

This is the accepted manuscript made available via CHORUS. The article has been published as:

# Asymptotic scaling and continuum limit of pure SU(3) lattice gauge theory

Bernd A. Berg

Phys. Rev. D **92**, 054501 — Published 8 September 2015

DOI: [10.1103/PhysRevD.92.054501](https://doi.org/10.1103/PhysRevD.92.054501)

# Asymptotic scaling and continuum limit of pure SU(3) lattice gauge theory

Bernd A. Berg

*Department of Physics, Florida State University, Tallahassee, FL 32306-4350, USA*

(Dated: August 5, 2015)

Recently the Yang-Mills gradient flow of pure SU(3) lattice gauge theory has been calculated in the range from  $\beta = 6/g_0^2 = 6.3$  to 7.5 (Asakawa et al.), where  $g_0^2$  is the bare coupling constant of the SU(3) Wilson action. Estimates of the deconfining phase transition are available from  $\beta = 5.7$  to 6.8 (Francis et al.). Here it is shown that the entire range from 5.7 to 7.5 is well described by a power series of the lattice spacing  $a$  times the lambda lattice mass scale  $\Lambda_L$ , using asymptotic scaling in the 2-loop and 3-loop approximations for  $a\Lambda_L$ . In both cases identical ratios for gradient flows versus deconfinement observables are obtained. Differences in the normalization constants with respect to  $\Lambda_L$  give a handle on their systematic errors.

PACS numbers: 11.15.Ha

## I. INTRODUCTION

We consider pure SU( $N$ ),  $N = 3$ , lattice gauge theory (LGT) with the Wilson action (see, e.g., [1])

$$S = -\frac{\beta}{N} \text{Re} \sum_p \text{Tr} U_p, \quad \beta = \frac{2N}{g_0^2}, \quad (1)$$

where the sum is over all plaquettes of a 4D hypercubic lattice with periodic boundary conditions.  $U_p$  is the SU( $N$ ) plaquette variable,  $g_0^2$  the bare coupling constant and  $\beta$  the usual convention, which emphasizes the interpretation as a 4D statistical mechanics, but gives up the  $\beta = 1/(kT)$  relation with the physical temperature. Namely,  $T = 1/(aN_\tau)$  holds in LGT, where the integer  $N_\tau$  is the extension of the lattice in Euclidean time and  $a$  is the lattice spacing.

For every physical observable  $m$  with the dimensions of a mass the relation

$$m = c_m \Lambda_L \quad (2)$$

holds in the continuum limit  $a(\beta) \rightarrow 0$  for  $\beta \rightarrow \infty$ , where  $\Lambda_L$  sets the mass scale of the lattice regularization and  $c_m$  are calculable constants. Their actual computation faces difficulties, because one has to rely on simulations at finite lattice spacings  $a(\beta)$ , introducing corrections to the continuum relation. The subject of a good reference scale arises. This topic gained renewed interest after Lüscher [2] introduced the Yang-Mills gradient flow scale,  $\sqrt{t_0}$ , which comes by now in several variants. As anticipated by Sommer in his review of the subject [3], gradient scales allow for an unprecedented precision, when compared with traditional scales like  $r_0$  or  $r_c$  [4] defined by the force between static quarks at intermediate distance.

In recent work Asakawa et al. [5] pushed estimates for gradient scales in SU(3) gauge theory all the way up to  $\beta = 7.5$ . The SU(3) deconfining phase transition defines another precise scale, second only to gradient scales. Francis et al. [6] managed to extend estimates of the SU(3) transition temperature  $T_t$  from lattice sizes of previously  $N_\tau \leq 12$  up to  $N_\tau = 22$ ,  $\beta_t = 6.7986$  (65).

Remarkably, neither Asakawa et al. nor Francis et al. fit the  $\beta$  dependence of their estimates so that there is a  $\beta \rightarrow \infty$  continuum limit as predicted by the universal part of asymptotic scaling. Instead, a parametrization for a limited  $\beta$  range is used and the continuum limit of ratios is subsequently estimated by fits in variables like  $(a/r_0)^2$ ,  $(a/\sqrt{t_0})^2$  and so on. This is in accord with a majority of publications on the subject, which all have given up on approaching the asymptotic scaling limit.

Reasons for this, and why the decision to give up on asymptotic scaling may have been premature, are outlined in section II. Inspired by an earlier approach of Allton [7], we are led to write the corrections to the mass relation (2) as a simple Taylor series in the lattice spacing times the lambda lattice mass scale,  $a\Lambda_L$ . In section III this is seen to yield excellent results for fitting the data of Ref. [5] and [6] (see the abstract). Summary and conclusions follow in the final section IV.

## II. ASYMPTOTIC SCALING AND CONTINUUM LIMIT

The realization that the continuum limit of LGT may not just in theory but in practice be reached by computer simulations started with a paper by Creutz [8], where he observed for the SU(2) string tension  $\kappa$  a cross-over from its strong coupling behavior  $a^2\kappa = -\ln(\beta/4)$  to the 1-loop asymptotic scaling behavior  $a^2\kappa = c_\kappa \exp(-6\pi^2\beta/11)$ .

As the accuracy of Markov chain Monte Carlo calculations improved, it was soon realized that there were, in particular for SU(3) with the Wilson action, strong violations of the asymptotic scaling relation and this did not improve noticeably by moving from the 1-loop to the 2-loop relation

$$a\Lambda_L = f_{as}^0(g_0^2) = (b_0 g_0^2)^{-b_1/2b_0^2} \exp\left(-\frac{1}{2b_0 g_0^2}\right), \quad (3)$$

where  $b_0 = 11 N/(48\pi^2)$  and  $b_1 = (34/3) N^2/(16\pi^2)^2$  are, respectively, the universal 1-loop [9, 10] and 2-loop

[11, 12] coefficients of asymptotic freedom, called asymptotic scaling in our context. Universal means that all renormalization schemes lead to the same  $b_0$  and  $b_1$  coefficients.

Next, the hope appeared to be that the situation would improve by including further, non-universal, terms of the expansion of  $a\Lambda_L$ :

$$a\Lambda_L = f_{as}(g_0^2) = f_{as}^0(g_0^2) \left( 1 + \sum_{j=1}^{\infty} q_j g_0^{2j} \right). \quad (4)$$

Computing up to 3-loops, Allés et al. [13] calculated  $q_1$  for SU(N) LGT,

$$q_1 = 0.1896 \text{ for SU}(3). \quad (5)$$

But, the discrepancies between the asymptotic scaling equation and data for physical quantities did not improve.

Assuming that lattice artifacts are responsible for the disagreements, Allton [7] suggested to include such corrections while constraining them with results from perturbative expansions of the considered operators and actions. Doubting, due to uncertainties with the very definition of non-trivial continuum functional integrals, that perturbative information beyond Eq. (4) is reliable, a general Taylor series expansion in  $a\Lambda_L$  is proposed here for corrections to Eq. (2),

$$m = c_m \Lambda_L \left( 1 + \sum_{i=1}^{\infty} \hat{a}_i (a\Lambda_L)^i \right), \quad a\Lambda_L = f_{as}(g_0^2), \quad (6)$$

where one has to determine the normalization constants  $c_m$  and the expansion coefficients  $\hat{a}_i$  by computer simulations. This has the potential to eliminate the essential singularity of the perturbative expansion at  $g_0^2 = 0$ . However, the full sum (4) for  $f_{as}(g_0^2)$  is not available. Instead, we have to work with approximations and define for  $q = 0, 1, \dots$

$$a\Lambda_L^q = f_{as}^q(\beta) = f_{as}^0[g_0^2(\beta)] \left( 1 + \sum_{j=1}^q q_j [g_0^2(\beta)]^j \right), \quad (7)$$

where we have presently the  $q = 0$  (2-loop) and  $q = 1$  (3-loop) asymptotic scaling functions  $f_{as}^q$  at our disposal and a conjecture for  $q_2$  if we believe in the Padé approximation made in Ref. [14]. It is instructive to consider the deconfining temperature  $T_t$  as reference scale. Then  $a(\beta_t) = 1/[N_\tau(\beta_t)T_t]$  implies  $\Lambda_L^q(\beta) = f_{as}^q(\beta)N_\tau(\beta)T_t$ .

Now, if the analyticity (6) is true when using the full  $f_{as}$ , it cannot be true at finite  $q$ . This is, for instance, seen by assuming that the expansion (6) is correct for  $f_{as}^1$  and comparing it with the same expansion using  $f_{as}^0$ . The difference lies in terms of the form

$$(f_{as}^0)^i \left[ (1 + q_1 g_0^2)^i - 1 \right]. \quad (8)$$

Expressing  $g_0^2$  by  $f_{as}^0$  gives rise to powers of logarithms like  $1/\ln(f_{as}^0)$ ,  $\ln|\ln(f_{as}^0)|$  and so on, which are singular for  $f_{as}^0 \rightarrow \infty$ . Nevertheless, we continue to use (6) with  $f_{as}$  replaced by  $f_{as}^q$  and come back to these issues after presenting the fits.

In the following we consider observables with the dimension of a length,  $L \sim 1/m$  and rewrite (6) as

$$\frac{L_k}{a} = c_k \left[ a\Lambda_L \left( 1 + \sum_{i=1}^{\infty} \hat{a}_i (a\Lambda_L)^i \right) \right]^{-1} \quad (9)$$

$$= \frac{c_k}{f_{as}(g_0^2)} \left( 1 + \sum_{i=1}^{\infty} a_i [f_{as}(g_0^2)]^i \right), \quad (10)$$

where  $a_i$  are the parameters with which we deal in our fits. There is no strong reason for using the expansion (10) instead of (9). It just developed this way out of Ref. [7]. To determine the expansion parameters  $a_i$  by numerical calculations one has to truncate the sum at rather small values of  $i$ . For sufficiently large  $\beta$  this should work well because  $(a\Lambda_L^q)$  falls for all  $q$  exponentially off with  $\beta \rightarrow \infty$ . We define the truncated functions,

$$l_\lambda^{p,q}(\beta) = \frac{1}{f_{as}^q(\beta)} + \sum_{i=1}^p a_i^{p,q} [f_{as}^q(\beta)]^{i-1}, \quad (11)$$

with  $f_{as}^q$  given by (7) and fit data according to

$$\frac{L_k}{a c_k^{p,q}} = l_\lambda^{p,q}(\beta), \quad (12)$$

where the 2-loop ( $q = 0$ ) and 3-loop ( $q = 1$ ) asymptotic scaling functions,  $l_\lambda^{0,0}$  and  $l_\lambda^{0,1}$ , are explicitly known (7). The labels  $p, q$  on the normalization constants  $c_k$  and parameters  $a_i$  indicate that their values depend on the choice of  $p, q$ . For simplicity the labels will be dropped when the association is obvious.

For  $q = 0$  as well as for  $q = 1$  it turns out that excellent fits are obtained using  $p = 3$  parameters  $a_i$  besides the  $c_k$  normalization constants. In the following we present  $l_\lambda^{3,q}$ ,  $q = 0, 1$ , expansions for the Yang-Mills gradient flow data [5] and for the deconfining transition estimates [6].

### III. ANALYSIS OF THE NUMERICAL DATA

For the gradient length scale a dimensionless variable  $t^2\langle E(t) \rangle$  is measured as a function of  $t$ . Then  $t_X$  at which the observable takes a specific value  $X$  is used as reference scale. An operator whose  $t$  dependence has been extensively studied is  $E(t) = F_{\mu\nu}^a F_{\mu\nu}^a / 4$ , where  $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]$  is the field strength. In Ref. [5] solutions to the equations

$$t^2\langle E(t) \rangle|_{t=t_X} = X \quad \text{and} \quad t^2 \frac{d}{dt} t^2\langle E(t) \rangle \Big|_{t=w_X^2} = X \quad (13)$$

have been calculated for  $X = 0.2, 0.3$  and  $0.4$ . The associated length scales are  $\sqrt{t_{0.2}}, \sqrt{t_{0.3}}, \sqrt{t_{0.4}}$  and, introduced in [15],  $w_{0.2}, w_{0.3}, w_{0.4}$ . For adaption to Eq. (12)

TABLE I: Error bars in percent of the signal,  $100 \Delta L_k / L_k$ .

$\beta$	$L_1$	$L_2$	$L_3$	$L_4$	$L_5$	$L_6$	$\beta_t$	$N_\tau$
6.3	0.09	0.11	0.12	0.16	0.17	0.22	5.69275	0.07
6.4	0.07	0.09	0.08	0.11	0.12	0.14	5.89425	0.05
6.5	0.13	0.16	0.19	0.22	0.21	0.24	6.06239	0.06
6.6	0.12	0.14	0.16	0.19	0.21	0.23	6.20873	0.07
6.7	0.26	0.33	0.35	0.40	0.46	0.49	6.33514	0.06
6.8	0.18	0.22	0.25	0.27	0.30	0.32	6.4473	0.25
6.9	0.46	0.57	0.65	0.73	0.81	0.87	6.5457	0.54
7.0	0.14	0.17	0.19	0.21	0.25	0.26	6.6331	0.26
7.2	0.43	0.52	0.59	0.65	0.71	0.75	6.7132	0.34
7.4	0.30	0.34		0.41	0.50		6.7986	0.84
7.5	0.37			0.62				
$n_k$	11	10	9	11	10	9		10

they are renamed into  $L_1, \dots, L_6$  according to the first two rows of Table I. Their estimates are given in Table 1 of [5] and are not reproduced here. Instead, we give in our Table I error bars in percent of the signal,  $100 \Delta L_k / L_k$ , for the data tagged by a \* in their paper, i.e., used in their analysis.

Estimates of the SU(3) deconfining phase transition couplings  $\beta_t$  are given in Table I of Ref. [6]. Whenever (for smaller lattices) a comparison is possible their estimates are consistent with previous work [16, 17]. The lengths associated with the deconfining phase transition temperatures  $T_t$  are  $1/(aT_t) = N_\tau$ . However, the statistical errors are in  $\beta_t$  with  $N_\tau$  fixed. To allow for direct comparison with the other quantities, we attach to  $N_\tau$  error bars relying on the later estimated  $l_\lambda^{3,1}(\beta)$  scaling behavior from all data sets

$$\Delta N_\tau = N_\tau \left[ l_\lambda^{3,1}(\beta_t + \Delta\beta_t) - l_\lambda^{3,1}(\beta_t) \right] / l_\lambda^{3,1}(\beta_t). \quad (14)$$

Starting with a guess and iterating the fit, one finds rapid convergence to the relative errors compiled in the  $L_7$  column of Table I. They are less than 0.25 for  $\beta_t \leq 6.33514$  ( $N_\tau \leq 12$ ) and  $\geq 0.25$  for  $\beta_t \geq 6.4473$  ( $N_\tau = 14, \dots, 22$ ), implying that the fit parameters will be dominated by the smaller  $\beta_t$  values. This is not good as the truncated parts of our expansion (11) become more important at smaller  $\beta$ . Therefore, we adjust the  $L_7$  error bars for the lower  $N_\tau$  to  $100 \Delta L_7 / L_7 = 0.2$ , which is still smaller than the best of the relative errors at the higher  $N_\tau$  values.

For the gradient flow data the bias from smaller relative errors is less severe and with  $\beta = 6.3$  the smallest  $\beta$  is not so small. No adjustments are made in that case.

The  $\chi_{dof}^2$  values of our fits (12) to the seven length scales are compiled in Table II ( $n_{dof} = n_k - 4$  with  $n_k$  given in the last row of Table I). All fits are in very good agreement with the data. Actually, the fits of the gradient flows are in too good agreement. This could be an accident, measurements of  $L_1$  to  $L_6$  were performed on the same configurations so that they are all correlated,

TABLE II:  $\chi_{dof}^2$  for our fits to each of the length scales.

$q$	$L_1$	$L_2$	$L_3$	$L_4$	$L_5$	$L_6$	$L_7$
0	0.46	0.34	0.23	0.40	0.47	0.39	0.76
1	0.42	0.32	0.24	0.38	0.46	0.39	0.74

or their error bars are systematically somewhat too large.

For a visual presentation we have combined the entire  $n = n_1 + \dots + n_7 = 70$  data into two  $l_\lambda^{3,q}(\beta)$ ,  $q = 0, 1$ , fits for  $L_k/(ac_k)$ , which works astonishingly well. This is done with an extension of the method of [18]. The constants  $c_k$  are defined as functions  $c_k(a_1, a_2, a_3; data)$ , which give the exact minimum of the fit for the particular constants  $a_i$ , effectively reducing the fitting procedure to three parameters, though the  $c_k$  are still counting against the degrees of freedom. The number of  $a_i$  parameters is reduced by  $6 \times 3$  to 3 from the  $7 \times 3$   $a_i$  parameters used altogether for the fits of Table II.

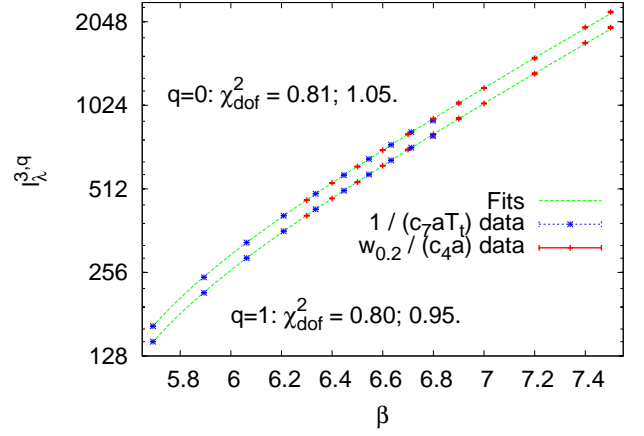


FIG. 1: Data for  $L_7/(c_7a)$  and  $L_4/(c_4a)$  versus the  $l_\lambda^{3,q}$  fits (12) using the 2-loop  $f_{as}^0$  ( $q = 0$ ) and the 3-loop  $f_{as}^1$  ( $q = 1$ ) asymptotic scaling functions.

In Fig. 1 the two fits are shown jointly with the data points ( $i = 1, \dots, n_k$ )

$$\frac{L_k}{a c_k}(i) \pm \frac{\Delta L_k}{a c_k}(i) \quad \text{for } k = 4, 7. \quad (15)$$

Both fits cover with splendid  $\chi_{dof}^2$  values the impressive range  $5.69275 \leq \beta \leq 7.5$ . One value of  $k$  is picked for the gradient flow, because on the scale of the figure the data for the other  $L_k/(ac_k)$  lie right on top of them. For each  $q$  the first  $\chi_{dof}^2$  value is for a fit that excludes the  $1/(aT_t)$  deconfinement data and the second  $\chi_{dof}^2$  value for the shown fit, which includes them. However, the increase from the  $\chi_{dof}^2$  values of Table II should be noted. This and the fact that the data of  $L_1$  to  $L_6$  are all correlated, as well as our “improvement” of the deconfinement data, may well obscure differences of the  $a_i$  parameters for distinct observables. In fact, it is obvious from Fig. 4 (right)

of Ref. [5] that correlations greatly reduce the error bars of ratios and that  $\sqrt{t_{0.3}}/w_{0.4}$  is not entirely flat as in our fits. To take these correlations into account one would best jackknife our fits, which requires the original time series. In the present context of simply demonstrating the almost identical scaling of all data graphically this would just be a distraction. Generally, one expects the  $a_1$  parameters to agree for all  $L_k$ , so that corrections to ratios are of order  $(a\Lambda_L^q)^2$ . There is no reason for  $a_2$  or  $a_3$  to agree for all  $L_k$ . Only, it can be enforced within the accuracy of the present data. When these fits are applied to a single data set there is then a small bias due to the input of the other data sets.

To make Fig. 1 reproducible, the fit parameters are given with high precision in Table III. More decent values are obtained when one redefines the expansion parameters  $a\Lambda_L^q$  by multiplicative constants, e.g., so that they become 1 at  $\beta = 6$ ,  $x^q(\beta) = f_{as}^q(\beta)/f_{as}^q(6)$ . The second row of Table III gives the fits parameters for this case with their error bars in the third row. The range covered by the  $x^q(\beta)$  goes from  $x^q(5.7) \approx 1.4$  down to  $x^q(7.5) \approx 0.18$ , so that  $x^q(7.5)^2 \approx 0.032$  and  $x^q(7.5)^3 \approx 0.0058$  become really small.

TABLE III: Fit parameters used for Fig. 1.

$a_1^{3,0}$	$a_2^{3,0}$	$a_3^{3,0}$	$a_1^{3,1}$	$a_2^{3,1}$	$a_3^{3,1}$
-155.559	24615.3	-5834850	-104.735	9926.28	-2673493
-0.365	0.135	-0.754	-0.292	0.773	-0.581
(13)	(20)	(82)	(13)	(42)	(82)

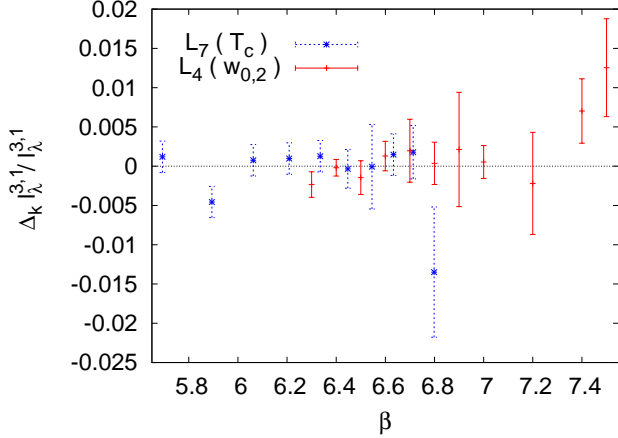


FIG. 2: Relative deviations (16) of the  $L_7$  and  $L_4$  data points from the  $l_\lambda^{3,1}$  fit.

Fig. 2 provides a visual impression for the quality of the fits, by plotting the deviations of the  $k = 4$  and  $7$  data points from the  $q = 1$  fit of Fig. 1 in the form

$$\frac{\Delta_k l_\lambda^{3,1}(i)}{l_\lambda^{3,1}(\beta_i)} \quad \text{with} \quad \Delta_k l_\lambda^{p,q}(i) = \frac{L_k}{a c_k}(i) - l_\lambda^{p,q}(\beta_i) \quad (16)$$

together with error bars  $\Delta L_k/(a c_k l_\lambda^{3,1})$ .

TABLE IV: Normalization constants  $c_k \times 100$ .

$q$	0	1	0	1
$c_1$	0.4918 (37)	0.5569 (41)	0.492 (06)	0.557 (07)
$c_2$	0.6198 (46)	0.7018 (52)	0.619 (11)	0.701 (12)
$c_3$	0.7152 (53)	0.8099 (59)	0.695 (25)	0.789 (28)
$c_4$	0.5392 (40)	0.6106 (45)	0.548 (10)	0.620 (11)
$c_5$	0.6304 (47)	0.7139 (52)	0.638 (16)	0.722 (17)
$c_6$	0.7028 (53)	0.7958 (59)	0.679 (34)	0.771 (38)
$c_7$	2.4404 (71)	2.7754 (79)	2.357 (46)	2.693 (51)

Perhaps surprisingly, instead of one satisfactory description of the data we got two (seven more pairs for the fits with their  $\chi_{dof}^2$  values listed in Table II). The quality of the fits does not care about the log corrections discussed after Eq. (8). Instead, the parameters adjust and the normalization constants  $c_1$  to  $c_7$  get shifted as shown in Table IV. Here the numbers in column 2 and 3 correspond to the joint fits of the six gradient flow operators, with exception of the last row, which corresponds to the fits displayed in Fig. 1 for which all seven operators are combined. Columns 4 and 5 give the results obtained from individual fits to which one should fall back when it comes to conservative estimates. Normalization constants of corresponding  $q = 0, 1$  fits differ by about 12%, while their statistical errors are much smaller.

TABLE V: Ratios of normalization constants.

$q$	0	1	0	1
$c_7^1$	0.19728 (22)	0.19724 (22)	0.209 (05)	0.207 (05)
$c_7^2$	0.24861 (28)	0.24856 (28)	0.263 (07)	0.260 (07)
$c_7^3$	0.28689 (33)	0.28683 (33)	0.295 (12)	0.293 (12)
$c_7^4$	0.21630 (26)	0.21625 (26)	0.233 (07)	0.230 (06)
$c_7^5$	0.25288 (32)	0.25283 (32)	0.271 (09)	0.268 (09)
$c_7^6$	0.28188 (37)	0.28182 (37)	0.288 (16)	0.286 (15)

For ratios,  $c_i^k = c_k/c_i$ , of the normalization constants these differences become tiny and are swallowed by the statistical error bars as is seen in Table V for  $c_7^k$  (columns are arranged as in Table IV). The deconfining transition is used as reference scale, because  $L_7$  is statistically independent from  $L_1$  to  $L_6$ . The estimates of the last row can be compared with Asakawa et al. [5]. Using  $q = 1$ , our values  $c_7^6 = w_{0.4}T_t = 0.28182(37)$  and  $0.286(15)$  are both well consistent with  $0.285(5)$  as given in their Table 3. Our value from column 3 is inconsistent with the precise estimate given in their Eq. (3.2),  $0.2826(3)$ . The discrepancy may be well explained by the small bias of our result and/or the fact that Asakawa et al. rely entirely on  $N_\tau = 12$ , whereas here a continuum fit is used that gives weight to all lattices, including  $N_\tau = 14$  to  $22$ .

The  $\chi_{dof}^2$  of the fits (12) are not sensitive to including or not including the  $q_1 g_0^2$  term into the scaling function

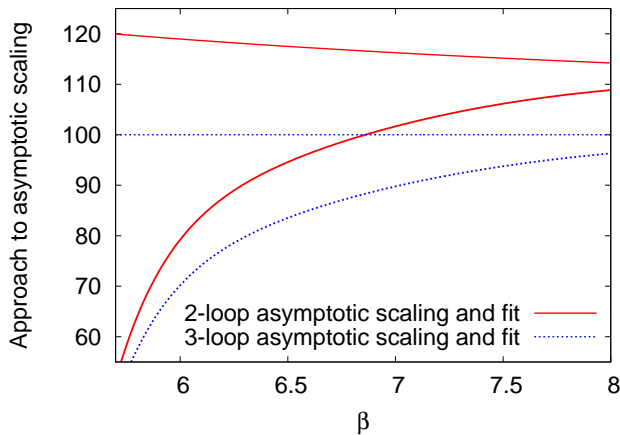


FIG. 3: Approach of the  $l_\lambda^{3,q}$  fits to the asymptotic 2-loop and 3-loop scaling functions  $l_\lambda^{0,q}$  ( $q = 0, 1$ ) times  $100/l_\lambda^{0,1}$ .

(7), while there is a remarkable shift in the normalization constants. It is then tempting, but entirely wrong, to argue that the  $g_0^2$  dependence is so weak that it does not matter and one could replace  $q_1 g_0^2$  by a constant, say  $q_1 g_0^2 \rightarrow q_1^c = 0.9 q_1$  for our  $\beta$  range. It is easy to see that with this, or any other  $q_1^c$ , the normalization constants of the  $q = 0$  fits will not change at all. So, the shift in the normalization constants comes entirely from the  $g_0^2$  dependence of the  $q_1$  term. These contributions re-sum in a way that they become for large  $\beta$  responsible for the difference between  $l_\lambda^{0,1}$  and  $l_\lambda^{0,0}$ .

We use our fits of all  $n = 70$  data to illuminate the situation by Fig. 3, where for  $q = 0, 1$  the inverse asymptotic scaling functions  $l_\lambda^{0,q}$  and their  $l_\lambda^{3,q}$  fits are plotted times  $100/l_\lambda^{0,1}$ , i.e., as fractions of the inverse 3-loop asymptotic scaling function  $l_\lambda^{0,1}$ . We see that the gap between the  $l_\lambda^{0,0}$  and  $l_\lambda^{0,1}$  asymptotic scaling functions narrows slowly and the fits  $l_\lambda^{3,0}$  and  $l_\lambda^{3,1}$  approach rapidly (exponentially fast for increasing  $\beta$ ) their respective asymptotic behaviors, where the  $l_\lambda^{3,1}$  fit stays closer to its asymptotic form than the  $l_\lambda^{3,0}$  fit:  $l_\lambda^{0,1}/l_\lambda^{3,1} \approx 0.8 l_\lambda^{0,0}/l_\lambda^{3,0}$  over the entire  $\beta$  range of the figure..

How does it come that the data cannot figure out whether the  $q = 0$  or  $q = 1$  fit is better? The answer lies in their ratios: If the ratio of the two fits is a constant, the difference between them will be entirely absorbed by the normalization. Defining the change in the ratios with respect to  $\beta = 6$  as reference point by

$$d^p(\beta) = 100 \left( 1 - \frac{l_\lambda^{p,0}(\beta)/l_\lambda^{p,1}(\beta)}{l_\lambda^{p,0}(6)/l_\lambda^{p,1}(6)} \right), \quad (17)$$

we find for the asymptotic scaling ( $p = 0$ ) functions a change by 3.2% at  $\beta = 7.5$ . With 0.16% it is twenty times smaller for the fits ( $p = 3$ ).

What is then the effect of including more and more  $q_j$  terms in the expansion (4) of  $f_{as}$ ? We may expect

convergence of the resulting normalization constants  $c_k$  towards their correct value. But how fast? Repeating the fits of all data with fake  $f_{as}^2$  functions (7) defined by  $q_2 = \pm 0.19$ , so that  $q_2$  has a similar absolute value as  $q_1$ , there is again no sensitivity of the  $\chi_{dof}^2$  of the fits for the additional term and corrections to the  $c_k$  normalization constants stay less than  $\pm 10\%$ . On this basis we end up with the result that our most reliable estimates of the  $c_k$  are those of column five of Table IV with a mainly systematic uncertainty of  $\pm 10\%$ . From  $c_7$  we get

$$\Lambda_L^1/T_t = c_7 \pm 10\% = 0.0269 \quad (27) \quad (18)$$

in good agreement with Francis et al. [6], who give  $T_t/\Lambda_{\overline{MS}} = 1.24(10)$ . Using standard relations between lambda scales [1] this becomes  $\Lambda_L/T_t = 0.0280$  (25). Similarly, our estimate for  $w_{0.4}\Lambda_L$ ,

$$L_6\Lambda_L^1 = c_6 \pm 10\% = 0.0077 \quad (9), \quad (19)$$

is in agreement with the one of Table 3 of Asakawa et al. [5] and the more accurate value of their Eq. (3.3), which translate, respectively, into  $w_{0.4}\Lambda_L = 0.00809$  (35) and  $w_{0.4}\Lambda_L = 0.00829$  (5).

When we believe in the Padé approximation of [14], we find  $q_2 = -0.02467$ , which is in magnitude almost ten times smaller than the range we allowed for our estimate of the systematic error. Using then fits with  $f_{as}^2(\beta)$  as reference, Eq. (18) and (19) improve to

$$\Lambda_L^2/T_t = 0.0266 \quad (9) \quad \text{and} \quad L_6\Lambda_L^2 = 0.00762 \quad (45), \quad (20)$$

where contributions of the statistical errors exceed now the systematic errors. So, it is difficult to understand why the error in Eq. 3.3 of Asakawa et al. is much smaller. Anyway, a small  $q_2$  suggests rapid convergence of the systematic errors of the normalization constants under increasing  $q$  for the used  $f_{as}^q$  functions.

#### IV. SUMMARY AND CONCLUSIONS

It appears that Eq. (6) is a natural parametrization of lattice spacing corrections to the continuum limit of SU(3) LGT. Incorporation of asymptotic scaling is still a viable alternative to other fitting methods for the approach to the continuum limit, which are utilized in [5, 6] and elsewhere. In a next step, our fitting procedure should be tested for other asymptotically free theories, in particular full QCD.

#### Acknowledgments

This work was in part supported by the US Department of Energy under contract DE-FG02-13ER41942. I would like to thank David Clarke for calculating  $q_2$  from the Padé approximation of [14].

- 
- [1] I. Montvay and G. Münster, *Quantum Fields on a Lattice*, Cambridge University Press, 1994.
  - [2] M. Lüscher, JHEP 08, 071 (2010); 03, 092(E) 2014.
  - [3] R. Sommer, POS (Lattice 2013) 015.
  - [4] S. Necco and R. Sommer, Nucl. Phys. B 622 (2002) 328.
  - [5] M. Asakawa, T. Hatsuda, T. Iritani, E. Itou, M. Kitazawa, and H. Suzuki, arXiv:1503.06516v2.
  - [6] A. Francis, O. Kaczmarek, M. Laine, T. Neuhaus, and H. Ohno, Phys. Rev. D 91, 096002 (2015).
  - [7] C.R. Allton, Nucl. Phys. B (Proc. Suppl.) 53, 867 (1997).
  - [8] M. Creutz, Phys. Rev. D **21**, 2308 (1980).
  - [9] D.J. Gross and F. Wilczek, Phys. Rev. Lett. 30, 1343 (1973).
  - [10] H.D. Politzer, Phys. Rev. Lett. 30, 1346 (1973).
  - [11] D.R.T. Jones, Nucl. Phys. B 75, 531 (1974).
  - [12] W. Caswell, Phys. Rev. Lett. 33, 244 (1974).
  - [13] B. Allés, A. Feo and H. Panagopoulos, Nucl. Phys. B 491, 498 (1997).
  - [14] M. Göckeler, R. Horsley, A.C. Irving, D. Pleiter, P.E.L. Rakow, G. Schierholz, and H. Stüben, Phys. Rev. D **73**, 014513 (2006).
  - [15] S. Borsányi, S. Dürr, Z. Fodor, C. Hoebling, S.D. Katz, S. Krieg, T. Kurth, L. Lellouch, T. Lippert, C. McNeile, and K.K. Szabó, JHEP 09, 010 (2012).
  - [16] G. Boyd, J. Engels, F. Karsch, E. Laermann, C. Legeland, M. Lütgemeyer and B. Peterson, Nucl. Phys. B 469, 419 (1996).
  - [17] B.A. Berg and H. Wu, Phys. Rev. D 88, 074507 (2013).
  - [18] B.A. Berg, arXiv:1505.07564, submitted to CPC.