Physical renormalization schemes and grand unification

Michael Binger* and Stanley J. Brodsky†
Stanford Linear Accelerator Center, Stanford University, Stanford, California 94309, USA
(Received 6 November 2003; published 19 May 2004)

In a physical renormalization scheme, gauge couplings are defined directly in terms of physical observables. Such effective charges are analytic functions of physical scales, and thus mass thresholds are treated with their correct analytic dependence. In particular, particles will contribute to physical predictions even at energies below their threshold. This is in contrast with unphysical renormalization schemes such as $\overline{\mathrm{MS}}$ where mass thresholds are treated as step functions. In this paper we analyze supersymmetric grand unification in the context of physical renormalization schemes and find a number of qualitative differences and improvements in precision over conventional approaches. The effective charge formalism presented here provides a template for calculating all mass threshold effects for any given grand unified theory. These new threshold corrections may be important in making the measured values of the gauge couplings consistent with unification.

DOI: 10.1103/PhysRevD.69.095007 PACS number(s): 12.10.Dm

I. INTRODUCTION

Precision measurements of the gauge couplings and their possible unification provide some of the few windows to the Planck scale. It is thus important to have a firm grasp of the theoretical ambiguities involved. This paper attempts to address some of these ambiguities.

In a physical renormalization scheme, gauge couplings are defined directly in terms of physical observables. Such effective charges are analytic functions of physical scales, and thus the thresholds associated with heavy particles are treated with their correct analytic dependence. This is in contrast to unphysical renormalization schemes such as the modified minimal subtraction ($\overline{\mathrm{MS}}$) scheme where mass thresholds are treated as step functions. In this paper we will analyze supersymmetric grand unification in the context of physical renormalization schemes with the goal of systematizing the effects of light and heavy mass thresholds and improving the precision of tests of unification compared with conventional approaches.

In Sec. II, we motivate physical renormalization schemes with a simple example and then present the notation and results used throughout the paper. In Sec. III, we look more carefully at the problem of decoupling heavy particles and the errors induced by unphysical schemes. In Sec. IV, we discuss the canonical self-energy-like effective charges for the minimal supersymmetric standard model (MSSM). These effective couplings run smoothly over spacelike momenta, have nonanalytic behavior only at the expected physical thresholds for timelike momenta, and more directly measure the strengths of the forces than the charges of unphysical schemes. The extraction of effective charges from low energy data is considered. We identify an important modification of the electromagnetic coupling $\alpha_{\mathrm{QED}}(M_Z)$ due to the proper inclusion of virtual $W^{\pm}$ loops, thus resulting in a 4σ change in its numerical value. Similar modifications are found for the weak mixing angle. As seen in Sec. V, these effective charges provide a more natural and physical framework for examining gauge coupling unification. In Sec. V A, we demonstrate the invalidity of neglecting heavy threshold corrections in analyzing grand unified models. The more rigorous treatment of light thresholds in physical schemes gives rise to new corrections, but these are numerically small for most sparticle spectra. The treatment of heavy thresholds with various unification boundary conditions is discussed in Sec. V B. In the simplest scenario, we find that the gauge couplings should unify at asymptotically large energies and the only heavy threshold corrections are logarithms of heavy mass ratios, corrections which can be obtained in unphysical schemes. An effective unification scale, defined in Sec. V C as the scale where quantum gravity corrections produce nonnegligible splittings between the gauge couplings, is found to be roughly $10^{17}-10^{18}$ GeV, depending on the specific grand unified theory (GUT) model used. Section V D considers more general unification boundary conditions with finite unification scale. The resulting heavy threshold corrections are given in Eq. (44). This result combined with the results of Sec. V B may be used to determine the experimental consistency of any given GUT model. This is the main result of the paper. An appendix discusses the details of constructing the effective charges. Throughout our analysis, we will find several attractive theoretical features of the supersymmetric regulator, dimensional reduction, which makes it the preferred regulator for physical effective charge schemes, even without supersymmetry.

There have been several previous works on threshold effects in grand unification. In the first such study [1], which appeared just after the discovery of grand unification, Ross uses form factors to define beta functions which are valid over all energy scales, including near-mass thresholds. The coupling constants run smoothly over all momenta, and nontrivial threshold corrections are found for grand unification. Despite this early significant work, most subsequent work on GUTs have ignored these threshold effects, perhaps due to the complexity of the Ross approach.

An exception from the late 1980s is the work of Kennedy and Lynn [2], who defined electroweak effective charges similar to the pinch technique charges used in this paper.

*Email address: mwbinger@stanford.edu
†Email address: sjbth@slac.stanford.edu

In several papers [3] by Kreuzer, Kummer, and Rebhan, the authors compared the Vilkovisky-DeWitt effective actioin (VDEA), the mass-shell momentum subtraction scheme (MMOM), and Weinberg's effective gauge theory (EGT). They wrote down explicit formula for the running charges which include analytic threshold behavior for all particles. In calculating predictions from grand unification, they assume asymptotic unification at energies much larger than heavy particles, so that the only threshold corrections from heavy particles come from finite constants which are independent of energy scale or masses. We find similar results in Sec. V B. Furthermore, we include the possibility of a finite unification scale in Sec. V D, which leads to more complicated corrections.

In [4,5], the authors include the effects of light supersymmetric scalar and fermion thresholds, although heavy thresholds and gauge bosons virtual effects are not treated. In [6], the authors include both light and heavy threshold corrections, although the treatment of gauge bosons is not adequate. In Refs. [7,8], the authors come to several conclusions similar to ours. However, their definition leads to gauge-parameter-dependent effective couplings.

II. PHYSICAL RENORMALIZATION SCHEMES AND EFFECTIVE CHARGES

In order to motivate the reanalysis of supersymmetric unification given in this paper, we will first discuss some general properties of renormalization schemes in the presence of massive fields and determine a criterion for consistent *physical renormalization schemes*. These criteria will not be satisfied by the schemes conventionally used in unification (and most perturbative calculations), $\overline{\text{MS}}$ and modified dimensional reduction ($\overline{\text{DR}}$), which have persistent threshold and matching errors. Heuristically, these errors can be understood by noting that such schemes implicitly integrate out all masses heavier than the physical energy scale until they are crossed, and then they are "clicked" on with a step function. Of course, integrating out heavy fields is only valid for energies well below their masses. This procedure is problematic since it does not correctly incorporate the finite probability that the uncertainty principle gives for a particle to be pair produced below threshold. Effective charge [9] schemes, derived from physical observables, naturally avoid such errors and are formally consistent.

A. Simple example

For the purpose of elucidating the benefits of physical renormalization schemes, we will give a simple toy example using QED with three fermions e, μ, and τ. Consider the amplitude for the process $e^-\mu^-\to e^-\mu^-$. This can be written as a dressed skeleton expansion—i.e., the dressed tree level graph plus the dressed box diagram plus the dressed double box, etc. The tree level diagram, dressed to all orders in perturbation theory, is equal to the tree level diagram with one modification: the QED coupling $\alpha=e^2/4\pi$ is replaced by the Gell-Man–Low–Dyson effective charge

$$\alpha(Q^2)=\frac{\alpha}{1+\Pi_{\gamma\gamma}(Q^2)-\Pi_{\gamma\gamma}(0)}. \tag{1}$$

Hence, from measurements of the cross section, one can measure the effective charge at two different scales $\alpha(Q_h^2)$ and $\alpha(Q_l^2)$. Suppose the value of the electron charge is not known, and we are trying to test the predictions of QED. The way to proceed is to use one measurement—say, at the low scale Q_l—as an input to determine e. Now the prediction at the high scale Q_h is well defined and represents a test of the theory. More directly, we could just write $\alpha(Q_h^2)$ in terms of $\alpha(Q_l^2)$, leading to the same prediction. Since the cross section $\sigma_{e^-\mu^-\to e^-\mu^-}(Q^2)$ is proportional to $[\alpha(Q^2)]^2$, we are clearly relating one observable to another. The procedure just outlined is simply an on-shell renormalization scheme if $Q_l=0$. More generally, we will refer to such a scheme as an effective charge scheme, since we are writing a given observable, here just $\sigma_{e^-\mu^-\to e^-\mu^-}(Q_h^2)$ [or $\alpha(Q_h^2)$], in terms of an effective charge $\alpha(Q_l^2)$ defined from a measurement of the cross section at the scale Q_l. One could equally well write any observable in terms of this effective charge. Note that this approach to renormalization works for arbitrary scales, even if the low scale lies below some threshold—say, $Q_l<m_\tau$, while $Q_h>m_\tau$. Decoupling and the smooth "turning on" of the τ are manifest.

Now we will compare with the results obtained by using the conventional implementation of $\overline{\text{MS}}$, which is as follows. First, the cross section is calculated at Q_l using the rules of $\overline{\text{MS}}$, which allows only the electrons and muons to propagate in loops, since $Q_l<m_\tau$. Comparing the observed cross section to this result will fix the value of the $\overline{\text{MS}}$ coupling for two flavors $\hat{\alpha}_2(Q_l)$. To predict the result of the same experiment at scale $Q_h>m_\tau$, we need to evolve $\hat{\alpha}_2$ to the tau threshold using the two-flavor beta function, match with a three-flavor coupling $\hat{\alpha}_3$ through the relation $\hat{\alpha}_2(m_\tau)=\hat{\alpha}_3(m_\tau)$, and then evolve $\hat{\alpha}_3(m_\tau)$ to Q_h using the three-flavor beta function. We will now have a prediction for $\sigma_{e^-\mu^-\to e^-\mu^-}(Q_h^2)\propto[\alpha(Q_h^2)]^2$. One might expect, from the general principle of renormalization group (RG) invariance of physical predictions that this result should be the same as the prediction derived using the physical effective charge scheme above. However, there is a discrepancy arising from the incorrect treatment of the threshold effects in $\overline{\text{MS}}$. A detailed discussion of this problem will be given in Sec. III. In any case, the result can be obtained by straightforwardly applying the procedure outlined above. One finds that the ratio of the cross section derived using $\overline{\text{MS}}$ with the cross section derived using effective charges, to first order in perturbation theory, is given by

$$\frac{\hat{\sigma}(Q_h^2)}{\sigma(Q_h^2)}=1+2\,\frac{\alpha(Q_l)}{3\pi}[L_\tau(Q_l/m_\tau)-5/3], \tag{2}$$

where L_τ is a logarithmlike function (the high energy limit is a logarithm) given by

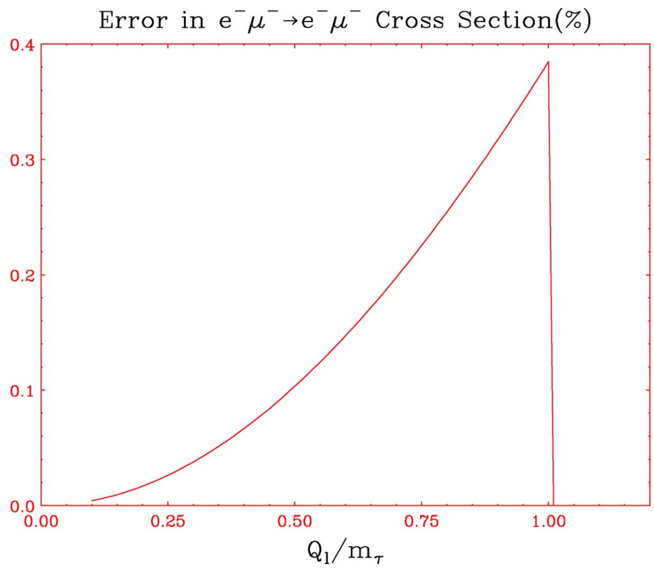

FIG. 1. Error in the $\overline{\mathrm{MS}}$-based prediction for the scattering cross section, $100\% \times [\hat{\sigma}(Q_h^2)/\sigma(Q_h^2)-1]$, plotted against the reference subtraction scale Q_l for the choice $\alpha(Q_l)\approx 0.1$.

$$L_\tau(Q/m)=\int_0^1 dx\, 6x(1-x)\log\left(1+\frac{Q^2}{m^2}x(1-x)\right)+5/3$$
$$=[\beta\tanh^{-1}(\beta^{-1})-1](3-\beta^2)+2, \tag{3}$$

where $\beta=\sqrt{1+4m^2/Q^2}$. It satisfies the property $L_\tau(0)=5/3$, so that there is no discrepancy when the low reference scale Q_l is much lower than the tau mass threshold. This reflects the important, but often overlooked, fact that unphysical schemes, such as $\overline{\mathrm{MS}}$, are formally consistent only in desert regions where particle masses can be neglected. The error is plotted in Fig. 1. Notice that in this example there is an error only for $Q_l<m_\tau$. However, in the more general case of multiple-flavor thresholds, there will be errors from both high and low scales. Similar discrepancies will be found in our analysis of grand unification.

B. General properties of effective charges and physical renormalization schemes

Effective charges [9] may be defined for any perturbatively calculable observable

$$\mathcal{O}(Q)\equiv A^{\mathcal{O}}+a_1^{\mathcal{O}}\alpha^{\mathcal{O}}(Q) \tag{4}$$

by absorbing all of the radiative corrections into the effective charge $\alpha^{\mathcal{O}}$. To one-loop order using dimensional regularization (DREG) or dimensional reduction (DRED), it is straightforward to show that any unrenormalized effective charge may be parametrized as[1]

TABLE I. Characteristic self-energy constants.

	Scalars	Fermions	Massive gauge bosons
η_p	8/3	5/3	40/21

$$\alpha^{\mathcal{O}}(Q)=\alpha_0-\frac{\alpha_0^2}{4\pi}\sum_p \beta_p\Bigg(L_p(Q/m_p)-\eta_p^{\mathcal{O}}(Q/m_p)$$
$$-C_{UV}+\log\frac{m_p^2}{\mu^2}\Bigg)+\cdots, \tag{5}$$

where the sum is over all particles p in the fundamental theory which contribute to the running of the effective charge. In the QED example above, the sum proceeds over e, μ, τ and $L_e=L_\mu=L_\tau$, and the function $\eta_p^{\mathcal{O}}(Q/m_p)=5/3$ is a constant for the simple observable $\mathcal{O}=\sqrt{\sigma_{e^-\mu^-\to e^-\mu^-}}$. In Eq. (5), $C_{UV}=1/\epsilon-\gamma E+\log 4\pi$ is the divergence and associated constants, μ is the regularization scale, α_0 is the bare gauge coupling, and β_p is the contribution of each particle to the one-loop beta function coefficient. The $L_p(Q/m_p)$ are logarithmiclike functions which are characteristic of the spin of each particle and are given exactly in Eq. (23). They may be approximated to within a few percent[2] by

$$L_p(Q/m)\approx\log\left(e^{\eta_p}+\frac{Q^2}{m_p^2}\right) \tag{6}$$

and have the limits

$$L_p(Q/m)\overset{Q\gg m}{\approx}\log\frac{Q^2}{m^2},\quad L_p(Q/m)\overset{m\gg Q}{\approx}\eta_p, \tag{7}$$

where the constants η_p have values given in Table I. We will see that these constants are of central importance in physical renormalization schemes. These logarithmlike functions characterize the self-energy-like effect of each particle, including the finite spread of the wave functions near thresholds due to the uncertainty principle, and may be calculated in several different ways, as will be discussed in Sec. IV and the Appendix. Figure 2 shows the L_p functions for spacelike momenta.

The functions $\eta_p^{\mathcal{O}}(Q/m_p)$ are characteristic of each observable, with a nontrivial functional form indicating deviations from self-energy-like behavior. For a general observable $\mathcal{O}$, the function $\eta_p^{\mathcal{O}}(Q/m_p)$ is nontrivial. We will show in Sec. IV that the constants η_p correspond to a particularly simple and canonical observable, called the pinch-technique (PT) self-energy.

[1]This follows from considering the high energy limit and requiring renormalizability. Note also that our parametrization can be easily extended to effective charges which have particles with different masses running together in the loops, and the results are similar. In any case, we will not have use for such charges in this paper.

[2]To be precise, the approximations reproduce the exact functions L_0, $L_{1/2}$, and L_1 (the subscripts refer to the spin of the massive field) with maximum error of 3.5%, 0.8%, and 2.2%, respectively, over the entire range of Q.

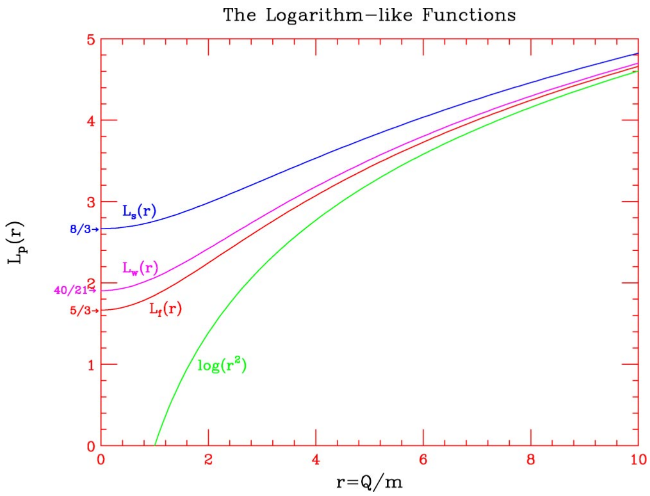

FIG. 2. The logarithmlike functions for massive particles of spin 0, 1/2, and 1 are denoted by L_s, L_f, and L_W, respectively.

The effective coupling renormalized in the most general scheme R is

$$\alpha^{\mathcal{O}}(Q)=\alpha_R(Q_0)-\frac{[\alpha_R(Q_0)]^2}{4\pi}\sum_p \beta_p[L_p(Q/m_p)-L_p(Q_0/m_p)-\eta_p^{\mathcal{O}}(Q/m_p)+\eta_p^R(Q_0/m_p)], \tag{8}$$

where the functions $\eta_p^R(Q_0/m_p)$ contain all of the information about the scheme. Here R can be any mathematical scheme for defining the couplings. In the case of $\overline{\text{MS}}$, we have $\eta_p^{\overline{\text{MS}}}(Q_0/m_p)=L_p(Q_0/m_p)-\log(Q_0^2/m_p^2)$ so that only-logarithms of the renormalization scale Q_0 are subtracted.[3]

It is straightforward to relate observables to each other:

$$\alpha^{\mathcal{O}_1}(Q_1)=\alpha^{\mathcal{O}_2}(Q_2)-\frac{[\alpha^{\mathcal{O}_2}(Q_2)]^2}{4\pi}\sum_p \beta_p[L_p(Q_1/m_p)-L_p(Q_2/m_p)-\eta_p^{\mathcal{O}_1}(Q_1/m_p)+\eta_p^{\mathcal{O}_2}(Q_2/m_p)]. \tag{9}$$

This satisfies the transitivity property of the physical renormalization group. As before, the sum over p runs over all particles in the fundamental theory which contribute to the effective charges.

For consistency, very massive particles must decouple properly and must not contribute to physical predictions. Taking the $m_p\to\infty$ limit in Eqs. (8) and (9) yields a fundamental requirement of renormalization schemes and observables:

$$\eta_p^R(0)=\eta_p^{\mathcal{O}_1}(0)=\eta_p^{\mathcal{O}_2}(0)\overset{\text{DRED}}{=}\eta_p. \tag{10}$$

This consistency requirement holds for all schemes R, observables $\mathcal{O}_1$, $\mathcal{O}_2$, and for each massive particle p. These are universal constants for each spin and, when DRED is used, are equal to the η_p given in Table I above, as can be verified through explicit calculations. If, for example, DREG was used instead of DRED, the last equality of Eq. (10) would hold only for fermions and scalars, but for spin-1 fields there is an additional constant.[4] Renormalization schemes that satisfy (the first two equalities of) Eq. (10) will henceforth be referred to as physical renormalization schemes, and those that do not will be called unphysical renormalization schemes, for reasons that will become clear.

The above discussion implies a unique decoupling limit $(Q/m\to 0)$ for observables. It is interesting that there is also a restriction on the high energy behavior $(Q/m\to\infty)$, which holds only for supersymmetric theories and takes the form of a sum rule. It is given by

$$\frac{\Sigma_{p\in S}\beta_p(G)\,\eta_p^{\mathcal{O}}(\infty)}{\Sigma_{p\in S}\beta_p(G)}=K^{\mathcal{O}}, \tag{11}$$

where $K^{\mathcal{O}}$ is a constant that depends only on the observable, not on the gauge group G or the supermultiplet S. The $\eta_p^{\mathcal{O}}(Q/m_p)$ are calculated using DRED; otherwise, the sum rule is true only for differences $\eta_p^{\mathcal{O}_1}(\infty)-\eta_p^{\mathcal{O}_2}(\infty)$ between observables. Further, the result holds for any number of supersymmetries, which may be broken or unbroken at low energies. This can be proved inductively given the result for $N=1$. It is easy to check using Table I above and the corresponding result for massless gauge bosons given below Eq. (19) that $K^{\text{PT}}=2$. The sum rule just expresses the fact that there is no resolution within a supermultiplet at high energies and is motivated from conformal invariance and physical renormalization scheme invariance. Such a sum rule may provide a powerful link between the contributions of various spin fields to any observable, particularly if a multiloop or nonperturbative generalization was found.

III. UNPHYSICAL RENORMALIZATION SCHEMES AND THE PROBLEM OF DECOUPLING

Notice that the physical renormalization scheme requirement Eq. (10) is not met by $\overline{\text{MS}}$, $\overline{\text{DR}}$, their massive extensions, or similar schemes. It is well known that $\overline{\text{MS}}$ by itself does not constitute a complete scheme; rather, one must truncate the sum over p to include only particles with masses less than the scale of the problem. For each region between thresholds a different scheme is implemented and one must translate between schemes when crossing thresholds. Hence $\overline{\text{MS}}$ is really a set of schemes related to each other. We will call such a set an artificial decoupling scheme (ADS) and now discuss the most general case at one loop. This will give

[3]The term in parentheses in Eq. (8) becomes $L_p(Q/m_p)-\log(Q_0^2/m_p^2)-\eta_p^{\mathcal{O}}(Q/m_p)$. Note that in most calculations the first term is taken to be a logarithm and mass corrections are systematically added, in order to approximate the full threshold dependence of $L_p(Q/m_p)$. However, the $\log(Q_0^2/m_p^2)$ term does not have the correct threshold dependence, as we will be discussing.

[4]This is explained below Eq. (19) and below Eq. (A11).

us an idea of the discrepancies one may expect in ADSs when compared to the physical renormalization scheme approach.

Let $S=\{m_1,m_2,\dots\}$ be the spectrum of massive particles of the fundamental theory ordered from lightest to heaviest, let S_0 be the set of massless particles, and let $S^N=S_0\oplus\{m_1,m_2,\dots,m_N\}$ be some subset up to a given mass scale. For any given renormalization scheme R, let the Nth phase of R, denoted R^N, be the scheme used to renormalize observables at energy scales Q such that

$$m_1,\dots,m_N<Q<m_{N+1}. \tag{12}$$

To use R^N, one simply renormalizes the contributions from particles $p\in S^N$ in the usual way dictated by scheme R and then entirely neglects the contributions from all $p\in S-S^N$. This is a formal statement of the usual implementation of ADSs using step functions.

For $m_1,\dots,m_N<Q,Q'<m_{N+1}$, the gauge coupling of the R^N scheme flows by

$$\alpha_{R^N}(Q)=\alpha_{R^N}(Q')-\frac{[\alpha_{R^N}(Q')]^2}{4\pi}\sum_{p\in S^N}\beta_p[L_p(Q/m_p)-L_p(Q'/m_p)-\eta_p^R(Q/m_p)+\eta_p^R(Q'/m_p)], \tag{13}$$

and the most general matching condition between schemes R^{N-1} and R^N takes the form

$$\alpha_{R^{N-1}}(\mu_N)=\alpha_{R^N}(\mu_N)+\frac{\alpha^2_{R^N}(\mu_N)}{4\pi}\beta_N A_N(\mu_N/m_N), \tag{14}$$

where $A_p(\mu_p/m_p)$ is arbitrary now, but will be specified below by minimizing errors and may depend only on the matching scale μ_p for each threshold m_p, for reasons discussed below.

For $Q_h>m_{N+n}$ and $m_1,\dots,m_N<Q_l<m_{N+1}$ (the h and l stand for heavy and light scales, respectively) one may relate observables by flowing through n thresholds using the above formulas to obtain

$$\alpha^{\mathcal{O}_1}(Q_h)=\alpha^{\mathcal{O}_2}(Q_l)-\frac{[\alpha^{\mathcal{O}_2}(Q_l)]^2}{4\pi}\Bigg[\sum_{p\in S^{N+n}}\beta_p[L_p(Q_h/m_p)-L_p(Q_l/m_p)-\eta_p^{\mathcal{O}_1}(Q_h/m_p)+\eta_p^{\mathcal{O}_2}(Q_l/m_p)]+\sum_{p\in S^{N+n}-S^N}\beta_p[A_p(\mu_p/m_p)-L_p(\mu_p/m_p)+\eta_p^R(\mu_p/m_p)+L_p(Q_l/m_p)-\eta_p^{\mathcal{O}_2}(Q_l/m_p)]\Bigg]. \tag{15}$$

Now let us compare this to the relation, Eq. (9), obtained in the previous section for the tracking of two observables. For the case of a high scale desert region ($Q_h\ll m_{N+n+1}$), the first sum reduces to Eq. (9). However, when $Q_h\lesssim m_{N+n+1}$, there are errors of one-loop order which are proportional to $L_p(Q_h/m_p)-\eta_p^{\mathcal{O}_1}(Q_h/m_p)$ and occur for each neglected threshold[5] $p\in S-S^{N+n}$ such that $Q_h\lesssim m_p$. These errors are naturally remedied in physical renormalization schemes, leading to heavy threshold corrections which will be of importance when grand unification is discussed later. There are also analogous light threshold corrections. The second sum in Eq. (15) contains extra terms which arise from the artificial decoupling and matching conditions at each threshold. These terms are also generally of order of the one-loop corrections and must cancel if the ADS is to be consistent (in the sense of giving reliable physical predictions in relations between two observables). In general, these terms do not cancel, since the Q_l-dependent terms cannot be cancelled by the choice of A_p, which depends only on the ratio μ_p/m_p. Suitably choosing A_p [see Eq. (16) below] leaves a term $L_p(Q_l/m_p)-\eta_p^{\mathcal{O}_2}(Q_l/m_p)$ for all $p\in S^{N+n}-S^N$. Hence we see that the high scale and low scale threshold corrections have exactly the same form; indeed, they have the same origin—namely, the necessity of using an ADS—which arises from improper decoupling.

In the case of a low scale desert region $Q_l\ll m_p\forall p\in S^{N+n}-S^N$, one finds that $L_p(Q_l/m_p)-\eta_p^{\mathcal{O}_2}(Q_l/m_p)\to 0$ [by Eqs. (7), (10)], and the light threshold errors are eliminated through the choice

$$A_p(\mu_p/m_p)=L_p(\mu_p/m_p)-\eta_p^R(\mu_p/m_p). \tag{16}$$

In this case, notice that once the A_p are chosen suitably, the matching scale μ_p exactly cancels and there is no need to fix its value. However, for $\overline{\text{MS}}$, this choice is equivalent to using $A_p=0$ and $\mu_p=m_p$, which is the matching scale typically used at one loop. One may object that even in nondesert regions the known anomalous matching threshold errors could be systematically subtracted off for each physical process considered [equivalently allowing $A_p=A_p(Q_l)$]. This is untenable, as it is the same as using a different coupling for each process, thus losing the remnants of universality left by ADSs (i.e., universality in each desert region).

We have identified two potential problems in artificial decoupling schemes, which arise solely from the failure of the decoupling requirement given in Eq. (10), regardless of whether or not the scheme R has analytic threshold dependence.[6] The low scale errors come from the matching conditions and are exhibited in the last two terms of Eq. (15). These are significant only when $Q_l\lesssim m_{N+1}$. The high scale

[5] Similar errors occur if $Q_l\lesssim m_{N+n+1}$.

[6] Proper analytic threshold dependence may be defined by $\eta_p^R(Q/m)$ going to a constant for both small and large Q/m. Consider the analytic extension of $\overline{\text{MS}}$ into the region of mass thresholds, which we call massive $\overline{\text{MS}}$ $\text{M}\overline{\text{MS}}$ (similar to that in [10]). This is defined by $\eta_p^{\text{M}\overline{\text{MS}}}(Q_0/m_p)=0$ so that the full logarithmiclike functions L_p are subtracted and trivially the conditions for smooth threshold dependence are satisfied. Nonetheless, $\text{M}\overline{\text{MS}}$ has matching errors, which result from the failure of Eq. (10) and the subsequent need to construct an ADS from $\text{M}\overline{\text{MS}}$.

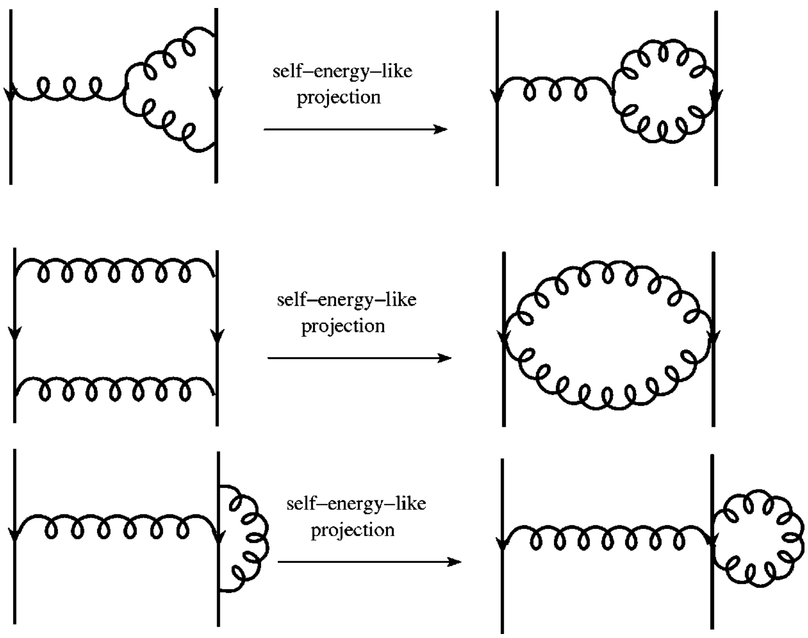

FIG. 3. Pinch technique for QCD at one loop. The unique gluonic self-energy-like projection of the vertex and box graphs yields terms which must be added to the conventional self-energy to get the PT effective charge.

errors occur when one is calculating an observable at an energy $Q_h \lesssim m_{N+n+1}$ that is slightly less than masses that should contribute, but are cut off in an ADS. These two errors will give rise to light and heavy threshold corrections in unification, as discussed in Sec. IV.

In practice, both types of errors can often be eliminated through a "threshold shifting" procedure. This involves modifying the definition of R^N by replacing Eq. (12) with $m_1,\ldots,m_N < aQ < m_{N+1}$ and making similar subsequent replacements and by choosing $a>1$ to be large enough so that the desired thresholds that are slightly above Q_l or Q_h are "moved" below aQ_l, so that no matching need occur for those thresholds, since they are already implicitly included in the couplings. The limit of this procedure as $a\to\infty$ leads to a formally consistent scheme where no matching or artificial decoupling is used, but as a result of the failure of decoupling, it requires inclusion of contributions from every particle in the (unknown) fundamental theory. This is the exact situation that caused us to introduce an ADS in the first place, since we did not want unknown and arbitrarily massive fields contributing to every physical observable (written in terms of the ADS scheme charge). It is true that such unknown contributions cancel in relations between observables, but the utility of the intermediate ADS scheme is lost since it's coupling is ill defined.[7] In many calculations in unphysical schemes such as $\overline{\text{MS}}$, the "threshold shifting" approach may be used to yield physical predictions which are arbitrarily accurate by choosing a sufficiently large a. However, the usefulness of this procedure depends on the details of the mass spectrum. There is no universal algorithm that applies to any field theory. The complicated nature of such artificial fixes to the decoupling problem are reflections of the unphysical nature of the schemes and couplings. See [11] for another approach to fixing $\overline{\text{MS}}$.

Thus we have shown that the $\overline{\text{MS}}$ and $\overline{\text{DR}}$ schemes suffer errors unless one is restricted to observables at energies Q which lie far between mass thresholds. In addition, complicated matching conditions must be applied when crossing thresholds to maintain consistency for such desert scenarios. In principle, these schemes are only valid for theories where all particles have zero or infinite mass or if one knows the full field content of the underlying physical theory.

IV. CANONICAL PHYSICAL EFFECTIVE CHARGES FOR THE MSSM

The difficulties associated with unphysical schemes are circumvented in physical renormalization schemes (PRSs) based upon effective charges. So far, we have given consistency conditions which are satisfied PRSs, but explicit examples have not been given. This is the topic we now take up.

For any observable $\mathcal{O}$, we define an effective charge scheme $R^{\mathcal{O}}$, by

$$\eta_p^{R^{\mathcal{O}}}(Q/m_p) = \eta_p^{\mathcal{O}}(Q/m_p), \tag{17}$$

which, after using Eq. (8), is equivalent to

$$\alpha^{\mathcal{O}}(Q) = \alpha_{R^{\mathcal{O}}}(Q), \tag{18}$$

thus motivating the terminology "effective charge." Here $R^{\mathcal{O}}$ is the physical subset of all possible mathematical schemes. The canonical example for using an effective charge as a scheme is furnished in QED by the Gell-Mann–Low–Dyson charge, which can be measured directly from scattering experiments. The extension of this concept to non-Abelian gauge theories is nontrivial [12], due to the self-interactions of the gauge bosons which make the usual self-energy gauge

[7]We might as well always write observables in terms of other observables; this is precisely the philosophy of effective-charge-inspired physical renormalization schemes.

dependent. However, systematically implementing the Ward identities of the theory allows one to project out the unique self-energy of each *physical* particle, resulting in a self-energy that is gauge independent, may be resummed to define an effective charge, and may be related via the optical theorem to appropriate cuts of differential cross sections. The algorithm for performing the calculation at the diagrammatic level is called the pinch technique [13–16]. The procedure is illustrated in Fig. 3 for the case of massless gauge theory, where momentum factors from internal gauge boson lines or vertices combine with gamma matrices to cancel internal fermion propagators, yielding a gluon self-energy-like graph. This is then added to the usual self-energy to yield the full PT self-energy.

The PT procedure is unambiguous and is merely an application of the Ward identities of the theory, which becomes more transparent in a dispersive derivation from physical cross sections $\sigma(q\overline{q}\to gg)$ (see [16] for such a construction for the electroweak sector). The generalization of the pinch technique to higher loops has recently been investigated [17–22]. In the work of Binosi and Papavassiliou [20–22], the authors prove the consistency of the pinch technique to all orders in perturbation theory, making clear how to define the QCD and electroweak effective charges at higher orders.

The PT charge, labeled by $\tilde{\alpha}$, written in terms of the bare coupling α_0 may be calculated for arbitrary gauge theory, broken or unbroken, to be[8]

$$\tilde{\alpha}(Q)=\alpha_0-\frac{\alpha_0^2}{4\pi}\sum_p \beta_p[L_p(Q/m_p)-\eta_p-C_{UV}+l_m]+\cdots, \tag{19}$$

where $\eta_p^{\text{PT-DRED}}(Q/m_p)=\eta_p$ are the constants given in Table I for massive fields and $\eta_g=64/33$ for massless spin-1 fields.[9] The fact that these $\eta_p^{\mathcal{O}}(Q/m)$ functions are constants is what makes the PT observable the most simple and natural choice for defining an effective charge scheme. More general physical effective charge schemes [see Eqs. (8), (9), (10)] have more complicated running due to the $\eta_p^{\mathcal{O}}(Q/m_p)$ terms. The calculation of $\tilde{\alpha}(Q)$ has been performed using dimensional reduction, rather than dimensional regularization. We will let $\overline{\text{PT}}$ stand for the renormalization scheme associated with the PT observable regularized using DRED. The results in DREG for $s=0,1/2$ fields are the same, but for spin-1 fields are given by the replacement $\eta_W\to\eta_W+2/21$, where the additional constant is due to the so-called ϵ ghosts. Although the constant terms will cancel in relations between physical observables, DRED is the more natural choice, even for nonsupersymmetric theories. This is because decoupling is manifest (including for massive spin-1 fields) individually for each observable without relying on cancellation terms arising from relations to another observable. This is just a statement of the fact that $\eta_p^{\text{PT-DRED}}=L_p(0)\forall p$. In contrast, with DREG one has $\eta_W^{\text{PT-DREG}}\neq L_W(0)$. Also, the supersymmetric sum rule in Eq. (11) becomes manifest in DRED.

Using the above results, it is straightforward to write down the effective charges for the standard model through

$$\widetilde{\alpha_1}(Q^2)=\frac{5}{3}\,\frac{\tilde{\alpha}(Q^2)}{1-\tilde{s}^2(Q^2)},$$

$$\widetilde{\alpha_2}(Q^2)=\frac{\tilde{\alpha}(Q^2)}{\tilde{s}^2(Q^2)},$$

$$\widetilde{\alpha_3}(Q^2)=\widetilde{\alpha_s}(Q^2), \tag{22}$$

where the effective couplings $\tilde{\alpha}$ and $\tilde{s}^2$ are defined from PT self-energies $\tilde{\Pi}_{\gamma\gamma}$ and $\tilde{\Pi}_{\gamma Z}$, respectively [16], as is detailed in the Appendix. It is convenient to write $\widetilde{\alpha_1}$ and $\widetilde{\alpha_2}$ in terms of $\tilde{\alpha}$ and $\tilde{s}^2$ since the latter contain the contributions from the mass eigenstate fields. One could use Eq. (19) directly, although the Higgs sector requires care.

Several subtleties should be addressed before the numerical values of the $\overline{\text{PT}}$ couplings are given.

An important difference between the physical effective charges and the unphysical $\overline{\text{MS}}$ couplings is a distinction between timelike and spacelike momenta. In conventional approaches, thresholds are treated in a step function approximation, and hence the running is always logarithmic. The analytic continuation from spacelike to timelike momenta is trivial, yielding $i\pi$ imaginary terms on the timelike side. Thus the real parts of such couplings are the same modulo three-loop $(i\pi)^2$ corrections. In contrast, the $\overline{\text{PT}}$ couplings on timelike and spacelike sides have considerable differences at one loop. To see this we need the exact expressions for the logarithmiclike functions of a particle of spin s, which can be written as

$$L_s(Q/m)=2\left[[\beta\tanh^{-1}(\beta^{-1})-1]\left(\frac{4S^2-\beta^2}{4S^2-1}\right)+1\right], \tag{23}$$

where $S^2=s(s+1)$ is the total spin-squared eigenvalue, $\beta=\sqrt{1+4m^2/Q^2}$, and the momenta is spacelike $(Q^2>0)$. This formula is merely a compact way to write the results for massive spin-0, -1/2, and -1 fields and has not been explicitly verified for higher spins. For example, $L_{s=1}$ is calculated from the sum of the usual gauge boson self-interaction loop, the ghost loops, the appropriate loops of Goldstone bosons that are absorbed, and the pinched parts of the vertex and box graphs (see the Appendix for details). In contrast, $L_{s=1/2}$

---

[8]If particles of different mass propagate together in the loops, this formula is trivially modified.

[9]We will use W or "1" subscripts to denote massive spin-1 fields and a g subscript for massless spin fields. The constants 64/33 and 40/21 are related straightforwardly. In general, for a massive gauge boson W in the representation R of group G that is being considered and representations R' in additional group factors G', we have

$$\beta_W=\frac{11}{3}C(R)d(R')-\frac{1}{6}C(R)d(R')=\frac{7}{2}C(R)d(R') \tag{20}$$

and

$$\eta_W=\frac{1}{\beta_W}\left[\frac{11}{3}C(R)d(R')\left(\frac{64}{33}\right)-\frac{1}{6}C(R)d(R')\left(\frac{8}{3}\right)\right]=\frac{40}{21}. \tag{21}$$

is simply related to the usual fermion vacuum polarization graph, and $L_{s=0}$ comes from the usual scalar contribution to the gauge boson self-energy. It is interesting that such a simple compact form is obtained, considering the seemingly different derivations of the three L_s functions. This may suggest a more efficient formulation of the perturbative dynamics of quantum fields which treats the various spins in a unified manner [23]. Notice that

$$\lim_{m\to\infty} L_s(Q/m)=\frac{8}{3}\left[\frac{1-3s(s+1)}{1-4s(s+1)}\right], \tag{24}$$

corresponding to the results of Table I. The analytic continuation of Eq. (23) to timelike momenta below threshold, $0<q^2=-Q^2<4m^2$, is obtained by replacing

$$\beta\to\bar{\beta}, \tag{25}$$

where

$$\bar{\beta}=\sqrt{\frac{4m^2}{q^2}-1},\quad \tanh^{-1}(\beta^{-1})\to -i\tan^{-1}(\bar{\beta}^{-1}).$$

Above threshold $q^2>4m^2$, one should replace

$$\tanh^{-1}(\beta^{-1})\to\tanh^{-1}(\beta)+i\frac{\pi}{2}, \tag{26}$$

where

$$\beta=\sqrt{1-\frac{4m^2}{q^2}}.$$

From these results it is clear that significant differences will arise between the spacelike and timelike couplings evaluated at scale M_Z^2, mainly due to the $W^{\pm}$ boson threshold asymmetry.

As has been discussed, another distinction of effective couplings is that they are automatically sensitive to light supersymmetry (SUSY) thresholds near M_Z, since the L_s functions are not zero below threshold (on the spacelike side nor on the timelike side). The effects of light SUSY thresholds on the values of the couplings at the Z pole will depend on the method of extraction from the data. The key question is whether or not the light sparticles are implicitly included in the measured values of the couplings at M_Z. For $\tilde{\alpha}(M_Z)$, which is extracted by running the precisely known fine structure constant from $Q=0$ to M_Z, we should include corrections from virtual effects of sparticles (with model-dependent mass), in the self-energy term $\Pi_{\gamma\gamma}(M_Z)$. However, these threshold corrections will cancel in any unification prediction, since then one is essentially running from $Q=0$ to $Q=M_{\rm GUT}$ and the light SUSYs are either fully decoupled or fully turned on. For the strong and weak couplings we use data from the Z pole, and thus no unknown sparticle thresholds must be accounted for since they are already implicitly contained in the measured values. When these couplings are run to the unification scale the induced light threshold corrections will not cancel. Of course, linear combinations of the electromagnetic and weak couplings [Eq. (22)] are used for unification, which complicates the matter further, since different methods of extraction are used for each. It would be unpleasant to quote a different value of $\tilde{\alpha}^{-1}(M_Z)$ for each different SUSY spectra considered. However, this approach has the advantage that the values of the couplings used are the values that one would directly measure in an experiment at M_Z if a given sparticle spectrum were the correct one. For convenience, we will quote the QED coupling extracted assuming a fully decoupled SUSY. When calculating detailed unification predictions in given models, however, the appropriate terms will be included in the determination of $\tilde{\alpha}^{-1}(M_Z)$. It should be emphasized that the above complications are only numerically significant for light sparticle spectra.

TABLE II. Coupling initial values in the various schemes.

	$\overline{\rm MS}$	$\overline{\rm DR}$	$\overline{\rm PT}_+$	$\overline{\rm PT}_-$
$\alpha^{-1}(M_Z)$	127.934(27)	127.881(27)	129.076(27)	128.830(27)
$s^2(M_Z)$	0.23114(20)	0.23030(20)	0.23130(20)	0.22973(20)
$\alpha_3(M_Z)$	0.118(4)	0.119(4)	0.140(5)	0.140(5)

The initial values may be extracted from experimental data and are given in Table II, where spacelike and timelike effective couplings are denoted with a "+" and "−," respectively. The $\overline{\rm MS}$ and $\overline{\rm DR}$ couplings are on the timelike side.

See the Appendix for detailed formulas for the effective couplings.

Notice that the value of the $\overline{\rm PT}_+$ electromagnetic inverse coupling, $\tilde{\alpha}^{-1}(M_Z)=129.076(27)$, does not correspond to the usual value of about 128.968(27). This discrepancy arises because $\tilde{\alpha}^{-1}(M_Z)$ includes the virtual effects of W^+W^- loops, whereas the usual construction of $\alpha_{\rm QED}(M_Z)$ entirely ignores the virtual effects of the massive gauge bosons. The proximate cause of this consistent oversight in the literature is the difficulty in extracting a gauge invariant self-energy-like contribution to the running couplings for non-Abelian theories, a problem which is resolved through the pinch technique, in particular, and more generally, in any effective charge scheme. Clearly, the present approach yields a coupling which more accurately reflects the strength of the electromagnetic force. Similar comments apply to the weak mixing angle.

It should be emphasized that although we have chosen to discuss a particular physical renormalization scheme (PRS), it will be shown in the next section that all predictions associated with unification are PRS invariant, as they should be. However, a definite scheme must be chosen for explicit calculations, and the $\overline{\rm PT}$ scheme is the simplest choice. As expected, we will find that PRS invariance does not extend to unphysical schemes such as $\overline{\rm MS}$ or $\overline{\rm DR}$, because of errors associated with the incorrect treatment of light and heavy thresholds.

V. UNIFICATION IN PHYSICAL RENORMALIZATION SCHEMES

Now we are ready to discuss unification. In Sec. V A, we will consider only the light spectrum given by the standard model fields and their $N=1$ superpartners. This gives a model-independent starting point for discussing unification and makes clear exactly what model-dependent heavy threshold corrections are needed for consistency with the unification hypothesis. New light threshold corrections, in addition to the usual light mass corrections, are evident, although they are numerically important for only a small range of parameter space corresponding to light sparticles. In Sec. V B, *asymptotic unification* is introduced, leading to substantial qualitative changes in the usual picture of gauge unification. This particular choice of unification boundary conditions will lead to corrections from logarithms of superheavy mass ratios, just as would be obtained by implementing $\overline{\text{DR}}$ with the step function approximation. This sheds light on the nature of the approximation of the $\overline{\text{DR}}$ approach. In Sec. V C, an effective unification scale is derived that is considerably higher than the usual unification scale. In Sec. V D, more general nonasymptotic boundary conditions are considered, and the new nontrivial thresholds corrections are found to be important.

In performing the analysis, the exact analytic one-loop formulas discussed in Sec. IV will be used, as well as the leading two-loop corrections. The analytic mass-dependent two-loop corrections are not known, but these can be estimated to be numerically small and well within the error bars, and hence can be neglected [24].

We will treat the SUSY spectrum as entirely arbitrary, rather than assume a particular model or theoretical bias. The advantage of this approach is that importance of various spectra parameters becomes transparent, and irrelevant details can be ignored.

A. (In)validity of neglecting heavy thresholds

In this subsection only, heavy thresholds will be entirely neglected.

The usual test of unification is to predict $\alpha_3(M_Z)$ contingent upon unification. Compared with the conventional $\overline{\text{DR}}$ framework, we expect to see improvements due to the correct treatment of light thresholds. To be precise, the corrections we are discussing are to the difference between the $\widetilde{\alpha}_3(M_Z)$ prediction obtained from the following two methods: (a) using the $\overline{\text{PT}}_+$ scheme throughout and (b) using $\overline{\text{DR}}$ (with the artificial decoupling and theta function treatment of light thresholds) to predict $\hat{\alpha}_3(M_Z)$, which is then translated to a prediction for $\widetilde{\alpha}_3(M_Z)$. Both approaches capture the leading light threshold effects, which appear as logarithms of light masses. The additional corrections in the $\overline{\text{PT}}$ scheme are from what we will call *analytic light threshold corrections*, since they arise from correctly and smoothly interpolating between thresholds. These are largest when there are light supersymmetric partners near or below M_Z. For most values of the sparticle masses, they fall inside the error bars. However, such corrections may become more important as the experimental values of the couplings are determined more precisely. The exact form of the new corrections will be shown explicitly in Sec. V B, Eqs. (31), (35).

Now let us compare the $\overline{\text{PT}}$ unification predictions with experiment. The predictions for the $\overline{\text{PT}}$ strong coupling $\widetilde{\alpha_3}$ [obtained through method (a)] are displayed in Fig. 4 against the SUSY scale and in Fig. 5 against the mass ratio of the gluino and W-ino. These are the two SUSY spectrum parameters to which the $\widetilde{\alpha_3}$ prediction is most sensitive.

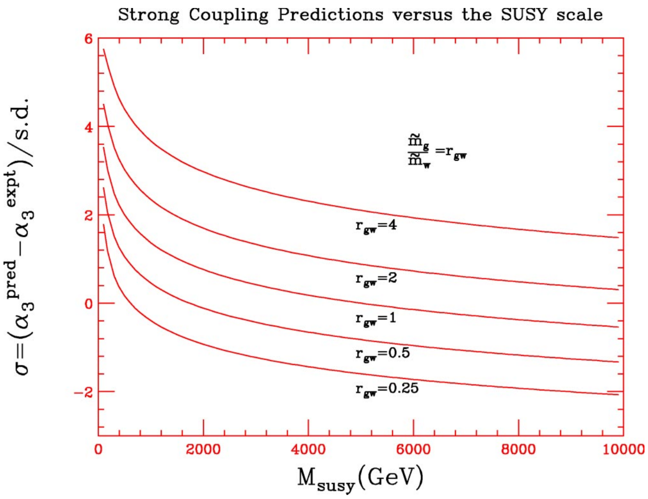

FIG. 4. The error in the prediction for $\widetilde{\alpha_3}(M_Z)$ is plotted against the typical SUSY mass scale, with different lines corresponding to values of the ratio of the gluino mass to the W-ino mass. The relative mass spectrum is roughly the same as most sparticle spectrum models, including supergravity models, with M_s setting the overall scale. The experimental standard deviation (s.d.) is 0.0055 for the PT strong coupling.

Only light gluino scenarios with $m_{\tilde{g}} \lesssim m_{\tilde{w}}$ are able to correctly predict the strong coupling for natural SUSY scales (less than about a TeV). However, it is generally expected that the gluino is several times heavier than the W-ino for most realistic models of supersymmetry breaking and spectra. Hence we reproduce the known result [25] that, at two

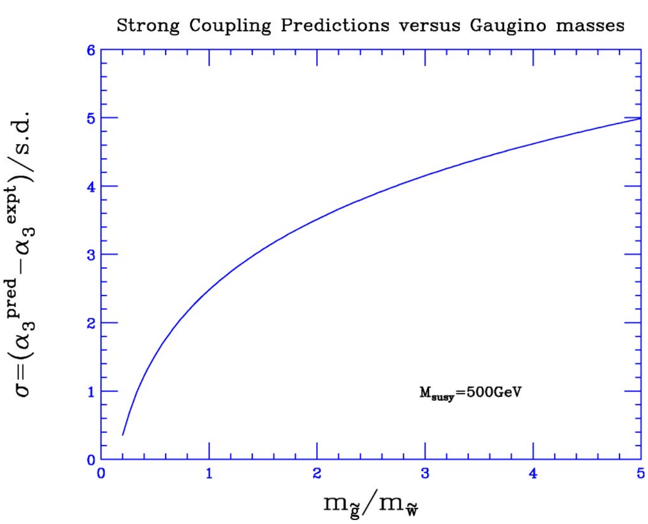

FIG. 5. The error in the prediction for $\widetilde{\alpha_3}(M_Z)$ is plotted against the ratio of the gluino mass to the W-ino mass, which is the sparticle spectrum parameter to which $\widetilde{\alpha_3}(M_Z)$ predictions are most sensitive. The spectrum is fully specified by the ratio and $\sqrt{m_{\tilde{g}} m_{\tilde{w}}} = 500\ \text{GeV} = M_{\text{SUSY}}$, where M_{SUSY} is the mass of all other sparticles.

loops and neglecting heavy thresholds, gauge coupling unification fails by several standard deviations. Except for the light gluino escape route, this points to the need for large heavy threshold corrections if unification is to be achieved.

B. Heavy thresholds and asymptotic unification

Henceforth, the complete heavy threshold behavior will be included in the running of the effective couplings. The form of the subsequent corrections will depend on the particular unification boundary conditions that are chosen, and the numerical values of the corrections will depend on the details of the GUT model. In this section we will choose the simplest boundary conditions, since it will reproduce known results. Later, more general cases will be considered.

Generally, there are four parameters which specify the unification boundary conditions. These are the unification scale M_U and the values of the couplings at that scale, $\widetilde{\alpha_i}(M_U)$ for $i=1,2,3$. For our purposes, we will always assume standard normalizations and take the couplings to be equal at some scale. In this case, the only free parameter is M_U. The two distinct cases are for finite M_U and infinite M_U. The so-called *asymptotic unification* considered in this section corresponds to the latter choice: namely, $M_U\to\infty$ and $\alpha_1^{-1}(M_U)=\alpha_2^{-1}(M_U)=\alpha_3^{-1}(M_U)$. The asymptotic unification conditions would be appropriate if the standard model group G_{SM} is embedded in a simple Lie group G which is fully restored before gravitational or other string interactions become relevant and neglecting any other exotic phenomena. Hence this choice is somewhat simple and naive, but it is very instructive.

We will find that asymptotic unification reproduces the same heavy threshold corrections which can be obtained by unphysical renormalization schemes ($\overline{\rm DR}$) with *finite* unification scale. The reason is that in taking $M_U\to\infty$ one is essentially looking at an observable (the unification requirement) in a desert region, which, as we have seen, unphysical schemes are capable of treating without error. At first sight, it may seem strange that the infinite unification scale predictions of physical schemes correspond to finite unification scale predictions of unphysical schemes. However, this is dictated by the nature of unphysical schemes where masses are turned on and off with a step function.

The paradigmatic improvement over conventional methods is summarized in Fig. 6, where asymptotic unification of the couplings occurs at very large energy. For demonstrative purposes, the parameters are chosen so that unification occurs.

Now let us derive the analytic formulas for the unification predictions. We will discuss the most general case of an $N=1$ supersymmetric $G_{\rm SM}=\mathrm{U}(1)_Y\otimes\mathrm{SU}(2)_L\otimes\mathrm{SU}(3)_C$ embedded in a larger gauge group G using any physical renormalization scheme (all others are inconsistent), which we label by its associated observable $\mathcal{O}$.

In general, the running of the couplings can be expressed in the form

$$\alpha_{\mathcal{O}_i}^{-1}(Q)=\alpha_{\mathcal{O}_i}^{-1}(Q_0)+\tilde{\Pi}_i^{\mathcal{O}}(Q,Q_0)-\theta_i(Q,Q_0), \tag{27}$$

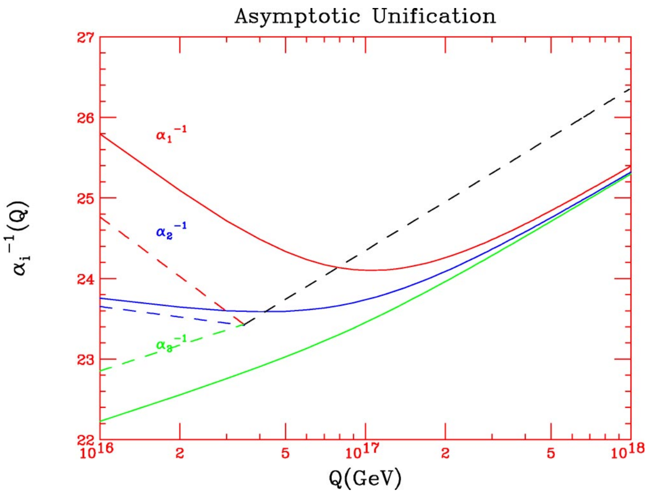

FIG. 6. Asymptotic unification. The solid lines are the analytic $\overline{\rm PT}$ effective couplings, while the dashed lines are the $\overline{\rm DR}$ couplings. For illustrative purposes, $\alpha_3(M_Z)$ has been chosen so that unification occurs at a finite scale for $\overline{\rm DR}$ and asymptotically for the $\overline{\rm PT}$ couplings. Here $M_{\rm SUSY}=200$ GeV is the mass of all light superpartners except the W-ino and gluino which have values $\frac{1}{2}m_{\tilde{g}}=M_{\rm SUSY}=2m_{\tilde{w}}$. For illustrative purposes, we use SU(5).

where the two-loop corrections[10] are contained in $\theta_i(Q,Q_0)$ and we have defined

$$\tilde{\Pi}_i^{\mathcal{O}}(Q,Q_0)=\frac{1}{4\pi}\sum_{p\in G}\beta_i^{(p)}[L_p(Q/m_p)-L_p(Q_0/m_p) - \eta_p^{\mathcal{O}_i}(Q/m_p)+\eta_p^{\mathcal{O}_i}(Q_0/m_p)], \tag{28}$$

which contains all of the one-loop corrections. Now we separate the sums over the light and heavy spectra, $L=G_{\rm SM}$ ($G_{\rm SM}$ means the standard model fields plus SUSY partners) and $H=G-G_{\rm SM}$, take $Q_0=M_Z$, and let $Q=M_U$ be some energy much larger than the mass of all fields, including the heavy fields; i.e., $M_U\gg m_p\ \forall p\in L+H$. The functions $\tilde{\Pi}_i^{\mathcal{O}}$ can then be written as

$$\tilde{\Pi}_i^{\mathcal{O}}(M_U\to\infty,M_Z)=\beta_G l_U-\Delta_i^L-\delta_i^H-\beta_i^H l_X-S_{L,i}^{\mathcal{O}}(\infty)+S_{L,i}^{\mathcal{O}}(M_Z)-S_{H,i}^{\mathcal{O}}(\infty), \tag{29}$$

where

$$\beta_G=\Sigma_{p\in G}\beta_p,\quad l_U=(1/2\pi)\log(M_U/M_Z),$$

$$l_X=(1/2\pi)\log(M_X/M_Z),$$

$$S_{L,i}^{\mathcal{O}}(Q)=\sum_{l\in L}\frac{1}{4\pi}\beta_i^{(l)}\eta_l^{\mathcal{O}}(Q/m_l), \tag{30}$$

[10]See the Appendix for details.

$$\Delta_i^L = \sum_{l\in L} \frac{1}{4\pi}\beta_i^{(l)}\left[L_l\left(\frac{M_Z}{m_l}\right) - \log\frac{M_Z^2}{m_l^2}\right], \tag{31}$$

and

$$\delta_i^H = \sum_{h\in H} \frac{1}{4\pi}\beta_i^{(h)} \log\frac{m_h^2}{M_X^2}. \tag{32}$$

The exact one-loop analytic light threshold corrections are contained in Δ_i^L, while the heavy threshold splittings are contained in δ_i^H, with some arbitrarily chosen heavy mass M_X which is conveniently taken to be the mass of heavy gauge bosons.

It is useful to verify that predictions for l_X and $\alpha_3(M_Z)$ are invariant under the choice of physical renormalization scheme. In performing the calculation, one must use the fact that the $\eta_p^{\mathcal{O}}$ functions do not depend on the gauge group or representation of p, only the spin. These are necessary (but not sufficient) conditions for the sum rule in Eq. (11). This scheme equivalence does not extend to unphysical schemes such as $\overline{\mathrm{DR}}$, though the errors are quantifiable.

As a result of the physical renormalization scheme invariance, we may choose the simplest scheme, which is the $\overline{\mathrm{PT}}$ scheme discussed earlier. Because the $\eta_p^{\overline{\mathrm{PT}}}$ functions are constants equal to $\eta_p = L_p(0)$, the expressions for the unification predictions are simple and compact when written in terms of the $\overline{\mathrm{PT}}$ charges $\tilde{\alpha}_i$.

From $\tilde{\alpha}_1(M_U) = \tilde{\alpha}_2(M_U)$, the heavy gauge boson mass M_X is given by

$$\frac{\log\left(\frac{M_X}{M_Z}\right)+1}{2\pi} = \frac{\tilde{\alpha}_2^{-1}(M_Z) - \tilde{\alpha}_1^{-1}(M_Z) + \Delta_{12}}{\beta_{12}} \tag{33}$$

where $\Delta_{12} = \Delta_1 - \Delta_2$, $\beta_{12} = \beta_1 - \beta_2$, and $\Delta_1 = \Delta_1^L + \delta_1^H + \theta_1$. Notice that M_X can be determined explicitly only for the (unlikely) case of a degenerate heavy spectrum when $\delta_i^H = 0$: otherwise, the expression is transcendental in M_X. In the degenerate case, the gauge boson mass $M_X = M_U^{\not H}/e$ is equal to the unification scale determined by entirely neglecting heavy thresholds (denoted by $\not H$), divided by $e = 2.718\,28\ldots$. This result relies on use of the sum rule in Eq. (11) which gives rise to the $1/2\pi$ term on the left-hand side (LHS) of Eq. (33). The generalization to arbitrary physical renormalization scheme is $M_X = M_U^{\not H} e^{-K^{\mathcal{O}}/2}$, where $K^{\mathcal{O}}$ is defined in Eq. (11). Neglecting the light and heavy analytic threshold corrections, the gauge boson mass prediction is the same as the unification scale prediction in the $\overline{\mathrm{DR}}$ scheme. Also, the "unification" scale $M_U^{\not H}$ depends on the particular scheme, which makes sense since different schemes correspond to different observables. In contrast, M_X is scheme independent.

The strong coupling prediction is

$$\tilde{\alpha}_3^{-1}(M_Z) = \tilde{\alpha}_1^{-1}(M_Z) + \Delta_{31} + \frac{\beta_{13}}{\beta_{12}}[\tilde{\alpha}_2^{-1}(M_Z) - \tilde{\alpha}_1^{-1}(M_Z) + \Delta_{12}], \tag{34}$$

which differs from the prediction obtained by neglecting heavy thresholds by only the terms $\delta_{12}^H, \delta_{13}^H$, which reflect the heavy splitting.

In order to explicitly compare with the $\overline{\mathrm{DR}}$ approach, the artificial decoupling treatment of thresholds should be employed, as described in Sec. II. This involves using a step function through each light field $l \in L$ with mass greater than M_Z and through every superheavy field $h \in H$. Then one must impose the unification condition such that the three gauge couplings are equal at the maximum mass of heavy fields, $M_X = \max\{m_h, h\in H\}$. At energies above this maximum mass, the three couplings run identically according to the beta function for the unified group G; hence, there is no arbitrariness in the choice of the unification scale. Next, the prediction for the $\overline{\mathrm{DR}}$ strong coupling should be translated to the $\overline{\mathrm{PT}}$ strong coupling. Doing this, one finds the exact same form of Eq. (34), except that Δ_i^L is replaced by

$$\Delta_i^L \rightarrow \sum_{m_l<M_Z} \frac{1}{4\pi}\beta_i^{(l)}\left[L_l\left(\frac{M_Z}{m_l}\right) - \log\frac{M_Z^2}{m_l^2}\right] + \sum_{m_l>M_Z} \frac{1}{4\pi}\beta_i^{(l)}\left(\eta_l - \log\frac{M_Z^2}{m_l^2}\right). \tag{35}$$

Notice that there are only light threshold corrections beyond the theta function approximation for particles of mass above M_Z, since those below M_Z are already implicitly accounted for. This formula is in agreement with Eqs. (15), (16), since there is a residual error proportional to $L_p(M_Z/m_p) - \eta_p$ for each crossed threshold. The analogous corrections for the heavy thresholds do not arise in the asymptotic unification scenario, since we are essentially comparing observables at energy scales $M_Z \sim m_l$, which is of the same order of magnitude as the light thresholds, and $M_U \gg m_h$, which is much greater than all thresholds when asymptotic unification conditions are assumed. The latter scale is a "desert" scale, and so the step function method has no errors, giving the same result obtained above in the δ_i^H.[11] For the more general unification conditions considered in the next subsection there will be additional heavy threshold corrections.

Equation (34) is a useful result, as it allows one to constrain the heavy spectrum, given a light SUSY spectrum. Up to two-loop finite threshold corrections, which we have estimated to be small, and assuming that Eq. (34) will yield the experimental value of the strong coupling given some appro-

[11]It should be emphasized that this is only the case when the $\overline{\mathrm{DR}}$ is correctly implemented by choosing the unification scale to be equal to the heaviest threshold in the theory. Different choices are sometimes made in the literature. For example, in [26], the authors advocate defining the unification scale to be the geometric mean of the heavy masses.

TABLE III. Constants which characterize each representations contribution to the $\alpha_3(M_Z)$ prediction.

	$\overline{B}_h(R_i)$		$\overline{B}_h(R_i)$
$R_1=(\mathbf{3},\mathbf{2},1/6)$	$-3/2$	$R_5=(\mathbf{1},\mathbf{1},1)$	$3/7$
$R_2=(\mathbf{3},\mathbf{1},-1/3)$	$9/14$	$R_6=(\mathbf{8},\mathbf{1},0)$	3
$R_3=(\mathbf{3},\mathbf{1},2/3)$	$15/14$	$R_7=(\mathbf{1},\mathbf{3},0)$	$-24/7$
$R_4=(\mathbf{1},\mathbf{2},1/2)$	$-9/14$	$R_8=(\mathbf{3},\mathbf{2},5/6)$	$3/14$

priate full GUT theory (i.e., assuming the asymptotic unification hypothesis is true), we can write

$$\epsilon^H \equiv \tilde{\alpha}_3^{-1}(M_Z)^{\text{pred}}_{\not H} - \tilde{\alpha}_3^{-1}(M_Z)^{\text{expt}} \approx \delta_{13}^H + \frac{\beta_{13}}{\beta_{12}}\delta_{21}^H, \quad (36)$$

where $\tilde{\alpha}_3^{-1}(M_Z)^{\text{pred}}_{\not H}$ is the predicted value of the strong coupling obtained by neglecting heavy thresholds, as illustrated in Figs. 4 and 5. We should emphasize the assumptions leading to this result. First, the standard normalizations of the couplings are assumed, so that Eq. (37) does not hold for higher affine levels or nonstandard hypercharge normalizations, as often occur in string models. Second, we assume that the gluino is somewhat heavier than the chargino, so that there are serious discrepancies, as in Fig. 4, which must be explained by heavy threshold corrections. Finally, we are using the paradigm of asymptotic unification, wherein the full gauge group G in which the SM is embedded is restored before other Planck scale physics becomes relevant. With these assumptions and noting that heavy thresholds were neglected in Figs. 4 and 5, we find a typical value of

$$\epsilon^H \overset{\text{expt}}{\approx} -1 \overset{\text{theory}}{\approx} -\frac{1}{4\pi}\sum_{h\in H} B_h \log\frac{m_h^2}{M_X^2}, \quad (37)$$

where we have defined[12]

$$B_h \equiv \beta_{31}^{(h)} + \frac{12}{7}\beta_{12}^{(h)}. \quad (38)$$

Values of B^h can be compiled for the heavy representations any unified gauge group and, hence, may be used with heavy mass ratios to exclude or provide evidence for a given GUT theory.

To calculate B_h, we first write $B_h=\overline{\beta}_{s_h}\overline{B}_h$, where $\overline{\beta}_{s_h}=-1/3,-2/3,11/3$ for spin 0, 1/2, 1 fields and the remaining group theory factor is $\overline{B}_h=\frac{5}{7}T_1(R)-\frac{12}{7}T_2(R)+T_3(R)$ for a representation R. It is necessary to decompose all representations in terms of their $U(1)_Y\otimes SU(2)_L\otimes SU(3)_C$ content. Here $T_1(R)=\frac{3}{5}\sum_{p\in R}Y_p^2$ and $T_i(R)\delta^{ab}=\sum_{p\in R}\mathrm{tr}_i(t_R^a t_R^b)$, $i=2,3$. For many grand unified theories of interest, all multiplets can be decomposed in terms of only eight different standard model multiplets (plus their conjugate representations, which have the same B_h, and a singlet which has $B_h=0$), which are given in Table III along with the value of $\overline{B}_h$.

These same constants will also govern the corrections from analytic heavy threshold corrections which will be discussed later in Sec. V D. Notice that, by definition, the constants satisfy the constraint that the sum over all heavy multiplets vanishes,

$$\sum_{h\in H} B_h = 0, \quad (39)$$

which equivalently reflects the arbitrariness in the choice of which heavy mass scale M_X one chooses to be canonical [see Eq. (37)], $\partial\epsilon^H/\partial M_X=0$. A similar relation also holds for any complete representation of the grand unified group. For example, the **24** of SU(5) decomposes into a singlet plus $R_6+R_7+R_8+\overline{R}_8$. From the table, we have $B_h(R_6)+B_h(R_7)+2B_h(R_8)=0$.

As a simple example, let us explore the (unlikely) possibility wherein the only heavy field with significantly different mass than the heavy gauge boson mass M_X is the **5**-dimensional Higgs supermultiplet in which the light Higgs doublets are embedded. The triplet components of the two Higgs supermultiplets contributes $-2/5$, 0, -1 to β_1, β_2, β_3, and hence $B_h(\mathbf{3}+\overline{\mathbf{3}})=-9/7$. Using Eq. (37), this leads to $M_3\approx M_X\exp(-14\pi/9)$, which is of order $M_X/100$. Such a large splitting is unnatural and difficult to accommodate in a theory. In general, "natural" splittings do not lead to ϵ^H values of the correct magnitude in SU(5). This is not terribly surprising, since minimal SUSY SU(5) is already known to be strongly disfavored.

In general, the large discrepancies in Figs. 4 and 5 imply a large splitting in the heavy spectrum, which, in turn, may imply a multistep unification scenario—e.g., $SO(10)\to G_{224}\to G_{SM}$. The reason is that for the heavy fields to contribute to $\tilde{\alpha}_3(M_Z)$, they must not only have a mass splitting compared to some reference heavy gauge boson X, but also must have different first beta function coefficients since only the differences $\beta_1^{(h)}-\beta_2^{(h)}$ and $\beta_1^{(h)}-\beta_3^{(h)}$ appear in the corrections.

Before moving to more general unification boundary conditions, we shall give a simple way to define an effective unification scale in the asymptotic unification scenario.

C. Effective unification scale

Because the couplings formally unify at infinite energy in the paradigm of asymptotic unification, there is no apparent unification scale. However, we suspect that in reality quantum gravitational fluctuations will affect the couplings as they approach the Planck energy. Hence one can define an effective unification scale to be where the splittings between the gauge couplings are of the same order as those induced by gravitational effects. To be precise, define a dimensionless gravitational coupling which classically runs with energy as

$$G_N(Q)=\frac{Q^2}{M_{\text{Pl}}^2}, \quad (40)$$

[12]Note that $\beta_{13}/\beta_{12}=\frac{12}{7}$.

where $M_{\rm Pl} \approx 1.22 \times 10^{19}$ GeV. The leading gravitational corrections to the running gauge couplings $\alpha_i(Q)$ will be proportional to $G_N(Q)\alpha_i(Q)$. Hence the effective asymptotic unification scale $M_{\rm eff}$ can be defined as the scale where the splittings in the gauge couplings are of order the gravitational corrections, $|\alpha_i(M_{\rm eff}) - \alpha_j(M_{\rm eff})| \approx b^2 G_N(M_{\rm eff})\alpha_U$ or, equivalently,

$$|\alpha_i^{-1}(M_{\rm eff}) - \alpha_j^{-1}(M_{\rm eff})| \approx b^2 G_N(M_{\rm eff})\alpha_U^{-1} \equiv \delta_g(M_{\rm eff}), \tag{41}$$

where we take $\alpha_U^{-1} \sim 24$ to be the typical gauge coupling near unification. The unknown parameter b^2 should be of order 1. Estimating $M_{\rm eff}$ using a simple SU(5) model, we find a typical effective unification scale of $1-5 \times 10^{17}$ GeV. This is only intended to be a very rough approximation since a naively simple SU(5) model was used. Nevertheless, more complicated and realistic GUT models yield a unification scale in the same ballpark. It is generally true that our effective unification scale is about an order of magnitude or more greater than what is typically called the unification scale ($\sim 2 \times 10^{16}$ GeV).

It may seem that our definition of an effective unification scale is rather artificial. However, it may be physically motivated by the following considerations. If indeed the standard model is embedded in some unified theory of gravity and gauge forces, then there may exist a phase at energies below the Planck scale which consists of a simple Lie group containing the supersymmetric standard model. In the absence of any gravitational corrections, the running gauge couplings certainly unify asymptotically, as this is the only case in which the higher group symmetry is fully realized up to arbitrarily high energies. Hence the running couplings should only deviate from asymptotic unification by the gravitational corrections parameterized above. So by the above reasoning, the effective unification scale should roughly correspond to the physical unification scale when the full (quantum gravitational) theory is considered.

These results may have consequences for the paradigm of string unification. In particular, one problem of string unification [27] is that the couplings seem to unify at a scale ($M_G^{\overline{\rm DR}} \approx 2 \times 10^{16}$ GeV) about 20 times lower than the scale predicted by four-dimensional heterotic string models ($M_G^{\rm string} \approx 5 \times 10^{17}$ GeV). In the approach presented here, heavy threshold effects seem to push the effective asymptotic unification scale to roughly $M_G^{\rm string}$. Despite the apparent success, this coincidence cannot be taken seriously until several questions are addressed in regards to this so-called string gauge coupling problem. First, the calculation of $M_G^{\rm string}$ [28] was performed in the $\overline{\rm DR}$ scheme, with the field theory step function treatment generalized to strings. An analogous calculation for physical renormalization schemes is lacking, so it is difficult to compare our results with string predictions. Second, the asymptotic unification boundary conditions are probably not valid for many string models, and so the unification scale will be further changed by more general boundary conditions, as discussed in the next section. See [29] for a recent string calculation of threshold corrections to grand unification.

D. More general boundary conditions

The discussion of this section concerns the next-simplest boundary conditions after asymptotic unification. In particular, we will impose $\alpha_1(M_U) = \alpha_2(M_U) = \alpha_3(M_U)$ at scale $M_U \sim M_h$, for some $h \in H$. As discussed in the previous section, one might expect M_U to roughly correspond to the asymptotic unification scale, which we found to be roughly 5×10^{17} GeV. However, we will consider M_U as an input and find the corrections for the strong coupling and gauge boson mass predictions.

Before proceeding, there is a subtle point that should be addressed. Notice that in the previous section, we assumed that unification *would have* occurred asymptotically were it not for gravitational corrections. Hence starting with a finite unification scale and then neglecting gravitational corrections, as we do in this section, does not seem logically consistent with what was done in the previous section. This observation is entirely correct, but the point is that indeed we are considering two orthogonal scenarios: one where a finite unification scale is obtained from gravity and another where finite unification scale is obtained from nontrivial threshold corrections. The latter case may have its origin in stringy or gravitational physics, but nevertheless becomes manifest through purely field theoretic mechanisms.

The corrections from imposing finite unification scale are straightforward to derive and can be stated in terms of the $\Delta_i = \Delta_i^L + \delta_i^H + \theta_i$, which we defined earlier. This gains an additional contribution and can now be written

$$\Delta_i = \Delta_i^L + \delta_i^H + \theta_i - \Delta_i^H, \tag{42}$$

where

$$\Delta_i^H = \sum_{h \in H} \frac{1}{4\pi} \beta_i^{(h)} \left[L_h\left(\frac{M_U}{m_h}\right) - \log \frac{M_U^2}{m_h^2} \right], \tag{43}$$

which is of exactly the same form expected from Eq. (15).[13] Evidently, these are finite heavy threshold corrections in addition to the corrections from the heavy threshold splittings. Hence the ϵ^H defined earlier will get an additional contribution from the Δ_i^H's and is now

$$\begin{aligned} \epsilon^H \overset{\rm theory}{\approx} & -\frac{1}{4\pi} \sum_{h \in H} B_h \left\{ \log \frac{m_h^2}{M_X^2} - \left[L_h\left(\frac{M_U}{m_h}\right) - \log \frac{M_U^2}{m_h^2} \right] \right\} \\ & = \frac{1}{4\pi} \sum_{h \in H} B_h L_h\left(\frac{M_U}{m_h}\right). \end{aligned} \tag{44}$$

Experimentally, $\epsilon^H \approx -1$, as seen in Figs. 4 and 5 for typical gluino to W-ino mass ratios; this value can be easily adjusted

[13] This is not obvious; one must work through the derivation to see that indeed the expected $L_p - \eta_p$ correction terms do arise.

for nonstandard sparticle spectra. This is our final formula which may be used to assess the experimental validity of gauge coupling unification in any specific GUT model where the gauge group, superheavy mass ratios, and light SUSY masses are given.

Let us now consider the numerical size of these new threshold corrections. From Eq. (6) one finds that

$$L_h\left(\frac{M_U}{m_h}\right)-\log\frac{M_U^2}{m_h^2}\approx\log\left(1+\frac{m_h^2}{M_U^2}e^{\eta_h}\right), \tag{45}$$

which can be larger than heavy splitting corrections $\log(m_h^2/M_X^2)$ for values of M_U that are not too large. Hence such corrections cannot be neglected.

The value of M_U is not fixed *a priori* and corresponds to the physically meaningful energy where the couplings become equal due to the new nontrivial heavy threshold corrections. This complicates the analysis of unification by introducing another parameter beyond those that are usually needed. However, this is to be expected, since a new physical phenomena (corrections arising from the virtuality of very massive particles) has been included.

VI. CONCLUSIONS

We have developed a new way of looking at detailed predictions of gauge coupling unification which is more physically motivated than conventional approaches. In addition to a dramatic paradigmatic improvement, novel heavy and light threshold corrections are obtained, and the resulting corrections to unification predictions are presented for a general GUT model. A natural extension of this work is a thorough analysis and classification of various unified theories. By calculating the B_h constants and the heavy spectrum, one may exclude or verify the gauge unification of a given model.

ACKNOWLEDGMENTS

We wish to thank Ratin Akhoury and Eduardo de Rafael for reading a draft of this paper and providing useful suggestions.

APPENDIX: THE PINCH TECHNIQUE SELF-ENERGY-LIKE EFFECTIVE CHARGES

Here we will give explicit formulas for the pinch technique effective couplings regularized using dimensional reduction, which will be denoted with a tilde. These effective charges will be similar to those constructed in [13,15] for QCD and in [14,16] for the electroweak sector. However, we will extend these to the minimal supersymmetric case, which involves explicitly including another Higgs doublet, and regulating the loop integrals with dimensional reduction, as opposed to dimensional regularization, which is used in most nonsupersymmetric settings. It is well known that DRED preserves both supersymmetry and gauge symmetry. Also, the effective charges presented in [16] were in the on-shell subtraction scheme ($Q_0=0$), whereas here we will need the result for arbitrary renormalization scale. In the appropriate limits our results reduce to those given in [15] and [16].

The charges are constructed using the pinch-technique, which allows one to extract the universal self-energy function in non-Abelian gauge theories, thus leading to gauge-invariant effective couplings which (i) contain explicit and complete mass-threshold behavior and (ii) reproduce the conventional massless beta function in the limit where masses can be neglected.

At one-loop, the spin-1/2 and spin-0 PT self-energies are trivially just the usual transverse vacuum polarization graphs. Only the graph with a gauge boson loop needs to have the self-energy-like part projected, as described briefly in Sec. III and in more detail in Refs. [15,16]. In calculating the following, we used both the direct diagrammatic pinch technique algorithm [15] and the dispersive derivation from physical cross sections [16].

The PT effective charges naturally measure the self-energy-like propagation of a gauge boson and hence can be interpreted as measuring the real force between two fermions of arbitrary mass, analogous to the QED effective charge. The PT charge includes finite mass recoil effects that are missed in the heavy quark effective charge (the V scheme). In fact, one may obtain the heavy quark potential in the appropriate kinematical limit of the pinch technique effective charge [15]. The difference between the two are due to finite mass test-charge effects that are not present in the (V) charge, but are in the (PT) charge. Different extensions of the PT effective charge beyond one loop have been put forth [17–19], although it seems that the approach of [19] most closely matches the philosophy used here. A multiloop generalization of this algorithm remains to be constructed.

QCD effective charges

The $\overline{\mathrm{PT}}$ self-energy function for supersymmetric QCD, $\tilde{\Pi}_3$, can be used to define the effective coupling for supersymmetric QCD by

$$\tilde{\alpha}_3(Q)=\frac{\tilde{\alpha}_3(Q_0)}{1+\tilde{\Pi}_3(Q,Q_0)}. \tag{A1}$$

The function $\tilde{\Pi}_3$ can be written down straightforwardly using Eqs. (5), (9), and the unsubtracted result is given by

$$\begin{aligned}\tilde{\Pi}_3(Q)=\frac{\tilde{\alpha}_3(Q_0)}{4\pi}\Bigg[&\frac{11}{3}N_c\left(\log\frac{Q^2}{\mu^2}-C_{UV}-64/33\right)\\&-\frac{2}{3}N_c[L_{1/2}(Q/m_{\tilde{g}})+\log(m_{\tilde{g}}^2/\mu^2)-C_{UV}-5/3]\\&-\sum_q\frac{2}{3}[L_{1/2}(Q/m_q)+\log(m_q^2/\mu^2)-C_{UV}-5/3]\\&-\sum_{\tilde{q}}\frac{1}{3}[L_0(Q/m_{\tilde{q}})+\log(m_{\tilde{q}}^2/\mu^2)-C_{UV}-8/3]\Bigg].\end{aligned} \tag{A2}$$

The four terms correspond respectively to the gluons, gluinos ($\tilde{g}$), Dirac quarks (q), and to complex squark doublets ($\tilde{q}$). For the scalars we will take the left and right components to be degenerate in mass since such complications do not change the unification predictions to any numerical significance. In any case, one may trivially treat the two separately.

To relate the resulting effective charge to other schemes or observables one needs to use Eq. (9).

Equation (A2) can be written in a more useful once subtracted form by relating the effective charges at different scales, leading to an expression governing the running of the charge given by

$$\begin{aligned}\tilde{\Pi}_3(Q,Q_0)\equiv\tilde{\Pi}_3(Q)-\tilde{\Pi}_3(Q_0)=\frac{\tilde{\alpha}_3(Q_0)}{4\pi}\Bigg[&\frac{11}{3}N_c\left(\log\frac{Q^2}{Q_0^2}\right)\\&-\frac{2}{3}N_c[L_{1/2}(Q/m_{\tilde{g}})-L_{1/2}(Q_0/m_{\tilde{g}})]\\&-\sum_q\frac{2}{3}[L_{1/2}(Q/m_q)-L_{1/2}(Q_0/m_q)]\\&-\sum_{\tilde{q}}\frac{1}{3}[L_0(Q/m_{\tilde{q}})-L_0(Q_0/m_{\tilde{q}})]\Bigg].\end{aligned}\tag{A3}$$

Though the gluon contribution in Eq. (A2) looks simple, it is actually the most difficult piece to compute. As discussed in [15], the pinch technique self-energy which contributes to the effective charge is gauge and scale independent, and indeed reproduces the pure gauge term of the β function coefficient. This of course is not the case for the full pure gluon vacuum polarization, which is gauge and scale dependent and does not reproduce the correct β function. The nontrivial and important part of the result is the constant 64/33. This constant may be obtained using the pinch technique and DRED. To translate to DREG one just subtracts 1/11 (from the so-called epsilon ghosts) to get the constant 67/33. For comparison, the heavy quark potential effective charge α_V replaces this constant with 28/33 and 31/33 when using DRED and DREG, respectively. Consequently, the V scheme does not satisfy the decoupling criterion of Eq. (10). This is just a reflection of the fact that infinitely heavy external quarks are used in the V-scheme calculation, thus rendering meaningless the limit where internal virtual gauge bosons acquire very large mass.

Notice that in the appropriate limit the above reduces to the standard RG β function coefficient:

$$\lim_{m_i\to 0}\tilde{\Pi}_3=\frac{\alpha_s}{4\pi}(9-n_f)\log\frac{Q^2}{Q_0}.\tag{A4}$$

Electroweak sector

The effective QED charge and the effective weak-mixing angle are obtained by diagonalizing the electroweak neutral currents and are given by [16]

$$\tilde{\alpha}(Q)=\frac{\tilde{\alpha}(Q_0)}{1+\tilde{\Pi}_{\gamma\gamma}(Q,Q_0)}\tag{A5}$$

and

$$\tilde{s}_w^2(Q)=\tilde{s}_w^2(Q_0)\left(1+\frac{\tilde{c}_w(Q_0)}{\tilde{s}_w(Q_0)}\frac{\tilde{\Pi}_{\gamma Z}(Q,Q_0)}{1+\tilde{\Pi}_{\gamma\gamma}(Q,Q_0)}\right).\tag{A6}$$

For the matter sector, we will write only the subtracted $\overline{\mathrm{PT}}$ self-energies, as it is now clear how to translate between schemes using the η_p constants as described before. The quarks (q) and leptons (l), along with their scalar superpartners ($\tilde{q},\tilde{l}$), yield

$$\begin{aligned}\tilde{\Pi}_{\gamma\gamma}(\text{matter})=\frac{\tilde{\alpha}(Q_0)}{4\pi}\Bigg[&-\sum_q\frac{4}{3}N_ce_q^2[L_{1/2}(Q/m_q)-L_{1/2}(Q_0/m_q)]\\&-\sum_{\tilde{q}}\frac{2}{3}N_ce_{\tilde{q}}^2[L_0(Q/m_{\tilde{q}})-L_0(Q_0/m_{\tilde{q}})]\\&-\sum_l\frac{4}{3}[L_{1/2}(Q/m_l)-L_{1/2}(Q_0/m_l)]\\&-\sum_{\tilde{l}}\frac{2}{3}[L_0(Q/m_{\tilde{l}})-L_0(Q_0/m_{\tilde{l}})]\Bigg].\end{aligned}\tag{A7}$$

The electric charge of a particle p is denoted e_p. The analogous contributions of individual Dirac mass eigenstate matter fields to the γZ self-energy are given by the relation

$$\tilde{\Pi}_{\gamma Z}^{(p)}=\left(\frac{1}{4|e_p|}-\tilde{s}_w^2\right)\frac{1}{\tilde{s}_w\tilde{c}_w}\tilde{\Pi}_{\gamma\gamma}^{(p)},\tag{A8}$$

where p denotes any of the fermions or scalars above.

The contribution of the charged vector bosons to the self-energies is more complicated than the matter multiplets. Similar to the QCD case, the non-Abelian nature of the theory implies that W^+W^- loops (along with possible gauge-dependent ghost and Goldstone boson loops) do not yield a gauge-invariant result and do not give the appropriate contribution to the electroweak beta functions. The proper treatment involves calculating the self-energy-like part of the one-loop $e^+e^-\to e^+e^-$ amplitude (or using any other fermions due to universality), including vertex and box corrections involving neutrinos. These contribute pinched parts which make the self-energy-like part gauge invariant and transverse. This calculation was first performed in [14] and then with dispersion relations in [16] for $n_H=1$ Higgs doublets and renormalized in the on-shell scheme at $Q_0=0$. Here we need to extend these results to arbitrary n_H and Q_0, and would like to have the finite constants in the unrenormalized expression, including constant terms arising from using DRED instead of DREG. The most efficient way to do this is to use the Feynman gauge $\xi=1$, where $W^\pm$ bosons, $-G^\pm$

Goldstone bosons, and $\eta^{\pm}$ ghosts all propagate with $-ig_{\mu\nu}/(p^2-M_W^2)$. Hence the only factors of transverse momenta arise from the three-boson vertex, and so the box graph and several of the vertex graphs may be neglected, as they do not have pinched parts. Here the dependence on the Higgs doublets comes only from the unphysical charged Goldstone scalars for the $n_H=1$ case and also on charged Higgs scalars for $n_H>1$. The result for the $\mathrm{SU}(2)_L\otimes\mathrm{U}(1)_Y$ electroweak theory is

$$\tilde{\Pi}_{\gamma\gamma}^{(W,H)}=\frac{\tilde{\alpha}(Q_0)}{4\pi}\Bigg[2\,\frac{11}{3}[\overline{L_1}(Q/M_W)+\log(M_W^2/\mu^2)-C_{UV} -64/33]+\sum_{a=1}^{n_H}\left(-\frac{1}{3}\right)[L_0(Q/M_a)+\log(M_a^2/\mu^2) -C_{UV}-8/3]\Bigg], \tag{A9}$$

where the constant 64/33 is the same as appeared for the gluon self-energy. The sum in the second line will be over mass eigenstate charged Higgs scalars; there will be one of these for each Higgs doublet in the theory. The first scalar ($a=1$ in the sum) is an unphysical Goldstone boson that is absorbed by the $W^{\pm}$, and hence one identifies its mass to be $M_1=M_W$ (in the Feynman gauge). The second charged scalar ($a=2$) is conventionally denoted by $H^{\pm}$ in the MSSM, with mass $M_2=M_{H^{\pm}}$. Additional Higgs doublets beyond the MSSM are not considered here, so we can take $n_H=2$. The function

$$\overline{L_1}(Q/m)=\frac{2}{11}[\beta\tanh^{-1}\beta^{-1}(12-\beta^2)+\beta^2-1], \tag{A10}$$

with $\beta=\sqrt{1+4m^2/Q^2}$, comes from the W^+W^- and ghost loops, the $W^+G^-+W^-G^+$ loops, and the pinched self-energy-like part of the γWW vertex where the internal neutrino line is pinched. The L_0 comes from the charged Goldstones and Higgs loops. As might be anticipated from the fermions and scalars, where, for example, $\lim_{m\to\infty}L_{1/2}(Q/m)=5/3$ is the same constant as appears in the self-energy, we also have the nice property that

$$\lim_{M_W\to\infty}\overline{L_1}(Q/M_W)=64/33. \tag{A11}$$

Notice that in DREG this does not cancel the constant, which in that case is 67/33. With DRED regularization, all massive particles decouple, modulo divergent pieces, from the unsubtracted self-energy-like expression.

Letting the W bosons absorb the Goldstones by performing simple algebra in Eq. (A9), one finds the result written in terms of physical degrees of freedom:

$$\tilde{\Pi}_{\gamma\gamma}^{(W,H)}=\frac{\tilde{\alpha}(Q_0)}{4\pi}\Bigg[7[L_1(Q/M_W)+\log(M_W^2/\mu^2)-C_{UV} -40/21]+\sum_{a=2}^{n_H}\left(-\frac{1}{3}\right)[L_0(Q/M_a)+\log(M_a^2/\mu^2) -C_{UV}-8/3]\Bigg]. \tag{A12}$$

The contribution of the physical massive gauge boson is characterized by $7L_1=(22/3)\overline{L_1}+(-1/3)L_s$, explicitly given by

$$L_1(Q/m)=2\beta\tanh^{-1}\beta^{-1}[1-(\beta^2-1)/7]+(2/7)(\beta^2-1). \tag{A13}$$

As expected,

$$\lim_{m\to 0}L_W(Q/m)=\log\frac{Q^2}{m^2},\quad \lim_{m\to\infty}L_W(Q/m)=40/21. \tag{A14}$$

Notice that Eq. (A13) precisely corresponds to Eq. (23) for $s=1$.

The separation of pure gauge effects and those arising in the broken phase of the theory is useful and allows us to immediately write down the analogous result for γZ without further calculation:

$$\tilde{\Pi}_{\gamma Z}^{(W,H)}=\frac{\tilde{\alpha}(Q_0)}{4\pi c_w s_w}\Bigg\{2\,\frac{11}{3}c_w^2[\overline{L_1}(Q/M_W)+\log(M_W^2/\mu^2) -C_{UV}-64/33]+\sum_{a=1}^{n_H}\left[-\frac{1}{3}\left(c_w^2-\frac{1}{2}\right)\right](L_0(Q/M_a) +\log(M_a^2/\mu^2)-C_{UV}-8/3)\Bigg\}. \tag{A15}$$

Finally, the W-ino and charged Higgsino, whose mixing is neglected, contribute

$$\tilde{\Pi}_{\gamma\gamma}^{\tilde{H},\tilde{W}}=\frac{\tilde{\alpha}(Q_0)}{4\pi}\Bigg[-\frac{4}{3}[L_{1/2}(Q/m_{\tilde{H}})-L_{1/2}(Q_0/m_{\tilde{H}})] -\frac{4}{3}[L_{1/2}(Q/m_{\tilde{W}})-L_{1/2}(Q_0/m_{\tilde{W}})]\Bigg] \tag{A16}$$

and

$$\tilde{\Pi}_{\gamma Z}^{\tilde{H},\tilde{W}}=\frac{\tilde{\alpha}(Q_0)}{4\pi c_w s_w}\Bigg[-\frac{4}{3}c_w^2[L_{1/2}(Q/m_{\tilde{H}})-L_{1/2}(Q_0/m_{\tilde{H}})] -\frac{4}{3}(c_w^2-1/2)[L_{1/2}(Q/m_{\tilde{W}})-L_{1/2}(Q_0/m_{\tilde{W}})]\Bigg]. \tag{A17}$$

The $SU(2)_L \otimes U(1)_Y$ effective couplings constructed from the above results are

$$\widetilde{\alpha}_i(Q) = \frac{\widetilde{\alpha}_i(Q_0)}{1+\widetilde{\Pi}_i(Q,Q_0)}, \tag{A18}$$

for $i=1,2$ and $\widetilde{\alpha}_i(Q)$ given in Eq. (22). The PT self-energies are related by

$$\widetilde{\Pi}_1 = \widetilde{\Pi}_{\gamma\gamma} - \frac{s_w}{c_w}\widetilde{\Pi}_{\gamma Z},$$

$$\widetilde{\Pi}_2 = \widetilde{\Pi}_{\gamma\gamma} + \frac{c_w}{s_w}\widetilde{\Pi}_{\gamma Z}. \tag{A19}$$

Notice that $\widetilde{\Pi}_{\gamma\gamma} = \Sigma_p \widetilde{\Pi}^{(p)}_{\gamma\gamma}$ and $\widetilde{\Pi}_{\gamma Z} = \Sigma_p \widetilde{\Pi}^{(p)}_{\gamma Z}$ (as well as $\widetilde{\Pi}_i$) have the correct beta function coefficients, which are summarized below, and smoothly interpolate between all mass thresholds. The full mass-dependent beta functions may be obtained by differentiating the above expressions, but we will just give the massless limits, in order to make clear our conventions.

The one-loop beta function coefficients are defined by the relations

$$\frac{d\alpha}{d\log Q^2} = -\alpha\frac{d\widetilde{\Pi}_{\gamma\gamma}}{d\log Q^2} = -\frac{\alpha^2}{4\pi}\beta_{\gamma\gamma}, \tag{A20}$$

$$\frac{ds_w^2}{d\log Q^2} = s_w c_w \frac{d\widetilde{\Pi}_{\gamma Z}}{d\log Q^2} = \frac{\alpha}{4\pi}\beta_{\gamma Z}, \tag{A21}$$

$$\beta_1 = \frac{3}{5}(c_w^2\beta_{\gamma\gamma} - \beta_{\gamma Z}), \tag{A22}$$

$$\beta_2 = s_w^2\beta_{\gamma\gamma} + \beta_{\gamma Z}. \tag{A23}$$

One finds

$$\beta_{\gamma\gamma} = -\frac{16}{3}N_g + 6 - n_H,$$

$$\beta_{\gamma Z} = -2N_g + \frac{16}{3}s_w^2 N_g + 6c_w^2 - n_H\left(\frac{1}{2} - s_w^2\right), \tag{A24}$$

thus leading to the MSSM beta function coefficients

$$\begin{pmatrix}\beta_1\\ \beta_2\\ \beta_3\end{pmatrix} = \begin{pmatrix}0\\ 6\\ 9\end{pmatrix} - N_g\begin{pmatrix}2\\ 2\\ 2\end{pmatrix} - n_H\begin{pmatrix}3/10\\ 1/2\\ 0\end{pmatrix}. \tag{A25}$$

The two-loop effects contribute to the running of the couplings through the terms $\theta_i(Q,Q_0)$ in Eq. (27). Since we do not have the full mass-dependent contributions, we will have to settle with the using the usual massless limits. These are explicitly determined by solving the two-loop renormalization group equations and are given by

$$\theta_i(Q,Q_0) = -\frac{1}{4\pi}\sum_{j=1}^{3}\frac{\beta_{ij}}{\beta_j}\log\left(1+\alpha_j(Q_0)\frac{\beta_j}{4\pi}\log(Q^2/Q_0^2)\right), \tag{A26}$$

where the beta matrix is

$$\beta_{ij}^{\mathrm{MSSM}} = -\begin{pmatrix}7.96 & 5.4 & 17.6\\ 1.8 & 25 & 24\\ 2.2 & 9 & 14\end{pmatrix}. \tag{A27}$$

[1] D. A. Ross, Nucl. Phys. **B140**, 1 (1978).

[2] D. C. Kennedy and B. W. Lynn, Nucl. Phys. **B322**, 1 (1989); B. W. Lynn, ''Running Couplings In SU(2)-LXU(1),'' SLAC-PUB-5077; D. C. Kennedy and B. W. Lynn, ''Grand Unification At Slc/Lep,'' SLAC-PUB-4608.

[3] M. Kreuzer, W. Kummer, and A. Rebhan, Z. Phys. C **44**, 479 (1989); M. Kreuzer and W. Kummer, Phys. Lett. B **197**, 389 (1987); **208**, 530(E) (1988).

[4] A. E. Faraggi and B. Grinstein, Nucl. Phys. **B422**, 3 (1994).

[5] J. Bagger, K. Matchev, and D. Pierce, Phys. Lett. B **348**, 443 (1995).

[6] L. Clavelli and P. W. Coulter, ''SUSY Unification with Smooth Threshold Behavior,'' hep-ph/9507261.

[7] M. Bastero-Gil and J. Perez-Mercader, Nucl. Phys. **B450**, 21 (1995).

[8] M. Bastero-Gil, ''Perturbative Unification and Low Energy Precision Data,'' hep-ph/9610340.

[9] G. Grunberg, Phys. Rev. D **29**, 2315 (1984).

[10] S. J. Brodsky, M. S. Gill, M. Melles, and J. Rathsman, Phys. Rev. D **58**, 116006 (1998).

[11] S. D. Bass, R. J. Crewther, F. M. Steffens, and A. W. Thomas, Phys. Rev. D **68**, 096005 (2003); **66**, 031901(R) (2002).

[12] For an early attempt at constructing electroweak effective charges see L. Baulieu and R. Coquereaux, Phys. Lett. **96B**, 345 (1980).

[13] J. M. Cornwall, Phys. Rev. D **26**, 1453 (1982).

[14] G. Degrassi and A. Sirlin, Phys. Rev. D **46**, 3104 (1992).

[15] N. J. Watson, Nucl. Phys. **B494**, 388 (1997).

[16] J. Papavassiliou, E. de Rafael, and N. J. Watson, Nucl. Phys. **B503**, 79 (1997).

[17] N. J. Watson, Nucl. Phys. B (Proc. Suppl.) **74**, 341 (1999).

[18] N. J. Watson, Nucl. Phys. **B552**, 461 (1999).

[19] J. Papavassiliou, Phys. Rev. Lett. **84**, 2782 (2000).

[20] D. Binosi and J. Papavassiliou, Phys. Rev. D **66**, 111901(R) (2002).

[21] D. Binosi and J. Papavassiliou, Nucl. Phys. B (Proc. Suppl.) **121**, 281 (2003).

[22] D. Binosi, ''Electroweak pinch technique to all orders,'' hep-ph/0401182.

[23] J. Schwinger, *Particles, Sources, and Fields* (Addison-Wesley, Redwood City, CA, 1989), Vol. 1.
[24] One way to see this is by looking at the mass-dependent analytic two-loop results for the related heavy quark effective charge: M. Melles, Phys. Rev. D **62**, 074019 (2000); S. J. Brodksy, M. Melles, and J. Rathsman, *ibid.* **60**, 096006 (1999).
[25] Nir Polonsky, Ph.D. dissertation, University of Pennsylvania, 1994.
[26] V. Lucas and S. Raby, Phys. Rev. D **54**, 2261 (1996).
[27] K. R. Dienes, Phys. Rep. **287**, 447 (1997).
[28] V. S. Kaplunovsky, Nucl. Phys. **B307**, 145 (1988).
[29] T. Friedmann and E. Witten, Adv. Theor. Math. Phys. **7**, 577 (2003).