Instantons and chiral Lagrangian

Kiwoon Choi
Department of Physics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

C. W. Kim
Department of Physics and Astronomy, The Johns Hopkins University, Baltimore, Maryland 21218

(Received 20 March 1989)

We derive the chiral Lagrangian including instanton corrections in a consistent way. *CP* violation associated with the θ term in QCD is discussed in terms of this chiral Lagrangian. A particular emphasis is given to a phenomenological determination of m_u and θ. It is shown that, contrary to the previous conclusion that $m_u = 0.56 m_d$ and $\theta = 0$, all three distinct cases (i) $0.56 \geq m_u/m_d > 0, \theta = 0$, (ii) $m_u = 0$, (iii) $0.56 \geq m_u/m_d > 0, \theta = \pi$ are phenomenologically acceptable depending on the strength of the instanton corrections. In cases (i) and (iii), the ratio m_u/m_d can be much smaller than the usually quoted value 0.56. As a result, the models of spontaneous *CP* violation which predict $|\theta| \gg 10^{-9}$ can be phenomenologically viable. Also in the axion models, because of the possibility of m_u/m_d being arbitrarily small, the cosmological upper bound on the axion decay constant can be relaxed.

I. INTRODUCTION

In QCD of three flavors of massless quarks, the quark part of the classical QCD Lagrangian is invariant under a global flavor symmetry $G_f = \mathrm{U}_A(1) \times \mathrm{SU}(3) \times \mathrm{SU}(3)$. The axial $\mathrm{U}_A(1)$ symmetry, while being valid classically, is broken at the quantum level by the anomaly and instanton effects and thus the $\mathrm{U}_A(1)$ problem is resolved.[1] If the flavor symmetry G_F is not an exact symmetry even at the classical level due to the nonzero intrinsic quark masses, $M \neq 0$, instantons provide not only $\mathrm{U}_A(1)$-breaking effects but also the pieces which violate $\mathrm{SU}(3) \times \mathrm{SU}(3)$, e.g., effective current masses of quarks, combined with the insertion of the intrinsic current masses.[2] Also in this case a peculiar role is played by the instantons in connection with the *CP* violation via the θ term,[3] viz., the strong *CP* violation is the result of instanton effects.[4] Therefore it is essential to include instanton effects carefully in studying physics associated with the flavor-symmetry breaking or with the strong *CP* violation.

Recently it has been observed that instantons can change the previous conclusion on the value of the intrinsic current mass of the u quark.[5] Instantons generate an effective mass of quarks of the form $(\det M/\Lambda_{\mathrm{QCD}})\bar{q}M^{-1}q$, where $M = \mathrm{diag}(m_u, m_d, m_s)$ denotes the intrinsic current mass matrix which appears in the renormalizable QCD Lagrangian. This instanton-induced mass does provide an explicit breaking of G_f which is second order in M and thus corresponds to an effective current mass. Note that it is the result of combined effects of both the explicit breaking of $\mathrm{U}_A(1)$ by instantons and the explicit breaking of G_f by the intrinsic current mass M. Even though it is still a subject with no definite answer, it has long been suspected that instantons do play some role in spontaneous chiral-symmetry breaking.[6] We emphasize here that the instanton-induced effective current mass under our consideration is completely independent of the details of spontaneous chiral-symmetry breaking, in particular it is independent of whether or not instantons are responsible for the nonzero quark-antiquark condensate, and can be clearly distinguished from the constituent mass which is the result of the spontaneous chiral-symmetry breaking. For example, as we will see, it gives a contribution to the pseudo-Goldston-boson masses (i.e., the pseudoscalar-meson masses) in exactly the same way as the intrinsic current mass does, while the constituent mass does not.

Even though second order in M, the above instanton-induced effective current mass can be important for the amplitudes concerning m_u. Note that the instanton-induced u-quark mass $(\det M/\Lambda_{\mathrm{QCD}})(M^{-1})_{uu} = m_d m_s/\Lambda_{\mathrm{QCD}}$ can be even larger than the original intrinsic mass m_u because m_s is not much smaller than Λ_{QCD}. This implies that we should include instanton effects carefully in estimating m_u by studying the explicit breaking of the flavor symmetry in hadronic amplitudes.

In addition to the classic example of the application of instantons to the $\mathrm{U}_A(1)$ problem,[1] some other phenomenological implication of instanton physics, e.g., to the pseudoscalar-meson masses in the framework of QCD sum rules,[7] to the baryon masses in the framework of the nonrelativistic quark model,[8] and to the $\Delta I = \frac{1}{2}$ rule,[9] have been considered. In this paper we consider the chiral Lagrangian which includes the instanton effects that explicitly break G_f in a consistent way for a systematic study of the phenomenological implications of instantons. In particular, following the analysis of Ref. 10, we include the above-mentioned instanton-induced mass which is second order in M without worrying about the other second-order effects. The main observation from our chiral Lagrangian is that the $\mathrm{SU}(3) \times \mathrm{SU}(3)$ breaking in the *CP*-conserving hadronic amplitudes is described by the effective current mass

$$M_{\rm eff}=\mathrm{diag}(\overline{m}_u,\overline{m}_d,\overline{m}_s)=\mathrm{diag}((-)^n(1+\lambda_2)m_u+\lambda_1 m_d,(1+\lambda_2)m_d+(-)^n\lambda_1 m_u,(1+\lambda_2)m_s)+(-1)^n\lambda_1 m_u m_d/m_s\ , \tag{1.1}$$

while the CP-violating amplitudes are proportional to $\Delta\theta(m_u m_d/(-)^n m_u+m_d)$, viz., are proportional to the intrinsic current masses of the u and d quarks. [Here λ_1 and λ_2 are parameters characterizing the strength of instanton effects. n and $\Delta\theta$ are defined as $\theta=n\pi+\Delta\theta(n=0$ or $1, |\Delta\theta|\leq 1)$.]

The CP-conserving hadronic amplitudes determine the effective current mass $M_{\rm eff}$; e.g., they give $\overline{m}_u/\overline{m}_d=0.56$, $\overline{m}_s/\overline{m}_d=20$. However both n and m_u/m_d which are relevant for the phenomenological determination of θ depend strongly on the magnitudes of λ_1 and λ_2. In Appendix A we provide a numerical estimation of the strength of instanton corrections within the semiclassical instanton-gas picture supplemented by the phenomenological constraints coming from the η-η' mass matrix. Then we find that, even in the semiclassical instanton gas regime, the instanton correction can be strong enough to completely change the previous conclusion on m_u/m_d and θ. For example, the case of $\theta=\pi$ (Ref. 11) which has been argued as being inconsistent with low-energy phenomenology[12] can be phenomenologically viable due to the instanton effects. Also m_u/m_d can take an arbitrarily small value without any difficulty with phenomenology. Even though we draw our conclusion via the chiral Lagrangian, we emphasize here that the observation does hold for all hadronic amplitudes which may be used to determine m_u/m_d or θ phenomenologically as long as we take into account instanton effects carefully.

The possibility of m_u/m_d being much smaller than the usually quoted value 0.56 readily indicates that the bound on $\Delta\theta$ from the neutron electric dipole moment can be relaxed because the neutron electric dipole moment rather gives a bound on $\Delta\theta(m_u/m_d)$, not directly the bound on $\Delta\theta$ (Refs. 12–16). This implies that many of the models of spontaneously broken CP, which have been considered as being inconsistent with phenomenology because of their prediction of relatively large value of $\Delta\theta$, can now be phenomenologically acceptable.[17] Another implication of smaller values of m_u/m_d appear in axion phenomenology.[4] As we will see, the axion mass squared m_a^2 is also proportional to m_u/m_d. As a result, for a given axion decay constant, both the nonzero- and zero-temperature axion mass become smaller. An immediate consequence would be the relaxation of the cosmological upper bound on f_a (Ref. 18), viz., $f_a\leq 10^{12}$ GeV, coming from a consideration of the axion energy density.

The organization of this paper is as follows. In Sec. II we present the flavor-symmetry-violating interactions of quark fields, including the instanton-induced ones. The corresponding chiral Lagrangian is derived in Sec. III and its phenomenological applications are described in Sec. IV. Appendix A is devoted to the numerical estimation of the strength of instanton effects and Appendix B provides the phenomenological determination of various strong-interaction parameters which appear in our chiral Lagrangian.

II. INSTANTON-INDUCED INTERACTIONS OF LIGHT-QUARK FIELDS

Consider the renormalizable QCD Lagrangian of three flavors of light quarks $q=(u,d,s)$:

$$\mathcal{L}_{\rm QCD}=-\tfrac{1}{4}F_{\mu\nu}F^{\nu\mu}+\overline{q}i\not{D}q-(\overline{q}_L M q_R+\text{H.c.})+\frac{g^2\theta}{32\pi^2}F_{\mu\nu}\tilde{F}^{\mu\nu}\ , \tag{2.1}$$

where $F^{\mu\nu}$ denotes the gluon field strength, $\tilde{F}^{\mu\nu}=\frac{1}{2}\epsilon^{\mu\nu\pi\sigma}F_{\pi\sigma}$ is its dual. Under the global flavor symmetry $G_f=\mathrm{U}_A(1)\times\mathrm{SU}(3)\times\mathrm{SU}(3)$, quark fields transform as

$$q_L\rightarrow e^{-i\alpha}Lq_L, q_R\rightarrow e^{i\alpha}Rq_R\ . \tag{2.2}$$

Here L and R are independent SU(3) matrices and α generates $\mathrm{U}_A(1)$ rotations. G_f is spoiled by the QCD anomaly as well as nonzero quark masses. One may consider the transformation group G_f' which is defined as the following transformation of parameters together with the above G_f transformation of quark fields:

$$M\rightarrow e^{-2i\alpha}LMR^\dagger,\quad \theta\rightarrow\theta+6\alpha\ . \tag{2.3}$$

Then the QCD Lagrangian, including the contribution from the path-integral measure of quark fields which gives rise to $\mathrm{U}_A(1)$ anomaly, is invariant under G_f'. Even though G_f' is not a symmetry on the physical Hilbert space (because it changes parameters), it plays a useful role in identifying the low-energy effective Lagrangian. The effective Lagrangian should also be invariant under G_f'.

One can do the same for the parity symmetry P. Under P, the quark and gluon fields transform as

$$q_{L,R}(x)\rightarrow q_{R,L}(\tilde{x}),\quad A_\mu(x)\rightarrow A^\mu(\tilde{x})\ , \tag{2.4}$$

where $\tilde{x}_\mu=x^\mu$ with metric $\eta_{\mu\nu}=(+,-,-,-)$. The corresponding P' which is the analog of G_f' includes the transformation of parameters

$$M\rightarrow M^\dagger,\quad \theta\rightarrow-\theta\ . \tag{2.5}$$

Then the QCD Lagrangian of Eq. (2.1) and the resulting low-energy effective Lagrangian are also invariant under P'.

As is well known by now,[4] the CP (or P) violation in strong interactions via the θ term in the Lagrangian (2.1) is the result of the QCD anomaly and instanton effects. The most convenient way to include instanton effects is the effective Lagrangian method in which the instantons are integrated out first and their effects appear explicitly in the effective Lagrangian of quark fields. In this section we consider the effective Lagrangian of light quarks including terms induced by instantons. The chiral Lagrangian as a low-energy realization of this effective Lagrangian of light-quark fields will be the subject of the next section.

The relevant interactions of quark fields mediated by instantons can be found in the literature.[1,2] The CP violation associated with θ appears through amplitudes which violate the flavor symmetry G_f explicitly because θ can be rotated away in the case when one of the light quarks (say, u quark) is massless. After integrating out instantons, the part of the effective Lagrangian which explicitly violates G_f (but is invariant under both G_f' and P') can be written as

$$\Delta\mathcal{L}_{\rm QCD}=-\bar{q}_L\left[M+e^{-i\theta}\frac{y_1}{\Lambda}(\det M^\dagger)(M^\dagger)^{-1}\right]q_R +e^{-i\theta}\frac{y_2}{\Lambda^3}\det\Phi[{\rm tr}M^\dagger(\Phi)^{-1}] -e^{-i\theta}\frac{y_3}{\Lambda^5}\det\Phi+{\rm H.c.}\ , \qquad (2.6)$$

where Φ is defined as $\Phi_{ij}=(\bar{q}_{iL}w)(\bar{w}q_{jR})$ for a color-triplet spinor w depending on the instanton orientation in the color space. Here the terms with the phase $e^{-i\theta}$ are the ones generated by instantons and

$$y_1=\frac{4\pi^2}{3}\Lambda\int\frac{d\mu}{\mu^2}D(\mu)Z^{-1}(\mu/\Lambda)\ ,$$
$$y_2=8\pi^4\Lambda^3\int\frac{d\mu}{\mu^4}D(\mu)Z(\mu/\Lambda)\ , \qquad (2.7)$$
$$y_3=\frac{32\pi^6}{3}\Lambda^5\int\frac{d\mu}{\mu^6}D(\mu)Z^3(\mu/\Lambda)\ .$$

$Z(\mu/\Lambda)=[\alpha(\Lambda)/\alpha(\mu)]^{4/9}$ is a multiplicative renormalization factor of the intrinsic quark mass and $D(\mu)$ is the instanton density whose explicit form will be given later [see Eq. (A2) in Appendix A]. The quark fields and masses in Eq. (2.6) are renormalized at the scale Λ. Notice $\Delta\mathcal{L}_{\rm QCD}$ is invariant under the renormalization-group transformation

$$M(\Lambda)\to M(\Lambda')=Z(\Lambda/\Lambda')M(\Lambda)\ ,\quad \bar{q}q\to Z^{-1}(\Lambda/\Lambda')\bar{q}q \qquad (2.8)$$

and thus is independent of our choice of the renormalization point. For the terms in $\Delta\mathcal{L}_{\rm QCD}$ containing Φ, the average over the instanton orientation is taken and then we have[2]

$$\det\Phi=\tfrac{2}{9}\det(\bar{q}_Lq_R)+[\text{terms containing color-SU(3) generators}]\ ,$$
$$\det\Phi[{\rm tr}M^\dagger(\Phi)^{-1}]=\tfrac{1}{9}\epsilon_{ijk}\epsilon_{lmn}M^\dagger_{il}(\bar{q}_{jL}q_{mR})(\bar{q}_{kL}q_{nR})+[\text{terms containing color-SU(3) generators}]\ . \qquad (2.9)$$

Among the instanton-induced terms in Eq. (2.6), the six-quark operator $\det\Phi$ is invariant under SU(3)$\times$SU(3) and violates $U_A(1)$ only. However, we note that the other two terms have a dependence of the intrinsic quark mass M and therefore they violate not only $U_A(1)$ but also SU(3)$\times$SU(3). It has been argued that the instanton-induced mass $y_1(\det M)M^{-1}/\Lambda$, even though it is second order in M, can play an important role in the isospin-violating low-energy amplitudes.[5,10] Furthermore a careful analysis of the full second-order effects with respect to M (including electromagnetic corrections) indicates that only this particular form of second-order effects can significantly change naive first-order results.[10] Therefore we include this instanton-induced mass term in our analysis without worrying about other second-order effects. For the six-quark operator, it is zeroth order in M and thus one might worry that its coefficient cannot serve as a useful expansion parameter. However it turns out that one can still consider an expansion with respect to its coefficient in the following sense.

For this purpose let us define the renormalization-group-invariant (i.e., Λ-independent) dimensionless parameters λ_a $(a=1,2,3)$ as

$$\lambda_1=y_1m_s/\Lambda\ ,$$
$$\lambda_2=-2y_2\langle 0|\bar{q}_Lq_R|0\rangle/9\Lambda^3\ , \qquad (2.10)$$
$$\lambda_3=2y_3(\langle 0|\bar{q}_Lq_R|0\rangle)^2/9\Lambda^5m_s\ ,$$

where $\langle 0|\bar{q}_Lq_R|0\rangle$ denotes the chiral condensate of each quark flavor. Then the effective Lagrangian of Eq. (2.6) can be written as

$$\Delta\mathcal{L}_{\rm QCD}=[M+e^{-i\theta}\lambda_1(\det M^\dagger)(M^\dagger)^{-1}/m_s]_{ij}\Gamma^{(1)}_{ij}+e^{-i\theta}\lambda_2M^\dagger_{ij}\Gamma^{(2)}_{ij}+e^{-i\theta}\lambda_3m_S\Gamma^{(3)}\ , \qquad (2.11)$$

where

$$\Gamma^{(1)}_{ij}=-\bar{q}_{iL}q_{jR}\ ,$$
$$\Gamma^{(2)}_{ij}=[\tfrac{1}{2}\epsilon_{ikl}\epsilon_{jmn}(\bar{q}_{kL}q_{mR})(\bar{q}_{lL}q_{nR})/\langle 0|\bar{q}_Lq_R|0\rangle]+[\text{terms containing color-SU(3) generators}]\ , \qquad (2.12)$$
$$\Gamma^{(3)}=-[\det(\bar{q}_Lq_R)/(\langle 0|\bar{q}_Lq_R|0\rangle)^2]+[\text{terms containing color-SU(3) generators}]\ .$$

Now it is not unreasonable to assume that the effects of $\Gamma^{(a)}$'s $(a=1,2,3)$ to the low-energy hadron amplitudes are the same in the order of magnitudes. For example, within the vacuum-insertion approximation[2] we have

$$\langle 0|\Gamma^{(1)}_{ij}|0\rangle\simeq\langle 0|\Gamma^{(2)}_{ij}|0\rangle\simeq\delta_{ij}\langle 0|\Gamma^{(3)}|0\rangle\ . \qquad (2.13)$$

Then the success of the first-order SU(3)$\times$SU(3) chiral perturbation with respect to M readily implies that, as long as $\lambda_2,\lambda_3<1$, the results which are first order in the coefficients of $\Gamma^{(a)}$'s can be a good approximation. In fact our estimation of λ_a suggests (see Appendix A)

$$\lambda_3<\lambda_2\leq 0.4 \tag{2.14}$$

and therefore we can restrict ourselves to the first-order results. Note that within this approximation, there is no ambiguity of double counting the instanton effects. In the next section we will consider the chiral Lagrangian up to terms which are first order in $\Delta\mathcal{L}_{\rm QCD}$.

III. CHIRAL LAGRANGIAN WITH INSTANTON EFFECTS

The chiral Lagrangian which describes the low-energy interactions of the pseudoscalar mesons of the baryon octet can be written as

$$\mathcal{L}_{\rm chiral}=\mathcal{L}_0+\Delta\mathcal{L}\ , \tag{3.1}$$

where $\mathcal{L}_0$ is invariant under the flavor symmetry $G_f={\rm U}_A(1)\times{\rm SU}(3)\times{\rm SU}(3)$ and $\Delta\mathcal{L}$ is the G_f-breaking piece which contains the meson potential $\Delta\mathcal{L}_M$ and the nonderivative couplings of mesons to baryons, $\Delta\mathcal{L}_B$. As was noted in the previous section, this chiral Lagrangian should be invariant under both G_f' and P'.

To obtain an explicit expression of $\mathcal{L}_{\rm chiral}$, let us define the transformation law of mesons and baryons under flavor symmetry G_f and parity P (Ref. 19). The unitary matrix

$$\Sigma=\exp(i\pi) \tag{3.2}$$

of the meson nonet $\pi=\pi^\alpha\lambda^\alpha/f+2\phi/\sqrt{6}f_0$ [λ^α=Gell-Mann matrix, $f=93$ MeV=pion decay constant, f_0=decay constant of the SU(3)-singlet meson ϕ] can be identified as a long-wavelength fluctuation of the chiral condensate of quark fields, viz., $\frac{1}{2}\langle\bar{q}_j(1-\gamma_f)q_i\rangle=v^3\Sigma_{ij}$, and thus it transforms under G_f and P as

$$\Sigma(x)\xrightarrow{G_f}e^{-2i\alpha}L\Sigma(x)R^\dagger,\quad \Sigma(x)\xrightarrow{P}\Sigma^\dagger(\tilde{x})\ . \tag{3.3}$$

The G_f transformation law of the baryon octet B is a little bit more complicated. It is represented by an SU(3) matrix A as

$$B\longrightarrow ABA^\dagger\ , \tag{3.4}$$

where A is defined through the transformation of ξ which is a square root of Σ, viz.,

$$\xi\longrightarrow e^{i\alpha}L\xi A^\dagger=e^{-i\alpha}A\xi R^\dagger\ , \tag{3.5}$$

where $\xi^2=\Sigma$. The parity transformation of B is the usual one for the Dirac spinor.

In the leading approximation of the derivative expansion, the G_f-invariant part $\mathcal{L}_0$ can be written as

$$\mathcal{L}_0=\tfrac{1}{2}(\partial_\mu\phi)^2+\tfrac{1}{4}f^2\,{\rm tr}\partial_\mu U\partial^\mu U^\dagger+i\,{\rm tr}\bar{B}\not{\partial}B+i\,{\rm tr}\bar{B}\gamma_\mu[V^\mu,B]+F\,{\rm tr}(\bar{B}\gamma_\mu\gamma_5[A^\mu,B])$$
$$+D\,{\rm tr}[\bar{B}\gamma_\mu\gamma_5\{A^\mu,B\}]+\frac{1}{\sqrt{6}}\frac{S}{f_0}\partial^\mu\phi\,{\rm tr}\bar{B}\gamma_\mu\gamma_5B-m_B\,{\rm tr}\bar{B}B\ , \tag{3.6}$$

where U is an SU(3) matrix of the meson octet, i.e., $U=\exp[i(\pi^\alpha\lambda^\alpha/f)]$, and

$$V^\mu=\tfrac{1}{2}(\eta\partial^\mu\eta^\dagger+\eta^\dagger\partial^\mu\eta),\quad A^\mu=\frac{i}{2}(\eta\partial^\mu\eta^\dagger-\eta^\dagger\partial^\mu\eta)\ , \tag{3.7}$$

for the SU(3) matrix η defined as $\eta^2=U$. The meson potential which is invariant under G_f' and P' and is the first-order result of $\Delta\mathcal{L}_{\rm QCD}$ of Eq. (2.11) can be obtained by using the vacuum-insertion approximation for the condensates of multiquark operators,[2] e.g.,

$$\langle(\bar{q}_Lq_R)(\bar{q}_Lq_R)\rangle\simeq\langle(\bar{q}_Lq_R)\rangle\langle(\bar{q}_Lq_R)\rangle=v^6\Sigma\Sigma\ .$$

The resulting meson potential is free of the ${\rm U}_A(1)$ problem and takes the form

$$\Delta\mathcal{L}_M=v^3(\,{\rm tr}\{[M+e^{-i\theta}\lambda_1(\det M^\dagger)(M^\dagger)^{-1}/m_s+e^{i\theta}\lambda_2(\det\Sigma)M]\Sigma^\dagger\}+e^{i\theta}\lambda_3m_s\det\Sigma+{\rm H.c.}\,)\ . \tag{3.8}$$

The nonderivative meson coupling to baryons, $\Delta\mathcal{L}_B$, which is also invariant under G_f' and P' and is first order in $\Delta\mathcal{L}_{\rm QCD}$, is given by

$$\Delta\mathcal{L}_B=-\Big[S_V[({\rm tr}\xi^\dagger\overline{M}_{SV}\xi^\dagger)+e^{i\theta}\lambda_Vm_s\det\Sigma+{\rm H.c.}\,]\,{\rm tr}\bar{B}B$$
$$+S_A[({\rm tr}\xi^\dagger\overline{M}_{SA}\xi^\dagger)+e^{i\theta}\lambda_Am_s\det\Sigma-{\rm H.c.}\,]\,{\rm tr}\bar{B}\gamma_5B+F_V\,{\rm tr}\bar{B}[(\xi^\dagger\overline{M}_{FV}\xi^\dagger+{\rm H.c.}),B]$$
$$+F_A\,{\rm tr}\bar{B}\gamma_5[(\xi^\dagger\overline{M}_{FA}\xi^\dagger-{\rm H.c.}),B]+D_V\,{\rm tr}\bar{B}\{(\xi^\dagger\overline{M}_{DV}\xi^\dagger+{\rm H.c.}),B\}$$
$$+D_A\,{\rm tr}\bar{B}\gamma_5\{(\xi^\dagger\overline{M}_{DA}\xi^\dagger-{\rm H.c.}),B\}\Big]\ , \tag{3.9}$$

where

$$M_I=\{M+e^{-i\theta}\lambda_1(\det M^\dagger)(M^\dagger)^{-1}/m_s\}+e^{i\theta}\lambda_I(\det\Sigma)M\quad(I=SV,SA,FV,FA,DV,DA)\ . \tag{3.10}$$

Here the terms with the coefficients λ_I's ($I = SV, \ldots, DA$) are the low-energy results of the four-quark operator $\Gamma^{(2)}$ (with the coefficient λ_2) in $\Delta\mathcal{L}_{\rm QCD}$ of Eq. (2.11) and terms with $\lambda_{A,V}$ come from the six-quark operator $\Gamma^{(3)}$ (with the coefficient λ_3). Then with this observation we are lead to expect $\lambda_I \simeq \lambda_2$ and $\lambda_{A,V} \simeq \lambda_3$.

Even though $\Delta\mathcal{L}_B$ given above is the most general one which is invariant under G_f' and P', and is first order in $\Delta\mathcal{L}_{\rm QCD}$, it contains too many (in principle calculable) parameters which spoil the predictability of our chiral Lagrangian. Therefore we improve the situation by assuming $\lambda_{SV} = \lambda_{SA} = \cdots = \lambda_{DA} = \lambda_2$. In fact any physical result from $\Delta\mathcal{L}_B$ of Eq. (3.9) does not depend much on this assumption as long as λ_I's are less than one. Then most of the strong-interaction parameters, i.e., $S_V, S_A, \ldots, D_A$, can be determined by the SU(3) breaking in the baryon masses and in the meson-baryon couplings (see Appendix B for details).

Flavor-symmetry breaking or the strong CP violation in the amplitudes involving the pseudoscalar mesons and the baryons can be studied with the chiral Lagrangian given above. Note that our chiral Lagrangian manifestly shows that there is no CP violation in the limit of vanishing instanton effects,[20] i.e., $\lambda_a \to 0$ ($a = 1,2,3$) which corresponds to the limit $N_c \to \infty$ (N_c denotes the number of color), or in the limit when one of the eigenvalues of M is zero. In these limits the potentially CP-violating phase in M or $e^{i\theta}$ can be rotated away by the G_f transformation of Σ and B.

For the study of CP-violating amplitudes it is convenient to remove the tadpole of the pseudoscalar-meson field π. From now on, let us put M as being real-diagonal and semipositive definite. One can always make M to be in such a form by the G_f transformation of fields and then the only CP-violating phase is $e^{i\theta}$. The vacuum expectation value of π can be determined by minimizing the meson potential $V_{\rm eff} = -\mathcal{L}_M$. As we will see in the next section the vanishingly small neutron electric dipole moment implies that $|\Delta\theta| \ll 1$ as long as $m_u/m_d \gg 10^{-10}$, where $\Delta\theta$ is defined as

$$\theta = n\pi + \Delta\theta \quad (n = 0,1) \ . \tag{3.11}$$

Then with the meson potential of Eq. (3.8) we find

$$\langle \pi \rangle = 2X = {\rm diag}(x_u, x_d, x_s) \ , \tag{3.12}$$

where

$$\begin{aligned} x_u &= -n\pi - \Delta\theta \left[\frac{m_d}{(-)^n m_u + m_d} \right] \left[1 + O\left[\frac{m_d}{m_s} \right] \right] , \\ x_d &= -\Delta\theta \left[\frac{(-)^n m_u}{(-)^n m_u + m_d} \right] \left[1 + O\left[\frac{m_d}{m_s} \right] \right] , \\ x_s &= -\Delta\theta \left[\frac{(-)^n m_u m_d}{[(-)^n m_u + m_d] m_s} \right] \\ &\quad \times \left[\frac{2\lambda_2^2 + \lambda_3(1+\lambda_2-\lambda_1)}{\lambda_2 + \lambda_3(1+\lambda_2)} + O\left[\frac{m_d}{m_s} \right] \right] . \end{aligned} \tag{3.13}$$

In deriving the above meson tadpole, the inequalities $m_d > m_u$, $m_s \gg m_d$ have been used. Another quantity whose explicit expression is useful in discussing the CP violation is

$$\begin{aligned} \theta + 2\,{\rm tr}X &= \Delta\theta \left[\frac{(-)^n m_u m_d}{[(-)^n m_u + m_d] m_s} \right] \\ &\quad \times \left[\frac{1-\lambda_1-\lambda_2(1+2\lambda_2)}{\lambda_2 + \lambda_3(1+\lambda_2)} + O\left[\frac{m_d}{m_s} \right] \right] . \end{aligned} \tag{3.14}$$

Note that in our convention, the intrinsic current-quark masses are all semipositive definite.

The above meson tadpole can be removed by the chiral rotation of meson field:

$$\Sigma \to e^{iX} \Sigma e^{iX} \ . \tag{3.15}$$

In fact, this chiral rotation of meson field and the corresponding transformation of baryons are the low-energy realization of Baluni's chiral rotation[13] of quark fields which has been considered to be suitable for the chiral perturbation of CP-violating amplitudes. However in Ref. 13, the instanton effects which violate not only $U_A(1)$ but also SU(3)$\times$SU(3) (i.e., terms with coefficient λ_1 or λ_2) has been neglected. Furthermore it is implicitly assumed that the $U_A(1)$ breaking by instantons is much stronger than the generic chiral-symmetry breaking due to the light-quark masses (including m_s). Therefore in our language, Baluni's analysis corresponds to the limit $\lambda_1, \lambda_2 \ll 1$ and $\lambda_3 \gg 1$. Note that in this limit our result of Eq. (3.13) reproduces Baluni's result $2X = -\theta M^{-1}/{\rm tr}M^{-1}$ (in the case of $n = 0$). However our numerical estimate of λ_a's (see Appendix A) indicates $\lambda_1 > \lambda_2 > \lambda_3$. Because of this discrepancy, our final formula for CP-violating amplitudes would be slightly different from the ones derived by using Baluni's result.

After the chiral rotation defined by Eq. (3.15) we finally obtain

$$\begin{aligned} \Delta\mathcal{L} = \Delta\mathcal{L}_M + \Delta_{\mathcal{L}_B} &= v^3[{\rm tr}(\overline{M}\Sigma^\dagger) + \lambda_3 m_s \det\Sigma e^{i(\theta+2\,{\rm tr}X)} + {\rm H.c.}] - S_V({\rm tr}\xi^\dagger \overline{M}\xi^\dagger + \lambda_V m_s \det\Sigma e^{i(\theta+2\,{\rm tr}X)} + {\rm H.c.})\,{\rm tr}\overline{B}B \\ &\quad - S_A({\rm tr}\xi^\dagger \overline{M}\xi^\dagger + \lambda_A m_s \det\Sigma e^{i(\theta+2\,{\rm tr}X)} - {\rm H.c})\,{\rm tr}\overline{B}\gamma_5 B \\ &\quad - F_V\,{\rm tr}\overline{B}[(\xi^\dagger \overline{M}\xi^\dagger + {\rm H.c.}), {\rm B}] - {\rm F}_A\,{\rm tr}\overline{B}\gamma_5[(\xi^\dagger \overline{M}\xi^\dagger - {\rm H.c.}), B] \\ &\quad - D_V\,{\rm tr}\overline{B}\{(\xi^\dagger \overline{M}\xi^\dagger + {\rm H.c.}), B\} - D_A\,{\rm tr}\overline{B}\gamma_5\{(\xi^\dagger \overline{M}\xi^\dagger - {\rm H.c.}), B\} \ , \end{aligned} \tag{3.16}$$

where

$$\overline{M}=(1+\lambda_2\,\det\Sigma e^{i(\theta+2\,\mathrm{tr}X)})M_x +\lambda_1[(\det M_x^\dagger)(M_x^\dagger)^{-1}/m_s]e^{-i(\theta+2\,\mathrm{tr}X)}\,,$$
$$M_x=e^{-iX}Me^{-iX}\,. \tag{3.17}$$

We emphasize here that in the above expression of chiral Lagrangian, the meson field π has vanishing tadpole and thus one can do the usual perturbation with the expansion $\Sigma=1+i\pi+O(\pi^2)$. In the next section we will use this chiral Lagrangian as a starting point of the phenomenological determination of the mass ratio m_u/m_d and of θ.

IV. APPLICATIONS: PHENOMENOLOGICAL DETERMINATION OF m_u/m_d, θ, AND THE AXION MASS

Flavor-symmetry-violating amplitudes (both *CP*-conserving and *CP*-violating ones) involving the pseudoscalar mesons and the baryons can be studied by using the chiral Lagrangian of Eq. (3.16). For this purpose, it is convenient to put $\det\Sigma=1$ and thus make the SU(3)-singlet component ϕ disappear. Then $\overline{M}$ which appears in our chiral Lagrangian can be written as

$$\overline{M}\simeq M_{\rm eff}+i\Delta\theta(1+\lambda_2-\lambda_1)\left[\frac{(-)^n m_u m_d}{(-)^n m_u+m_d}\right]1\,, \tag{4.1}$$

where

$$M_{\rm eff}=\mathrm{diag}(\overline{m}_u,\overline{m}_d,\overline{m}_s)\,,$$
$$\overline{m}_u=(-)^n(1+\lambda_2)m_u+\lambda_1 m_d\,,$$
$$\overline{m}_d=(1+\lambda_2)m_d+(-)^n\lambda_1 m_u\,, \tag{4.2}$$
$$\overline{m}_s=(1+\lambda_2)m_s+(-)^n\lambda_1 m_u m_d/m_s\,.$$

A remarkable thing here is that the flavor-symmetry breaking in the *CP*-conserving amplitudes is represented by *the effective current mass* $M_{\rm eff}$ while the *CP*-violating amplitudes which come from the nonvanishing $\mathrm{Im}\overline{M}$ are proportional to the determinant of *the intrinsic current-quark mass matrix M*. (In fact m_s of $\det M$ does not appear explicitly in $\mathrm{Im}\overline{M}$ because we have used the mass hierarchy $m_u<m_d\ll m_s$ in deriving the expression of $\mathrm{Im}\,\overline{M}$.)

The usual current-algebra analysis without taking into account the corrections which are the results of the combined effects of both the instantons and the intrinsic current-quark mass insertion [i.e., SU(3)$\times$SU(3)-breaking λ_1 and λ_2 terms in our Lagrangian] has produced the mass ratios $m_u/m_d\simeq0.56$ and $m_s/m_d\simeq20$ (Ref. 21). In our case of including instanton corrections in a consistent way, the same analysis would give rise to

$$\overline{m}_u/\overline{m}_d=[(-)^n(1+\lambda_2)m_u+\lambda_1 m_d]/[(1+\lambda_2)m_d+(-)^n\lambda_1 m_u]\simeq0.56\,,$$
$$\overline{m}_s/\overline{m}_d=[(1+\lambda_2)m_s+(-)^n\lambda_1 m_u m_d/m_s]/[(1+\lambda_2)m_d+(-)^n\lambda_1 m_u]\simeq20\,. \tag{4.3}$$

These phenomenological equations for the effective current mass $M_{\rm eff}$ give the following relation for the intrinsic quark mass M:

$$\frac{m_u}{m_d}=\frac{(-)^n(0.56+0.56\lambda_2-\lambda_1)}{1+\lambda_2-0.56\lambda_1}\,,$$
$$\frac{m_s}{m_d}=\frac{20(1+\lambda_2+\lambda_1)(1+\lambda_2-\lambda_1)}{(1+\lambda_2)(1+\lambda_2-0.56\lambda_1)}\,. \tag{4.4}$$

Note that if we simply neglect the instanton-mediated SU(3)$\times$SU(3) breaking and thus put $\lambda_1=\lambda_2=0$, then we have $(-1)^n m_u/m_d=0.56$ which implies that $n=1$ (i.e., $\theta=\pi$) or $m_u=0$ is phenomenologically not allowed.[12] (Note that in our convention, all intrinsic current masses are semipositive definite.) However as we can easily see, the instanton corrections parametrized by λ_1 and λ_2 can completely change this phenomenological conclusion, depending on the magnitudes of λ_1 and λ_2. For example, if $\lambda_1>0.56\,(1+\lambda_2)$, then the value of n which is determined by the low-energy phenomenology is $n=1$ (i.e., $\theta=\pi$), and $m_u=0$ can be possible for $\lambda_1=0.56\,(1+\lambda_2)$. In the case of $\lambda_1<0.56\,(1+\lambda_2)$ we are led to the conclusion $n=0$ (i.e., $\theta=0$) but still the numerical value of m_u/m_d does sensitively depend on the magnitudes of λ_1 and λ_2.

Any reliable estimate of λ_a's requires a quantitative understanding of the infrared QCD dynamics which is out of reach for us at present time. Note that the main contribution comes from the instantons of size around QCD scale. As an illustration we provide an estimate of λ_a's in Appendix A based on the semiclassical instanton-gas picture which is supplemented by the phenomenological constraints from $\eta-\eta'$ mass matrix and the numerical results are given in Table I. Because of the various assumptions on the infrared QCD dynamics adopted in our numerical estimation, the result should be understood as an order of magnitude estimate.

A remarkable result of our analysis is that, for a reasonable choice of m_s, $\Lambda_{\overline{\rm MS}}$ ($\overline{\rm MS}$ denotes the modified minimal subtraction scheme), and of the gluon condensate $\langle 0|(\alpha/\pi)FF|0\rangle$, instanton effects can be strong enough to completely change the previous conclusions on θ and m_u even in the semiclassical instanton-gas regime. For example with λ_a's in Table I, even though $m_s/m_d\simeq20$, all the following three distinct cases of m_u and θ: (i) $0.34\geq m_u/m_d\gg10^{-10}$, $\theta=0$; (ii) $m_u=0$; and (iii) $0.22\geq m_u/m_d\gg10^{-10}$, $\theta=\pi$, can be consistent with the phenomenological relations of Eq. (4.3). Although the upper limits on m_u/m_d are somewhat different from

TABLE I. Numerical values of dimensionless instanton parameters λ_1, λ_2, and λ_3 as functions of the infrared cutoff Λ.

Λ(MeV)	λ_1	λ_2	λ_3
620	0.959	0.381	0.154
630	0.813	0.329	0.136
640	0.693	0.284	0.119
650	0.592	0.245	0.104
660	0.508	0.212	0.090
670	0.437	0.183	0.078
680	0.378	0.158	0.068
690	0.327	0.137	0.058
695	0.305	0.127	0.054

the standard value 0.56, the difference is insignificant in view of the approximation involved in estimating λ_a's. In fact m_u/m_d can be arbitrarily small. In other words, it implies that we cannot draw any firm conclusion on the values of m_u and θ due to the ambiguities in the strength of the instanton corrections.

The importance of instanton corrections in the phenomenological determination of m_u and θ can be easily understood as the result of the explicit symmetry breaking due to the anomaly and instantons. As an illustrative example let us imagine the world in which $m_u=0$. (From the arguments given above, such a world can be a realistic one.) Then at the classical level, the axial-vector current $\bar{u}\gamma_\mu\gamma_5 u$ is exactly conserved with the corresponding axial U(1) symmetry. However at the quantum level, this U(1) symmetry is explicitly broken by the anomaly and instantons, and as a result, a finite but nonzero effective current mass $\overline{m}_u=\lambda_1 m_d$ is generated. The appearance of m_d in $\overline{m}_u$ is due to the d quark zero mode under the instanton background and the effects of heavier quarks, i.e., the s, c quarks, etc., reside in the parameter λ_1. In all CP-conserving amplitudes we see the effective current mass $\overline{m}_u$ and thus there is no clear sign of vanishing m_u as long as λ_1 is not much less than one. However this anomalous U(1) symmetry of the axial transformation of u quark, even though explicitly broken by instantons, guarantees that θ can be rotated away by the axial transformation of a u quark and thus there is no CP violation. Note that the CP violation due to the θ term is proportional to det$M$, the determinant of the intrinsic current-quark matrix.

Until now we have considered the CP-conserving part which gives information on m_u and n. The neutron electric dipole moment (NEDM) D_n provides the most stringent bound on the strength of CP violation due to the nonzero $\Delta\theta$. As in other CP-violationg amplitudes, it is proportional to $\Delta\theta m_u m_d(\Delta\theta \ll 1)$. As was noticed, with instanton corrections, any value of m_u/m_d smaller than $\overline{m}_u/\overline{m}_d=0.56$ (if we adopt the range of λ_a's given in Table I, m_u/m_d is smaller than 0.34) is consistent with the low-energy phenomenology. An immediate consequence is the relaxation of the bound on $|\Delta\theta|$ because the NEDM gives a bound on $m_u m_d|\Delta\theta|$, not on $|\Delta\theta|$ itself.

In the literature, the contributions to D_n from the low-lying baryon resonances,[13] the nucleon-pion intermediate states,[12] the CP-odd nucleon mass,[14] and the mixing between the scalar and pseudoscalar mesons[15] have been considered. Among these contributions, the scalar-pseudoscalar mixing has been claimed to give the largest contribution. However the phenomenological model of Ref. 15 within which the NEDM via the scalar-pseudoscalar mixing has been computed does not satisfy the proper anomalous Ward identity, e.g., CP-violating amplitudes do not vanish in the limit $m_u=0$ even though they do vanish in the different limit $m_\pi=0$. Apart from this ambiguous contribution of the CP-odd scalar-pseudoscalar mixing, it has been argued that the CP-odd nucleon mass gives a dominant contribution to the NEDM (Ref. 14). Here we evaluate, as another application of our chiral Lagrangian of Eq. (3.16), the NEDM induced from the CP-odd nucleon mass and pion-nucleon couplings.

Our method of evaluating the NEDM from the pion-nucleon intermediate state is slightly different from the one used in Ref. 12 in the sense that we use the axial-vector coupling for the CP-conserving pion-nucleon coupling, while the pseudoscalar coupling has been used in Ref. 12. However we find that both prescriptions of the CP-conserving pion-nucleon coupling give rise to the essentially same result. Note that the equivalence of the axial-vector coupling and the pseudoscalar coupling is nontrivial in our case due to the off-shell propagation of a nucleon in the diagram [see Fig. 1(b)] responsible for the NEDM (Ref. 22). For the NEDM from the CP-odd nucleon mass, our method of estimating the size of the CP-odd nucleon mass within the chiral Lagrangian is completely different from the one used in Ref. 14 and thus would be an independent check of the result of Ref. 14.

The CP-conserving part of the effective Lagrangian describing the interactions of the nucleon doublet $N=\binom{p}{n}$, the isotriplet pions $\pi_a(a=1,2,3)$, and the electromagnetic fields includes

$$\frac{g_A}{2f}\partial_\mu\pi^a\overline{N}\gamma^\mu\gamma_5\tau^a N+\frac{e}{4m_N}F^{\mu\nu}\overline{N}\sigma_{\mu\nu}ZN\ , \tag{4.5}$$

where $F^{\mu\nu}$ is the electromagnetic field strength, m_N being the nucleon mass, $Z=\text{diag}(\mu_p,\mu_n)=\text{diag}(1.79,-1.91)$ denotes the anomalous magnetic moments, and $g_A=D+F=1.25$ is the pion-nucleon axial-vector coupling constant. The CP-violating interactions relevant to the

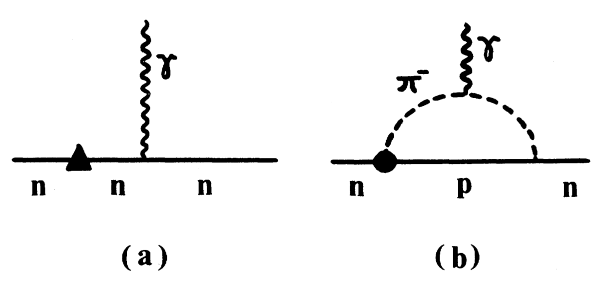

FIG. 1. Diagrams generating NEDM. The dark triangle is the CP-odd neutron mass and the dark blob denotes the CP-violating πNN coupling.

NEDM can be obtained from the chiral Lagrangian of Eq. (3.16) and then we find

$$\delta L_{CP} \supset -\Delta\theta \left[\frac{(-)^n m_u m_d}{(-)^n m_u + m_d} \right] \left[h_1 \bar{N} i\gamma_5 N + h_2 \frac{\pi^a}{f} \bar{N}\tau^a N \right] , \tag{4.6}$$

where

$$h_1 = 2(1+\lambda_2-\lambda_1)[(3+\epsilon)S_A + 2D_A] ,$$
$$h_2 = 2(1+\lambda_2-\lambda_1)(F_V+D_V) . \tag{4.7}$$

Here $\epsilon = \lambda_A(1-\lambda_1-\lambda_2-2\lambda_2^2)/(\lambda_2+\lambda_3+\lambda_2\lambda_3)(1+\lambda_2-\lambda_1)$ and $|\epsilon| < 0.5$ for the values of λ_a's in Table I (with the assumption $\lambda_A \simeq \lambda_3$), and thus this will be neglected in the following. Our chiral Lagrangian also gives the phenomenological relations (see Appendix B for the details and notations)

$$F_V + D_V \simeq \frac{m_\Xi - m_\Sigma}{2(1+\lambda_2)m_s} \simeq \frac{65 \text{ MeV}}{(1+\lambda_2)m_s} ,$$
$$3S_A + 2D_A \simeq \frac{f}{2(1+\lambda_2)m_s} \left[3G_{N\Sigma K} + \frac{3\sqrt{3}}{2} G_{NN\eta} - \sqrt{3}\, G_{N\Lambda K} + 3.2 \right]$$
$$\simeq \frac{1}{(1+\lambda_2)m_s} \times (-150\text{–}1300) \text{ MeV} , \tag{4.8}$$

which give rise to

$$h_1 \simeq \frac{1+\lambda_2-\lambda_1}{(1+\lambda_2)m_s} \times (-300\text{–}2600) \text{ MeV} ,$$
$$h_2 \simeq \frac{1+\lambda_2-\lambda_1}{(1+\lambda_2)m_s} \times 130 \text{ MeV} . \tag{4.9}$$

The *CP*-violationg interactions of Eq. (4.6), combined with the *CP*-even interactions given in Eq. (4.5), give rise to the NEDM through the diagrams of Fig. 1. Following Refs. 12 and 14 we obtain

$$|D_n^a| = \left| \Delta\theta \left[\frac{m_u m_d}{(-)^n m_u + m_d} \right] \frac{e h_1 \mu_n}{2m_N^2} \right| ,$$
$$|D_n^b| = \left| \Delta\theta \left[\frac{m_u m_d}{(-)^n m_u + m_d} \right] \frac{e h_2 g_A \ln(m_N/m_\pi)}{4\pi^2 f^2} \right| , \tag{4.10}$$

where D_n^a and D_n^b denote the NEDM from diagrams (1a) and (1b), respectively.

The Goldberger-Treiman relation $g_{\pi NN} f = g_A m_N$ assures that D_n^b computed via the axial-vector pion-nucleon coupling is the same as the result of Ref. 12 which was obtained by using the pseudoscalar pion-nucleon coupling. In Ref. 14, the size of *CP*-odd nucleon mass was estimated based on the assumption of η'-pole dominance. Then it has been observed that, with a reasonable choice of the Yukawa coupling of η' to the nucleon, h_1 can be as large as 20. Even though the range of h_1 given in Eq. (4.9) contains a region which is in broad agreement with the result of Ref. 14, our method of estimating h_1 via the chiral Lagrangian indicates that h_1 can be much smaller than the value obtained in Ref. 14. We will not discuss the ambiguities in both approaches of estimating h_1 any more because here we are more interested in the intrinsic current mass dependence of the NEDM. Then we observe that both D_n^a and D_n^b are proportional to

$$\Delta\theta \left[\frac{m_u}{(-)^n m_u + m_d} \right] \left[\frac{m_d}{m_s} \right] \left[\frac{1+\lambda_2-\lambda_1}{1+\lambda_2} \right]$$
$$= \frac{\Delta\theta}{20} \left[\frac{m_u}{(-)^n m_u + m_d} \right] \left[\frac{1+\lambda_2-0.56\lambda_1}{1+\lambda_2+\lambda_1} \right] . \tag{4.11}$$

Now the experimental bound of D_n (Ref. 23), viz., $|D_n| \le 10^{-25}$ ecm, can be used to obtain the bound on $\Delta\theta m_u/[(-)^n m_u + m_d]$. As the most stringent bound possible in view of the values of h_1 and h_2 determined as in Eq. (4.9) we find

$$\left| \Delta\theta \left[\frac{m_u}{(-)^n m_u + m_d} \right] \left[\frac{1+\lambda_2-0.56\lambda_1}{1+\lambda_2+\lambda_1} \right] \right| \le 3\times 10^{-11} . \tag{4.12}$$

If we consider the limit of negligible instanton corrections, i.e., $\lambda_1 = \lambda_2 = 0$ and thus $n = 0$, $m_u/m_d = 0.56$, the experimental bound on the NEDM would imply $|\Delta\theta| \le 10^{-10}$. However as we have discussed already, m_u/m_d can be arbitrarily small due to the instanton corrections. Therefore the extremely small NEDM does not necessarily require $|\Delta\theta|$ to be smaller than 10^{-10}. For example, in the case of $m_u = 0$ which can be perfectly consistent with the low-energy phenomenology in view of our previous discussions, $\Delta\theta$ (or $\theta = n\pi + \Delta\theta$) can take any value and its effect is not observable. An immediate consequence of this relaxation of the phenomenological bound on $\Delta\theta$, is that many of the models of spontaneous *CP* violation[17] which predict $|\Delta\theta| \gg 10^{-10}$ and thus have been considered to be not viable phenomenologically can now be acceptable.

Finally let us briefly consider the implication of the arbitrarily smaller value of m_u/m_d to the axion models.[24] Of course if $m_u/m_d \le 10^{-10}$, axion is no longer motivated because in this case we do not have any fine-tuning problem for θ. However in principle we can contemplate the situation in which $m_u/m_d \gg 10^{-10}$ and thus still $\Delta\theta$ is required to be very small. As is well known, in the axion models, $\theta = n\pi + \Delta\theta$ is dynamically relaxed down to zero.[4] By solving the equation of motion for the axion field a which is defined as $\theta = a/f_a$ where f_a denotes the axion decay constant, we always arrive at the *CP*-invariant vacuum $\langle\theta\rangle = \langle a/f_a\rangle = 0$.

The axion potential $V_{\text{eff}}[a]$ can be obtained from the meson potential of Eq. (3.8) by replacing the meson field Σ by its axion-dependent vacuum expectation value $\langle\Sigma\rangle = \exp(i\langle\pi\rangle) = \exp(2iX)$ [see Eq. (3.12)]. Then we find

$$V_{\rm eff}[a]=-\left\{\lambda_0\Lambda\,{\rm det}Me^{i\theta}+v^3\left[(1+\lambda_2e^{i(\theta+2\,{\rm tr}X)})\,{\rm tr}(Me^{-2iX})+\lambda_1\left[\frac{{\rm det}M}{m_s}\right]{\rm tr}(M^{-1}e^{-i(\theta+2X)})\right.\right.$$
$$\left.\left.+\lambda_3 m_s e^{i(\theta+2\,{\rm tr}X)}\right]+{\rm H.c.}\right\}, \tag{4.13}$$

where

$$2X=-\left[\frac{a}{f_a}\right]{\rm diag}\left[\frac{m_d}{m_u+m_d},\frac{m_u}{m_u+m_d},\frac{m_u m_d}{(m_u+m_d)m_s}\left[\frac{2\lambda_2^2+\lambda_3(1+\lambda_2-\lambda_1)}{\lambda_2+\lambda_3(1+\lambda_2)}\right]\right] \tag{4.14}$$

and

$$\lambda_0=\frac{1}{\Lambda}\int d\mu\,D(\mu)Z^{-3}(\mu/\Lambda)\ . \tag{4.15}$$

Note that the first term of the axion potential, with the coefficient λ_0, has not been included in the meson potential because it has nothing to do with the meson dynamics. The axion mass from this axion potential is

$$m_a^2=\frac{2}{f_a^2}\left[\frac{m_u m_d}{m_u+m_d}\right]\left[\lambda_0\Lambda m_s(m_u+m_d)+v^3\left[w^2(1+\lambda_2)+(3-2w)\lambda_1+z^2(\lambda_2+\lambda_3)\frac{m_u m_d}{(m_u+m_d)m_s}\right]\right], \tag{4.16}$$

where

$$\begin{aligned} w&=[\lambda_2+\lambda_3(1+\lambda_2)]\Big/\left[\lambda_2+\lambda_3(1+\lambda_2)+\frac{m_u m_d}{(m_u+m_d)m_s}(1-\lambda_1-\lambda_2+\lambda_3+\lambda_3\lambda_2-\lambda_3\lambda_1)\right],\\ z&=[1-\lambda_1-\lambda_2(1+2\lambda_2)]\Big/\left[\lambda_2+\lambda_3(1+\lambda_2)+\frac{m_u m_d}{(m_u+m_d)m_s}(1-\lambda_1-\lambda_2+\lambda_3+\lambda_3\lambda_2-\lambda_3\lambda_1)\right]. \end{aligned} \tag{4.17}$$

In fact one can obtain the temperature-dependent axion mass by considering the temperature dependence of λ_a's ($a=0,1,2,3$) which is determined by the temperature-dependent instanton density $D(\mu{:}T)$ (Ref. 25) and also of $v(T)$ which is the order parameter of the spontaneous chiral-symmetry breaking in QCD at the temperature T. For example, at $T=0$, with $\lambda_2+\lambda_3(1+\lambda_2)\gg m_u m_d/(m_u+m_d)m_s$, we get the usual formula for the axion mass[4]

$$\begin{aligned} m_a^2&=\left[\frac{2v^3}{f_a^2}\right]\left[\frac{m_u m_d}{m_u+m_d}\right](1+\lambda_1+\lambda_2)\\ &=\frac{m_u m_d}{(m_u+m_d)^2}\left[\frac{f m_\pi}{f_a}\right]^2, \end{aligned} \tag{4.18}$$

where the formula for the pion mass

$$\begin{aligned} m_\pi^2&=2v^3(\overline{m}_u+\overline{m}_d)/f^2\\ &=2v^3(m_u+m_d)(1+\lambda_1+\lambda_2)/f^2 \end{aligned}$$

is used for the second line of the above equation. At the high-temperature limit of the chiral-symmetry restoration, viz., $v(T)=0$, the axion mass is simply given as $m_a^2=2\Lambda m_u m_d m_s \lambda_0(T)/f_a^2$. In any case, the axion mass squared is proportional to m_u/m_d for all temperature range. Therefore if m_u/m_d is much smaller than the usually quoted value 0.56, the cosmological upper bound on f_a (Ref. 18), i.e., $f_a\leq 4\times 10^{12}$ GeV, obtained from the consideration of cosmological axion energy density can be relaxed up to the value much larger than 10^{12} GeV. Note that in the limit of $m_u=0$, even though the axion is not strongly motivated in this case, the axion is massless and thus there does not exist any cosmological upper bound on f_a. This relaxation of the cosmological upper bound on f_a due to the very small value of m_u/m_d may be applied for some superstring models which predict the existence of the axion with the decay constant $f_a\simeq 10^{16}$ GeV (Ref. 26). A detailed discussion on the cosmological axion energy density for an arbitrary value of m_u/m_d will be discussed elsewhere.[27]

ACKNOWLEDGMENTS

We would like to thank J. Choe for helping with the numerical estimate of instanton parameters. One of the authors (K.C.) thanks L. F. Li, A. Santamaria, and L. Wolfenstein for useful conversations. This work has been supported in part by DOE Contract No. DE-AC02-76ER03066 and by the National Science Foundation.

APPENDIX A: NUMERICAL ESTIMATE OF THE STRENGTH OF INSTANTON EFFECTS

In this appendix we provide an example of the numerical estimate of the renormalization-group-invariant parameters

$$\lambda_1=\frac{4\pi^2}{3}m_s(\Lambda)\int\frac{d\mu}{\mu^2}D(\mu)Z^{-1}(\mu/\Lambda)\ ,$$
$$\lambda_2=\left[\frac{4\pi^2}{3}\right]^2 v^3(\Lambda)\int\frac{d\mu}{\mu^4}D(\mu)Z(\mu/\Lambda)\ , \qquad \text{(A1)}$$
$$\lambda_3=\left[\frac{4\pi^2}{3}\right]^3\frac{v^6(\Lambda)}{m_s(\Lambda)}\int\frac{d\mu}{\mu^6}D(\mu)Z^3(\mu/\Lambda)\ ,$$

where $v^3(\Lambda)\delta_{ij}=-\langle 0|\overline{q}_{iL}q_{jR}|0\rangle$ for the quark fields renormalized at the scale Λ. For the evaluation we use the instanton calculus, based on the semiclassical instanton-gas picture with the background gluon condensate, whose physical meaning has been elegantly explained in Ref. 2. Because of the lack of quantitative understanding of the infrared behavior of QCD, our estimate is not reliable enough so that it should be understood as an order-of-magnitude estimate. However we believe it provides a useful guideline for the numerical values of λ_a's.

The instanton density $D(\mu)$ which includes the effect of gluon condensate was obtained by Shifman, Vainshtein, and Zakharov as

$$D(\mu)=3.64\times10^{-3}\left[\frac{2\pi}{\alpha(\mu)}\right]^6\exp\left[-\frac{2\pi}{\alpha(\mu)}\left[1-\frac{\pi^3}{16\alpha(\mu)}\frac{\left\langle\frac{\alpha}{\pi}FF\right\rangle}{\mu^4}\right]\right]\ , \qquad \text{(A2)}$$

where $\alpha(\mu)$ is the running QCD fine-structure constant and $\langle(\alpha/\pi)FF\rangle$ denotes the gluon condensate. A naive use of the above expression over all scales gives rise to an enormous infrared divergence. Therefore it is important that we restrict ourselves to instantons of relatively small size. Here we introduce an infrared cutoff for the size of relevant instantons and put our renormalization point Λ as this infrared cutoff scale. It is assumed that at energies below Λ, the semiclassical picture of instantons is no more valid and the large size ($\gtrsim 1/\Lambda$) instantons are effectively destroyed. In Ref. 2 it was argued that the instanton density of Eq. (A2) is valid (at least qualitatively) up to scales above 500 MeV and thus we expect Λ to be not far from 500 MeV.

The numerical values of λ_a's strongly depend on our choice of the infrared cutoff Λ. In Table I we present the values of λ_a's for 620 MeV $\le\Lambda\le$ 695 MeV which is believed to be the most probable range of Λ in view of our later discussion. For numerical analysis we use

$$\alpha(\mu)=[2\pi/9\ln(\mu/\Lambda_{\overline{\rm MS}})]\ ,$$
$$Z(\mu/\Lambda)=[\alpha(\Lambda)/\alpha(\mu)]^{4/9}\ ,$$
$$\left\langle 0\left|\frac{\alpha}{\pi}FF\right|0\right\rangle=(330\ \text{MeV})^4\ ,$$
$$\langle 0|\overline{q}q|0\rangle=2v^3(\Lambda)=f_\pi^2m_\pi^2/(\overline{m}_u+\overline{m}_d)\ , \qquad \text{(A3)}$$
$$\overline{m}_u=(\overline{m}_s/36)=(1+\lambda_2)m_s/36\ ,$$
$$\overline{m}_d=(\overline{m}_s/20)=(1+\lambda_2)m_s/20\ ,$$
$$m_s(1\ \text{GeV})=200\ \text{MeV},\quad \Lambda_{\overline{\rm MS}}=200\ \text{MeV}\ .$$

The value of the gluon condensate is taken from Ref. 28 which estimates it from the charmonium decay.

The reasoning which leads to the range of infrared cutoff scale 620 MeV $\le\Lambda\le$ 695 MeV is as follows. The allowed range of λ_a's can be constrained by the η-η' mass matrix (upon neglecting the mixing with π_0)

$$\begin{bmatrix} A & \Gamma \\ \Gamma & B\end{bmatrix}\ . \qquad \text{(A4)}$$

Then from the chiral Lagrangian of Eq. (3.16) we find

$$\frac{\Gamma^2}{AB}=\frac{(1-2\lambda_2)^2}{(1+\lambda_2)(1+4\lambda_2+9\lambda_3)}\ . \qquad \text{(A5)}$$

This quantity can also be expressed in terms of m_η, m'_η, and the η-η' mixing angle Θ as

$$\frac{\Gamma^2}{AB}=\frac{(m_{\eta'}^2-m_\eta^2)^2\sin^2\Theta\cos^2\Theta}{(m_\eta^2\cos^2\Theta+m_{\eta'}^2\sin^2\Theta)(m_{\eta'}^2\cos^2\Theta+m_\eta^2\sin^2\Theta)}\ . \qquad \text{(A6)}$$

Phenomenologically the most favored value of Θ (Ref. 29) is $-20°$. Then by inserting $m_\eta=549$ MeV, $m_{\eta'}=958$ MeV, $\Theta=-20°$ to Eq. (A6), we obtain $\Gamma^2/AB=0.124$.

However the value of Γ^2/AB obtained from Eq. (A6) is very sensitive to Θ whose value is also sensitive to the potential corrections which are ignored in our approximation. For example, within our approximation, we have the Gell-Mann–Okubo relation

$$A=\tfrac{4}{3}m_K^2-\tfrac{1}{3}m_\pi^2=m_\eta^2\cos^2\Theta+m_{\eta'}^2\sin^2\Theta \qquad \text{(A7)}$$

which gives rise to $\Theta=-10°$. This discrepancy between the value of Θ determined from phenomenology and the one obtained by using the Gell-Mann–Okubo relation indicates that much of the observed mixing angle ($=-20°$) is the result of the higher-order corrections ignored in our approximation. Note, for example, that the potential chiral loop corrections to the off-diagonal mass term Γ, which is expected to be of the order of 10% of A, can change Θ by the amount of the order of 10°. Therefore based on the above observation, here we will effectively take into account the whole ambiguity of our approximation by considering the following range of Θ:

$$-35°\le\Theta\le-5° \qquad \text{(A8)}$$

while fixing m_η and $m_{\eta'}$ as their experimental values. Then for the above range of Θ, Eq. (A6) gives

$$0.01 \leq \frac{\Gamma^2}{AB} \leq 0.23 \ . \tag{A9}$$

This range of Γ^2/AB can be realized via Eq. (A5) for the range of the infrared cutoff scale

$$620 \text{ MeV} \leq \Lambda \leq 695 \text{ MeV} \ , \tag{A10}$$

where the corresponding values of λ_a's appear in Table I.

For this range of Λ, let us check the validity of our instanton calculus at $\Lambda \geq 620$ MeV, which can be tested by evaluating the quantity

$$D_{\text{eff}}(\Lambda) = \left[\prod_{i=u,d,s} \left[\frac{m_i}{\Lambda} - \frac{2\pi^2}{3} \frac{\langle 0|q_i q_i|0\rangle}{\Lambda^3} \right] \right] D(\Lambda) \ . \tag{A11}$$

This quantity corresponds to the number of instantons of size Λ^{-1} within the space-time volume Λ^{-4}. If $D_{\text{eff}}(\Lambda) \simeq 1$, then the instantons are closely packed and thus the semiclassical instanton-gas picture would be no longer valid. In our case we have $D_{\text{eff}}(\Lambda \geq 620$ MeV) $\leq D_{\text{eff}}(\Lambda = 620$ MeV) $\simeq 3\times 10^{-2}$ which implies that the semiclassical instanton-gas picture is a reasonable approximation.

Finally we consider the dependence of our results on the values of the input parameter m_s, $\Lambda_{\overline{\text{MS}}}$, and $\langle(\alpha/\pi)FF\rangle$. Our choice $m_s = 200$ MeV, $\Lambda_{\overline{\text{MS}}} = 200$ MeV, $\langle(\alpha/\pi)FF\rangle$ (330 MeV)^4 is a reasonable one in view of the various estimations of these input parameters and is taken to provide a suggestive example showing that the instanton corrections can be strong enough to completely change the usual conclusion on m_u/m_d and θ which has been obtained without taking into account instanton effects. However the magnitude of these input parameters are uncertain to a somehow large extent. Then one can easily see that for larger (smaller) input parameters, the resulting instanton corrections become stronger (weaker). Recent studies[30] of the QCD sum rule indicate that the gluon condensate is larger than (330 MeV)^4 which is the value used here. Therefore for the values m_s and Λ_{QCD} less than 200 MeV (as long as they are still reasonable), we can obtain a similar conclusion on the strength of instanton corrections.

APPENDIX B. STRONG-INTERACTION PARAMETERS IN CHIRAL LAGRANGIAN

Here we briefly discuss the phenomenological determination of various strong-interaction parameters which appear in the chiral Lagrangian. As is well known, F and D can be determined by the nuclear β decay and the semileptonic hyperon decay which yield[31]

$$F = 0.45, \quad D = 0.8 \ . \tag{B1}$$

Also recent European Muon Collaboration (EMC) data can be used to determine S as[32]

$$S = 0.15 \pm 0.3 \ . \tag{B2}$$

For the SU(3)-breaking parameters, e.g., F_V, D_V, etc., we are interested in the parameters which are relevant to our discussion on the electric dipole moment of the neutron. With the chiral Lagrangian of Eqs. (3.6) and (3.16) we find

$$\begin{aligned}
&2(D_V + F_V)(1+\lambda_2)m_s = m_\Xi - m_\Sigma \ , \\
&2(D_V - F_V)(1+\lambda_2)m_s = m_N - m_\Sigma \ , \\
&\sqrt{3} f G_{NN\eta} = (3F - D)m_N + 4m_s(1+\lambda_2)(S_A + D_A - F_A) \ , \\
&2\sqrt{3} f G_{N\Lambda K} = (D + 3F)(m_N + m_\Lambda) \\
&\qquad\qquad - 2m_s(1+\lambda_2)(D_A + 3F_A) \ , \\
&2fG_{N\Sigma K} = (D - F)(m_n + m_\Sigma) - 2m_s(1+\lambda_2)(D_A - F_A) \ ,
\end{aligned} \tag{B3}$$

where m_N, m_Σ, and m_Ξ denote the corresponding baryon masses; $f = 93$ MeV is the pion decay constant; and $G_{NN\eta}$, $G_{N\Lambda K}$, and $G_{N\Sigma K}$ denote the coupling constants of Yukawa interactions $\eta \bar{N} i\gamma_5 N$, $K_0 \bar{n} i\gamma_5 \Lambda$, and $K_0 \bar{n} i\gamma_5 \Sigma$, respectively. These Yukawa coupling constants are defined at the Born approximation and thus include the contributions from both the SU(3)-invariant meson-baryon axial-vector couplings and the SU(3)-breaking meson-baryon pseudoscalar couplings in the chiral Lagrangian of Eq. (3.16).

The phase shift analysis of the baryon-baryon scattering amplitudes within the Born approximation gives rise to the following values of the Yukawa coupling constants:[33]

$$G_{NN\eta} = 7-10, \quad G_{N\Lambda K} = 13-16, \quad G_{N\Sigma K} = 1-7 \ . \tag{B4}$$

Then, together with the baryon masses, the above phenomenological information allows us to express the strong-interaction parameters in terms of $(1+\lambda_2)m_s$. For example, we obtain the following expression for the strong-interaction parameters which appear in the CP-violating interactions of Eq. (4.6):

$$\begin{aligned}
&F_V + D_V \simeq (m_\Xi - m_\Sigma)/2(1+\lambda_2)m_s \simeq \frac{65 \text{ MeV}}{(1+\lambda_2)m_s} \ , \\
&3S_A + 2D_A \simeq \frac{f}{2(1+\lambda_2)m_s} \left[3G_{N\Sigma K} + \frac{3\sqrt{3}}{2} G_{NN\eta} - \sqrt{3}\, G_{N\Lambda K} + 3.2 \right] \simeq \frac{(-150-1300) \text{ MeV}}{(1+\lambda_2)m_s} \ .
\end{aligned} \tag{B5}$$

[1]G.'t Hooft, Phys. Rev. Lett. **37**, 8 (1976); Phys. Rev. D **14**, 3432 (1976); Phys. Rep. **142**, 357 (1986).

[2]M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, Nucl. Phys. **B163**, 46 (1980); **B165**, 45 (1980).

[3]R. Jackiw and C. Rebbi, Phys. Rev. Lett. **37**, 172 (1976); C. G. Callan, Jr., R. F. Dashen, and D. J. Gross, Phys. Lett. **63B**, 334 (1976).

[4]For recent review on the strong-*CP* problem, see J. E. Kim, Phys. Rep. **150**, 1 (1987); H. Y. Cheng, *ibid.* **158**, 1 (1988); R. D. Peccei, DESY Report No. DESY 88-109, 1988 (unpublished).

[5]K. Choi, C. W. Kim, and W. K. Sze, Phys. Rev. Lett. **61**, 794 (1988); A similar idea has been considered in the unpublished work by H. Georgi and I. N. McArthur, Report No. HUTP-81/A011, 1981 (unpublished).

[6]D. G. Caldi, Phys. Rev. Lett **39**, 121 (1977); C. G. Callan, R. Dashen, and D. J. Gross, Phys. Rev. D **17**, 2717 (1978); **19**, 1826 (1979); R. D. Carlitz and D. B. Creamer, Ann. Phys. (N.Y.) **116**, 429 (1979); E. V. Shuryak, Nucl. Phys. **B302**, 599 (1988).

[7]E. V. Shuryak, Nucl. Phys. **B214**, 237 (1983).

[8]E. V. Shuryak and J. L. Rosner, Phys. Lett. B **218**, 72 (1989).

[9]J. E. Kim, University of Pennsylvania Report No. UPR-0106T, 1979 (unpublished); K. Konish and S. Ranfone, Phys. Lett. B **172**, 417 (1986).

[10]D. B. Kaplan and A. V. Manohar, Phys. Rev. Lett. **56**, 2004 (1986).

[11]Recent study of wormhole effects in quantum gravity implies that $\theta=\pi$ is the favored one in the so-called wormhole calculus. For details, see K. Choi and R. Holman, Phys. Rev. Lett. **62**, 2575 (1989); J. Preskill, S. P. Trivedi, and M. B. Wise, Caltech Report No. CALT-89-1539, 1989 (unpublished).

[12]R. Crewther, P. DiVecchia, G. Veneziano, and E. Witten, Phys. Lett. **89B**, 123 (1979).

[13]V. Baluni, Phys. Rev. D **19**, 2227 (1979).

[14]P. Cea and G. Nardulli, Phys. Lett **144B**, 115 (1984).

[15]E. P. Shabalian, Yad. Fiz. **36**, 981 (1982) [Sov. J. Nucl. Phys. **36**, 575 (1982)].

[16]M. A. Morgan and G. A. Miller, Phys. Lett. B **179**, 379 (1986); H. Schnitzer, Phys. Lett. **139B**, 217 (1984); M. M. Musakhanov and Z. Z. Israilov, *ibid.* **137B** 419 (1984); K. Kanaya and M. Kobayashi, Prog. Theor. Phys. **66**, 2173 (1981).

[17]A. Nelson, Phys. Lett. **136B**, 387 (1984); S. M. Barr, Phys. Rev. Lett. **53**, 329 (1984); H. Georgi, A Nelson, and M. Shin, Phys. Lett. **150B**, 306 (1985).

[18]J. Preskill, M. B. Wise, and F. Wilczek, Phys. Lett. **120B**, 127 (1983); L. F. Abbott and P. Sikivie, *ibid.* **120B**, 133 (1983); M. Dine and W. Fischler, *ibid.* **120B**, 137 (1983).

[19]For an elegant review on chiral Lagrangians, see H. Georgi, *Weak Interactions and Modern Particle Theory* (Benjamin-Cummins, New York, 1985).

[20]C. Bernard, Phys. Rev. D **19**, 3013 (1979).

[21]S. Weinberg, Trans. N. Y. Acad. Sci. **38**, 185 (1977); J. Gasser and H. Leutwyler, Phys. Rep. **87**, 77 (1982).

[22]For example, for the axion-nucleon bremsstrahlung process which has been used to obtain a lower bound on the axion decay constant from the SN 1987A, a discrepancy between the use of the pseudoscalar coupling and the use of the axial-vector coupling does appear. For the details, see G. Raffelt and D. Seckel, Phys. Rev. Lett. **60**, 1793 (1988); R. P. Brinkmann and M. S. Turner, Phys. Rev. D **38**, 2338 (1988); K. Choi, K. Kang, and J. E. Kim, Phys. Rev. Lett. **62**, 849 (1989).

[23]N. F. Ramsay, Phys. Rep. **43**, 409 (1977); I. S. Altarev *et al.*, Pis'ma Zh. Eksp. Teor. Fiz. **44**, 360 (1986) [JETP Lett. **44**, 460 (1986)].

[24]R. D. Peccei and H. R. Quinn, Phys. Rev. Lett. **38**, 1440 (1977); S. Weinberg, *ibid.* **40**, 223 (1978); F. Wilczek, *ibid.* **40**, 279 (1978); J. E. Kim, *ibid.* **43**, 103 (1979); M. Dine, W. Fischler, and M. Srednicki, Phys. Lett. **104B**, 199 (1981).

[25]D. J. Gross, R. D. Pisarski, and L. G. Yaffe, Rev. Mod. Phys. **53**, 43 (1981).

[26]K. Choi and J. E. Kim, Phys. Lett. **154B**, 393 (1985).

[27]K. Choi, C. W. Kim, and W. P. Lam, in preparation.

[28]M. Shifman, A. Vainshtein, and V. I. Zakharov, Nucl. Phys. **B147**, 385 (1979); **B147**, 448 (1979).

[29]J. F. Donoghue, B. R. Holstein, and Y-C. R. Lin, Phys. Rev. Lett. **55**, 2766 (1985); F. J. Gilman and R. Kauffman, Phys. Rev. D **36**, 2761 (1987).

[30]R. A. Bertlmann, Report No. UWThPh-1988-26, 1988 (unpublished), and references cited therein.

[31]R. Schrock and L.-L. Wang, Phys. Rev. Lett. **41**, 1692 (1978).

[32]S. Brodsky, J. Ellis, and M. Karliner, Phys. Lett. B **206**, 309 (1988); D. Kaplan and A. Manohar, Report No. HUTP-88-8021, 1988 (unpublished).

[33]M. M. Nagels *et al.*, Nucl. Phys. **B147**, 189 (1979).