nature

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41586-018-0161-8

In the format provided by the authors and unedited.

A per-cent-level determination of the nucleon axial coupling from quantum chromodynamics

C. C. Chang[1,2], A. N. Nicholson[1,3,4], E. Rinaldi[1,5,6], E. Berkowitz[6,7], N. Garron[8], D. A. Brantley[1,6,9], H. Monge-Camacho[1,9], C. J. Monahan[10,11], C. Bouchard[9,12], M. A. Clark[13], B. Joó[14], T. Kurth[1,15], K. Orginos[9,16], P. Vranas[1,6] & A. Walker-Loud[1,6]*

[1]Nuclear Science Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. [2]Interdisciplinary Theoretical and Mathematical Sciences (iTHEMS) Program, RIKEN, Saitama, Japan. [3]Department of Physics, University of California, Berkeley, CA, USA. [4]Department of Physics and Astronomy, University of North Carolina, Chapel Hill, NC, USA. [5]RIKEN-BNL Research Center, Brookhaven National Laboratory, Upton, NY, USA. [6]Physics Division, Lawrence Livermore National Laboratory, Livermore, CA, USA. [7]Institut für Kernphysik and Institute for Advanced Simulation, Forschungszentrum Jülich, Jülich, Germany. [8]Theoretical Physics Division, Department of Mathematical Sciences, University of Liverpool, Liverpool, UK. [9]Department of Physics, The College of William and Mary, Williamsburg, VA, USA. [10]Physics Department and Astronomy, Rutgers, The State University of New Jersey, Piscataway, NJ, USA. [11]Institute for Nuclear Theory, University of Washington, Seattle, WA, USA. [12]School of Physics and Astronomy, University of Glasgow, Glasgow, UK. [13]NVIDIA Corporation, Santa Clara, CA, USA. [14]Scientific Computing Group, Thomas Jefferson National Accelerator Facility, Newport News, VA, USA. [15]NERSC, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. [16]Theory Center, Thomas Jefferson National Accelerator Facility, Newport News, VA, USA. *e-mail: awalker-loud@lbl.gov

SUPPLEMENTAL INFORMATION

S.1. Data and software availability

An HDF5 file containing the lattice QCD correlation functions generated for this work is made available with the publication of this Letter. Additionally, bootstrap distributions from the correlation function analysis are published in a comma-separated-values (CSV) file along with the Jupyter notebook and chiral-continuum extrapolation Python library used to perform the physical-point extrapolation analysis. These files along with installation instructions for the package may be found at https://github.com/callat-qcd/project_gA[31].

The software used for this work was built on top of the USQCD Chroma software suite[32] and the highly optimised QCD GPU library QUDA[33,34]. We utilise the highly efficient HDF5 I/O Library[35] with an interface to HDF5 in the USQCD QDP++ package, added with SciDAC 3 support (CalLat)[36], as well as the MILC software for solving for HISQ propagators and generating new ensembles. The HPC jobs were efficiently managed with a `bash` job manager, `METAQ`[37], capable of intelligently backfilling idle nodes in sets of nodes bundled into larger jobs submitted to HPC systems. `METAQ` was also developed with SciDAC 3 support (CalLat) and is available at the `git` repository https://github.com/evanberkowitz/metaq.

S.2. Correlation functions from the Feynman-Hellman theorem

The significant advancement of our work can be attributed to an unconventional computational method. This method was developed and implemented on one of the ensembles used in this work (a15m310)[11] and demonstrated on a larger subset in preliminary calculations[38,39]. We briefly summarise the method here as it is central to our final result.

The Feynman-Hellmann (FH) theorem relates matrix elements to linear variations in the energy eigenvalue with respect to an external source, $\partial E_n/\partial\lambda = \langle n|H_\lambda|n\rangle$, where the Hamiltonian of the system is modified appropriately as $H = H_0 + \lambda H_\lambda$. The Heisenberg representation illuminates the fact that correlation functions must be exponentially damped by the energy eigenvalue under time evolution in Euclidean spacetime. Therefore, on the lattice the energy can be determined from the *effective mass*,

$$m^{\text{eff}}(t) = \ln\left(\frac{C(t)}{C(t+1)}\right), \tag{S1}$$

where the spectral decomposition of the two-point correlation function is given by

$$C(t) = \sum_{n=0}^{N_{\text{states}}} z_n z_n^\dagger e^{-E_n t}. \tag{S2}$$

Applying the FH theorem to Eq. (S1) yields the correlation function we construct to obtain the results in this Letter, which we denote the FH ratio:

$$\left.\frac{\partial m_\lambda^{\text{eff}}(t)}{\partial\lambda}\right|_{\lambda=0} = \left[\frac{\partial_\lambda C_\lambda(t)}{C(t)} - \frac{\partial_\lambda C_\lambda(t+1)}{C(t+1)}\right]_{\lambda=0}. \tag{S3}$$

The path integral representation of $C(t)$ and $\partial_\lambda C(t)$ for this calculation may be derived by sourcing the nucleon and current operators into the generating functional, while applying the appropriate derivatives with respect to each source.

There are other implementations of methods motivated by the FH theorem[40–42] which are similar, but our method is the most economical implementation, at least for single nucleon properties. Our method directly calculates the $-\partial_\lambda C(t)$ correlation function[11], without the need to numerically implement the derivative or to disentangle the different orders of the response of the correlation function to the perturbation, as is required by other implementations.

While we are interested in the axial coupling of the ground state nucleon, the nucleon operators couple to an infinite tower of states, and therefore it is customary to filter out the ground state signal by exponentially damping the excited state signals at large time separations. By going to the Heisenberg representation, we derive the complete spectral decomposition of $-\partial_\lambda C_\lambda(t)|_{\lambda=0}$,

$$\begin{aligned}-\partial_\lambda C_\lambda(t) = &\sum_{n=0}^{N_{\text{states}}}\left[(t-1)z_n g_{nn} z_n^\dagger + d_n\right]e^{-E_n t} \\ &+ \sum_{n\neq m=0}^{N_{\text{states}}} z_n g_{nm} z_m^\dagger \frac{e^{-E_n t}e^{\frac{\Delta_{nm}}{2}} - e^{-E_m t}e^{\frac{\Delta_{mn}}{2}}}{e^{\frac{\Delta_{mn}}{2}} - e^{\frac{\Delta_{nm}}{2}}},\end{aligned} \tag{S4}$$

allowing us to analyse correlation functions even at small time separations. Here, g_{nn} for $n = 0$ is the ground state coupling of the nucleon, $z_n z_n^\dagger$ is the non-relativistically normalised relative probability of finding the nucleon in the n^{th} state and $\Delta_{nm} = E_n - E_m$. The linear enhancement of g_{nn} is a direct manifestation of the FH theorem, in which the first derivative of the spectrum (described by the two-point correlation function) is taken, thereby generating a linear moment. Additionally, excited state contributions in the linearly enhanced $n > 0$ terms are analogous to contamination present in standard two-point correlation functions, which are generically well under control. The remaining contamination from lattice artifacts, d_n and the sum over $n \neq m$, are not linearly enhanced and therefore are functionally distinct from the signal of interest and can be cleanly removed. The artifacts, d_n arise from contact terms where the current insertion is at the same time as the nucleon creation or annihilation operators, and also from the time region where the current is earlier or later than the nucleon creation or annihilation operators respectively[11]. At $t = 1$ the contribution from all terms aside from d_n exactly vanish, allowing for a robust estimate of the contributions to $-\partial_\lambda C_\lambda(t)$ from these undesired artifacts.

Inserting Eqs. (S2) and (S4) in to Eq. (S3), it is straightforward to show that in the long-time limit, we recover the ground-state matrix element of interest

$$\lim_{t\to\infty}\left.\frac{\partial m_\lambda^{\text{eff}}(t)}{\partial\lambda}\right|_{\lambda=0} = g_{00}. \tag{S5}$$

The difference in Eq. (S3) leads to an additional suppression of the excited states (and contact terms) beyond the standard exponential suppression by the mass gap. This allows us to make use of numerical data very early in Euclidean time, before the stochastic noise overwhelms the signal[43,44], providing an effective exponential enhancement of the signal for a fixed number of stochastic samples as compared to the standard methods. In Extended Data Fig. 1 and Supplemental Figs. 9–15, we demonstrate our ability to fit early in Euclidean time on all ensembles used in this work.

One potential drawback of our Feynman-Hellmann strategy is that it requires new calculations of Feynman-Hellmann quark propagators for each matrix element or momentum injection of interest. By contrast, the standard methods used in the literature provide the flexibility to study arbitrary quark bi-linear matrix elements between the proton and neutron, with arbitrary momentum injection by the current, without need for additional computational cost. On the other hand, our Feynman-Hellmann method comes with an additional flexibility not present in the standard methods: we can compute the matrix elements of the same quark bi-linear currents in various hadronic states, such as hyperons or multi-nucleon systems, without needing to recompute the quark propagators coupling to the current. Using the standard methods, the sink interpolating operator is fixed, so computing a new sequential propagator is necessary for each final state.

S.3. Lattice action

For this work, we have chosen a *mixed action* (MA) in which the discretisations for the generation of the gauge configurations and solution of the quark propagators differ.[21] Details of tuning

the action and its salient features are summarised here.

A. Action Details

To control the continuum limit, infinite volume and physical pion mass extrapolation, a set of LQCD ensembles with multiple lattice spacings, multiple volumes and near physical pion masses must be used. The only set of publicly-available gauge configurations that satisfy these criteria are the Highly Improved Staggered Quark (HISQ)[45] action ensembles with dynamical light, strange and charm quarks ($N_f = 2+1+1$) generated by the MILC Collaboration[30,46]. They were generated with near-physical values of the strange and charm quark masses and three values of the pion mass, $m_\pi \sim \{130, 220, 310\}$ MeV. For this work, we utilise ensembles with three different lattice spacings of $a \sim \{0.15, 0.12, 0.09\}$ fm. Formally the HISQ action has leading discretisation errors starting at $\mathrm{O}(\alpha_S a^2, a^4)$, however improved link-smearing greatly suppresses taste-changing interactions leading to numerically smaller discretisation errors. The gluons are simulated using the tadpole-improved[47], one-loop Symanzik gauge action[48] with leading discretisation errors starting at $\mathrm{O}(\alpha_S^2 a^2, a^4)$.

We performed a dedicated volume study at $a \sim 0.12$ fm $m_\pi \sim 220$ MeV with three volumes. To control the pion mass extrapolation, we generated six new HISQ ensembles with $m_\pi \sim \{350, 400\}$ MeV on the same three lattice spacings (these ensembles are available to any interested group upon request). Details of the HISQ ensembles are presented in Extended Data Table II.

The valence quark propagators are solved with the Möbius Domain Wall Fermion (MDWF) action[49–51] after applying the gradient-flow smoothing algorithm[52,53] to the HISQ configurations. The Möbius kernel with gradient-flow smoothing reduces the residual chiral symmetry breaking such that $m_{\mathrm{res}} < 0.1 \times m_l^{\mathrm{val.}}$ for moderate values of L_5 and thus the valence action has approximate chiral symmetry[54] as it satisfies the Ginsparg-Wilson relation[55] with small corrections. The values of b_5 and c_5 were chosen (Extended Data Table II) such that the Möbius kernel is a rescaled Shamir kernel[56,57] of the domain-wall action[58] ($b_5 - c_5 = 1$). The calculation of these MDWF propagators required the significant majority of computing cycles and were efficiently solved using the QUDA library[33] with parallel MPI Support[34]. After absorbing m_{res} into the quark mass through the PCAC relation, the MDWF action has discretisation errors beginning at $\mathrm{O}(a^2, \alpha_S a^2)$[59].

B. Computational Details

The valence quarks are tuned so that the MDWF pion mass matched the taste-5 HISQ pion mass within 2%, ensuring a unitary theory in the continuum limit. Multiple sources per configuration are used to increase our statistical samplings. On a given configuration, for a series of evenly spaced time-locations, a seeded random origin is chosen, (x_0, y_0, z_0, t_0) along with its antipode, $(x_0, y_0, z_0, t_0) + L/2(1, 1, 1, 0)$ (modulo the periodic spatial boundary conditions). At each point, a smeared source is generated to solve the MDWF quark propagators using the `SHELL_SOURCE` with the `GAUGE_INV_GAUSSIAN` routine in `Chroma`[32]. The correlation functions are constructed with two choices of sink smearing, a `SHELL_SINK` with parameters matching the source and a `POINT_SINK`. The proton and neutron correlation functions are constructed using the local interpolating operator with the largest overlap on the positive-parity nucleon states[60,61]. We double the statistics by generating analogous correlation functions for the negative parity nucleon. Under a time-reversal transformation the nucleon reverses parity, allowing us to average the forward propagating nucleon correlation functions with the backward propagating negative-parity nucleon correlation functions. Once constructed, all the correlation functions are shifted to $t_0 = 0$ and averaged. We observe no correlation between different sources, resulting in statistical uncertainties inversely proportional to $\sqrt{N_{src}}$ as first observed and studied in detail in previous work[62]. All parameter choices for the valence MDWF action are presented in Extended Data Table II.

We have demonstrated the gradient-flow time (t_{gf}) independence of M_N/F_π and F_K/F_π for our action[21]. In particular, we demonstrate, with a reduced data set, that the extrapolation of F_K/F_π to the physical point in the continuum is independent of flow time and also consistent with the FLAG determination[13]. In this work, we also study the flow-time dependence of g_A. In Figure 1, we show ratios of the axial over the vector FH ratios for the a15m310 and a09m310 ensembles with 196 configurations at a single source. The point-sink (squares) and smeared-sink (circles) FH ratios are plotted with $t_{gf} = \{1.0, 0.6, 0.2\}$ respectively from left to right. In both ensembles, we observe minimal flow time dependence in the ratio of correlators. Additionally the flow time is fixed to $t_{gf} = 1.0$ in lattice units on all gauge configurations, ensuring that any quantity extrapolated to the continuum limit will be flow-time independent. For $t_{gf} = 1.0$, we find it sufficient to solve the gradient flow diffusion equation with 40 integration steps using the Runge-Kutta algorithm. Furthermore, we observe smaller stochastic uncertainty at increasingly larger values of t_{gf} due to the gradient flow suppressing the ultraviolet noise. These conclusions are consistent with the results observed in previous work[21] for other hadronic quantities (*e.g.* Fig. 3 therein).

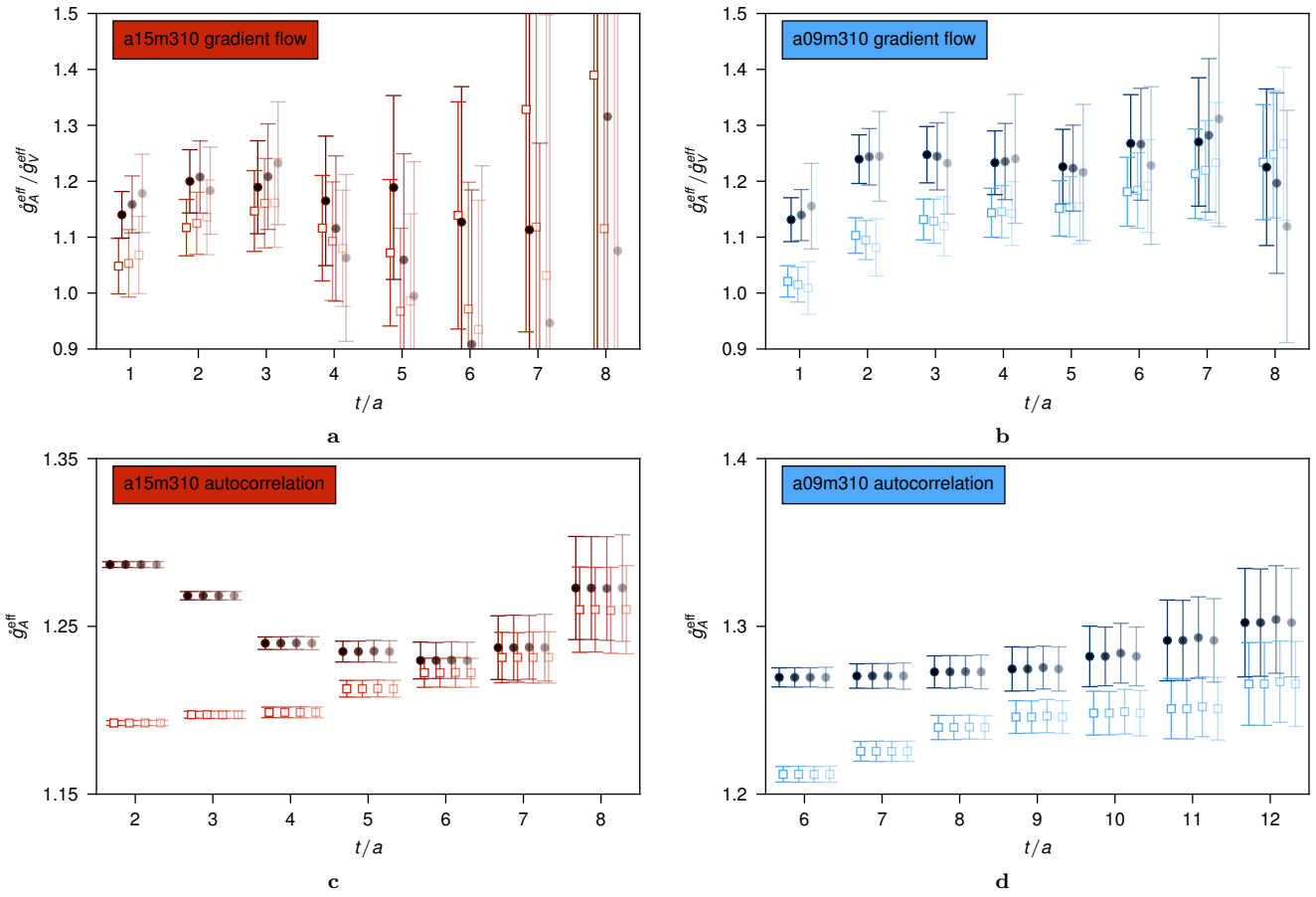
We also study possible autocorrelations in our data set by binning the FH ratio correlation functions for every ensemble used in this work. Extended Data Figure 1 shows a representative example of a binning study. We observe that the standard deviation of the raw correlation function is stable under binning for bin sizes up to four, demonstrating that no autocorrelations are present in the data. The complete binning study is presented in the Supplemental Figs. 22–24. We do not bin any of our data in this work.

S.4. Correlator analysis

The exact wavefunction for the ground state nucleon is unknown, so lattice correlation functions are constructed with interpolating operators for the initial and final states. Therefore, the correlation function describes a superposition of the ground state nucleon of interest and nucleon excited states. Disentangling the ground state from the excited state contamination requires careful analysis of the correlation functions, and has proved to be one of the major challenges for past calculations of g_A. As a result of the unique construction of the lattice correlation function though our Feynman-Hellmann strategy, we have access to measurements of the correlation function at both longer and shorter separations between the initial and final states, allowing for a more complete study of excited state contributions compared to previous works. Additionally, nucleon observables suffer from exponentially severe decay in signal-to-noise, posing a serious challenge for high precision calculations. Compared to previous works, exponentially more precise data (at early times) is leveraged in this analysis to combat the severe decay in the signal.

A. Analysis Strategy

For each ensemble, we perform a simultaneous fit to six correlation functions which include the nucleon two-point correlation function, the vector and axial-vector FH ratios [Eq. (S3)], and for each of these, two different correlation functions corresponding to two different choices of sink smearing for the quark fields. This greatly enhances the amount of correlated data when determining a large subset of shared parameters (*i.e.* E_n and z_n). Strategies for estimating the unknown parameters were previously discussed[11]. In the present work, we first perform a two-state Bayesian constrained fit to explore the parameter space in t, and then take the central value of the resulting posterior distribution as the initial guess to a final two-state unconstrained fit using non-linear χ^2 minimization. Preconditioning the unconstrained fit does not



Supplemental Data Figure 1 | Gradient flow time dependence and Monte Carlo autocorrelation time study. a, b, The flow-time dependence from $t_{gf} = \{1.0, 0.6, 0.2\}$ from left to right (dark to light). We observe flow time independence in the axial to vector correlation function ratios, and smaller stochastic noise at larger flow times. **c, d,** The uncertainty of the mean under bootstrap resampling with successively larger bin sizes of $N_{\text{bins}} = \{1, 2, 3, 4\}$ from left to right (dark to light). We observe the uncertainty of the mean to be unchanged under binning, indicating that there are no autocorrelations present in our data. For all plots, the circle and square data points correspond to the smeared- and point-sink correlation functions. Uncertainties are one s.e.m.

ensemble	t^C_{min}	t^C_{max}	t^A_{min}	t^A_{max}	t^V_{min}	t^V_{max}	$\mathring{g}_A$	$\mathring{g}_V$	χ^2/dof [dof]	P-value	$(Z_A/Z_V - 1) \times 10^5$
a15m400	6	11	3	11	5	11	1.213(06)	0.998(01)	1.3 [24]	0.14	4.20(17)
a15m350	4	12	2	12	4	12	1.195(13)	0.997(01)	1.3 [38]	0.11	3.00(19)
a15m310	4	14	4	10	4	15	1.216(11)	1.001(02)	1.1 [40]	0.32	1.77(24)
a15m220	3	11	4	9	4	9	1.275(13)	1.000(04)	1.5 [22]	0.07	3.7(1.3)
a15m130	2	10	2	6	2	7	1.262(53)	0.994(35)	1.6 [20]	0.05	1.87(52)
a12m400	5	10	5	11	7	11	1.237(10)	1.016(01)	1.5 [16]	0.07	4.22(13)
a12m350	7	14	5	11	5	14	1.255(14)	1.016(01)	0.93 [30]	0.57	2.75(10)
a12m310	5	12	4	12	6	12	1.239(13)	1.021(02)	1.5 [28]	0.06	1.86(13)
a12m220S	4	10	5	10	3	10	1.294(28)	1.018(03)	1.1 [22]	0.35	2.02(37)
a12m220	3	15	4	10	3	14	1.277(15)	1.015(02)	1.1 [44]	0.34	2.02(37)
a12m220L	4	12	4	12	5	10	1.277(21)	1.020(05)	1.4 [28]	0.09	2.02(37)
a12m130	2	13	3	13	2	12	1.318(29)	1.020(08)	1.1 [48]	0.24	0.45(77)
a09m400	8	18	6	15	7	15	1.238(08)	1.023(01)	0.98 [40]	0.50	5.10(10)
a09m350	7	16	8	16	7	14	1.258(15)	1.024(02)	1.3 [34]	0.09	3.40(14)
a09m310	9	16	3	12	7	17	1.266(11)	1.024(01)	1.0 [38]	0.40	2.09(16)
a09m220	9	15	3	14	6	12	1.280(09)	1.022(02)	1.3 [32]	0.09	1.86(17)

Supplemental Data Table I | Correlator fit region and results, and renormalisation coefficients Fit regions for the two-point correlation function $t^C_{\text{min}}, t^C_{\text{max}}$, the axial FH ratio $t^A_{\text{min}}, t^A_{\text{max}}$, and the vector FH ratio $t^V_{\text{min}}, t^V_{\text{max}}$ are given in lattice units. The resulting central value and standard deviation of the bare couplings are given in columns $\mathring{g}_A$ and $\mathring{g}_V$ along with the χ^2/dof and P-value of these fits. The last column gives the values of the ratio of renormalisation coefficients Z_A/Z_V at $\mu \approx 2.8$GeV in the γ_μ scheme.

change the final result, but serves as an effective method to explore large parameter spaces, and minimises the iteration count required for convergence. In principle, preconditioning the unconstrained fit from the posterior distribution obtained from Bayes' theorem provides a strategy for avoiding unphysical local minima in the χ^2 manifold. In hindsight however, the data is well-behaved with relatively sharp minima. Bayesian constrained fits with up to eight states were performed resulting in consistent results[11].

We assess the quality of the candidate fits by first considering only results with P-values greater than 0.05 in order to discriminate against fits of poor quality (e.g. Extended Data Fig. 1e,f). Next, we study the effects of excited state contamination by varying the fit regions over different time separation, and demand that the candidate fit lies in the region of stability (Extended Data Fig. 1e,f). Finally we quantify the uncertainty of our determination of the matrix element by drawing 5000 bootstrap samples, and accept candidates that are Gaussian distributed, as expected from the distribution of the (path integral) mean as a consequence of the central limit theorem (Extended Data Fig. 1e,f). The preferred fit is one which satisfies all the above requirements while sampling the largest fit region, such that we maximise the amount of information extracted from these numerically intensive calculations. As a final check, we overlay the preferred fit on top of the data and observe agreement between model and data.

The list of fit regions and preferred results are given in Supplemental Table I. The complete correlator fit study plots are shown in Supplemental Figs. 8–15, with the fit stability plots show in Supplemental Figs. 16–21 and the bootstrap distributions of the resulting values of ϵ_π in Supplemental Figures 25–27. All stages of the analysis are implemented using the Python library `lsqfit`[63].

B. *Discussion*

Studying ground state stability as a function of $t_{\rm min}$ is the most robust way to demonstrate understanding and control of excited state contributions. Such a study is only possible with the data set of the current work because the Feynman-Hellmann strategy makes all possible source-sink separation times accessible. In contrast, all previous calculations use conventional strategies to generate lattice correlation functions, and thus do not generate enough source-sink separation times to perform such a study.

The results of this study are shown in Extended Data Fig. 1e and Figs. 16–18. Because excited states are always heavier than the ground state, their contributions are more pronounced at smaller source-sink separations. As a result, we observe that choosing smaller values of $t_{\rm min}$ compared to the preferred fit (in solid black) leads to results that are sometimes in tension, indicating that the ground state signal is being contaminated by excited state artifacts at these times. In contrast, the preferred fits are always consistent with more conservative fits which only include data at larger values of $t_{\rm min}$. To further demonstrate the quality of the ground state determination, we numerically subtract (under bootstrap) the excited state contributions determined in the analysis from the raw correlation functions and plot these processed results, along with the asymptotic value of the ground state (Fig. 2, Extended Data Figs. 1a,b,c and Supplemental Figs. 8–15).

Surveying these figures reveals the existence of correlated statistical fluctuations at ~1 fm separation times and is the manifestation of the signal-to-noise problem suffered by nucleon observables. In all previous work, the nucleon axial and vector correlation functions (*e.g.* Extended Data Fig. 1b,c and similar Supplemental figures) are constructed in this intermediate time separation region in order to avoid contamination from excited state contributions, and may be affected by these uncontrolled statistical fluctuations. Furthermore, the limited number of source-sink separation times available in previous calculations makes the *a posteriori* identification of correlated fluctuations extremely difficult.

In contrast, this work analyses data at much smaller separation times, where the signal-to-noise has not yet degraded, in order to extract the ground state signal. As a result, this analysis strategy is robust against random statistical fluctuations. To provide more evidence, Extended Data Fig. 1f and Supplemental Figs. 19–21 demonstrate stability of the extracted ground state nucleon couplings under varying $t_{\rm max}$, showing that the result is insensitive to whether the fluctuations are included as part of the fitted data. Additionally, utilizing data at smaller time separations results in an exponentially more precise determination of the nucleon couplings and is the key to obtaining the sub-percent uncertainty presented in this work.

In summary, the Feynman-Hellmann strategy for constructing lattice correlation functions enables leverage over the full source-sink time dependence in the analysis. We observe stability of the ground state nucleon couplings under varying fit regions, demonstrating control over excited-state contamination. The asymptotic values of the ground state are in good agreement with the excited-state subtracted lattice correlation functions. The final bootstrapped results all yield nearly ideal Gaussian distributions. With this preponderance of evidence, we resolve the two major challenges identified from previous works, and demonstrate full control over systematic uncertainties at unprecedented levels of precision.

S.5. Renormalisation

Discretisation of the Dirac action leads to differences between the local current used in the calculation and the conserved current. We correct for this difference using the non-perturbative Rome-Southampton renormalisation procedure[64], with non-exceptional kinematics[65,66]. Explicitly we compute

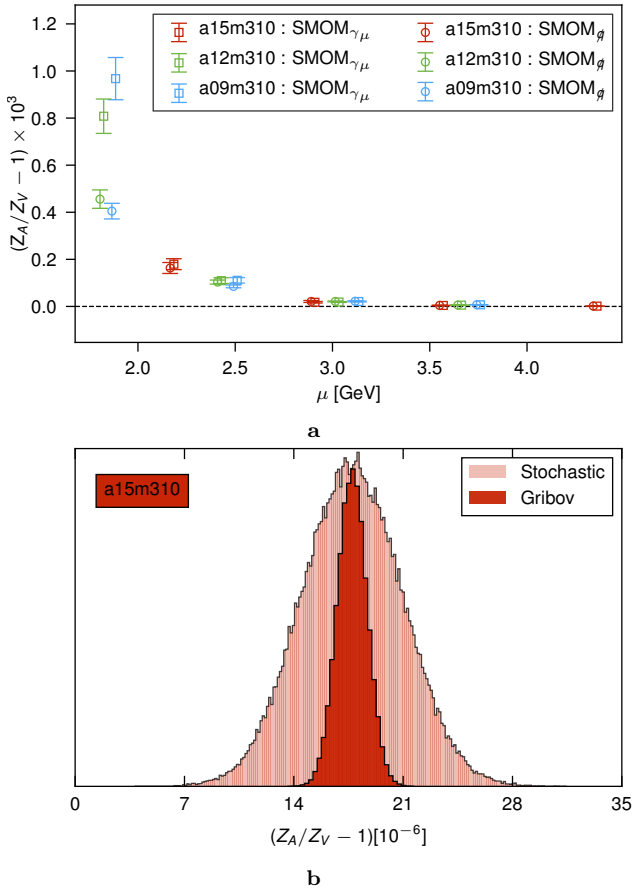
$$g_A = \frac{Z_A}{Z_V}\frac{\mathring{g}_A}{\mathring{g}_V}\,, \tag{S6}$$

where Z_A and Z_V are the renormalisation factors of the axial and vector current, while $\mathring{g}_A$ and $\mathring{g}_V$ correspond to the bare (unrenormalised) couplings. In Eq. (S6), we take advantage of the fact that $Z_V\mathring{g}_V = 1$. Furthermore, because of the good chiral property of our lattice discretisation, we expect $Z_A = Z_V$ up to small artifacts. We have computed these factors in RI-SMOM schemes and with momentum sources, as proposed in[67], resulting in high statistical precision. We observe the ratio of the renormalisation coefficients Z_A/Z_V to be commensurate with unity at one part in 10,000, indicating that the lattice action we use preserves chiral symmetry to very good approximation, see Supplemental Data Fig. 2. In this procedure, the renormalisation scale is given by $\mu = \sqrt{q^2}$ ($q^2 \geq 0$) and q is the vertex momentum transfer. This result, together with the improved stochastic uncertainty gained from simultaneously fitting the vector FH ratio, further reduces the final uncertainty of our result.

Since the quark bi-linear matrix elements used to determine the renormalisation coefficients are not gauge invariant, we perform these calculations in Landau gauge. Landau gauge fixing however, is *incomplete* and the resulting coefficients will be evaluated at one of many Gribov regions[68]. We sample the distribution of renormalisation coefficients over different Gribov regions by repeating the calculation after performing random global gauge transformations to the gauge fields. We observe that the systematic uncertainty from this effect is subdominant to the statistical uncertainty of the renormalisation coefficients (Supplemental Data Fig. 2). We performed a dedicated flow-time study of Z_A/Z_V on a subset of the ensembles and find this ratio is also flow-time independent.

S.6. Parameterization of the chiral, continuum and infinite volume extrapolations

The results of these calculations (Extended Data Table I) must be extrapolated to the physical point. g_A is a dimensionless quantity and therefore the entire extrapolation can be performed by



Supplemental Data Figure 2 | Renormalisation coefficients versus cut-off scale, and Gribov region study. a, Renormalisation coefficients from RI-SMOM for the $a \sim \{0.15, 0.12, 0.09\}$ fm $m_\pi \sim 310$ MeV ensembles as a function of the renormalisation scale μ for intermediate γ_μ and $\not{q}$ schemes in the Landau gauge. The axial coupling is itself a physical observable, and therefore its value is independent of scale. The scale dependence observed for $\mu \leq 2.5$ GeV comes from the infrared (IR) contamination of a light meson, and in principle at high scales (UV) the coefficient receives large $\mathrm{O}(a\mathbf{p})$ corrections. The intermediate region free of IR and UV contaminations is coined the Rome-Southampton window. **b,** The statistical uncertainty of the renormalisation coefficient for the a15m310 ensemble evaluated at $\mu = 2.86$ GeV in the γ_μ scheme is shown by the light red histogram. The Landau gauge admits remnant gauge degrees of freedom resulting in the Gribov distribution shown in dark red. Random global gauge transformations are applied to the gauge fields, and the RI-SMOM perscription is repeated to obtain the Gribov distribution. We observe that the systematic uncertainty coming from Landau gauge fixing to be smaller than the statistical uncertainty. Uncertainties are one s.e.m.

using ratios of physical quantities that form dimensionless variables without the need for performing a scale setting. On each ensemble, we determine the three quantities

$$\epsilon_a^2 = \frac{1}{4\pi}\frac{a^2}{w_0^2}, \quad m_\pi L, \quad \epsilon_\pi = \frac{m_\pi}{4\pi F_\pi}, \tag{S7}$$

which are used to parameterise the continuum, infinite volume and physical pion mass extrapolations. w_0 is a gradient-flow scale that can be precisely and accurately determined[69] and F_π is the pion decay constant (with $F_\pi \sim 92$ MeV normalization). EFT methods can be used to parameterise the dependence upon these variables.

A. *χPT through N3LO*

For a static quantity such as g_A, Heavy Baryon χPT (HBχPT)[18] can be used to parameterise the pion mass dependence. The convergence issues of $SU(3)$ HBχPT[70–74] require two-flavor HBχPT[75] to be used for a controlled extrapolation. The complete pion mass dependence of g_A is known through $\mathrm{O}(m_\pi^3)$[75–77], which is next-to-next-to-leading order (NNLO) in the chiral expansion. In terms of ϵ_π, the pion mass dependence is given by

$$g_A = g_0 + c_2\epsilon_\pi^2 - \epsilon_\pi^2(g_0 + 2g_0^3)\ln(\epsilon_\pi^2) + g_0 c_3 \epsilon_\pi^3, \tag{S8}$$

where g_0, c_2 and c_3 are low-energy constants (LECs) that must be determined in the analysis. In this expression, we have set the χPT renormalisation scale to $\mu = 4\pi F_\pi$. The corrections to using a fixed renormalisation scale enter at $\mathrm{O}(\epsilon_\pi^4)$ which can be seen by expanding F_π/F in the above expression, where $F = \lim_{m_\pi \to 0} F_\pi$.

The complete next-to-next-to-next-to-leading order (N3LO) calculation of g_A has not been determined, however, the $\ln^2(\epsilon_\pi)$ corrections have been determined with a renormalisation group analysis[78]. Even though the complete calculation has not been performed, the full parameterisation of these corrections is given by

$$\delta_\chi^{(4)} g_A = \epsilon_\pi^4 \Big[c_4 + \tilde{\gamma}_4 \ln(\epsilon_\pi^2) + \left(\frac{2}{3}g_0 + \frac{37}{12}g_0^3 + 4g_0^5\right)\ln^2(\epsilon_\pi^2)\Big]. \tag{S9}$$

The extrapolation formula was provided formula in terms of m_π/F and so the difference in the coefficient of the $\ln^2(\epsilon_\pi^2)$ term given here and previously[78] is attributed to using F_π rather than F in the expression. The LEC $\tilde{\gamma}_4$ differs from γ_4[78]. Our data set is not sufficient to use the full N3LO expression as it contains a total of 5 unknown LECs, and we have results at 5 different values of m_π. We do, however, include partial corrections from N3LO, like the c_4 counter term (*i.e.* NNLO+ct), to check the stability of the analysis, and the Bayesian Framework allows us to use the full expression.

B. *Including explicit delta degrees of freedom*

The ϵ_π dependence described above stems from the chiral Lagrangian with only pions and nucleons as explicit degrees of freedom. There are many publications in the literature advocating for the explicit inclusion of the delta resonances in the theory in order to accurately describe properties of the nucleon. While the delta states are strong resonances, in the large-N_c limit[79,80], the splitting between them and the nucleons vanishes. Further, the deltas are strongly coupled to the nucleons and the mass gap between them ($\Delta \equiv M_\Delta - M_N$) is comparable to the pion mass, such that contributions from the delta states to nucleon quantities can be poorly captured without explicitly including them as dynamical states in the EFT[81,82]. In lattice QCD calculations of nucleon quantities, the pion masses are still generally heavier than in nature and for $m_\pi \gtrsim 290$ MeV, the deltas become stable, asymptotic states. Finally, it has been observed that including explicit deltas in the EFT leads to a milder pion mass dependence for g_A[83].

This observation follows straightforwardly from the large-N_c formalism[84–91]. Combining the large-N_c expansion with the chiral expansion leads to an improved perturbative expansion for many quantities, including the baryon spectrum[92], which has been observed numerically with lattice QCD calculations[73,74]. It has been shown that there are cancellations between nucleon and delta virtual corrections for g_A as well, which lead to the milder pion mass dependence[93].

In the present work, an extrapolation including the deltas explicitly is a *phenomenological* extrapolation as there are three new quantities that are required to perform the chiral extrapolation, which we have not determined in our calculation, and therefore, some knowledge from experiment must be used to constrain them.

One must know the delta-nucleon mass splitting, Δ, as well as two additional axial couplings, the $\Delta \to \Delta$ coupling and the $\Delta \to N\pi$ transition coupling which we denote $\mathring{g}_{\Delta\Delta}$ and $\mathring{g}_{N\Delta}$, respectively (the mathrings denote the chiral limit value of these couplings, just as g_0 is the chiral limit value of g_A). These quantities are particularly challenging to compute due to the resonant nature of the delta (for sufficiently light pion masses), and require calculations of not only the external states, but also the πN scattering phase shifts[94–96]. The first lattice QCD calculation of such $1 \to 2$ transitions has only recently been performed for mesons[97–100]. For our mixed-action calculation, this problem is further exacerbated by the non-unitary nature of the theory as these non-unitary effects can go on-shell in the πN scattering system, thus precluding the use of the known formalism[101].

The continuum, infinite volume extrapolation function including deltas in the $SU(2)$ chiral expansion was first determined at NLO, and is given by[83]

$$\begin{aligned} g_A = \; & g_0 + c_2^{\Delta}\epsilon_\pi^2 + \frac{32\pi}{27} g_0 g_{N\Delta}^2 \frac{\epsilon_\pi^3}{\epsilon_\Delta} \\ & - \epsilon_\pi^2 \ln(\epsilon_\pi^2)\left[g_0 + 2g_0^3 + \frac{2}{9} g_0 g_{N\Delta}^2 + \frac{50}{81} g_{\Delta\Delta} g_{N\Delta}^2 \right] \\ & - R\left(\frac{\epsilon_\pi^2}{\epsilon_\Delta^2}\right) g_{N\Delta}^2 \left[\epsilon_\pi^2 \frac{32}{27} g_0 + \epsilon_\Delta^2 \left(\frac{76}{27} g_0 + \frac{100}{81} g_\Delta \right) \right] \\ & - \epsilon_\Delta^2 \ln\left(\frac{4\epsilon_\Delta^2}{\epsilon_\pi^2}\right) \left[\frac{76}{27} g_0 g_{N\Delta}^2 + \frac{100}{81} g_\Delta g_{N\Delta}^2 \right], \end{aligned} \tag{S10}$$

where we have defined

$$\epsilon_\Delta \equiv \frac{\Delta}{4\pi F_\pi}, \tag{S11}$$

and the new non-analytic function is given by

$$R(z) \equiv \begin{cases} \sqrt{1-z} \ln\left(\frac{1-\sqrt{1-z}}{1+\sqrt{1-z}}\right) + \ln(4/z), & z \le 1 \\ 2\sqrt{z-1} \arctan z + \ln(4/z), & z > 1 \end{cases}. \tag{S12}$$

In order to use this extrapolation formula, we take the value of $\Delta \simeq 293$ MeV from the experimental splitting of the nucleons and deltas. As the pion mass is increased above the physical value, this splitting is known to reduce[70], but we do not account for this change in the extrapolation. At this order in the expansion, it is sufficient to pick a single value, with the difference appearing at higher orders in the expansion. In order to constrain the two new axial couplings, we set the central values to those predicted from the large-N_c expansion[88,89]

$$\mathring{g}_{N\Delta} = -\frac{6}{5} g_0, \qquad \mathring{g}_{\Delta\Delta} = -\frac{9}{5} g_0. \tag{S13}$$

Our numerical results are insufficient to constrain these couplings, so to study the sensitivity to them, we include them in the analysis under the Bayesian Framework with prior widths varying between 5% and 40%.

The sign difference between these two axial couplings and g_0 is what leads to the milder pion mass dependence as this sign difference results in axial coefficients of the non-analytic terms in Eq. (S10) being ~ 2 times smaller. Another interesting feature of the delta-full extrapolation is the presence of m_π^3 dependence with a fixed coefficient. This does not appear until NNLO in the delta-less expansion, and with an unknown LEC.

C. *Taylor expansion*

In addition to these two χPT-derived extrapolation functions, a simple Taylor expansion can also be considered. In particular, for a quantity with mild pion mass dependence, such as g_A, the Taylor expansion should provide an adequate description of the quark mass dependence[14]. The most natural parameter for performing the Taylor expansion is a reference light quark mass. Since the squared pion mass scales approximately as the quark mass, a corresponding Taylor expansion can be performed about ϵ_π^2. This is equivalent to dropping the non-analytic contributions to Eq. (S8).

One can also consider a Taylor expansion in the parameter ϵ_π. This is not a natural expansion parameter as it scales approximately as the the square root of the input quark masses. However, it has been observed that the nucleon mass displays a remarkably linear dependence upon the pion mass, such that phenomenological estimate of the nucleon mass in lattice calculations is given by $m_N \simeq 800 + m_\pi$ MeV[70,102,103]. This observation motivates us to consider this Taylor expansion as an alternative model for describing the pion mass dependence. As with the Taylor expansion in ϵ_π^2, we observe that the choice of the reference point to perform the expansion has insignificant impact on extrapolated value of g_A when the NNLO (ϵ_π^2 in this case) expansion is considered.

D. *Dependence upon $m_\pi L$*

The finite-volume corrections can be incorporated into the EFT through an infrared modification of the pion propagators[19]. In the asymptotically large volume limit, these corrections vanish at least as fast as $e^{-m_\pi L}$. The leading volume corrections to g_A can be parameterised as[104]

$$\begin{aligned} \delta_L &\equiv g_A(L) - g_A(\infty) \\ &= \frac{8}{3}\epsilon_\pi^2 \left[g_0^3 F_1(m_\pi L) + g_0 F_3(m_\pi L) \right] \end{aligned} \tag{S14}$$

where

$$\begin{aligned} F_1(x) &= \sum_{\mathbf{n}\neq 0} \left[K_0(x|\mathbf{n}|) - \frac{K_1(x|\mathbf{n}|)}{x|\mathbf{n}|} \right], \\ F_3(x) &= -\frac{3}{2} \sum_{\mathbf{n}\neq 0} \frac{K_1(x|\mathbf{n}|)}{x|\mathbf{n}|}. \end{aligned} \tag{S15}$$

$K_\nu(z)$ are modified Bessel functions of the second kind and g_0 is the leading order (LO) contribution to g_A in the chiral expansion. In the large $m_\pi L$ limit,

$$\delta_L = 8 g_0^3 \epsilon_\pi^2 \sqrt{2\pi} \frac{e^{-m_\pi L}}{\sqrt{m_\pi L}} + \mathrm{O}\left(e^{-\sqrt{2} m_\pi L}, \frac{1}{(m_\pi L)^{3/2}} \right). \tag{S16}$$

In order to asses the uncertainty arising from the FV corrections, we can also model higher order contributions. The NNLO contribution to g_A also arises from single loop diagrams (rather than two loops). It is therefore reasonable to model the finite volume corrections from these terms as similar to those arising from the NLO contributions, particularly the graph that gives rise to the F_1 correction. Therefore, we add an additional FV correction with the following form,

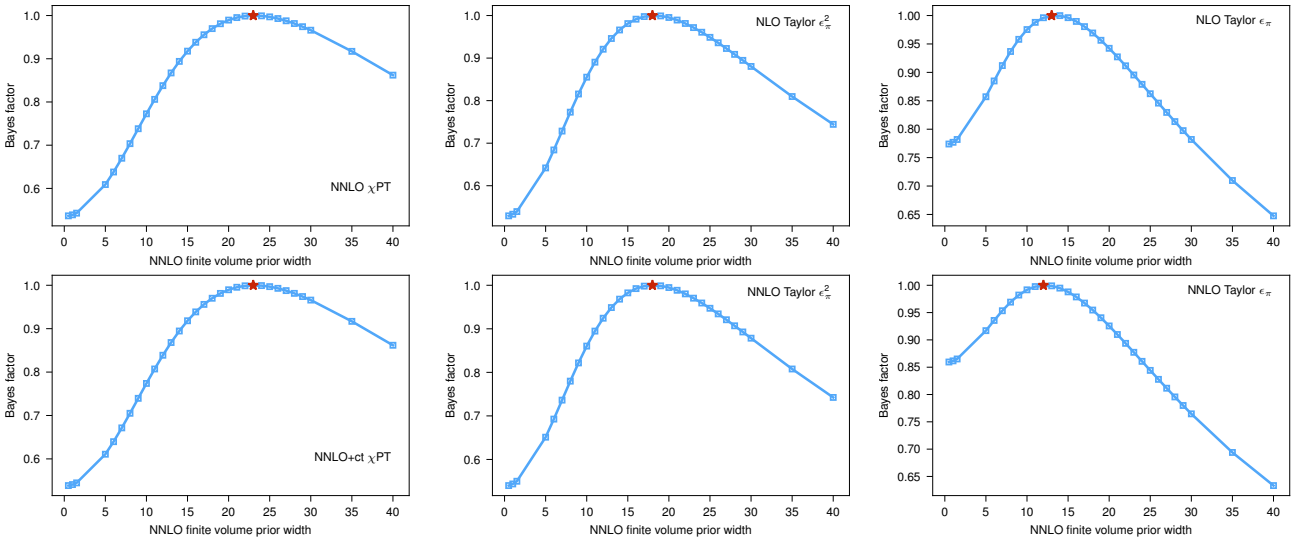
$$\delta_{L_3} \equiv f_3 \epsilon_\pi^3 F_1(m_\pi L) \tag{S17}$$

where f_3 is an unknown LEC.

We do not currently have any prior knowledge for what the value of f_3 is. Therefore, the prior central value for f_3 is set to zero, while the width is determined by an empirical Bayes study shown in Supplemental Data Fig. 3. For the six models that enter the final result, we vary the prior width from 0.5 to 40 and for each model, choose the value with the largest Bayes factor. The priors for f_3 are listed in Supplemental Data Table II. When compared to the coefficient of the leading finite volume discretisation correction $8g_0/3$, the width of f_3 is approximately 3 to 5 times wider when determined by the empirical Bayes analysis, and provides a conservative estimate for the finite volume uncertainty. The complete finite volume correction considered is defined to be

$$\delta_L' \equiv \delta_L + \delta_{L_3}. \tag{S18}$$

There has been some discussion in the literature that g_A may be particularly susceptible to finite-volume corrections such that the leading χPT prediction for the volume dependence is grossly insufficient to explain the observed volume dependence[105–108]. In Supplemental Data Fig. 4, we plot the resulting NLO χPT prediction of the volume dependence (determined in a NNLO fit to all 16



Supplemental Data Figure 3 | Empirical Bayes analysis for f_3 prior width. The prior width of the dimensionless LEC f_3 is plotted with the resulting Bayes factors for the six models that enter in the final result. The red star marks the maximum Bayes factor and sets the prior width for f_3 in the final analysis.

ensembles) as well as the estimated NNLO corrections, Eq. (S18), alongside the three a12m220 ensembles, which are in perfect accord.

E. *Dependence upon ϵ_a*

The discretisation corrections can be incorporated into *mixed-action* EFT (MAEFT)[109,110] which is known for our MDWF on HISQ action[111–118] through next-to-leading order (NLO) in the chiral and continuum expansion. Unfortunately, the MAEFT introduces new unknown coefficients that are not well constrained by our results. However, we observe our results are well described by a simple Taylor expansion in the discretisation scale, with no discernible pion-mass-dependent discretisation effects. We are therefore able to supplement the continuum HBχPT formula with corrections that parameterise the possible discretisation effects to NNLO in the Symanzik exapansion[20119],

$$\delta_a = a_2\epsilon_a^2 + b_4\epsilon_a^2\epsilon_\pi^2 + a_4\epsilon_a^4\,, \tag{S19}$$

where the first term is an NLO correction and the second and third term arise at NNLO in a power counting where $\epsilon_\pi^2 \sim \epsilon_a^2$. The coefficients are unknown constants which must be determined in the extrapolation analysis.

We also consider discretisation corrections of the form

$$\delta'_a = a_1\sqrt{4\pi}\epsilon_a + s_2\alpha_s\epsilon_a^2\,, \tag{S20}$$

where $a_1 = \mathrm{O}(m_{\mathrm{res}})$ and $s_2 = \mathrm{O}(1)$. The first term could arise from residual chiral symmetry breaking corrections arising from our use of local axial and vector currents and the second term originates from generic one-loop radiative gluon corrections at finite lattice spacing.

S.7. Extrapolation analysis

In the following sections, we discuss how the physical-point extrapolation is performed in order to obtain Eq. (1), the concluding result of this work, followed by a discussion on all sources of uncertainty, and end by studying the sensitivity of our final result under changes to different inputs of the analysis.

A. *Model averaging*

The extrapolation analysis uses the various Ansätze described in Sec. S.6. While χPT provides our best hope for a model-independent extrapolation in the pion mass, it is not known *a priori* whether the chiral expansion converges for a given quantity near the physical pion mass. For this reason, we include all models equally in our model average. Convergence of the chiral expansion will be discussed in Sec. S.8.

Our main result comes from the Bayesian model averaging[120] of a set of six different models. This procedure accounts for model selection uncertainty, and avoids over-confident inferences resulting from trusting any single model. The six extrapolation models used are:

$$\text{NNLO } \chi\text{PT}: \quad \text{Eq. (S8)} + \delta_a + \delta'_L \tag{S21a}$$
$$\text{NNLO+ct } \chi\text{PT}: \quad \text{Eq. (S8)} + c_4\epsilon_\pi^4 + \delta_a + \delta'_L \tag{S21b}$$
$$\text{NLO Taylor } \epsilon_\pi^2: \quad c_0 + c_2\epsilon_\pi^2 + \delta_a + \delta'_L \tag{S21c}$$
$$\text{NNLO Taylor } \epsilon_\pi^2: \quad c_0 + c_2\epsilon_\pi^2 + c_4\epsilon_\pi^4 + \delta_a + \delta'_L \tag{S21d}$$
$$\text{NLO Taylor } \epsilon_\pi: \quad c_0 + c_1\epsilon_\pi + \delta_a + \delta'_L \tag{S21e}$$
$$\text{NNLO Taylor } \epsilon_\pi: \quad c_0 + c_1\epsilon_\pi + c_2\epsilon_\pi^2 + \delta_a + \delta'_L \tag{S21f}$$

In Table II, we list all the priors used in the analysis of each of the extrapolations.

Under the Bayesian framework, the model averaged posterior distribution of g_A is determined by marginalising over the set of models $\{M_k\}$,

$$P(g_A|D) = \sum_k P(g_A|M_k, D)P(M_k|D), \tag{S22}$$

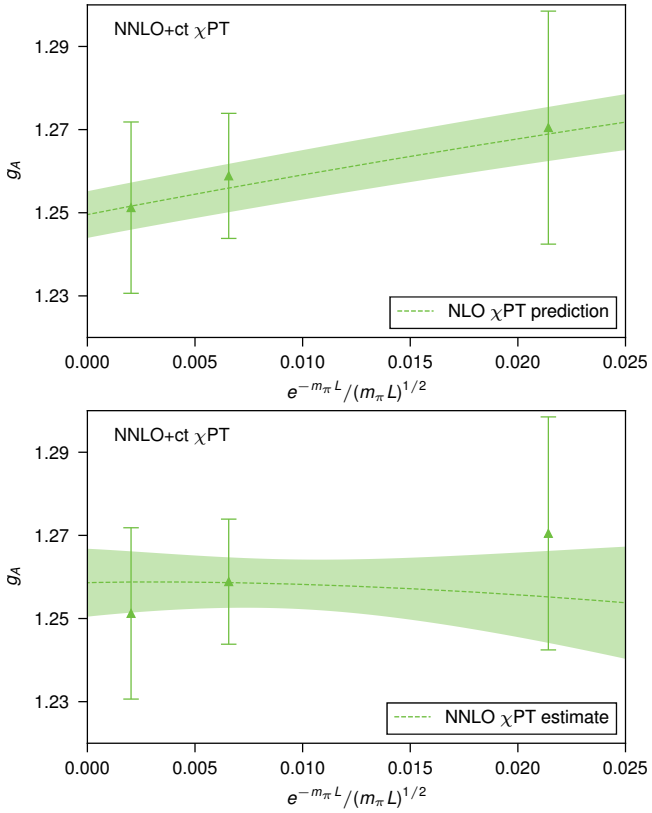
where $P(M_k|D)$ is the posterior distribution of model k given data D and is related to the likelihood of the model $P(D|M_k)$ through Bayes' Theorem,

$$P(M_k|D) = \frac{P(D|M_k)P(M_k)}{\sum_l P(D|M_l)P(M_l)}. \tag{S23}$$

In particular, the likelihood that the data is produced under model k is given by the marginalising over all (continuous) parameters θ_k

$$P(D|M_k) = \int P(D|\theta_k, M_k)P(\theta_k|M_k)d\theta_k \tag{S24}$$

and as a result $P(D|M_k)$ is now explicitly a number and not a



Supplemental Data Figure 4 | Finite volume dependence We plot the predicted NLO finite volume corrections (top) for the $m_\pi \sim 220$ MeV ensembles at $a \sim 0.12$ fm, with coefficients determined from an NNLO χPT analysis of all 16 ensembles, along with the numerical results on the a12m220S, a12m220 and a12m220L ensembles. In the bottom figure, we plot the same result except using the analysis with the estimated NNLO FV correction as well. Uncertainties are one s.e.m.

distribution. The act of marginalising over all parameters naturally penalises over-parameterised models. In our average, we choose agnostic priors $P(M_k)$ for all models listed in Eq. (S21a–S21f). Consequently, the posterior mean $\mathrm{E}[g_A]$ and variance $\mathrm{Var}[g_A]$ for the model weighted average follows

$$\mathrm{E}[g_A] = \sum_k \mathrm{E}[g_A|M_k]P(M_k|D), \tag{S25}$$

$$\mathrm{Var}[g_A] = \sum_k \mathrm{Var}[g_A|M_k]P(M_k|D) + \left\{\sum_k \mathrm{E}^2[g_A|M_k]P(M_k|D)\right\} - \mathrm{E}^2[g_A|D], \tag{S26}$$

where $\mathrm{Var}(g_A)$ is a direct consequence of the *law of total variance*. The first line of Eq. (S26) yields the *expected value of the process variance* which we refer to as the *model averaged variance* while the second line gives the *variance of the hypothetical means* which we refer to as the *model uncertainty*. The weighted average is performed with `lsqfit`[63].

The resulting physical point extrapolations are provided in Supplemental Data Tab. III and plotted in Extended Data Fig. 5, and the model average extrapolation is presented in Extended Data Fig. 4. The convergence of each model, the model average continuum, and infinite volume extrapolations are presented in Extended Data Figs. 6, 3, and 2 respectively.

We present our final physical point extrapolation as a function of pion mass in Fig. 3a. A comparison with other LQCD results[2–9] is presented in Fig. 3b

main parameters

Model	$\tilde{g}_0$	$\tilde{c}_0$	$\tilde{c}_2$	$\tilde{c}_3$	$\tilde{c}_4$	$\tilde{a}_2$	$\tilde{a}_4$	$\tilde{b}_4$
χPT	1(50)	–	0(50)	0(50)	0(1)	0(50)	0(1)	0(1)
Taylor ϵ_π^2	1.2(1.0)	1(50)	0(50)	–	0(1)	0(50)	0(1)	0(1)
Taylor ϵ_π	1.2(1.0)	1(50)	0(50)	–	0(1)	0(50)	0(1)	0(1)

alternate parameters

Model	$\tilde{g}_{N\Delta}$	$\tilde{g}_{\Delta\Delta}$	$\tilde{\gamma}_4$	$\tilde{a}_1$	$\tilde{s}_2$	$\tilde{f}_3$
χPT	−1.44(35)	−2.16(52)	0(50)	$0(10^{-3})$	0(1)	0(23)
Taylor ϵ_π^2	–	–	–	$0(10^{-3})$	0(1)	0(18)
Taylor ϵ_π	–	–	–	$0(10^{-3})$	0(1)	0(12.5)

Supplemental Data Table II | Priors used in extrapolation analysis. The priors used in the extrapolation analysis for all unknown constants. For the lower order coefficients, we use unconstraining priors. For the NNLO coefficients, we use O(1) priors. For the Taylor fits, the coefficient g_0 is used to parameterize the leading finite volume corrections. The impact of varying the prior widths is discussed in Sec. S.7 C and show in Extended Data Fig. 5. The choice of priors for the alternate parameters are discussed in Secs. S.7 B and S.8 B.

Fit	χ^2/dof	$\mathcal{L}(D\|M_k)$	$P(M_k\|D)$	$P(g_A\|M_k)$
NNLO χPT	0.727	22.734	0.033	1.273(19)
NNLO+ct χPT	0.726	22.729	0.033	1.273(19)
NLO Taylor ϵ_π^2	0.792	24.887	0.287	1.266(09)
NNLO Taylor ϵ_π^2	0.787	24.897	0.284	1.267(10)
NLO Taylor ϵ_π	0.700	24.855	0.191	1.276(10)
NNLO Taylor ϵ_π	0.674	24.848	0.172	1.280(14)
average				1.271(11)(06)

Supplemental Data Table III | Model selection analysis. We explore model uncertainties by choosing six different models and studying the variation in the extrapolation to the physical point. With the inclusion of priors, the *augmented* χ^2/dof listed assumes there to be 16 degrees of freedom (equal to the number of data points) for all six models. $\mathcal{L}(D|M_k)$ lists the log-likelihood distribution $\log P(D|M_k)$. The Taylor expansion fits are strongly favored over the χPT fits as measured by the posterior of the model. In `lsqfit`, $\mathcal{L}(D|M_k)$ is called the *log Gaussian Bayes Factor* (logGBF). The final result is given with two uncertainties: the first is the averaged variance (line one of Eq. (S26)) and the second is the model uncertainty (line two of Eq. (S26)).

B. *Uncertainty analysis*

The final uncertainty budget receives contributions from statistical uncertainty, extrapolation to the chiral, continuum and infinite volume limits, the model selection uncertainty, and isospin symmetry breaking.

Statistical uncertainty

Statistical uncertainty incorporates the correlated uncertainties of $\mathring{g}_A$, $\mathring{g}_V$, m_π, and F_π, as well as the uncorrelated uncertainty of m_π and F_π obtained from the PDG[12], which is used to evaluate the chiral-continuum extrapolation at the physical point.

Chiral and continuum extrapolation uncertainty

The chiral extrapolation uncertainty is determined from the uncertainty on the resulting LECs that control the ϵ_π dependence and do not vanish in the continuum limit. Uncertainty from the continuum extrapolation includes statistical uncertainty from ϵ_a and the resulting uncertainty on all LECs associated with corrections that vanish in the continuum limit. Additional generic one-loop and chirally-suppressed tree-level discretisation errors are investigated

in Fig. 5, labeled as '+O($\alpha_s a^2$) disc.' and '+O(a) disc.' respectively, yielding insignificant changes to the result and are therefore omitted from the final extrapolation.

Infinite volume extrapolation uncertainty

We include the finite volume correction given by Eq. (S18). Specifically, the leading volume correction to g_A derived from χPT[104] is used and higher-order finite volume corrections are estimated by Eq. (S17) with an unkown LEC f_3. The FV uncertainty is derived through the uncertainties on the LECs that determine the infinite volume dependence. In Extended Data Fig. 2**a**, we display the model averaged FV correction along with the raw a12m220 ensembles on the three volumes used. In panel **b**, we add to the model average extrapolation, black horizontal ticks to denote the central value of the renormalized values of g_A from Extended Data Table I. In all but two ensembles, the FV shift is significantly less than one sigma, with one data point shifting about one sigma and the other closer to two sigma. As a cross check of our FV uncertainty, we compare two additional analyses: one without FV corrections and one including only NLO corrections, which are displayed in Extended Data Fig. 5 as "omit FV" and "NLO FV" respectively. These both result in a relative difference with our final analysis consistent with our estimated FV uncertainty.

Isospin breaking uncertainty

Finally, isospin breaking estimates both strong and electromagnetic isospin breaking added in quadrature. Experimental results of g_A include radiative corrections up to one-loop[121,122], therefore, we estimate the two-loop radiative corrections to be $(\alpha_{\rm EM}/\pi)^2 \sim 0.0005\%$. Strong isospin breaking corrections can enter at $O(m_d - m_u)$ with the leading correction to the axial current in the chiral Lagrangian appearing at NLO[91,123,124]

$$\delta^{\rm NLO} j^{\mu}_{5,a} = \frac{b_1^{\delta}}{(4\pi F)^2} \bar{N}\{\tau_a^{\xi}, \chi_{\delta}\} S^{\mu} N \,, \tag{S27}$$

where S^{μ} is the spin operator, the spurion fields are

$$\tau_a^{\xi} = \frac{1}{2}\left(\xi \tau_a \xi^{\dagger} + \xi^{\dagger} \tau_a \xi\right) , \tag{S28}$$

$$\chi_{\delta} = \frac{1}{2}\left(\xi 2B\delta\tau_3 \xi + \xi^{\dagger} 2B\delta\tau_3 \xi^{\dagger}\right) , \tag{S29}$$

with $\xi^2 = \Sigma = \exp\{\sqrt{2} i\phi/F\}$, $2\delta = m_d - m_u$ and B is the chiral condensate (scaled by F^2) related to the pion mass by the Gell-Mann–Oakes–Renner relation[125]. However, this contribution vanishes for the $n \to p$ transition used to determine g_A as $\{\tau_+, \tau_3\} = 0$. In order to contribute to g_A, there needs to be two insertions of isospin breaking corrections, such that the anti-commutator of τ_+ and the isospin breaking contributions are non-vanishing. In QCD, the only isospin breaking parameter is the quark mass operator and so pure strong isospin breaking corrections enter as $O\left(\frac{(m_d-m_u)^2}{(m_d+m_u)^2}\epsilon_\pi^4\right) \sim 0.002\%$. There can also be mixed QED + QCD isospin breaking corrections which will scale as $O\left(\alpha_{\rm EM}\frac{m_d-m_u}{m_d+m_u}\epsilon_\pi^2\right) \sim 0.004\%$.

Finally, QED corrections modify the values of $m_\pi^{\rm nature}$ and $F_\pi^{\rm nature}$. The π^0 mass provides a good estimate of the pion mass in the isospin limit[13] and the QED corrections to F_{π^-} are given by

$$F_{\pi^-} = F_{\pi^-}^{\rm LQCD}\left(1 + \frac{\delta R_{\pi^-}}{2}\right) , \tag{S30}$$

where the correction has been estimated to be[126]

$$\delta R_{\pi^-} = 0.0169(15) \,, \tag{S31}$$

in good agreement with the χPT estimate[127,128]. We can make a conservative estimate of the uncertainty from these corrections by splitting the difference of our extrapolated answers by using

$$\epsilon_{\pi^-} = \frac{m_{\pi^-}}{4\pi F_{\pi^-}} \,, \tag{S32}$$

$$\epsilon_{\pi^0} = \frac{m_{\pi^0}(1 + \delta R_{\pi^-}/2)}{4\pi F_{\pi^-}} \,, \tag{S33}$$

resulting in an uncertainty estimate of

$$\left|\frac{g_A(\epsilon_{\pi^0}) - g_A(\epsilon_{\pi^-})}{2}\right| = 0.00038(14) \,, \tag{S34}$$

which is a 0.03% uncertainty.

Model uncertainty

The model selection uncertainty is determined as described in Sec. S.7 A.

Final uncertainty

The final model averaged uncertainty breakdown, including the uncertainty arising from the different extrapolation functions, is presented in the main text, Eq. (1). Broken down to the different contributions of statistical (s), chiral (χ), continuum (a), infinite volume (v), isospin breaking (I) and model (M), we have

$$g_A = 1.2711(103)^{\rm s}(39)^{\chi}(15)^{\rm a}(19)^{\rm v}(04)^{\rm I}(55)^{\rm M} \,.$$

The total uncertainty arises from adding these uncertainties in quadrature. More precise values at the physical pion mass will have the largest impact in simultaneously reducing the extrapolation and model-selection uncertainty. This demonstrates a straightforward path towards a sub-percent precision, which may be able to offer insight to the upward trending values of the measurements of g_A, including the most recent determination[129].

C. *Sensitivity analysis*

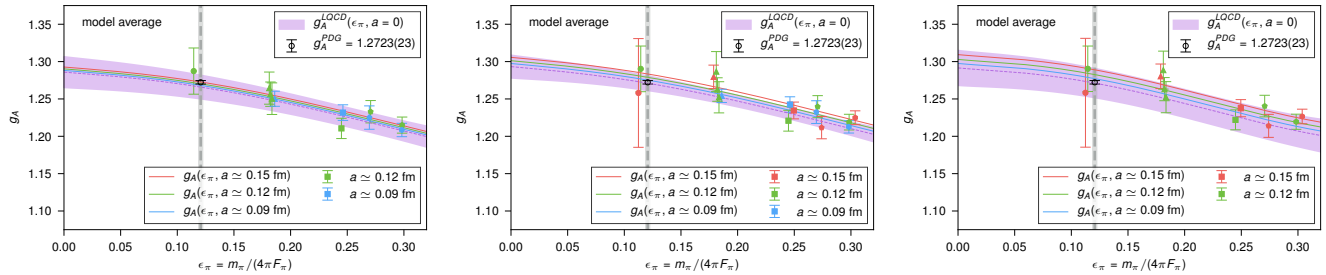
Robustness of the final result is tested under changes in the initial prior distributions, and data. Specifically, subsets of data are explored to quantify sensitivity for different regions of pion mass and lattice spacing.

Prior sensitivity

Our final result, displayed as the black square in Extended Data Fig. 5, is robust and extrapolated with unconstraining priors. In this figure, the set of results labeled '2× LO width' and '2× all widths' explores doubling the prior widths for the leading-order and all LECs respectively. We observe indiscernible change to the final result when doubling only the widths on the priors for the leading-order LECs, demonstrating these priors to be unconstraining. Doubling the prior widths of all LECs leads to insignificant changes in the final result, demonstrating that none of the priors are biasing the fit.

Pion mass sensitivity and posterior correlations

Next, we demonstrate that the heavy pion masses do not skew the results by analysing subsets of our full data. We study the variation of our final analysis if the two points close to the physical pion mass are removed. The results of these extrapolation analyses are presented in Extended Data Fig. 5. We observe that, as expected, when some of the results are removed, the uncertainty grows, but the resulting extrapolation is consistent with our main result within the 1-σ level. In Supplemental Table IV, we show the impact of cutting these heavy pion mass points for each of the models which enter our model average. We conclude there is no statistical justification for truncating the heavy pion mass points. The impact of including or excluding these heavier pion mass points on our final extrapolated uncertainty is already included through the use of varying orders in the various model extrapolation functions. We also observe that the growth in the uncertainty from the full



Supplemental Data Figure 5 | Sensitivity to cutting lattice spacings. We plot the resulting model average analysis when either the coarsest or finest ensembles are cut from those considered, with the full data set in the middle for comparison. Uncertainties are one s.e.m.

data set to the $m_\pi \lesssim 310$ MeV data set scales as one would expect when going from 5 to 3 different pion masses.

Additionally, to assess the influence of the heavy pion mass region on our extrapolated result, we use the resulting analyses to compute the correlation between between the extrapolated result at the physical pion mass and at $m_\pi = 400$ MeV. For simplicity, we compute the correlation in both cases in the continuum and infinite volume limits. This correlation allows for the determination of the conditional mean between these two points, which provides a measure of how much the extrapolated answer at one point (physical pion mass) would shift, given a fluctuation at the other point (heavy pion mass)

$$g_A(\epsilon_\pi^{(1)}, \epsilon_\pi^{(2)}) = \hat{g}_A(\epsilon_\pi^{(1)}) + C_{1,2}\ \sigma_1 \frac{g_A(\epsilon_\pi^{(2)}) - \hat{g}_A(\epsilon_\pi^{(2)})}{\sigma_2}\,, \qquad \text{(S35)}$$

where $\hat{g}_A(\epsilon_\pi)$ is the expected value of g_A at ϵ_π given the analysis, $g_A(\epsilon_\pi^{(2)}) - \hat{g}_A(\epsilon_\pi^{(2)})$ is the hypothetical fluctuation at point 2, σ_i are the continuum and infinite volume extrapolated uncertainties at the two points and the coefficient $C_{1,2}$ is the correlation coefficient between g_A at $\epsilon_\pi^{(1)}$ and at $\epsilon_\pi^{(2)}$.

The uncertainty of the continuum and infinite volume extrapolated result at the heavy pion mass, σ_{400} is approximately the same as the uncertainty on the input data points at that pion mass (~ 0.009 compared to 0.006, 0.010 and 0.008 on the a15m400, a12m400 and a09m400 ensembles respectively, see Extended Data Table I). Therefore, if we consider a hypothetical 1-σ fluctuation of the data at $m_\pi \sim 400$ MeV, the expected shift in our extrapolated value of g_A at the physical point is well approximated by

$$\delta g_A^{phys.} = C_{phys.,400}\ \sigma_{phys.}\,. \qquad \text{(S36)}$$

The model-averaged correlation coefficient, with weights given by the model-selection analysis (Supplemental Data Table III), is given by

$$C_{phys.,400}^{\text{model avg.}} = 0.37\,. \qquad \text{(S37)}$$

The shift in the extrapolated value of g_A at the physical pion mass, due to a hypothetical 1-σ fluctuation at $m_\pi \simeq 400$ MeV is given by the weighted average of $C_{phys.,400}^{(i)}\sigma_{phys.}^{(i)}$ over the models, i, resulting in

$$\delta g_A^{phys.} = 0.0030 \times \hat{g}_A(\epsilon_\pi^{phys.})\,, \qquad \text{(S38)}$$

where $\hat{g}_A(\epsilon_\pi^{phys.})$ is our final result, Eq. (1), or a 0.3% change. We conclude that the physical-point extrapolation from the model-averaged fit ansatz is relatively insensitive to fluctuations at larger pion masses, leading further evidence for the robustness of our final result.

Lattice spacing sensitivity

To further check the sensitivity of our results on the continuum limit, we perform the full analysis discarding, one at a time, each of the individual $a \sim 0.15$ fm and $a \sim 0.09$ fm ensembles. In Fig. 5, we display the resulting model average extrapolations with these two data cuts side-by-side with the extrapolation of the full data set. The extrapolated final results are also shown in Extended Data Fig. 5. Similar to the pion mass cuts, removing results from one of the discretisation scales leads to a larger, but consistent result.

m_π range	$\lesssim 400$ MeV	$\lesssim 350$ MeV	$\lesssim 310$ MeV
Fit	weight, g_A	weight, g_A	weight, g_A
NNLO χPT	0.033, 1.273(19)	0.044, 1.281(25)	0.076, 1.275(27)
NNLO+ct χPT	0.033, 1.273(19)	0.044, 1.280(25)	0.077, 1.275(27)
NLO Taylor ϵ_π^2	0.287, 1.266(09)	0.305, 1.278(12)	0.300, 1.282(15)
NNLO Taylor ϵ_π^2	0.284, 1.267(10)	0.306, 1.278(13)	0.300, 1.282(16)
NLO Taylor ϵ_π	0.191, 1.276(10)	0.156, 1.288(14)	0.125, 1.290(17)
NNLO Taylor ϵ_π	0.172, 1.280(14)	0.146, 1.288(16)	0.121, 1.290(18)
model average	1.271(11)(06)	1.281(15)(04)	1.283(18)(05)

Supplemental Data Table IV | Effect of cutting heavy pion mass points. The weights are determined from the log Bayes Factors as described in the text. The $m_\pi \lesssim 400$ MeV analysis is described in the text and listed in Extended Data Fig. 5 as **model avg**. The model average results here correspond to the same points in the figure.

S.8. χPT convergence and inclusion of the Δs

For sufficiently light pion masses, χPT provides a model independent description of low-energy QCD. What is not known *a priori* is the range of pion masses for which χPT is a converging, perturbative expansion about the chiral limit. As discussed in Sec. S.6 B, there is also theoretical evidence that the explicit inclusion of the delta degrees of freedom in the chiral Lagrangian will improve the convergence of g_A. We explore these extrapolations in more detail.

A. *Convergence of the χPT expansion*

We discuss the convergence of the χPT expansion without explicit delta degrees of freedom. The first fit that results in an acceptable $\chi^2_{\text{aug}}/\text{dof}$ is the NNLO χPT fit. This fit has 3 LECs determined from our 5 different pion mass points. The convergence of the fit is displayed in Extended Data Fig. 6a. Each curve is the sum of all contributions up to the order listed. One observes large cancellations between the NLO and NNLO contributions already at pion masses lighter than nature (grey vertical line). The strong curvature of the NLO curve is driven by competition between the counter term, c_2 and the $\ln(\epsilon_\pi^2)$ contribution. Because of this competition, it is more difficult to assess the convergence of the theory.

We are not able to perform a meaningful fit with the full N3LO formula as there are 5 unknown LECs, and we have only 5 pion mass points. However, we can check the convergence by adding just the local counter term contribution, $c_4\epsilon_\pi^4$, the NNLO+ct χPT fit. The convergence of this fit is depicted in Extended Data Fig. 6b.

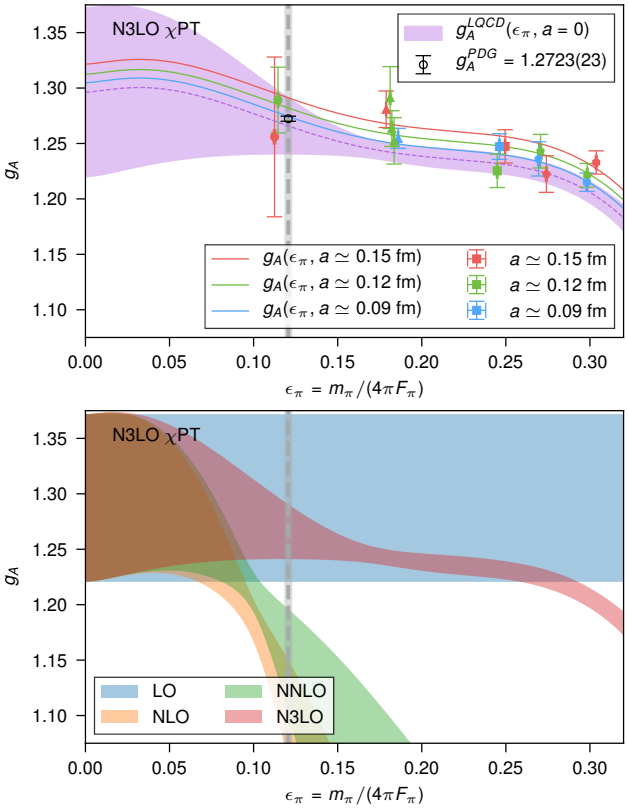
Order-by-order contribution				
order n	$\delta^{(n)} g_A$	% of total	LEC	value
0	+1.236(34)	97.1(2.7)	g_0	1.236(34)
2	−0.026(30)	− 2.0(2.4)	c_2	-23.0(3.5)
3	+0.062(14)	+ 4.9(1.1)	c_3	28.7(5.5)
4	+0.0000(2)	+ 0.0(0.0)	c_4	0.007(1.000)
Total	1.273(19)			

LEC correlation matrix				
LEC	g_0	c_2	c_3	c_4
g_0	1	-0.02010	-0.09365	0.03797
c_2	-0.02010	1	0.97231	-0.99050
c_3	-0.09365	0.97231	1	-0.99401
c_4	0.03797	-0.99050	-0.99401	1

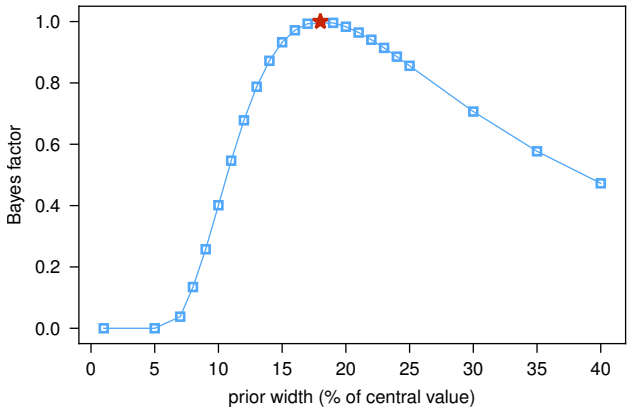
Supplemental Data Table V | NNLO+ct χPT analysis results, We provide the order-by-order contribution to g_A and the resulting LEC correlation matrix from the NNLO+ct χPT analysis.

We see that the addition of this term has negligible impact on the resulting extrapolated value of g_A for all pion masses depicted. The resulting order-by-order contributions of this analysis and the resulting correlation matrix for the LECs are given in Supplemental Data Table V.

Under the Bayesian Framework, we can perform the full N3LO fit. In Supplemental Data Fig. 6, we show the resulting fit as well as the convergence of the expansion. We first observe the uncertainty of the extrapolation begins to grow significantly outside the region of constraining data, which is a clear sign of overfitting. One also observes that the convergence of the expansion changes markedly from the NNLO+ct χPT extrapolation, where now, the NNLO result is also quickly dropping as a function of the pion mass for pion masses near and above the physical pion mass. The full extrapolation has a mild pion mass dependence, which also demonstrates there is a large cancellation between different terms in the expansion. This is not surprising given the large and positive coefficient of the $\ln^2(\epsilon_\pi)$ contribution appearing at N3LO, Eq. (S9). These results suggest that the χPT expansion may be particularly poor for g_A. However, a strong conclusion can not be drawn without having results at more pion mass points. In particular, having precise results in the lighter pion mass region is desirable. The resulting extrapolation is compared with our final result in Extended Data Fig. 5.



Supplemental Data Figure 6 | The N3LO χPT extrapolation. We plot the N3LO extrapolation of our results and resulting convergence of the expansion. Uncertainties are one s.e.m.



Supplemental Data Figure 7 | Bayes factors versus prior width. We plot Bayes factors of the resulting NLO χPT(Δ) fits versus the prior width (in %) given to the two new axial couplings. An Empirical Bayes analysis selects the 18% as the optimal prior width.

B. *Including the Δs*

We finally turn to an extrapolation including explicit Δ degrees of freedom. As discussed in Sec. S.6 B, it is expected the pion mass dependence of g_A predicted from χPT will be milder with the inclusion of these states, due to cancellations between the nucleon and delta virtual loops imposed by the large-N_c expansion. Including the deltas at NLO introduces two new axial coupling LECs and the extra mass splitting parameter, for a total of 4 LECs. An input parameter to the analysis is ϵ_Δ, defined in Eq. (S11). Our results do not directly constrain these two new axial couplings, and so the resulting fits have larger uncertainties. We do not include the analysis with explicit deltas in our set that are averaged as they require phenomenological input, unlike all the fits which go in the final analysis. Nevertheless, it is interesting to explore how the delta degrees of freedom impact the analysis.

We begin by exploring how varying the prior width on the new axial couplings impacts the analysis. We take the central values in Eq. (S13) and vary the prior width of these couplings from 1% to 40%. In Supplemental Data Fig. 7, we plot the resulting Bayes factors normalised to the maximum Bayes factor which occurs with a 18% prior width on $\mathring{g}_{N\Delta}$ and $\mathring{g}_{\Delta\Delta}$. It is interesting to note that the prior width motivated by Empirical Bayes is comparable to the nominal $1/N_c$ correction which would be $\sim 33\%$.

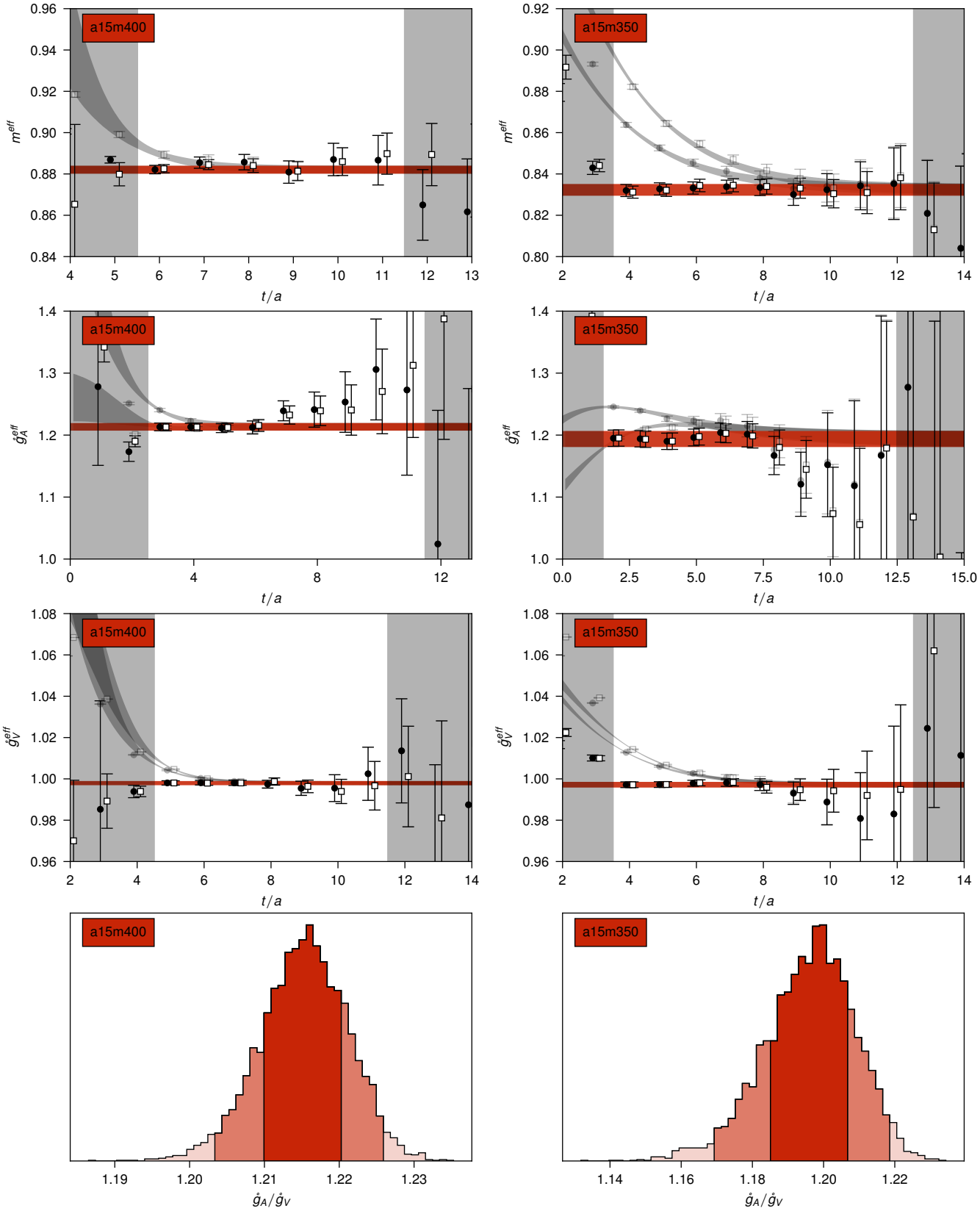
Next, we explore the convergence of the expansion when the delta degrees of freedom are included. While we are taking $\Delta \simeq 293$ MeV, ϵ_Δ depends upon ϵ_a and ϵ_π through the denominator of F_π. We have not parameterised this dependence, precluding our ability to plot the resulting fit, but we can observe the values of the LECs that are determined with and without explicit delta degrees of freedom to compare the size of the NLO contributions. This study can only be performed for the NLO LEC c_2, as we do

Order	with delta		without delta	
	c_2	$\delta g_A^{\rm NLO}$	c_2	$\delta g_A^{\rm NLO}$
NLO	-4.8(1.8)	0.28(13)	–	–
NNLO	12.1(8.2)	0.75(40)	-23.0(3.5)	-0.026(30)

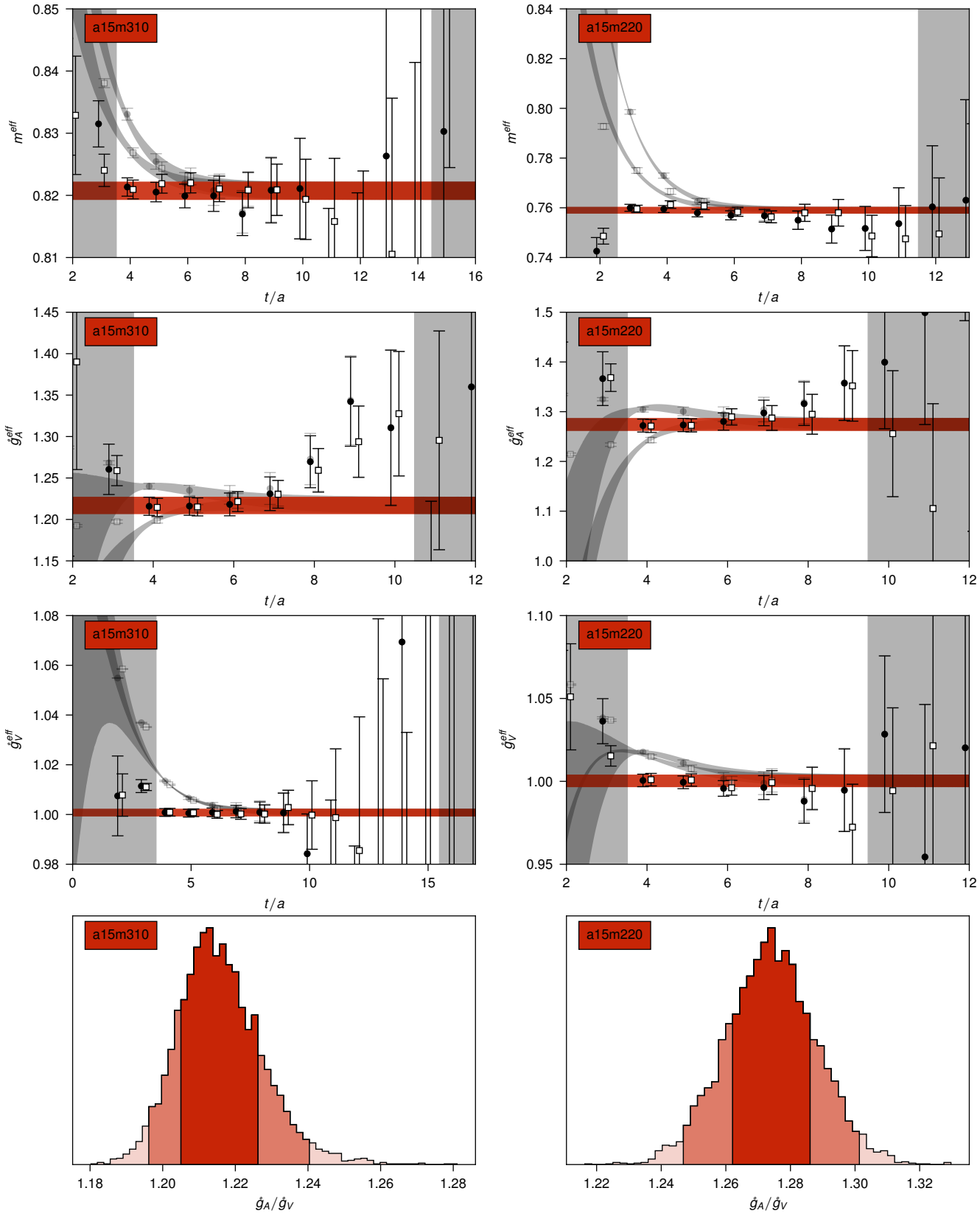
Supplemental Data Table VI | Effect of including delta degrees of freedom. We list the resulting value of c_2 and the NLO contribution to g_A at the physical point for two fits with and without the delta degrees of freedom. The NLO χPT($\not\Delta$) fit has a sufficiently poor $\chi^2_{\rm aug}$/dof we do not report the resulting LEC and $\delta g_A^{\rm NLO}$ values.

not have the extrapolation function with deltas beyond NLO. In Supplemental Data Table VI, we list the resulting value of c_2 and the size of the NLO contribution for the analysis with and without the delta. The NLO χPT($\not\Delta$) results in a very poor fit, so we do not report the values of c_2 or $\delta g_A^{\rm NLO}$. The NLO χPT(Δ) does result in a good fit with $\chi^2_{\rm aug}/{\rm dof} = 0.49$. The resulting fit is displayed in the bottom entry of Extended Data Fig. 5. Comparing the NNLO fits, which means either NLO χPT(Δ) and NLO χPT($\not\Delta$) plus NNLO χPT($\not\Delta$), we observe the value of $c_2(\Delta)$ is approximately half as big as $c_2(\not\Delta)$, indicating that the non-analytic terms are smaller when the deltas are included. This suggests that the convergence will be improved with the explicit inclusion of Δ degrees of freedom.

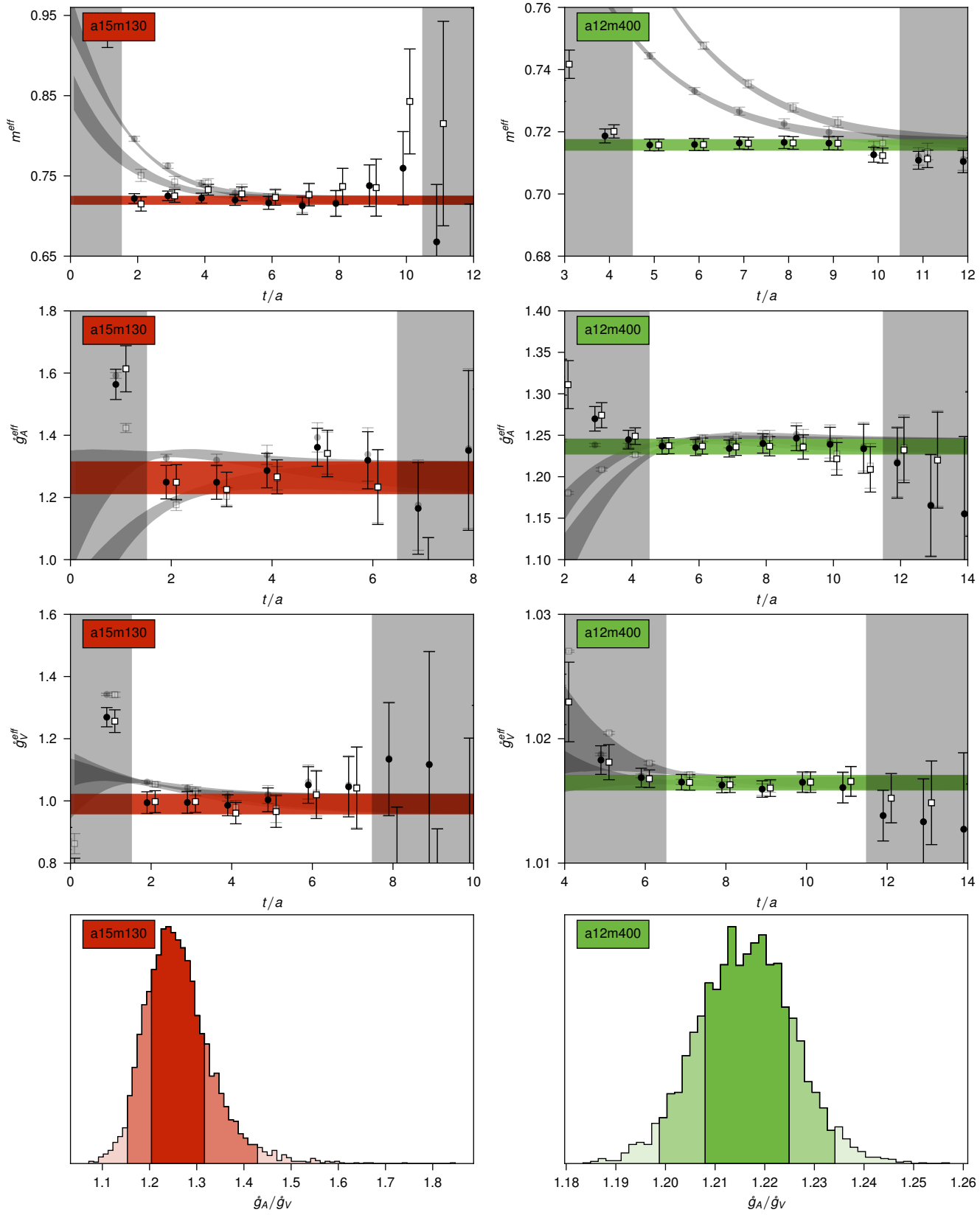
In order to fully constrain this fit and test the convergence of the expansion with and without delta degrees of freedom, first, a calculation including the $N \to \Delta$ and $\Delta \to \Delta$ axial matrix elements is needed, and second, the EFT including the deltas must be worked out to at least one higher order in the expansion. This is a particularly interesting question to resolve as the χPT($\not\Delta$) seems to be poorly converging, if at all, as observed in the prior section.



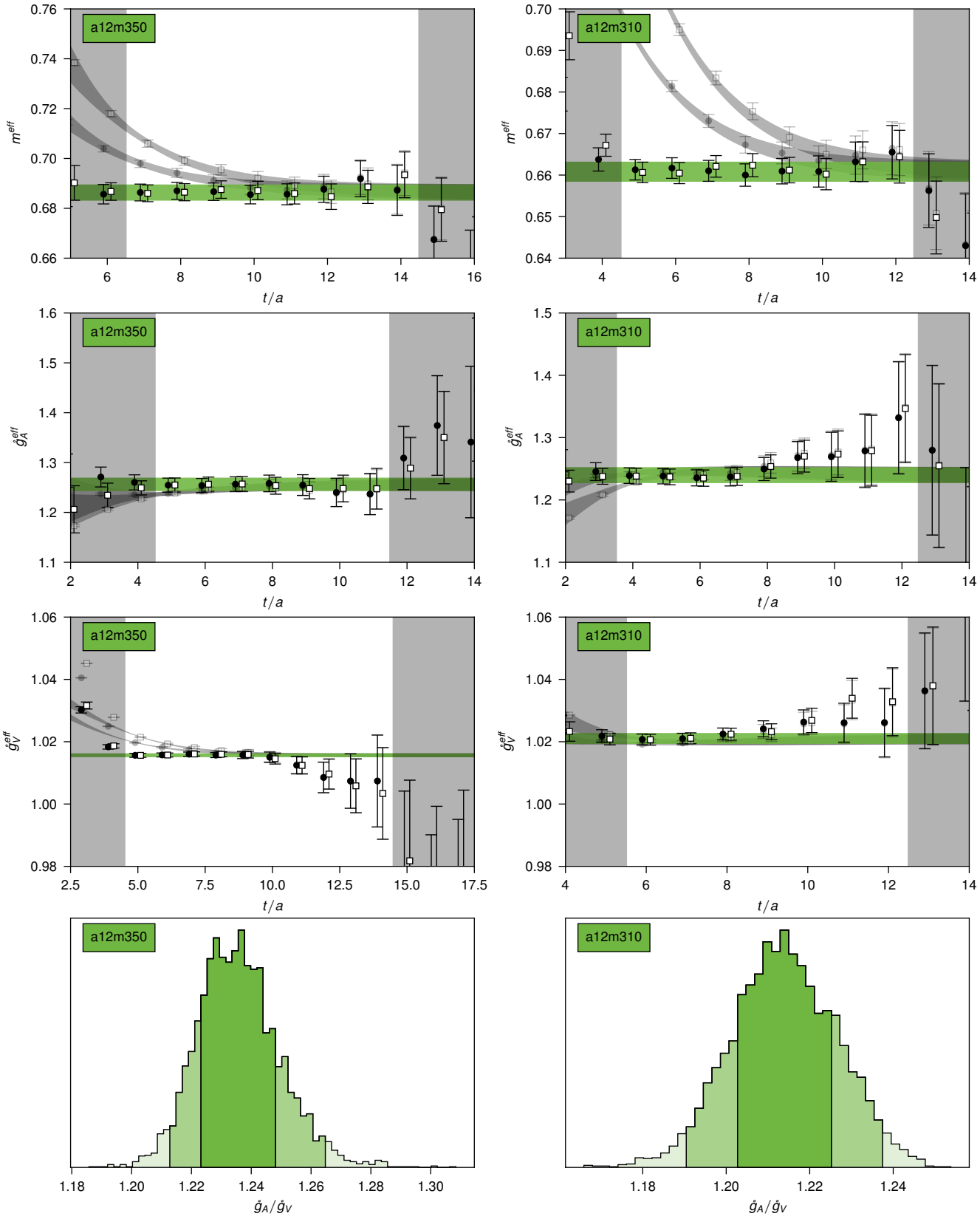
Supplemental Data Figure 8 | Correlator fit study I. Analogous to Extended Data Fig. 1a, b, c and d for the a15m400 and a15m350 ensembles. Unbiased bootstrap fit curves with 68% confidence intervals. Results from one simultaneous fit are represented in each column. The resulting biased bootstrap histograms for $\mathring{g}_A/\mathring{g}_V$ follow at the bottom. In the histograms, regions mark the 68% and 95% confidence interval. Uncertainties are one s.e.m.



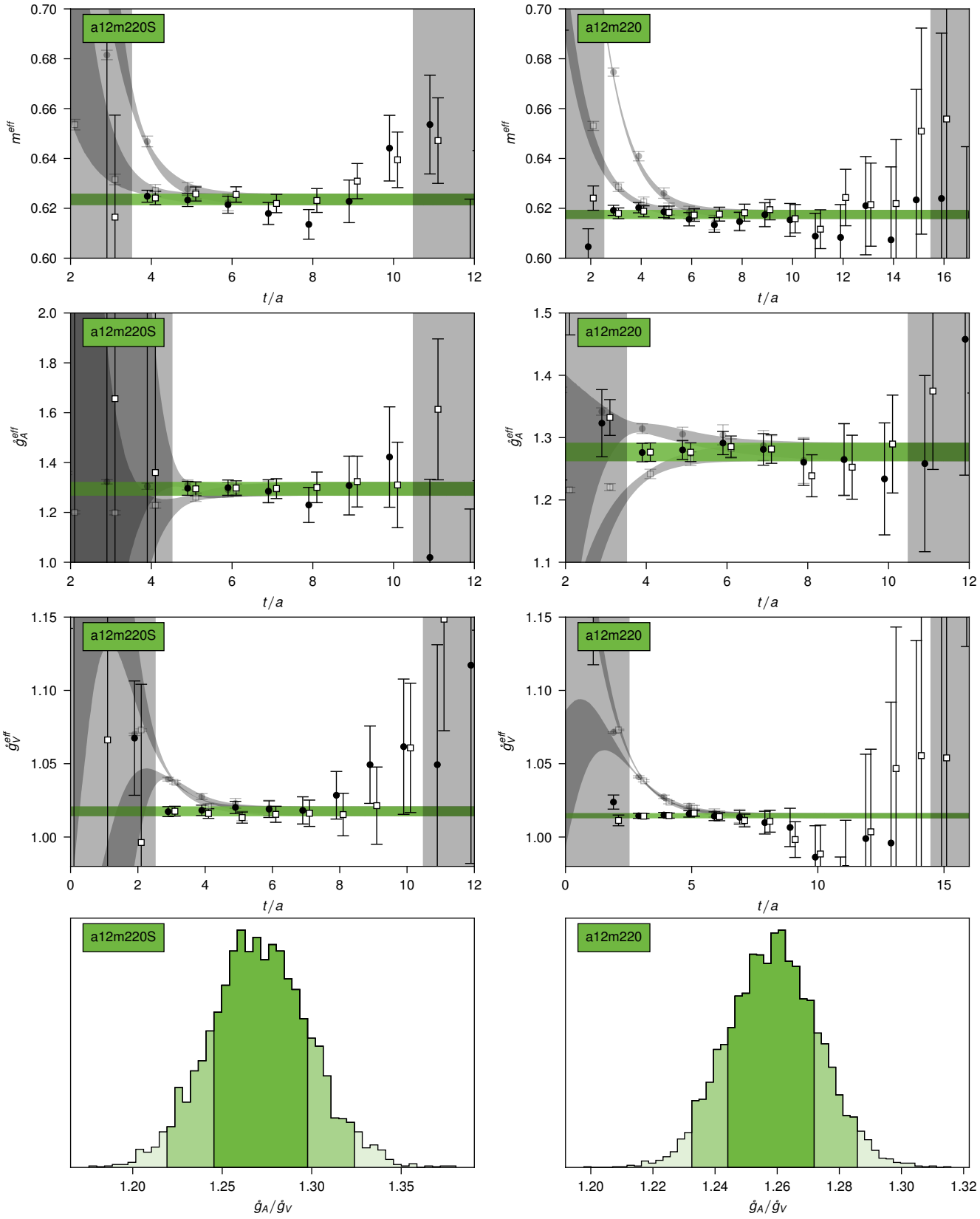
Supplemental Data Figure 9 | Correlator fit study II. Analogous to Extended Data Fig. 1a, b, c and d for the a15m130 and a12m400 ensembles. Uncertainties are one s.e.m.



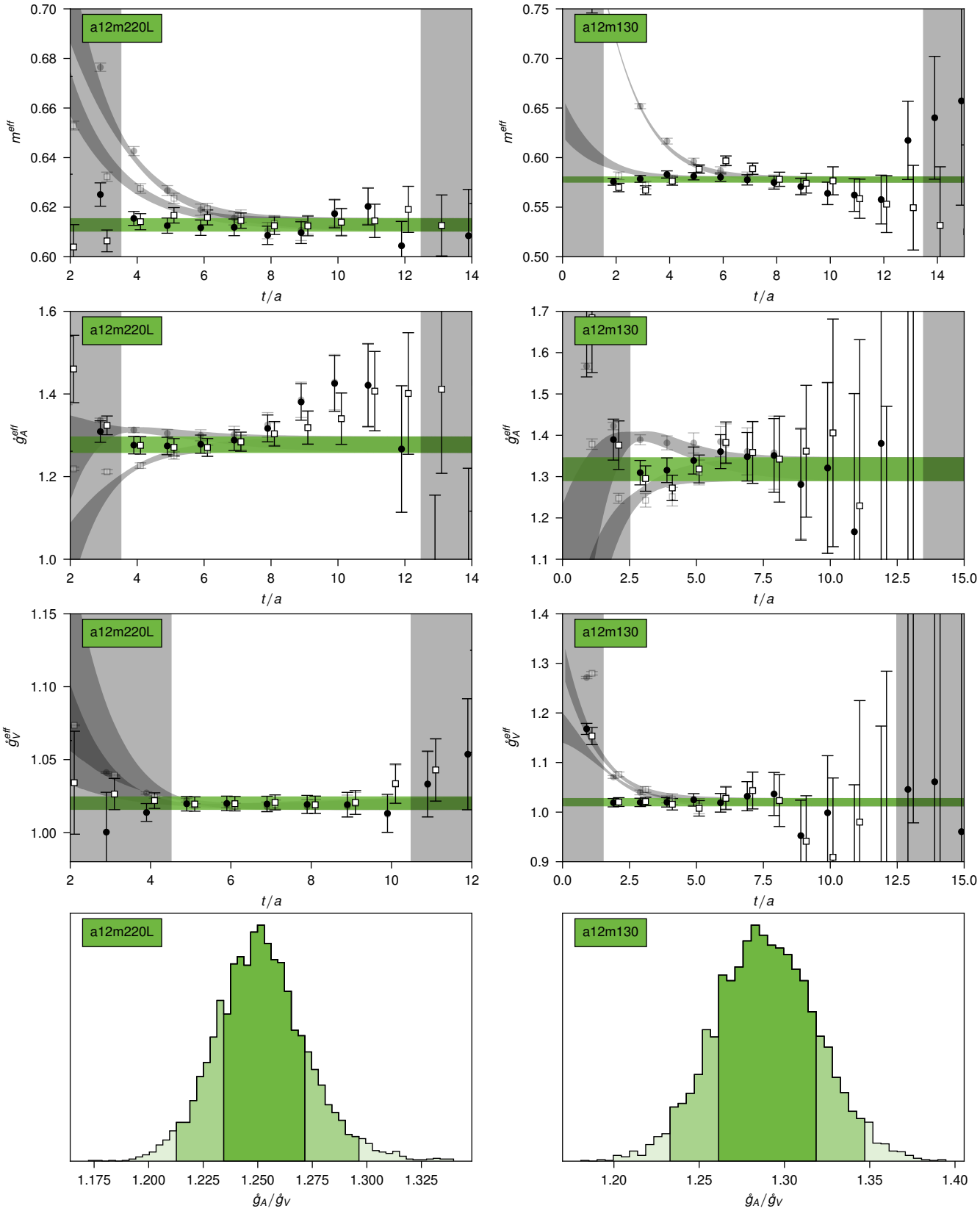
Supplemental Data Figure 10 | Correlator fit study III. Analogous to Extended Data Fig. 1a, b, c and d for the a15m130 and a12m400 ensembles. Uncertainties are one s.e.m.



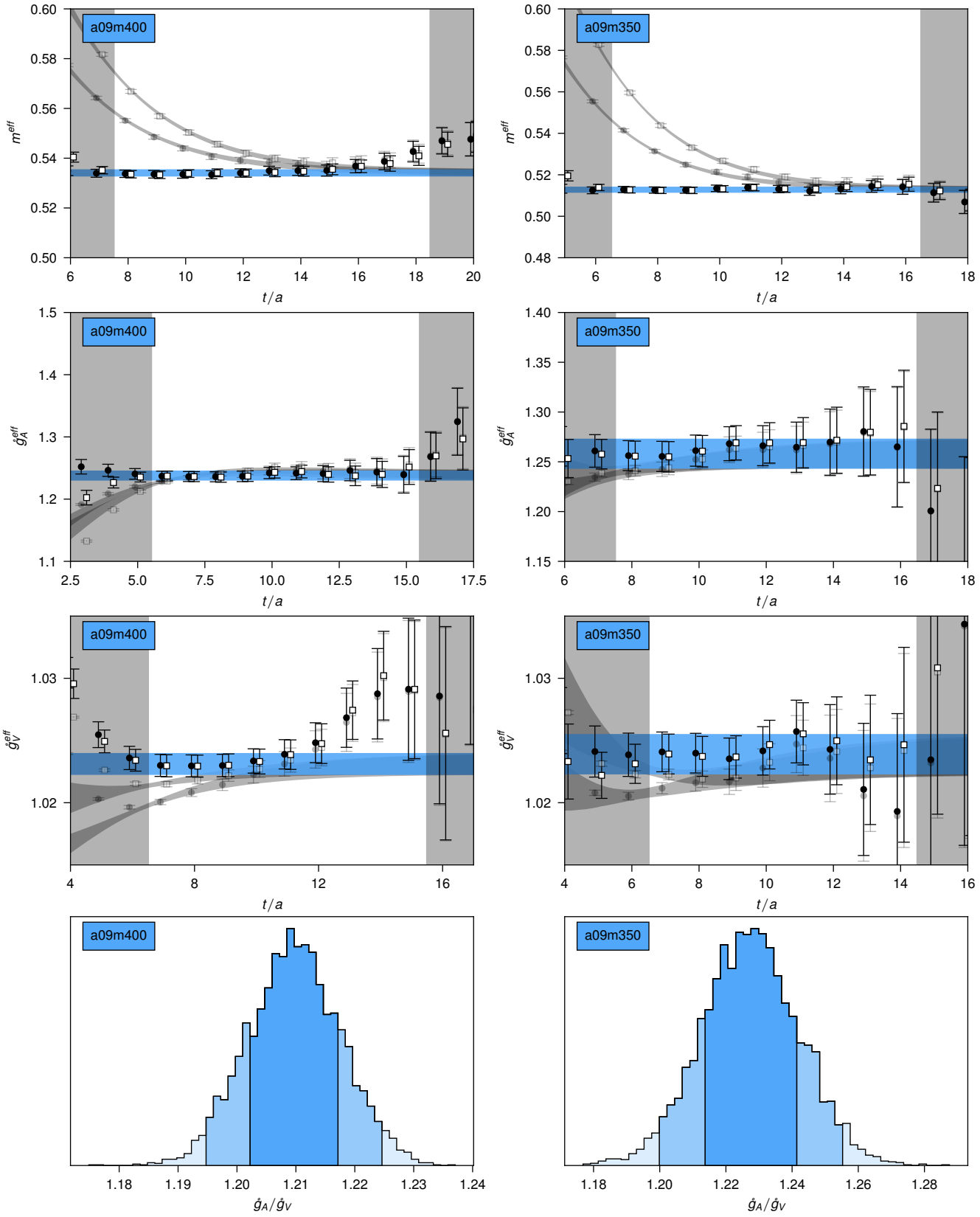
Supplemental Data Figure 11 | Correlator fit study IV. Analogous to Extended Data Fig. 1a, b, c and d for the a12m350 and a12m310 ensembles. Uncertainties are one s.e.m.



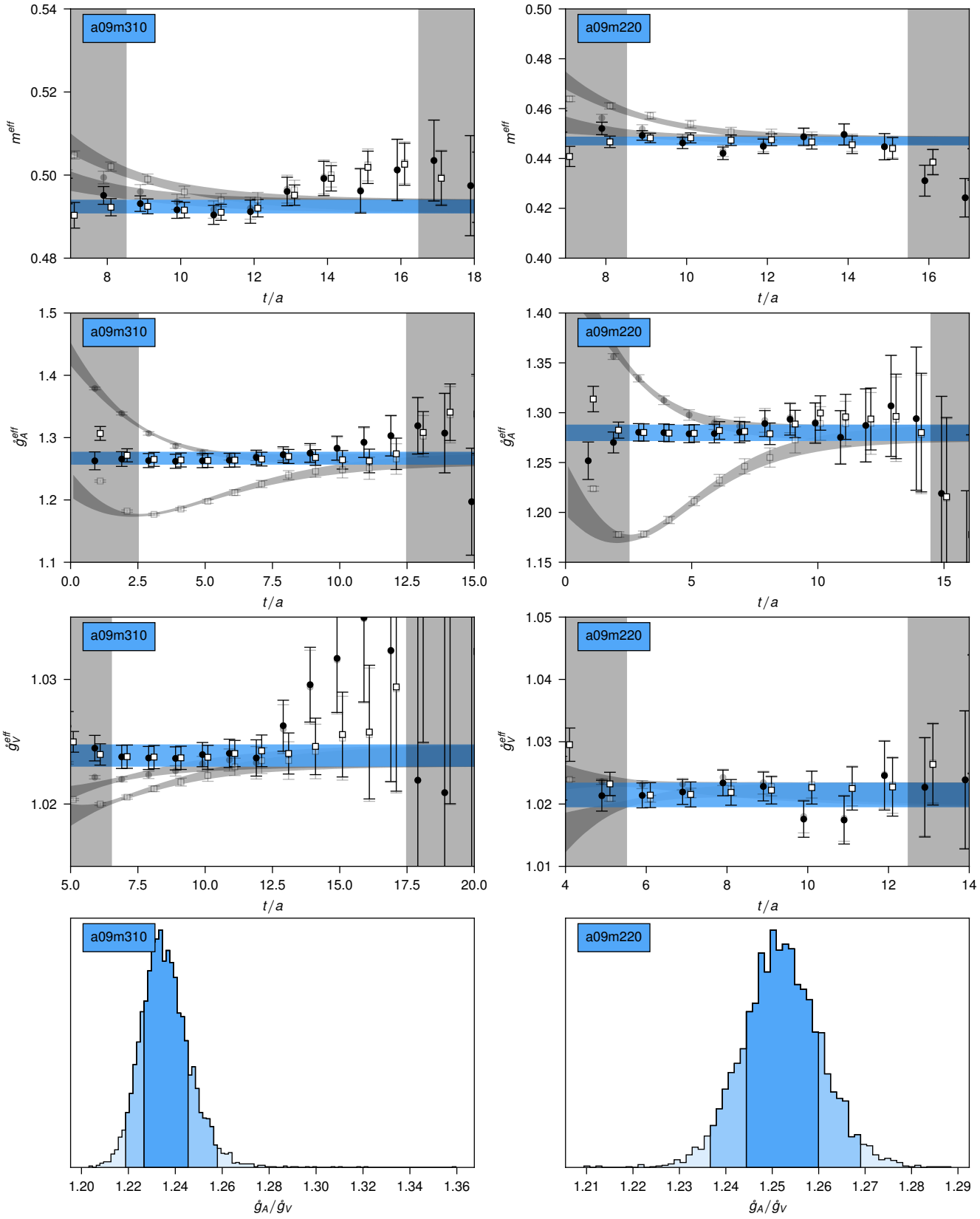
Supplemental Data Figure 12 | Correlator fit study V. Analogous to Extended Data Fig. 1a, b, c and d for the a12m220 and a12m220S ensembles. Uncertainties are one s.e.m.



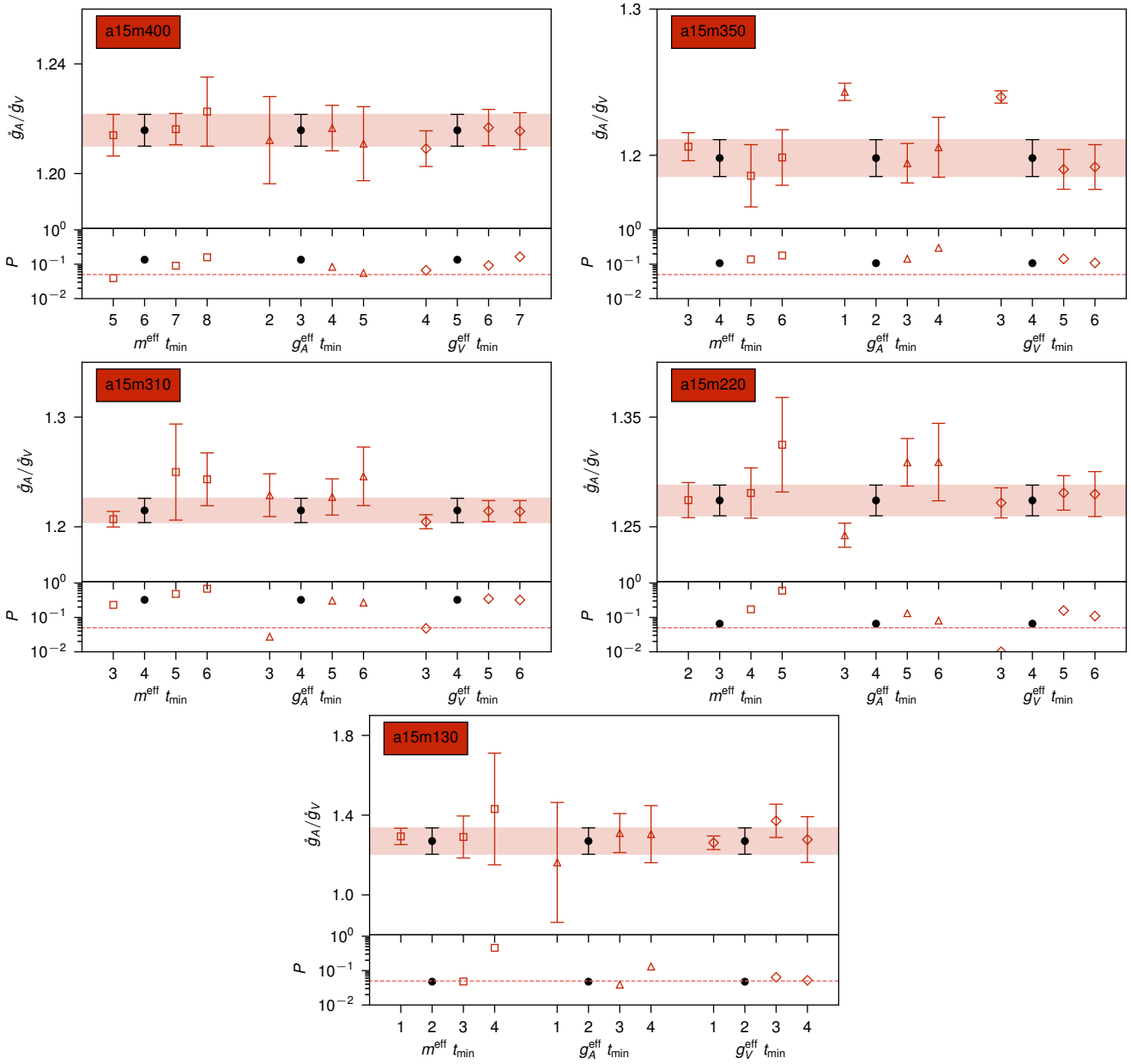
Supplemental Data Figure 13 | Correlator fit study VI. Analogous to Extended Data Fig. 1a, b, c and d for the a12m220L and a12m130 ensembles. Uncertainties are one s.e.m.



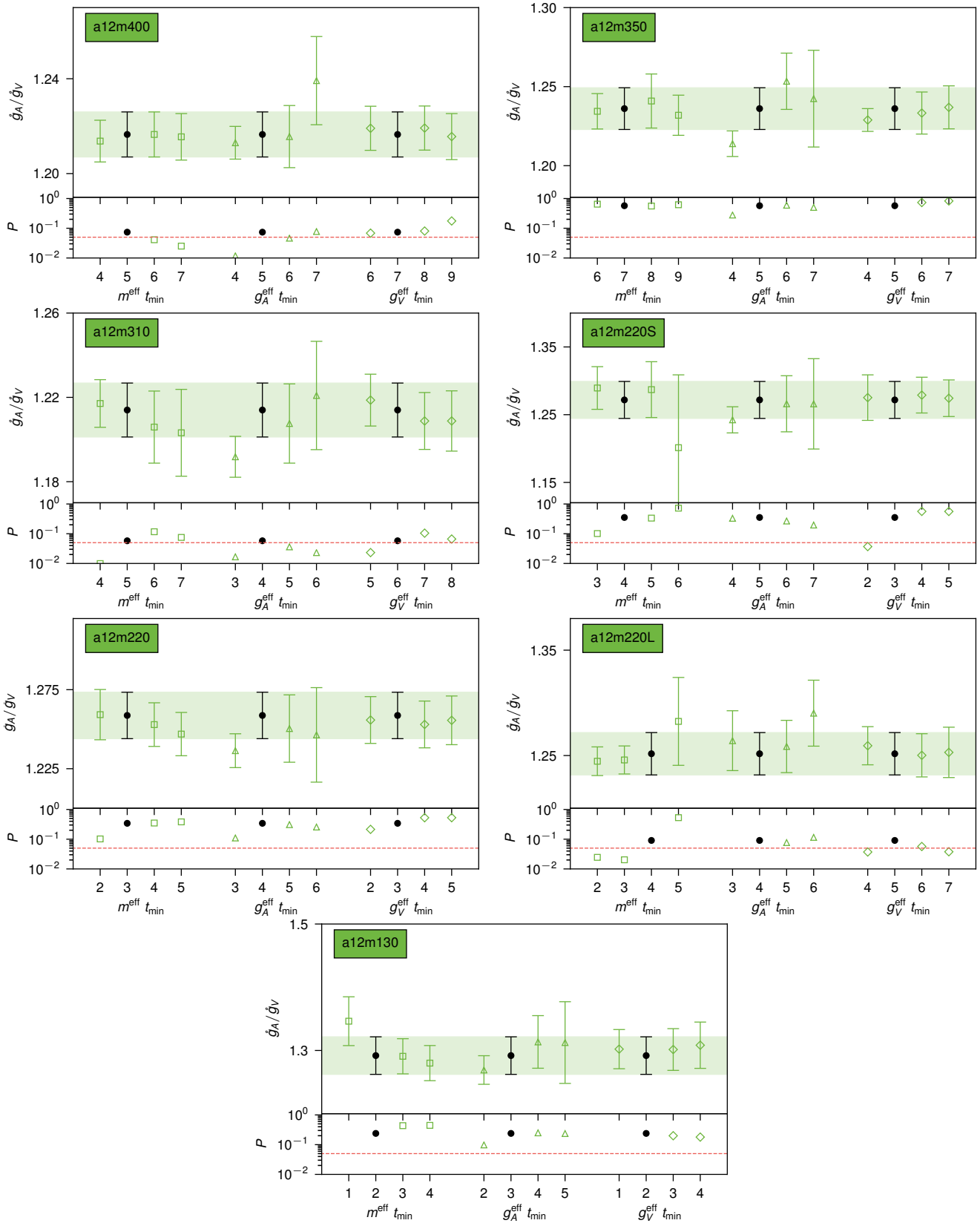
Supplemental Data Figure 14 | Correlator fit study VII. Analogous to Extended Data Fig. 1a, b, c and d for the a09m400 and a09m350 ensembles. Uncertainties are one s.e.m.



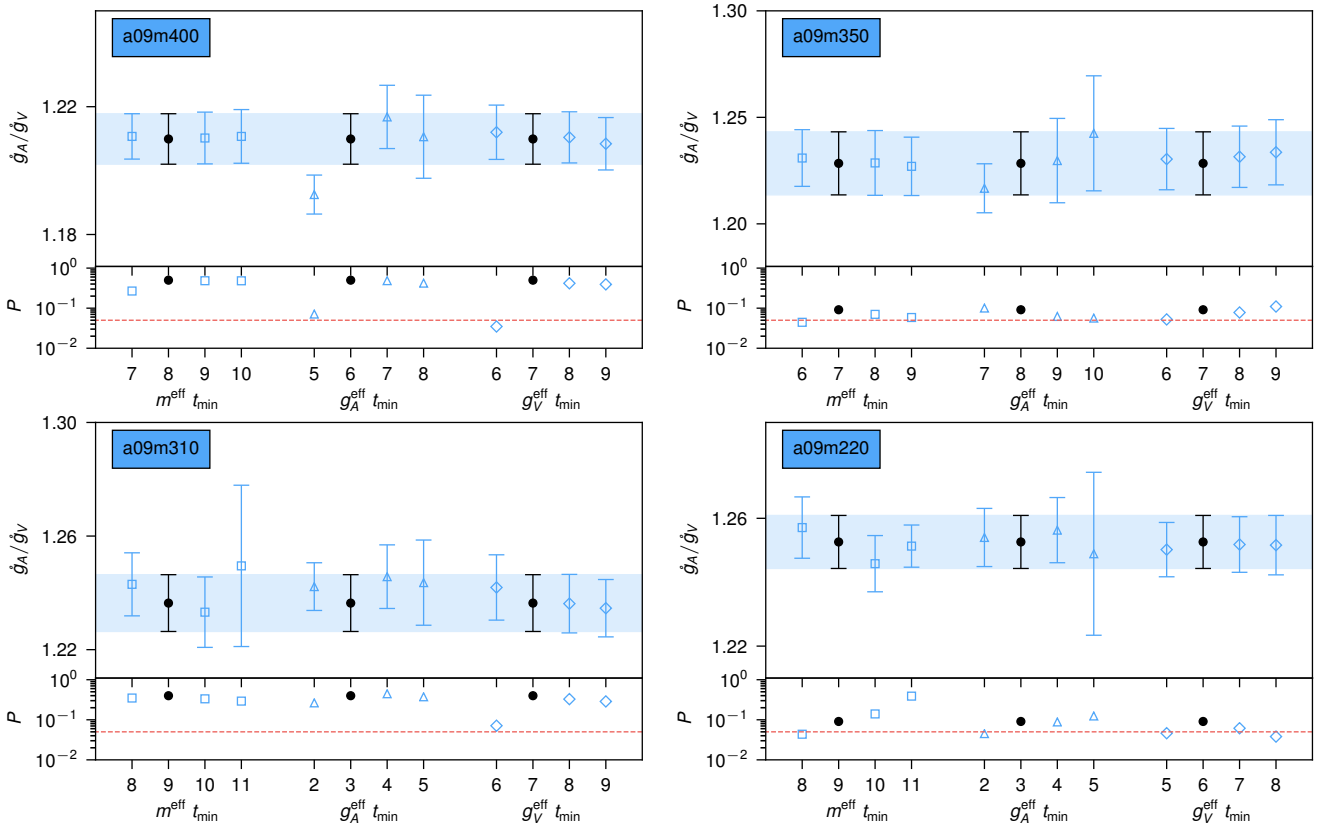
Supplemental Data Figure 15 | Correlator fit study VIII. Analogous to Extended Data Fig. 1a, b, c and d for the a09m310 and a09m220 ensembles. Uncertainties are one s.e.m.



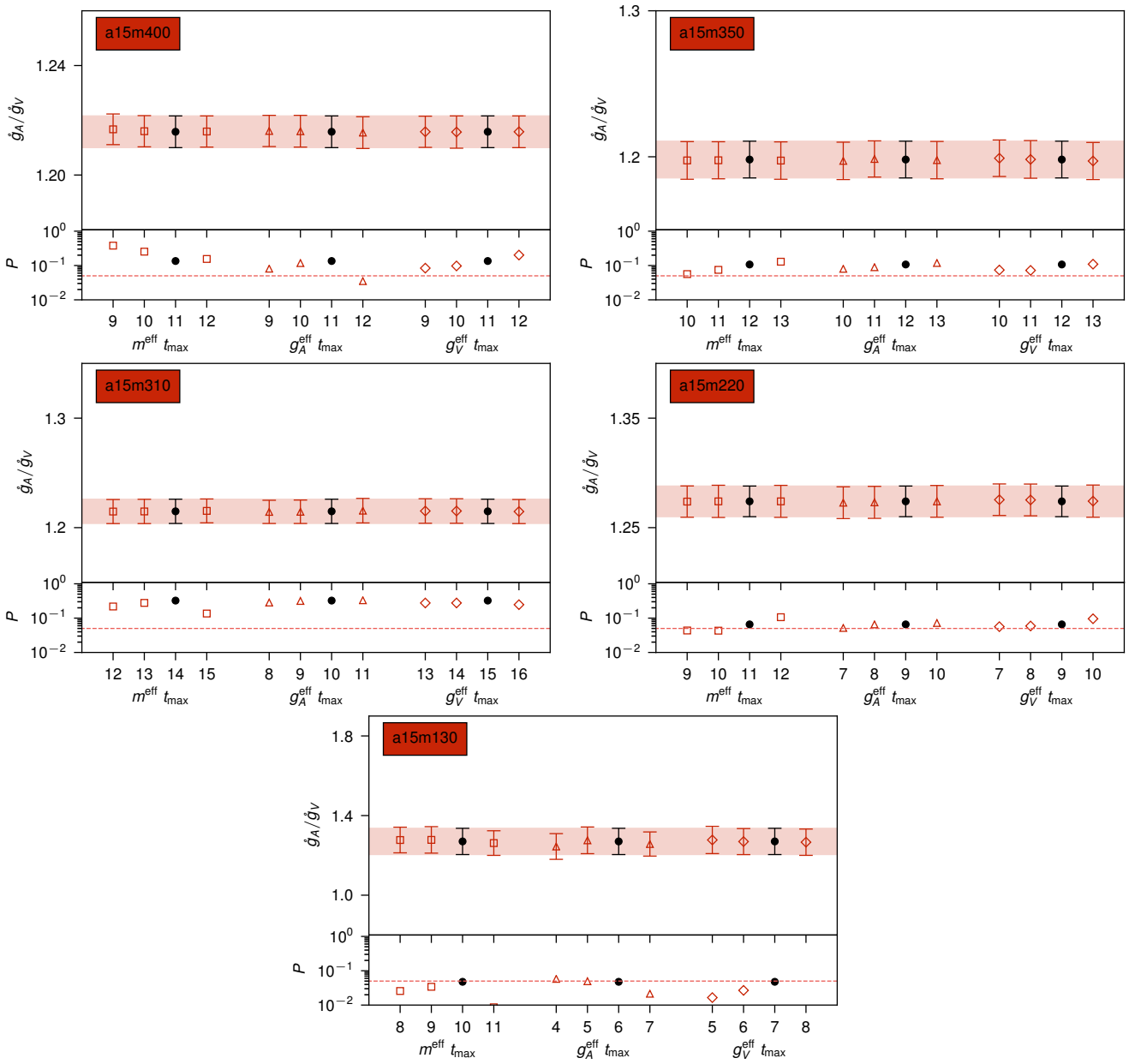
Supplemental Data Figure 16 | Correlator fit t_{min} stability study I. Analogous to Extended Data Fig. 1e. Solid circles accompanied by shaded bands are the preferred simultaneous fits. Varying fit regions for the two-point correlator (□), and axial (△), and vector (◇) effective derivatives are presented. Corresponding P-values are presented, with the dashed red line at $p = 0.05$ discriminating statistical significance of the fit results. Uncertainties are one s.e.m.



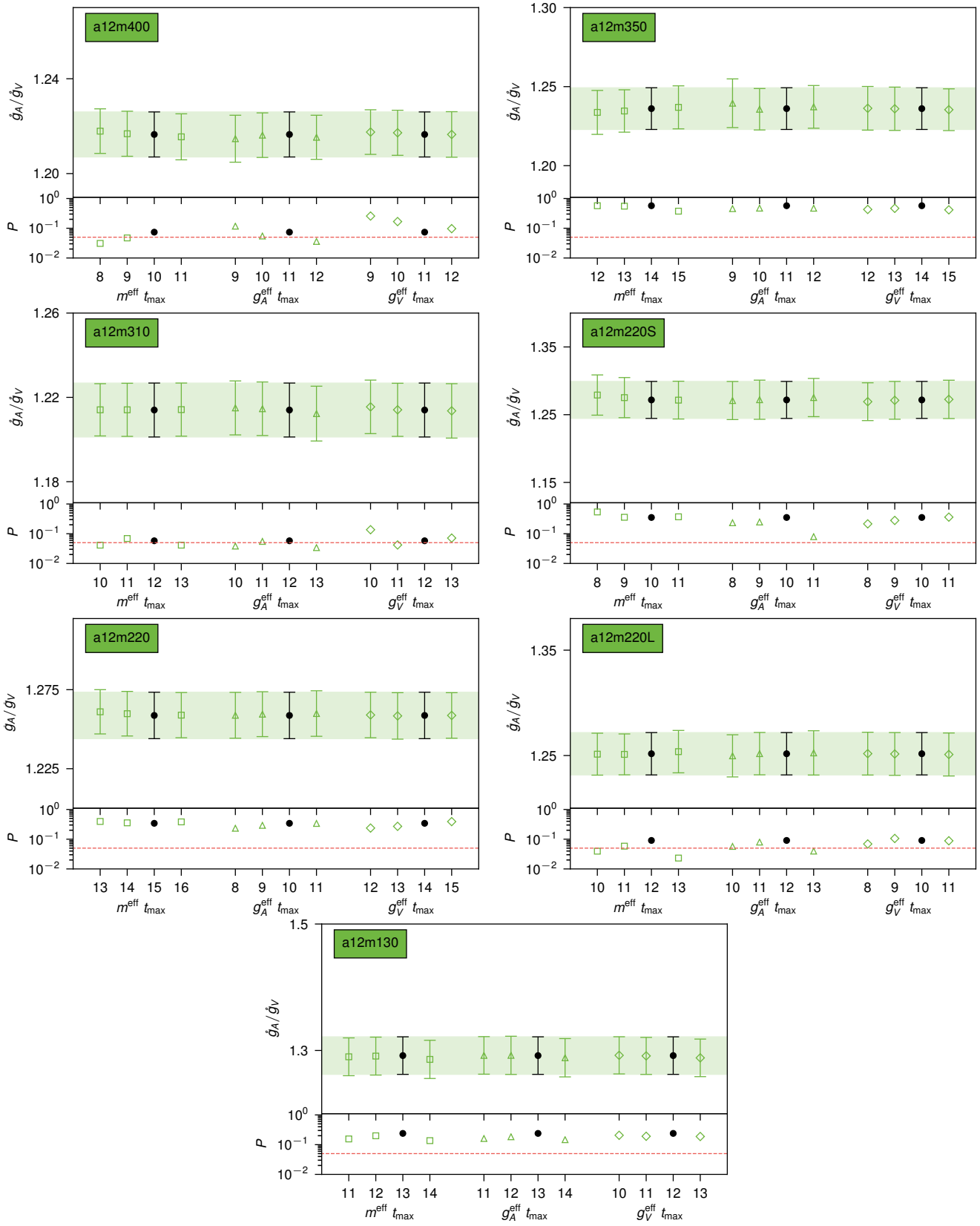
Supplemental Data Figure 17 | Correlator fit t_{min} stability study II. Analogous to Extended Data Fig. 1e and Supplemental Fig. 16 for the remaining ensembles. Uncertainties are one s.e.m.



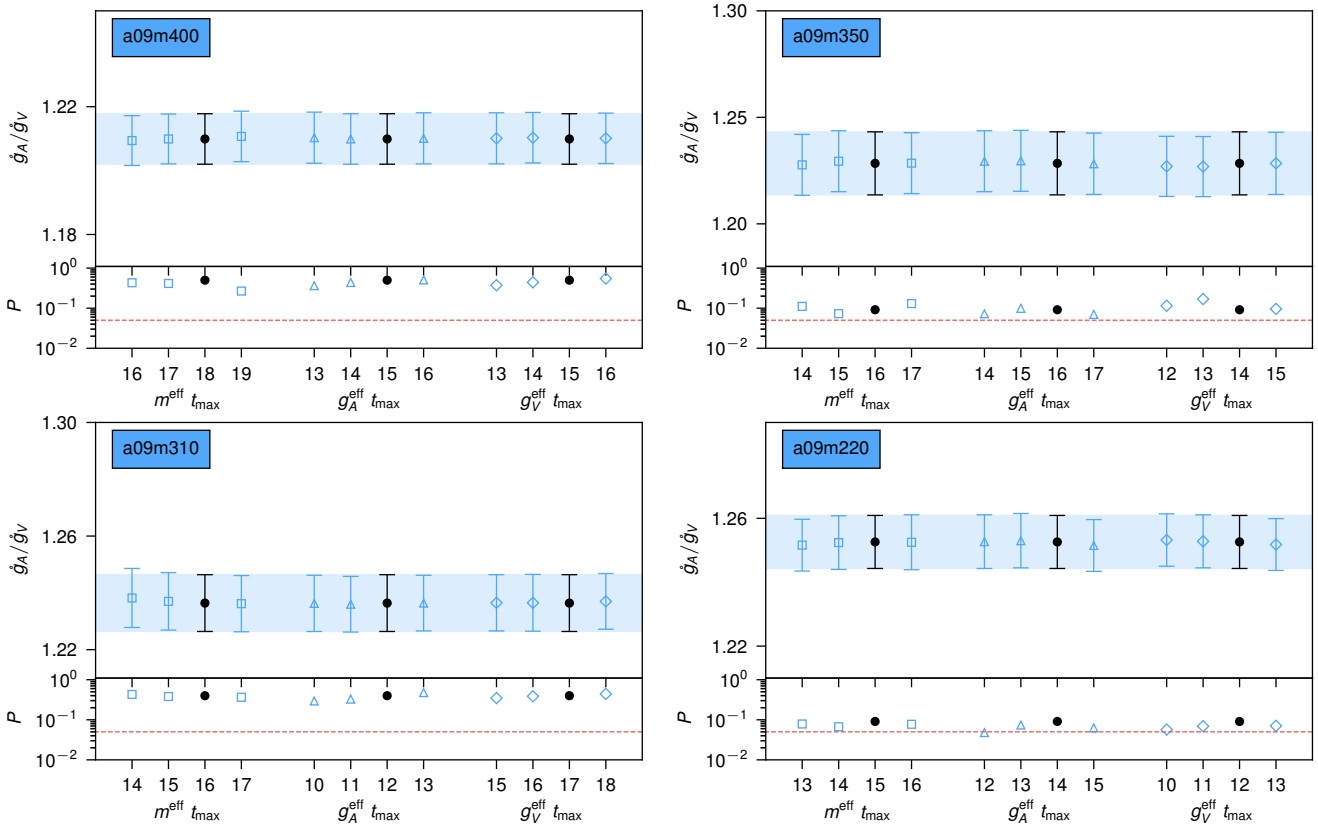
Supplemental Data Figure 18 | Correlator fit t_{min} stability study III. Analogous to Extended Data Fig. 1e and Supplemental Fig. 16 for the remaining ensembles. Uncertainties are one s.e.m.



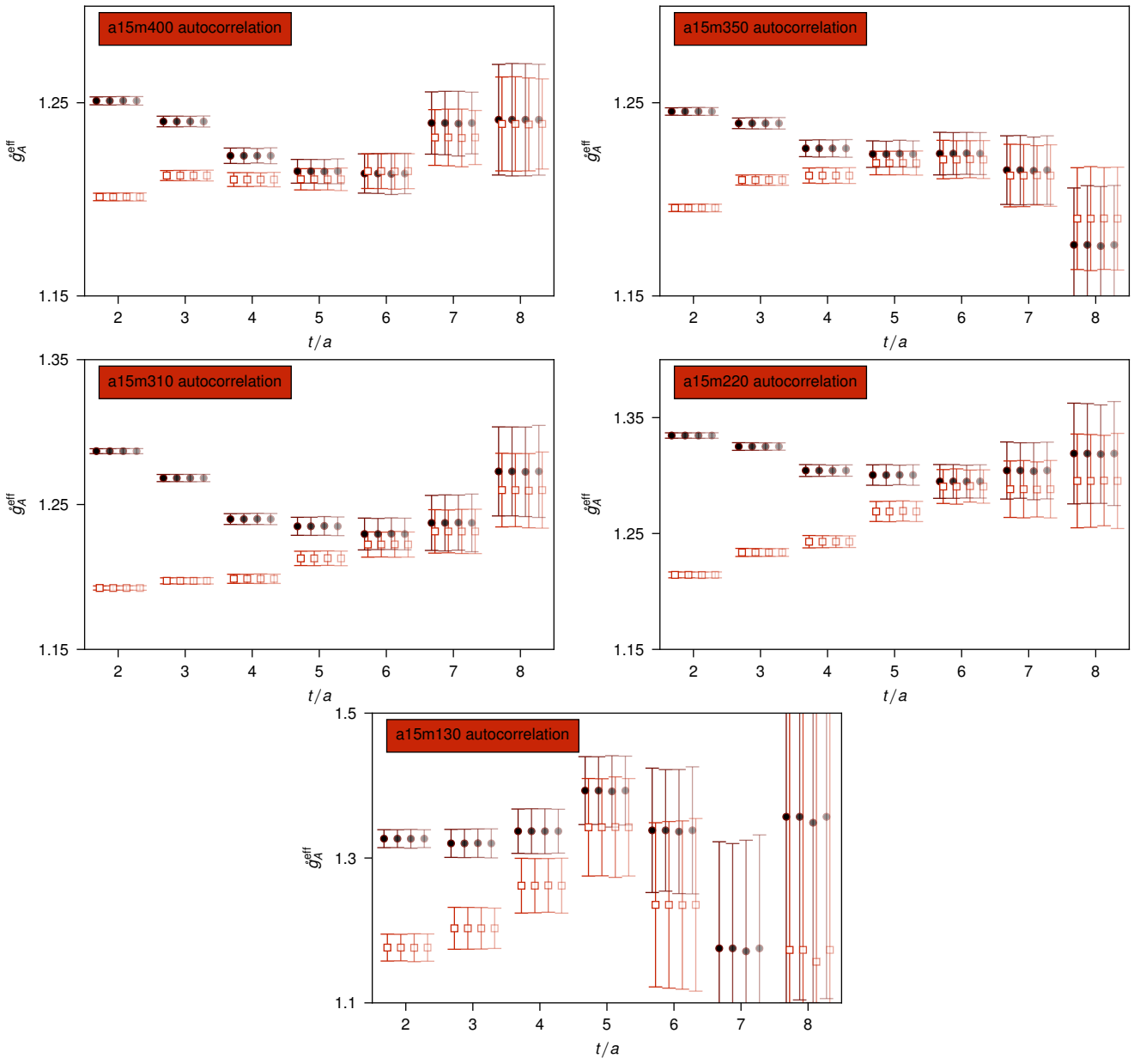
Supplemental Data Figure 19 | Correlator fit t_{max} stability study I. Analogous to Extended Data Fig. 1f. Solid symbols accompanied by shaded bands are the preferred simultaneous fits. Varying fit regions for the two-point correlator (□), and axial (△), and vector (◇) effective derivatives are presented. Corresponding P-values are presented, with the dashed red line at $p = 0.05$ discriminating statistical significance of the fit results. Uncertainties are one s.e.m.



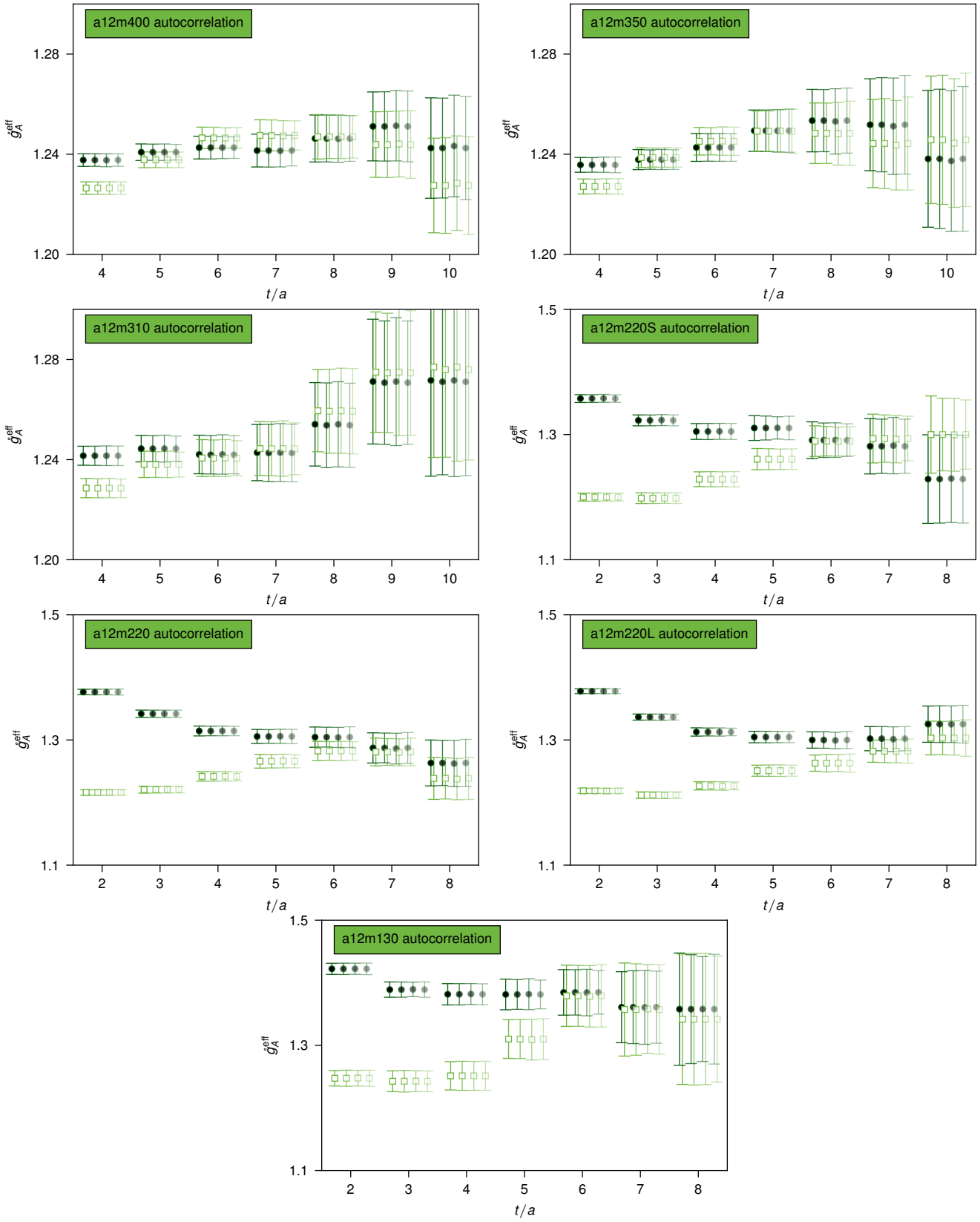
Supplemental Data Figure 20 | Correlator fit t_{max} stability study II. Analogous to Extended Data Fig. 1f and Supplemental Fig. 19 for the remaining ensembles. Uncertainties are one s.e.m.



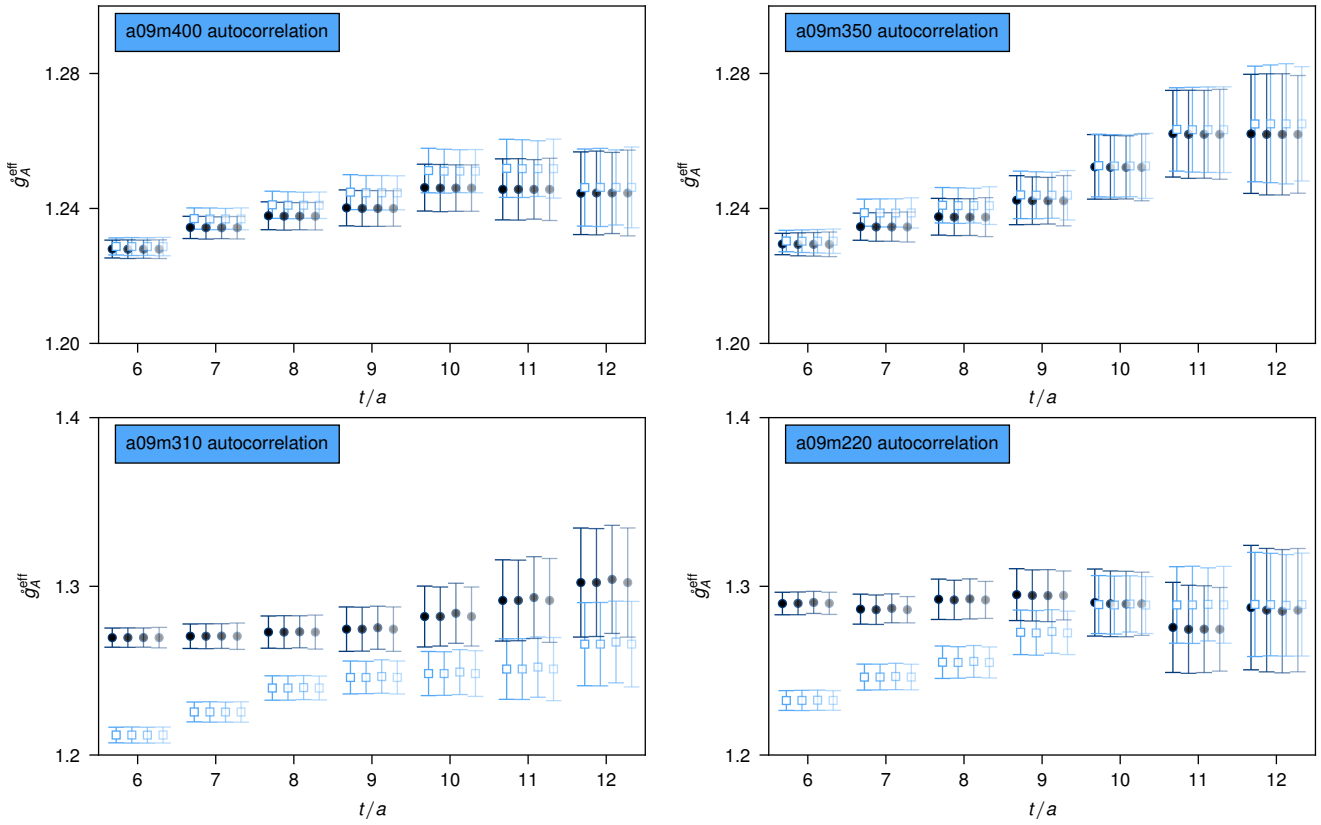
Supplemental Data Figure 21 | Correlator fit stability study III. Analogous to Extended Data Fig. 1f and Supplemental Fig. 19 for the remaining ensembles. Uncertainties are one s.e.m.



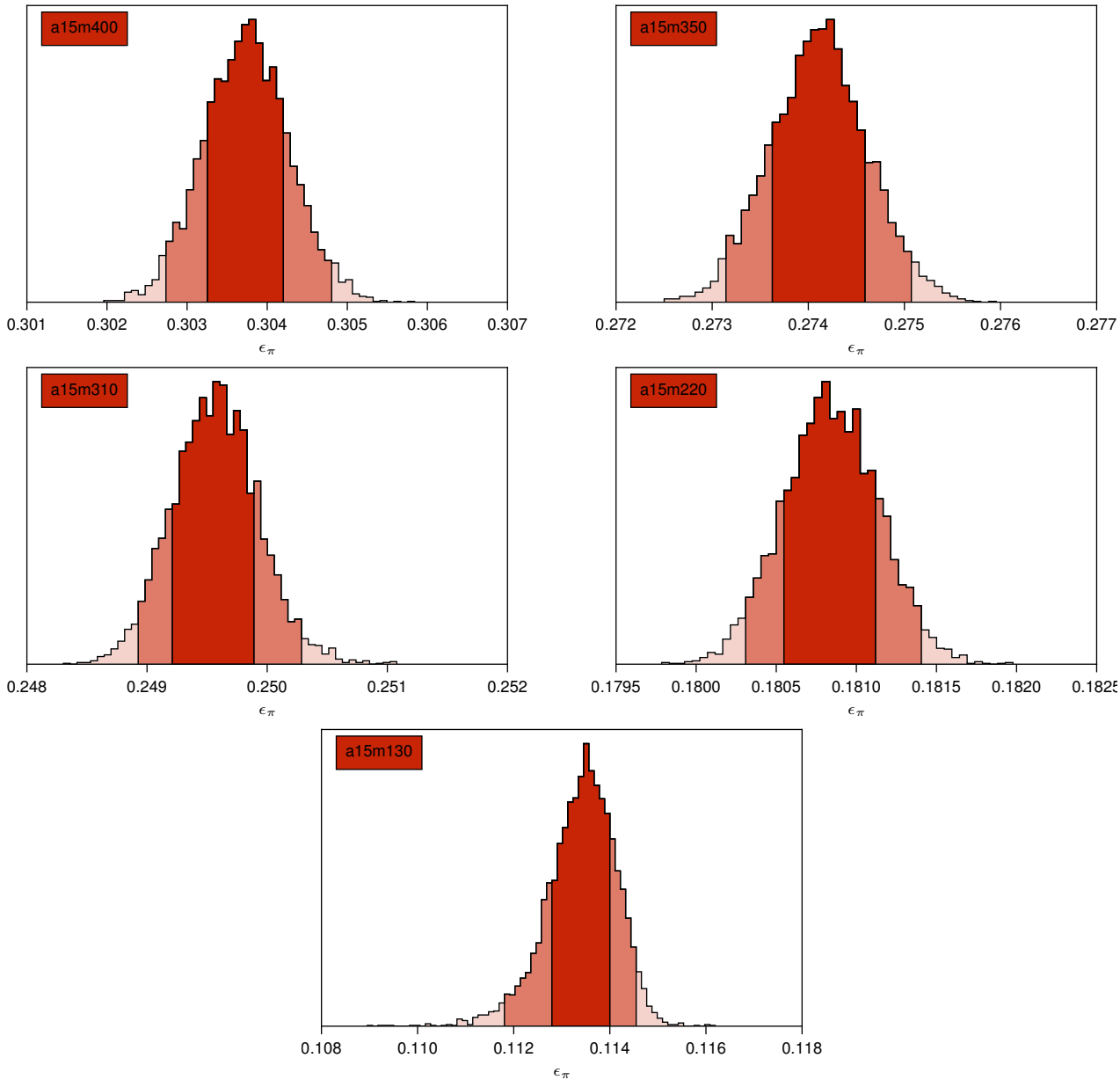
Supplemental Data Figure 22 | Autocorrelation study I. Analogous to Extended Data Fig. 1c and d. Uncertainties are one s.e.m.



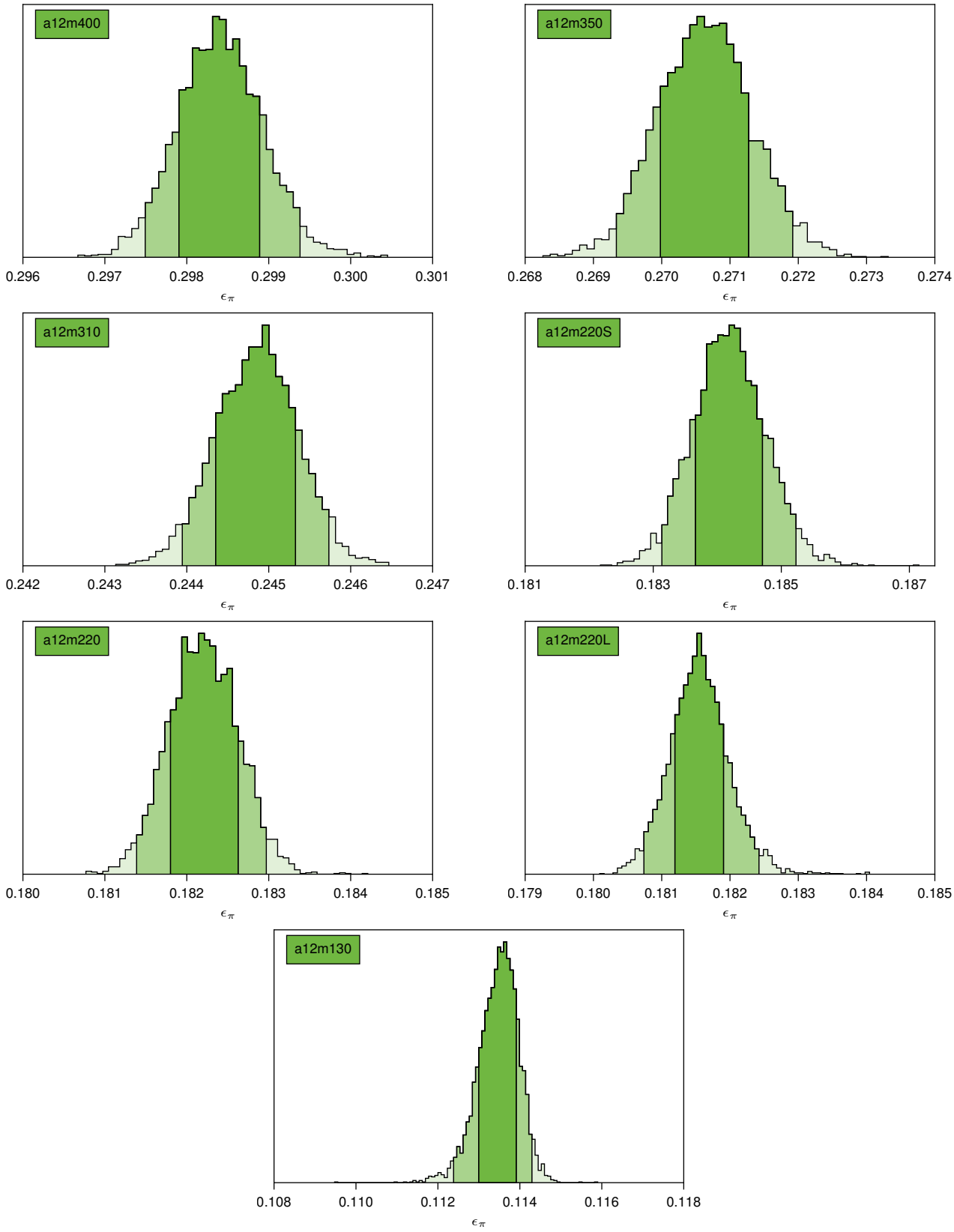
Supplemental Data Figure 23 | Autocorrelation study II. Analogous to Extended Data Fig. 1c and d. Uncertainties are one s.e.m.



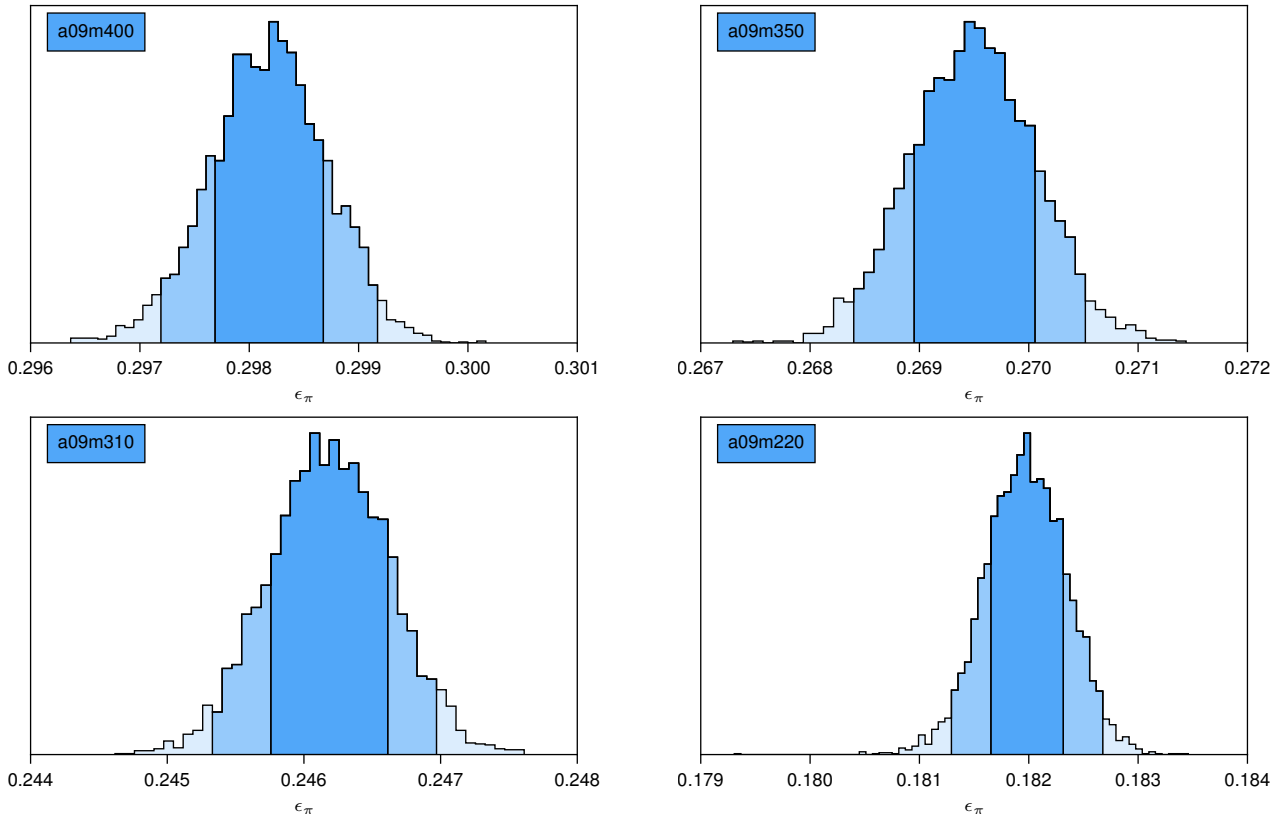
Supplemental Data Figure 24 | Autocorrelation study III. Analogous to Extended Data Fig. 1c and d. Uncertainties are one s.e.m.



Supplemental Data Figure 25 | Histograms I for ϵ_π. Analogous to Extended Data Fig. 1d. The inner shaded regions correspond to the 68% and 95% confidence intervals.



Supplemental Data Figure 26 | Histograms II for ϵ_π. Analogous to Extended Data Fig. 1d and Supplemental Fig. 25 for the remaining ensembles.



Supplemental Data Figure 27 | Histograms III for ϵ_π. Analogous to Extended Data Fig. 1d and Supplemental Fig. 25 for the remaining ensembles.

31. Chang, C. C. and Walker-Loud, A. `project_gA` v1.0. (2018) URL http://doi.org/10.5281/zenodo.1241374.
32. Edwards, R. G. and Joo, B. (SciDAC Collaboration, LHPC Collaboration, UKQCD Collaboration) The Chroma software system for lattice QCD. *Nucl.Phys.Proc.Suppl.* **140**, 832 (2005).
33. Clark, M., Babich, R., Barros, K., Brower, R., and Rebbi, C. Solving Lattice QCD systems of equations using mixed precision solvers on GPUs. *Comput.Phys.Commun.* **181**, 1517 (2010).
34. Babich, R., Clark, M., Joo, B., Shi, G., Brower, R., et al. Scaling Lattice QCD beyond 100 GPUs. (2011).
35. The HDF Group, Hierarchical Data Format, version 5. (1997-NNNN) http://www.hdfgroup.org/HDF5/.
36. Kurth, T., Pochinsky, A., Sarje, A., Syritsyn, S., and Walker-Loud, A. High-Performance I/O: HDF5 for Lattice QCD. *PoS* **LATTICE2014**, 045 (2015).
37. Berkowitz, E. METAQ: Bundle Supercomputing Tasks. (2017) https://github.com/evanberkowitz/metaq.
38. Berkowitz, E., Bouchard, C., Brantley, D. B., Chang, C., Clark, M., Garron, N., Joó, B., Kurth, T., Monahan, C., Monge-Camacho, H., Nicholson, A., Orginos, K., Rinaldi, E., Vranas, P., and Walker-Loud, A. An accurate calculation of the nucleon axial charge with lattice QCD. (2017).
39. Chang, C., Nicholson, A., Rinaldi, E., Berkowitz, E., Garron, N., Brantley, D. B., Monge-Camacho, H., Monahan, C., Bouchard, C., Clark, M., Joó, B., Kurth, T., Orginos, K., Vranas, P., and Walker-Loud, A. in *35th International Symposium on Lattice Field Theory (Lattice 2017) Granada, Spain, June 18-24, 2017* (2017) 1710.06523 URL http://inspirehep.net/record/1631340/files/arXiv:1710.06523.pdf.
40. Chambers, A. J. et al. (QCDSF/UKQCD, CSSM) Feynman-Hellmann approach to the spin structure of hadrons. *Phys. Rev.* **D90**, 014510 (2014).
41. Chambers, A. J. et al. Disconnected contributions to the spin of the nucleon. *Phys. Rev.* **D92**, 114517 (2015).
42. Savage, M. J., Shanahan, P. E., Tiburzi, B. C., Wagman, M. L., Winter, F., Beane, S. R., Chang, E., Davoudi, Z., Detmold, W., and Orginos, K. Proton-proton fusion and tritium β-decay from lattice quantum chromodynamics. (2016).
43. Parisi, G. The Strategy for Computing the Hadronic Mass Spectrum. *Phys. Rept.* **103**, 203 (1984).
44. Lepage, G. P. The analysis of algorithms for lattice field theory. (1989) proceedings of the TASI 1989, edited by T. Degrand and D. Toussaint, World Scientific, Singapore.
45. Follana, E., Mason, Q., Davies, C., Hornbostel, K., Lepage, G. P., Shigemitsu, J., Trottier, H., and Wong, K. (HPQCD, UKQCD) Highly improved staggered quarks on the lattice, with applications to charm physics. *Phys. Rev.* **D75**, 054502 (2007).
46. Bazavov, A. et al. (MILC) Scaling studies of QCD with the dynamical HISQ action. *Phys. Rev.* **D82**, 074501 (2010).
47. Lepage, G. P. and Mackenzie, P. B. Viability of lattice perturbation theory. *Phys. Rev. D* **48**, 2250 (1993).
48. Alford, M., Dimm, W., Lepage, G., Hockney, G., and Mackenzie, P. Lattice {QCD} on small computers. *Physics Letters B* **361**, 87 (1995) ISSN 0370-2693.
49. Brower, R. C., Neff, H., and Orginos, K. Mobius fermions: Improved domain wall chiral fermions. *Nucl. Phys. Proc. Suppl.* **140**, 686 (2005) [,686(2004)].
50. Brower, R. C., Neff, H., and Orginos, K. Mobius fermions. *Nucl. Phys. Proc. Suppl.* **153**, 191 (2006).
51. Brower, R. C., Neff, H., and Orginos, K. The Móbius Domain Wall Fermion Algorithm. (2012).
52. Lüscher, M. Properties and uses of the Wilson flow in lattice QCD. *JHEP* **1008**, 071 (2010).
53. Lohmayer, R. and Neuberger, H. Continuous smearing of Wilson Loops. *PoS* **LATTICE2011**, 249 (2011).
54. Lüscher, M. Exact chiral symmetry on the lattice and the Ginsparg-Wilson relation. *Phys. Lett.* **B428**, 342 (1998).
55. Ginsparg, P. H. and Wilson, K. G. A Remnant of Chiral Symmetry on the Lattice. *Phys. Rev.* **D25**, 2649 (1982).
56. Shamir, Y. Chiral fermions from lattice boundaries. *Nucl. Phys.* **B406**, 90 (1993).
57. Furman, V. and Shamir, Y. Axial symmetries in lattice QCD with Kaplan fermions. *Nucl. Phys.* **B439**, 54 (1995).
58. Kaplan, D. B. A Method for simulating chiral fermions on the lattice. *Phys. Lett.* **B288**, 342 (1992).
59. Allton, C. et al. (RBC-UKQCD) Physical Results from 2+1 Flavor Domain Wall QCD and SU(2) Chiral Perturbation Theory. *Phys. Rev.* **D78**, 114509 (2008).
60. Basak, S., Edwards, R. G., Fleming, G. T., Heller, U. M., Morningstar, C., Richards, D., Sato, I., and Wallace, S. Group-theoretical construction of extended baryon operators in lattice QCD. *Phys. Rev.* **D72**, 094506 (2005).
61. Basak, S., Edwards, R., Fleming, G. T., Heller, U. M., Morningstar, C., Richards, D., Sato, I., and Wallace, S. J. (Lattice Hadron Physics (LHPC)) Clebsch-Gordan construction of lattice interpolating fields for excited baryons. *Phys. Rev.* **D72**, 074501 (2005).
62. Beane, S. R., Detmold, W., Luu, T. C., Orginos, K., Parreno, A., Savage, M., Torok, A., and Walker-Loud, A. High Statistics Analysis using Anisotropic Clover Lattices: (I) Single Hadron Correlation Functions. *Phys.Rev.* **D79**, 114502 (2009).
63. Lepage, G. P. `lsqfit` v9.1.2. (2017) URL https://doi.org/10.5281/zenodo.60221.
64. Martinelli, G., Pittori, C., Sachrajda, C. T., Testa, M., and Vladikas, A. A General method for nonperturbative renormalization of lattice operators. *Nucl. Phys.* **B445**, 81 (1995).
65. Aoki, Y., Boyle, P., Christ, N., Dawson, C., Donnellan, M., et al. Non-perturbative renormalization of quark bilinear operators and B(K) using domain wall fermions. *Phys.Rev.* **D78**, 054510 (2008).
66. Sturm, C., Aoki, Y., Christ, N., Izubuchi, T., Sachrajda, C., et al. Renormalization of quark bilinear operators in a momentum-subtraction scheme with a nonexceptional subtraction point. *Phys.Rev.* **D80**, 014501 (2009).
67. Gockeler, M., Horsley, R., Oelrich, H., Perlt, H., Petters, D., et al. Nonperturbative renormalization of composite operators in lattice QCD. *Nucl.Phys.* **B544**, 699 (1999).
68. Giusti, L., Paciello, M. L., Parrinello, C., Petrarca, S., and Taglienti, B. Problems on lattice gauge fixing. *Int. J. Mod. Phys.* **A16**, 3487 (2001).
69. Borsanyi, S. et al. High-precision scale setting in lattice QCD. *JHEP* **09**, 010 (2012).
70. Walker-Loud, A. et al. Light hadron spectroscopy using domain wall valence quarks on an Asqtad sea. *Phys. Rev.* **D79**, 054502 (2009).
71. Torok, A. et al. (NPLQCD) Meson-Baryon Scattering Lengths from Mixed-Action Lattice QCD. *Phys. Rev.* **D81**, 074506 (2010).
72. Ishikawa, K. I. et al. (PACS-CS) SU(2) and SU(3) chiral perturbation theory analyses on baryon masses in 2+1 flavor lattice QCD. *Phys. Rev.* **D80**, 054502 (2009).
73. Jenkins, E. E., Manohar, A. V., Negele, J. W., and Walker-Loud, A. A Lattice Test of 1/N(c) Baryon Mass Relations. *Phys. Rev.* **D81**, 014502 (2010).
74. Walker-Loud, A. Evidence for non-analytic light quark mass dependence in the baryon spectrum. *Phys. Rev.* **D86**, 074509 (2012).
75. Bernard, V., Kaiser, N., Kambor, J., and Meißner, U. G. Chiral structure of the nucleon. *Nucl. Phys.* **B388**, 315 (1992).
76. Jenkins, E. E. and Manohar, A. V. Chiral corrections to the baryon axial currents. *Phys. Lett.* **B259**, 353 (1991).
77. Kambor, J. and Mojzis, M. Field redefinitions and wave function renormalization to O(p**4) in heavy baryon chiral perturbation theory. *JHEP* **04**, 031 (1999).
78. Bernard, V. and Meißner, U.-G. The Nucleon axial-vector coupling beyond one loop. *Phys. Lett.* **B639**, 278 (2006).
79. 't Hooft, G. A PLANAR DIAGRAM THEORY FOR STRONG INTERACTIONS. *Nucl. Phys.* **B72**, 461 (1974).
80. Witten, E. Baryons in the 1/n Expansion. *Nucl. Phys.* **B160**, 57 (1979).
81. Butler, M. N. and Savage, M. J. Electromagnetic polarizability of the nucleon in chiral perturbation theory. *Phys. Lett.* **B294**, 369 (1992).
82. Griesshammer, H. W., McGovern, J. A., Phillips, D. R., and Feldman, G. Using effective field theory to analyse low-energy Compton scattering data from protons and light nuclei. *Prog. Part. Nucl. Phys.* **67**, 841 (2012).
83. Hemmert, T. R., Procura, M., and Weise, W. Quark mass dependence of the nucleon axial vector coupling constant. *Phys. Rev.* **D68**, 075009 (2003).
84. Dashen, R. F. and Manohar, A. V. Baryon - pion couplings from large N(c) QCD. *Phys. Lett.* **B315**, 425 (1993).
85. Dashen, R. F. and Manohar, A. V. 1/N(c) corrections to the baryon axial currents in QCD. *Phys. Lett.* **B315**, 438 (1993).
86. Jenkins, E. E. Baryon hyperfine mass splittings in large N QCD. *Phys. Lett.* **B315**, 441 (1993).
87. Dashen, R. F., Jenkins, E. E., and Manohar, A. V. The 1/N(c) expansion for baryons. *Phys. Rev.* **D49**, 4713 (1994) [Erratum: Phys. Rev.D51,2489(1995)].
88. Dashen, R. F., Jenkins, E. E., and Manohar, A. V. Spin flavor structure of large N(c) baryons. *Phys. Rev.* **D51**, 3697 (1995).
89. Jenkins, E. E. Chiral Lagrangian for baryons in the 1/n(c) expansion. *Phys. Rev.* **D53**, 2625 (1996).
90. Jenkins, E. E. Large N(c) baryons. *Ann. Rev. Nucl. Part. Sci.* **48**, 81 (1998).
91. Fettes, N., Meissner, U.-G., Mojzis, M., and Steininger, S. The Chiral effective pion nucleon Lagrangian of order p**4. *Annals*

Phys. **283**, 273 (2000) [Erratum: Annals Phys.288,249(2001)].
92. Jenkins, E. E. and Lebed, R. F. Baryon mass splittings in the 1/N(c) expansion. *Phys. Rev.* **D52**, 282 (1995).
93. Calle Cordon, A. and Goity, J. L. Baryon Masses and Axial Couplings in the Combined $1/N_c$ and Chiral Expansions. *Phys. Rev.* **D87**, 016019 (2013).
94. Lellouch, L. and Luscher, M. Weak transition matrix elements from finite volume correlation functions. *Commun. Math. Phys.* **219**, 31 (2001).
95. Agadjanov, A., Bernard, V., Meißner, U. G., and Rusetsky, A. A framework for the calculation of the $\Delta N\gamma^\star$ transition form factors on the lattice. *Nucl. Phys.* **B886**, 1199 (2014).
96. Briceño, R. A., Hansen, M. T., and Walker-Loud, A. Multichannel $1 \to 2$ transition amplitudes in a finite volume. *Phys. Rev.* **D91**, 034501 (2015).
97. Owen, B. J., Kamleh, W., Leinweber, D. B., Mahbub, M. S., and Menadue, B. J. Transition of $\rho \to \pi\gamma$ in lattice QCD. *Phys. Rev.* **D92**, 034513 (2015).
98. Briceno, R. A., Dudek, J. J., Edwards, R. G., Shultz, C. J., Thomas, C. E., and Wilson, D. J. The resonant $\pi^+\gamma \to \pi^+\pi^0$ amplitude from Quantum Chromodynamics. *Phys. Rev. Lett.* **115**, 242001 (2015).
99. Briceño, R. A., Dudek, J. J., Edwards, R. G., Shultz, C. J., Thomas, C. E., and Wilson, D. J. The $\pi\pi \to \pi\gamma^\star$ amplitude and the resonant $\rho \to \pi\gamma^\star$ transition from lattice QCD. *Phys. Rev.* **D93**, 114508 (2016).
100. Gérardin, A., Meyer, H. B., and Nyffeler, A. Lattice calculation of the pion transition form factor $\pi^0 \to \gamma^*\gamma^*$. *Phys. Rev.* **D94**, 074507 (2016).
101. Briceño, R. A., Hansen, M. T., and Walker-Loud, A. Multichannel $1 \to 2$ transition amplitudes in a finite volume. *PoS* **LATTICE2014**, 095 (2015).
102. Walker-Loud, A. New lessons from the nucleon mass, lattice QCD and heavy baryon chiral perturbation theory. *PoS* **LATTICE2008**, 005 (2008).
103. Walker-Loud, A. Baryons in/and Lattice QCD. *PoS* **CD12**, 017 (2013).
104. Beane, S. R. and Savage, M. J. Baryon axial charge in a finite volume. *Phys. Rev.* **D70**, 074029 (2004).
105. Jaffe, R. L. Delocalization of the axial charge in the chiral limit. *Phys. Lett.* **B529**, 105 (2002).
106. Cohen, T. D. The Extraction of g(A) from finite volume systems: The Long and short of it. *Phys. Lett.* **B529**, 50 (2002).
107. Yamazaki, T., Aoki, Y., Blum, T., Lin, H. W., Lin, M. F., Ohta, S., Sasaki, S., Tweedie, R. J., and Zanotti, J. M. (RBC+UKQCD) Nucleon axial charge in 2+1 flavor dynamical lattice QCD with domain wall fermions. *Phys. Rev. Lett.* **100**, 171602 (2008).
108. Yamazaki, T., Aoki, Y., Blum, T., Lin, H.-W., Ohta, S., Sasaki, S., Tweedie, R., and Zanotti, J. Nucleon form factors with 2+1 flavor dynamical domain-wall fermions. *Phys. Rev.* **D79**, 114505 (2009).
109. Bar, O., Rupak, G., and Shoresh, N. Chiral perturbation theory at $\mathcal{O}(a^2)$ for lattice QCD. *Phys. Rev.* **D70**, 034508 (2004).
110. Sharpe, S. R. and Singleton, R. L., Jr Spontaneous flavor and parity breaking with Wilson fermions. *Phys. Rev.* **D58**, 074501 (1998).
111. Bar, O., Bernard, C., Rupak, G., and Shoresh, N. Chiral perturbation theory for staggered sea quarks and Ginsparg-Wilson valence quarks. *Phys. Rev.* **D72**, 054502 (2005).
112. Tiburzi, B. C. Baryons with Ginsparg-Wilson quarks in a staggered sea. *Phys. Rev.* **D72**, 094501 (2005) [Erratum: Phys. Rev.D79,039904(2009)].
113. Chen, J.-W., O'Connell, D., Van de Water, R. S., and Walker-Loud, A. Ginsparg-Wilson pions scattering on a staggered sea. *Phys. Rev.* **D73**, 074510 (2006).
114. Chen, J.-W., O'Connell, D., and Walker-Loud, A. Two Meson Systems with Ginsparg-Wilson Valence Quarks. *Phys. Rev.* **D75**, 054501 (2007).
115. Orginos, K. and Walker-Loud, A. Mixed meson masses with domain-wall valence and staggered sea fermions. *Phys. Rev.* **D77**, 094505 (2008).
116. Jiang, F.-J. Mixed Action Lattice Spacing Effects on the Nucleon Axial Charge. (2007).
117. Chen, J.-W., O'Connell, D., and Walker-Loud, A. Universality of mixed action extrapolation formulae. *JHEP* **04**, 090 (2009).
118. Chen, J.-W., Golterman, M., O'Connell, D., and Walker-Loud, A. Mixed Action Effective Field Theory: An Addendum. *Phys. Rev.* **D79**, 117502 (2009).
119. Symanzik, K. Continuum Limit and Improved Action in Lattice Theories. 2. O(N) Nonlinear Sigma Model in Perturbation Theory. *Nucl. Phys.* **B226**, 205 (1983).
120. Kass, R. E. and Raftery, A. E. Bayes factors. *Journal of the American Statistical Association* **90**, 773 (1995) ISSN 01621459.
121. Czarnecki, A., Marciano, W. J., and Sirlin, A. Precision measurements and CKM unitarity. *Phys. Rev.* **D70**, 093006 (2004).
122. Marciano, W. J. and Sirlin, A. Improved calculation of electroweak radiative corrections and the value of V(ud). *Phys. Rev. Lett.* **96**, 032002 (2006).
123. Arndt, D. and Savage, M. J. Chiral corrections to matrix elements of twist-2 operators. *Nucl. Phys.* **A697**, 429 (2002).
124. Beane, S. R. and Savage, M. J. Nucleons in two flavor partially quenched chiral perturbation theory. *Nucl. Phys.* **A709**, 319 (2002).
125. Gell-Mann, M., Oakes, R. J., and Renner, B. Behavior of current divergences under SU(3) x SU(3). *Phys. Rev.* **175**, 2195 (1968).
126. Lubicz, V., Martinelli, G., Sachrajda, C. T., Sanfilippo, F., Simula, S., Tantalo, N., and Tarantino, C. Electromagnetic corrections to the leptonic decay rates of charged pseudoscalar mesons: lattice results. *PoS* **LATTICE2016**, 290 (2016).
127. Cirigliano, V. and Neufeld, H. A note on isospin violation in Pl2(gamma) decays. *Phys. Lett.* **B700**, 7 (2011).
128. Rosner, J. L., Stone, S., and Van de Water, R. S. Leptonic Decays of Charged Pseudoscalar Mesons - 2015. *Submitted to: Particle Data Book* (2015).
129. Brown, M. A.-P., Dees, E. B., Adamek, E., Allgeier, B., Blatnik, M., Bowles, T. J., Broussard, L. J., Carr, R., Clayton, S., Cude-Woods, C., et al. (UCNA) New result for the neutron β-asymmetry parameter A_0 from UCNA. (2017).