that no very precise statement about the value of Δ can be made. The KMI data appear to exclude the σ-model value $\Delta = 2$. dChPT, which corresponds to $\Delta \to 4$ with fixed $\gamma_m = \gamma_*$, is consistent with the fits shown in Tables IV and V. An exception is the second window fit, fit 5B, which yields a result with a rather small error, $\Delta = 4.4(1)$. But clearly, this result does not account for the variation of Δ across all fits shown in Tables IV, V, and VI.

Our results are consistent with those of Ref. [32]. The main difference is that the KMI data, which were not considered in Ref. [32], present a much stronger lower bound on Δ.

As we show in Appendix B, for values of Δ not close to 4, no power counting exists for the low-energy theory with $\mathcal{L}_d$ replaced by $\mathcal{L}_\Delta$ of Eq. (4.1). However, we do not wish to imply that attempts to understand data in terms of models are not interesting. Fits to models, including Δ-dChPT (with Δ not constrained to be close to 4), can provide a

valuable "stress test" of dChPT. This is why we considered fits of the LSD and KMI data to Δ-dChPT; Ref. [32] can be seen as a similar exploration of only the LSD data.

Fits of the LSD data, comparing, in particular, the values $\Delta = 2$ and $\Delta \to 4$, were considered also in Ref. [8].[32] There, it was found that both dChPT and Δ-dChPT with $\Delta = 2$ provide good fits to data using all five of the LSD ensembles. This finding agrees with our fits in Table IV: fits 4A and 4C are consistent with $\Delta = 2$, but are less than $\sim 2\sigma$ away from $\Delta = 4$.

In summary, a precise determination of the favored value of Δ is not possible with presently available data. Taking the results based on fits to both the LSD and KMI data together, we arrive at an estimated range for Δ,

$$3.5 < \Delta < 4.5. \qquad (4.6)$$

Our lower bound is based on the four-ensemble fits to the LSD data, which favor a value around $\Delta \sim 3.5$, combined with the $\gamma\Delta$-dChPT scan of Table VI, which strongly disfavors values below 3.8. Any fit of the KMI dataset must somehow account for the running of γ_m. Including higher orders systematically is not an option here, because, as we prove in Appendix B, the claim of Ref. [32] that Δ-dChPT admits a systematic expansion is incorrect. The model alternatives are to use a fixed value of γ_m while limiting the mass range as in the "window" fits, or else to use an explicitly varying γ_m function. As for the window fits, Table V shows that Δ is rather insensitive to the mass range in the fit. Also, while both the five-ensemble fits to the LSD data and some of the window fits to the KMI data allow for $\Delta < 3.5$, the fits of Table VI to the KMI data strongly disfavor $\Delta < 3.8$. Based on all fits together, the σ-model value $\Delta = 2$ appears to be excluded. Once again, the caveats discussed in the previous section regarding the LSD and KMI datasets, and, in particular, the lack of

---

[32]See also Ref. [9] for related studies of the SU(3) theory with two sextet fermions, which also has a light flavor-singlet scalar. We recall, however, that dChPT is strictly speaking not applicable to this theory, as the Veneziano limit can be taken only for fermions in the fundamental representation.

information about scaling violations, apply also to our conclusions in this section.

V. CONCLUSION

Our main goal in this paper was to confront the EFT framework provided by dChPT with the KMI data for the eight-flavor SU(3) gauge theory [4]. The KMI simulations were performed at larger fermion masses than the LSD ones [3], taking the theory further away from conformality. Hence, even with the successful application of LO dChPT to the LSD data, which we reported on in Ref. [33], there is no guarantee that LO dChPT can also be applied to the KMI data.

Indeed, we found that the full fermion-mass range of the KMI data cannot be fitted to LO dChPT. The natural next step would be to attempt an NLO fit in dChPT. However, as we explained in Sec. III, this is not feasible with presently available data. First, the large number of parameters involved in any NLO dChPT fit requires extensive precision data for a successful fit. Moreover, the KMI dataset (and, likewise, the LSD dataset) has only a single lattice spacing, making a continuum extrapolation impossible.

Instead, we introduced γ-dChPT, a model extension of LO dChPT with a scale-dependent mass anomalous dimension, which can be interpreted as arising from partially resumming higher orders in the EFT expansion. We found that γ-dChPT provides a successful description of the KMI data over the entire mass range.

Given the success in describing the LSD data using LO dChPT [33], and the KMI data using γ-dChPT with a relatively simple *ansatz* for the γ_m function, the question arises whether γ-dChPT can be used to fit the LSD and KMI data simultaneously. Over the KMI mass range, γ_m would then have to increase as the fermion mass is decreased, eventually saturating to a constant when reaching the lower LSD mass range (see Fig. 3). Once again, however, the inability to take the continuum limit makes it impossible to carry out this program at this time. The lack of information on the lattice spacing dependence is even more severe when trying to consider the LSD and KMI datasets together, because they were produced with different lattice actions, and thus, their scaling violations for any given physical observable are different functions of the corresponding lattice spacing.

We also considered Δ-dChPT—another generalization of LO dChPT in which the dilaton potential is replaced by a class of potentials depending on a new parameter Δ. We emphasize that Δ-dChPT does not allow for a systematic power counting, and should thus be considered a model, except in the limit $\Delta \to 4$ where dChPT is recovered. Δ-dChPT was applied to the LSD data before [32], where it was found that it is difficult to determine the parameter Δ from these data. We confirmed this result, but found that the KMI data allow us to better constrain the value of Δ. We used both the "window" fits in which Δ-dChPT is applied

to subsets of the KMI ensembles, as well as a combination of the two extensions of LO dChPT, with the Δ class of dilaton potentials together with a varying γ_m. We concluded that the preferred range of our combined analysis of the LSD and KMI data is $3.5 < \Delta < 4.5$. This is centered around $\Delta = 4$, where Δ-dChPT reduces to LO dChPT.

Recently, LO dChPT has also been successfully applied to the light sector of the SU(3) gauge theory with four light and six heavy flavors [10]. dChPT provides for a systematic treatment of the pNGBs, the pions and the dilaton, of a near-conformal gauge theory, but it does rest on certain assumptions [17,20]. These initial successes are thus encouraging. We hope that, in the future, more extensive and refined data will become available, allowing for further and more stringent tests of dChPT.

ACKNOWLEDGMENTS

APPENDIX A: SCALE SETTING PRESCRIPTION

Any analysis of lattice data requires a scale setting prescription, and the basic choice is between mass-independent or mass-dependent prescriptions. In this paper, as in Ref. [33], we opted for a scale-independent prescription, and confirmed the self-consistency of this choice by checking that the values of $a\hat{B}_\pi$ on all ensembles agree within error (see Sec. III).

Here we discuss the alternative of using a mass-dependent prescription. In QCD simulations it is common nowadays to use the gradient flow scale t_0 for scale setting [42]. In particular, the ensemble value of t_0 can be used for a mass-dependent prescription. What makes t_0 particularly convenient for setting the scale is that it can be determined with high precision, and it admits a chiral expansion, with nonanalytic terms in the quark mass entering only at NNLO [43]. By contrast, as we showed in Ref. [33], in dChPT there is no (useful) chiral expansion for t_0. This implies that one cannot derive expansions for dimensionless quantities such as $\sqrt{t_0}M_\pi$ and $\sqrt{t_0}F_\pi$ using dChPT.

If we are interested in dChPT fits, we are thus unable to use t_0 for scale setting. Instead, we may consider using a physical quantity such as F_π for a mass-dependent scale setting prescription. As explained in Sec. II D, with our mass-independent scale setting, the basic fit has six parameters, two of which, namely d_2 and $\hat{f}_\pi$, have a mass dimension equal to one. On each ensemble, we fit aF_π and M_π^2/F_π^2 to Eqs. (2.8b) and (2.50), respectively. In addition, we treat the fermion mass am as a data point with a small

fictitious error, in order to determine the expectation value v of the dilaton field on each ensemble, via Eq. (2.47).

If we use F_π for mass-dependent scale setting, we may still fit M_π^2/F_π^2 to Eq. (2.50) as before. In addition, combining Eqs. (2.8b) and (2.47), we may fit m/F_π as

$$\frac{m}{F_\pi} = \frac{d_2}{\hat{f}_\pi} \frac{1}{\tilde{d}_1(3-\gamma_m)} v e^{F(v)}, \qquad (A1)$$

and use it to determine v, now as a function of m/F_π. This procedure gives us access only to the ratio $d_2/\hat{f}_\pi$, instead of to $a d_2$ and $a\hat{f}_\pi$ separately, as in the fitting procedure of Sec. II.D.

We tried to repeat the fits from Table II, using Eqs. (2.50) and (A1). The result was that almost all fit parameters remained completely undetermined. The nominal fit quality was always very high (p value ≥ 0.97), consistent with the failure of the fits to resolve the parameters.

We believe the problem is caused by the much smaller number of degrees of freedom which are available for the mass-dependent fitting procedure. As an example, consider fit 2A, which uses all 10 ensembles and has $2 \times 10 - 6 = 14$ degrees of freedom. As noted in Sec. III C, since the v_i's are determined in terms of the $a m_i$'s using Eq. (2.47), this does not change the number of degrees of freedom. By contrast, within the mass-dependent fitting procedure, we have in total only 20 relations to determine both the five fit parameters and the ten auxiliary v_i's. This leaves us with just 5 degrees of freedom, which apparently is just not enough to resolve the fit parameters.

Although we were unable to actually perform a fit with a mass-dependent prescription for scale setting, we may consider the following "thought experiment." Assume that the data allowed for fits with a mass-dependent prescription, and that the results of those fits are in agreement with Table II. This would mean, in particular, that the value of the new dimensionless fit parameter, $d_2/\hat{f}_\pi$, obtained from fitting Eq. (A1) is consistent with the results for $a\hat{f}_\pi$ and $a d_2$ reported in Table II. Now, while we determine all parameters in the large-mass regime, which is where both the LSD and KMI data are, $\hat{f}_\pi$ and d_2 are LECs that characterize the massless theory. As we discussed in Sec. III E, the values of $a\hat{f}_\pi$ and $a d_2$ extracted from the LSD and the KMI datasets appear to be in conflict, both with the chiral limit values of t_0 (determined from a phenomenological fit) and with each other. It is interesting to check what is the situation for the dimensionless ratio $d_2/\hat{f}_\pi$. Comparing the value of this ratio using the results in Table II to those from Ref. [33] reveals that there is still a significant conflict, of roughly the same size as for $a\hat{f}_\pi$, though smaller than for $a d_2$. We conclude that restricting ourselves to a mass-dependent scale setting prescription would not by itself alleviate the problem of the long extrapolation from the large-mass regime to the chiral limit.

APPENDIX B: POWER COUNTING

In Ref. [32] it was proposed that Δ-dChPT—in which the potential $\mathcal{L}_d$ of Eq. (2.6a) is replaced by $\mathcal{L}_\Delta$ of Eq. (4.1), with Δ a new free parameter—admits a systematic power counting. In this Appendix, we show that the arguments given in Ref. [32] are not correct.

The potential (4.1) was already considered in Refs. [24,25]. In those papers it was assumed that the Lagrangian of the underlying theory contains an operator with scaling dimension Δ, with some unspecified value of Δ, and a coupling which may be small. This naturally leads to the consideration of potentials such as Eq. (4.1) in the EFT describing the same theory at low energy.

By contrast, here the underlying theory is known: it is the asymptotically free SU(3) gauge theory with $N_f = 8$ Dirac fermions in the fundamental representation. This theory does not fall into the class of theories considered in Refs. [24,25].

It is instructive to briefly recall how the breaking of scale invariance is introduced into the (massless) quantum theory, and then, how this breaking translates to the EFT [17]. As can be seen in Eq. (2.16), regularizing the bare Lagrangian requires the introduction of a scale factor, μ_0^{d-4}, with the limit $d - 4 \to 0$ to be taken after renormalization. By letting μ_0 transform according to Eq. (2.17b), we may promote μ_0 to a spurion, formally restoring scale invariance.

In making the transition to the EFT, we will want to use the well-known fact that the EFT Lagrangian must be analytic in the spurion fields, if the underlying Lagrangian is analytic in the (same set of) spurions. Correlation functions can then be generated by differentiating the partition function of the EFT with respect to the spurion fields and compared with their counterparts in the underlying theory by applying the same derivatives again. This matching procedure fixes the LECs of the EFT order by order, according to the power counting.

A technical obstacle is that the action (2.16) is nonanalytic in the spurion μ_0. To overcome this problem, we introduce a new spurion field $\sigma(x)$ and replace

$$\mu_0 \Rightarrow \hat{\mu}_0 e^{\sigma(x)}. \qquad (B1)$$

The new scale transformation rules replacing Eq. (2.17b) are

$$\sigma(x) \to \sigma(\lambda x) + \log \lambda, \qquad (B2a)$$

$$\hat{\mu}_0 \to \hat{\mu}_0. \qquad (B2b)$$

Now $\hat{\mu}_0$ is invariant under a scale transformation, which in turn is "carried" by the constant mode of the new spurion field. Writing $\sigma(x) = \sigma_0 + \delta\sigma(x)$, with the constraint $\int d^d x \, \delta\sigma(x) = 0$, it follows that

$$\sigma_0 \to \sigma_0 + \log \lambda. \tag{B3}$$

With this replacement, the bare action (2.16) becomes

$$S = \hat{\mu}_0^{d-4} \int d^d x \, e^{(d-4)\sigma(x)} \mathcal{L}(x). \tag{B4}$$

Classically, the $\sigma(x)$ dependence vanishes for $d \to 4$, showing that any dependence of the renormalized theory on $\sigma(x)$ represents quantum breaking of scale invariance [17,19,38]. Since the underlying theory is now analytic in the spurion field $\sigma(x)$, so must be the EFT [17]. Note that if, instead, one were to use μ_0 as a scale spurion, there would be no reason for the EFT to be analytic in μ_0, for the simple reason that the underlying theory (2.16) is nonanalytic in μ_0.[33]

In Ref. [32], the starting point of the argument was to assume that the Lagrangian of the low-energy theory depends analytically on a spurion field $\tilde{\mu}(x)$, with the scale transformation rule[34]

$$\tilde{\mu}(x) \to \lambda^{4-\Delta} \tilde{\mu}(\lambda x). \tag{B5}$$

It is clear from the previous discussion that the underlying gauge theory does not accommodate such a spurion. Comparing transformation rules, one can, however, make the identification

$$\tilde{\mu}(x) \equiv e^{(4-\Delta)\sigma(x)}. \tag{B6}$$

As we have just shown, the correct EFT must be analytic in $\sigma(x)$, but not in $e^{\sigma(x)}$ (nor in any power of $e^{\sigma(x)}$). It follows immediately that the EFT must be analytic in $\log \tilde{\mu}(x)$, but not in $\tilde{\mu}(x)$ itself. This proves that the arguments of Ref. [32] are not valid, because the incorrect assumption that the EFT is analytic in $\tilde{\mu}(x)$ served as their starting point.

While this proves that the power counting claimed in Ref. [32] is unfounded, several comments are in order.

First, we draw the reader's attention to Sec. II C where we made use of the original spurion μ_0, instead of $\sigma(x)$. The reason is that our goal in Sec. II C was to derive the extension of LO dChPT to the case of a running γ_m. This requires mainly the consideration of renormalization-group and scale transformation properties, and, for this purpose, using μ_0 as a (constant) scale spurion is sufficient.[35] The γ-dChPT framework developed in Sec. II C does deviate from the strict power counting of dChPT [17],

though it can be viewed as a resummation of contributions from all orders under the assumption that these dominate. As for establishing the power counting itself, this necessitates the replacement of μ_0 by $\hat{\mu}_0 e^{\sigma(x)}$; cf. Eq. (B1). Correspondingly, the transformation rules (B2) take over the transformation rule of μ_0 in Eqs. (2.17) and (2.32). For the actual proof of the power counting, and a detailed discussion of the assumptions that it requires, we refer to Refs. [17,20].

A key step in constructing a power-counting scheme is the identification, in the underlying theory, of a small parameter in terms of which the EFT expansion is to be organized. In ordinary ChPT, the small parameter is the fermion mass m, which is also the "expectation value" of the chiral spurion, $\langle \chi(x) \rangle = m$. Chiral symmetry is restored for $m \to 0$, which in turn allows one to establish that the pion mass is parametrically small.

By contrast, in Ref. [17], the small parameter controlling the hard breaking of scale invariance was identified as $n_f - n_f^*$, serving as a proxy for the β function at the chiral symmetry breaking scale. More precisely, the hypothesis made in Ref. [17] is

$$\tilde{\beta} \sim |n_f - n_f^*|^\eta, \qquad n_f \nearrow n_f^*, \tag{B7}$$

for some $\eta > 0$, where

$$\tilde{\beta} = \frac{\mu}{4\alpha} \frac{\partial \alpha}{\partial \mu}, \tag{B8}$$

and α is the 't Hooft coupling, $\alpha = g^2 N / (4\pi)$, evaluated at the chiral symmetry breaking scale. While we often assume $\eta = 1$ for simplicity, including earlier in this paper, this assumption is not essential. The power counting is valid for any fixed $\eta > 0$; the η dependence is restored trivially via the substitution $|n_f - n_f^*| \Rightarrow |n_f - n_f^*|^\eta$.

The small parameter $|n_f - n_f^*|^\eta$ does not appear explicitly in the underlying Lagrangian, and, in particular, it is not identified with the expectation value of $\sigma(x)$. Indeed, unlike chiral symmetry, which is restored for $m = 0$, there is no fixed value of $\sigma(x)$ for which scale invariance is *not* broken. Rather, the expansion of correlation functions in powers of $\sigma(x)$ corresponds to an expansion in the number of insertions of the trace anomaly. In the massless limit, every such insertion is proportional to the β function at the chiral symmetry breaking scale, hence to $|n_f - n_f^*|^\eta$. For this argument to work, it is crucial to use the spurion field $\sigma(x)$, and not μ_0 or $\tilde{\mu}(x)$. The role of $\sigma(x)$, or of its constant mode σ_0, is analogous to that of the θ parameter in the large-N_c limit of ChPT in which the $U(1)_A$ symmetry is restored. For a detailed comparison, we refer to Ref. [17]. The upshot is that one cannot establish a relation between the expectation value of μ_0 or $\tilde{\mu}(x)$ and the β function of the underlying theory. Hence, even if one were to allow the low-energy

---

[33]The same statement applies if the constant spurion μ_0 is promoted to a field.

[34]The spurion $\tilde{\mu}(x)$ is denoted as $\lambda(x)$ in Ref. [32]; see Eq. (A1) therein. We have reserved λ for the scale transformation parameter, which in turn is denoted as e^ρ in Ref. [32].

[35]In accordance with our general reasoning, in Eq. (2.28), $\mathcal{L}_m$ indeed depends on $\log \mu_0$.

theory to depend only on integer powers of the $\tilde{\mu}(x)$ spurion, as was postulated in Ref. [32], there is no reason to assume that its expectation value should tend to zero when the conformal window is approached. Both assumptions of Ref. [32], analyticity in the spurion $\tilde{\mu}$, and its smallness in the conformal limit, are thus in conflict with the properties of the underlying theory, and, in general, not valid.

This concludes our discussion of the claims made in Ref. [32] with regard to power counting. But, a little more can be said about the connection of the potential $\mathcal{L}_\Delta$ with dChPT, which corresponds to the limit $\Delta \to 4$. According to the dChPT power counting developed in Ref. [17], the scale invariant dilaton potential $e^{4\tau}$ is multiplied by a potential $\tilde{V}_d(\tau)$ that breaks scale invariance,

$$\tilde{V}_d(\tau) = \sum_{n=0}^{\infty} \frac{\tilde{c}_n}{n!} \tau^n. \qquad (B9)$$

The LECs $\tilde{c}_n$ scale as $\tilde{c}_n \sim |n_f - n_f^*|^{n\eta}$, and with the power counting [compare Eq. (2.3)]

$$p^2 \sim m \sim |n_f - n_f^*|^\eta \sim 1/N, \qquad (B10)$$

it follows that the term $\frac{\tilde{c}_n}{n!} \tau^n$ can appear only at N^{n-1}LO in dChPT. In particular, the tree-level potential $V_d(\tau)$ of Eq. (2.6b) obtained after the τ shift corresponds to $c_1 = \tilde{c}_1 = -4\tilde{c}_0$.

When Δ is close to 4, we may identify $\tilde{V}_d(\tau) = V_\Delta(\tau)$, which, using Eq. (4.1b), implies that $\tilde{c}_0 = -c_1/\Delta$ and

$$\tilde{c}_n = (4c_1/\Delta)(\Delta - 4)^{n-1}, \qquad n \geq 1. \qquad (B11)$$

The first two terms in the expansion reproduce the LO potential $V_d(\tau)$.[36] It follows that, for any fixed value of Δ such that $|\Delta - 4| \sim |n_f - n_f^*|^\eta$ with any $\eta > 0$, the $V_\Delta(\tau)$ potential will inherit the power counting of dChPT. The same is not true for values of Δ not close to 4, and thus, it is also not true for the low-energy Lagrangian in which Δ is treated as a free parameter.

---

[36]Except for the innocuous replacement $c_1 \Rightarrow 4c_1/\Delta$.

[1] T. Appelquist *et al.* (LSD Collaboration), Lattice simulations with eight flavors of domain wall fermions in SU(3) gauge theory, Phys. Rev. D **90**, 114502 (2014).

[2] T. Appelquist *et al.*, Strongly interacting dynamics and the search for new physics at the LHC, Phys. Rev. D **93**, 114514 (2016).

[3] T. Appelquist *et al.* (Lattice Strong Dynamics Collaboration), Nonperturbative investigations of SU(3) gauge theory with eight dynamical flavors, Phys. Rev. D **99**, 014509 (2019).

[4] Y. Aoki *et al.* (LatKMI Collaboration), Light flavor-singlet scalars and walking signals in $N_f = 8$ QCD on the lattice, Phys. Rev. D **96**, 014508 (2017).

[5] Z. Fodor, K. Holland, J. Kuti, D. Nogradi, C. Schroeder, and C. H. Wong, Can the nearly conformal sextet gauge model hide the Higgs impostor?, Phys. Lett. B **718**, 657 (2012).

[6] Z. Fodor, K. Holland, J. Kuti, D. Nogradi, and C. H. Wong, Can a light Higgs impostor hide in composite gauge models?, *Proc. Sci.*, LATTICE2013 (2014) 062 [arXiv: 1401.2176].

[7] Z. Fodor, K. Holland, J. Kuti, D. Nogradi, and C. H. Wong, The twelve-flavor β-function and dilaton tests of the sextet scalar, EPJ Web Conf. **175**, 08015 (2018).

[8] Z. Fodor, K. Holland, J. Kuti, and C. H. Wong, Tantalizing dilaton tests from a near-conformal EFT, *Proc. Sci.*, LATTICE2018 (2019) 196 [arXiv:1901.06324].

[9] Z. Fodor, K. Holland, J. Kuti, and C. H. Wong, Dilaton EFT from p-regime to RMT in the ϵ-regime, *Proc. Sci.*, LATTICE2019 (2020) 246 [arXiv:2002.05163].

[10] T. Appelquist *et al.* (Lattice Strong Dynamics), Near-conformal dynamics in a chirally broken system, arXiv: 2007.01810.

[11] R. C. Brower, A. Hasenfratz, C. Rebbi, E. Weinberg, and O. Witzel, Composite Higgs model at a conformal fixed point, Phys. Rev. D **93**, 075028 (2016).

[12] T. DeGrand, Lattice tests of beyond Standard Model dynamics, Rev. Mod. Phys. **88**, 015001 (2016).

[13] D. Nogradi and A. Patella, Strong dynamics, composite Higgs and the conformal window, Int. J. Mod. Phys. A **31**, 1643003 (2016).

[14] C. Pica, Beyond the Standard Model: Charting fundamental interactions via lattice simulations, *Proc. Sci.*, LATTICE2016 (2016) 015 [arXiv:1701.07782].

[15] B. Svetitsky, Looking behind the Standard Model with lattice gauge theory, EPJ Web Conf. **175**, 01017 (2018).

[16] V. Drach, Composite electroweak sectors on the lattice, *Proc. Sci.*, LATTICE2019 (2020) 242 [arXiv:2005.01002].

[17] M. Golterman and Y. Shamir, Low-energy effective action for pions and a dilatonic meson, Phys. Rev. D **94**, 054502 (2016).

[18] M. Golterman and Y. Shamir, Effective field theory for pions and a dilatonic meson, *Proc. Sci.*, LATTICE2016 (2016) 205 [arXiv:1610.01752].

[19] M. Golterman and Y. Shamir, Effective pion mass term and the trace anomaly, Phys. Rev. D **95**, 016003 (2017).

[20] M. Golterman and Y. Shamir, Large-mass regime of the dilaton-pion low-energy effective theory, Phys. Rev. D **98**, 056025 (2018).

[21] T. V. Brown, M. Golterman, S. Krøjer, Y. Shamir, and K. Splittorff, The ϵ-regime of dilaton chiral perturbation theory, Phys. Rev. D **100,** 114515 (2019).

[22] T. Appelquist and Y. Bai, A light dilaton in walking gauge theories, Phys. Rev. D **82,** 071701 (2010).

[23] C. N. Leung, S. T. Love, and W. A. Bardeen, Aspects of dynamical symmetry breaking in gauge field theories, Nucl. Phys. **B323,** 493 (1989).

[24] W. D. Goldberger, B. Grinstein, and W. Skiba, Distinguishing the Higgs Boson from the Dilaton at the Large Hadron Collider, Phys. Rev. Lett. **100,** 111802 (2008).

[25] Z. Chacko and R. K. Mishra, Effective theory of a light dilaton, Phys. Rev. D **87,** 115006 (2013).

[26] S. Matsuzaki and K. Yamawaki, Dilaton Chiral Perturbation Theory: Determining the Mass and Decay Constant of the Technidilaton on the Lattice, Phys. Rev. Lett. **113,** 082002 (2014).

[27] M. Hansen, K. Langæble, and F. Sannino, Extending chiral perturbation theory with an isosinglet scalar, Phys. Rev. D **95,** 036005 (2017).

[28] T. Appelquist, J. Ingoldby, and M. Piai, Dilaton EFT framework for lattice data, J. High Energy Phys. 07 (2017) 035.

[29] T. Appelquist, J. Ingoldby, and M. Piai, Analysis of a dilaton EFT for lattice data, J. High Energy Phys. 03 (2018) 039.

[30] O. Catà, R. J. Crewther, and L. C. Tunstall, Crawling technicolor, Phys. Rev. D **100,** 095007 (2019).

[31] T. Appelquist et al. (LSD Collaboration), Linear sigma EFT for nearly conformal gauge theories, Phys. Rev. D **98,** 114510 (2018).

[32] T. Appelquist, J. Ingoldby, and M. Piai, The dilaton potential and lattice data, Phys. Rev. D **101,** 075025 (2020).

[33] M. Golterman, E. T. Neil, and Y. Shamir, Application of dilaton chiral perturbation theory to $N_f = 8$, SU(3) spectral data, Phys. Rev. D **102,** 034515 (2020).

[34] A. Bazavov et al. (MILC Collaboration), Nonperturbative QCD simulations with $2 + 1$ flavors of improved staggered quarks, Rev. Mod. Phys. **82,** 1349 (2010).

[35] M. Golterman, Applications of chiral perturbation theory to lattice QCD, in *Modern Perspectives in Lattice QCD: Quantum Field Theory and High Performance Computing, Proceedings, International School, 93rd Session, Les Houches, France, 2009* (2009), pp. 423–515 [arXiv:0912 .4042].

[36] A. Hasenfratz, A. Cheng, G. Petropoulos, and D. Schaich, Reaching the chiral limit in many flavor systems, arXiv: 1303.7129.

[37] G. Veneziano, Some aspects of a unified approach to gauge, dual and Gribov theories, Nucl. Phys. **B117,** 519 (1976); U(1) without instantons, Nucl. Phys. **B159,** 213 (1979).

[38] J. C. Collins, A. Duncan, and S. D. Joglekar, Trace and dilatation anomalies in gauge theories, Phys. Rev. D **16,** 438 (1977).

[39] W. J. Lee and S. R. Sharpe, Partial flavor symmetry restoration for chiral staggered fermions, Phys. Rev. D **60,** 114503 (1999).

[40] C. Aubin and C. Bernard, Pion and kaon masses in staggered chiral perturbation theory, Phys. Rev. D **68,** 034014 (2003).

[41] T. DeGrand, Y. Shamir, and B. Svetitsky, Infrared fixed point in SU(2) gauge theory with adjoint fermions, Phys. Rev. D **83,** 074507 (2011).

[42] M. Lüscher, Properties and uses of the Wilson flow in lattice QCD, J. High Energy Phys. 08 (2010) 071; Erratum, J. High Energy Phys. 03 (2014) 92.

[43] O. Bär and M. Golterman, Chiral perturbation theory for gradient flow observables, Phys. Rev. D **89,** 034505 (2014); Erratum, Phys. Rev. D **89,** 099905 (2014).