



Universal pattern in fermion mass ratios and implications for flavor models

Samir Varma^a

VS Asset Management, LLC, Cos Cob, CT 06807, USA

Received: 10 July 2025 / Accepted: 9 September 2025
© The Author(s) 2025

Abstract We show that a single analytic expression using only the labels that distinguish Standard Model mass eigenstates—the ordering index $d \in \{1, 2, 3\}$ and electric charge q —fits all fermion mass ratios with high precision and stable $\mathcal{O}(1)$ parameters. The functional form $\ln(m_1/m_2) = d_1^{\zeta_1} d_2^{\zeta_2} \kappa^{[(d_1|q_1|)^{\gamma(\mu)} - (d_2|q_2|)^{\gamma(\mu)}]}$ reproduces all 30 independent quark ratios from M_Z to 10^{12} GeV with $\chi^2/\text{dof} \lesssim 10^{-6}$ using per-scale fits (24 parameters), fixed- κ fits (18), or global fits (5–6). Parameters remain near $\kappa \simeq 2.3$, $\zeta_1 \simeq 1.16$, $\zeta_2 \simeq -0.8$, $\gamma \simeq 1.1$. Three orthogonal checks establish specificity and robustness: (i) a 120-permutation test where only the physical mapping of ratios to $(d, |q|)$ fits; (ii) a 720-basis scan in which the adjacent-ratio basis is uniquely preferred by AIC/BIC and many alternatives fail to converge; (iii) alternative functional forms are decisively disfavored by AIC/BIC. The same form, with natural parameters, describes charged-lepton and neutrino ratios; for $q = 0$ the charge term vanishes, leaving pure d^ζ scaling. We sketch three minimal mechanisms—modified Froggatt–Nielsen, partial compositeness, and a Randall–Sundrum overlap—that generate the additive $d-|q|$ dependence at the exponent level and reduce to the neutrino limit. The pattern predicts percent-level deviations from standard running near 10^{14} GeV and, with meV-scale anchors, $\Sigma m_\nu \simeq 0.06$ eV. To our knowledge, no previous empirical relation achieves simultaneous cross-sector consistency with such parameter stability; the result offers a compact target for flavor model building.

1 Introduction

The masses of the Standard Model fermions span more than 12 orders of magnitude, yet no accepted principle explains the observed pattern. Existing empirical relations—Koide’s charged-lepton formula [1], the Georgi–Jarlskog relation [2],

Froggatt–Nielsen textures [3] and many others—describe at most one sector at a time and fail when extended unchanged to the remaining sectors. Consequently, the literature contains *no* single analytic expression that fits *all* quark, charged-lepton, and neutrino mass ratios simultaneously.¹

We close this gap by exploiting a minimal and model-agnostic choice of *labels* for mass eigenstates: an *ordering index* $d \in \{1, 2, 3\}$ that ranks the states by mass within a sector, and the electric charge q in units of $|e|$. Only q is a conserved gauge quantum number; d is a bookkeeping index used to organize the data and fits (we return to this point in Sects. 4–5). If an underlying organizing principle governs fermion masses, it may act through these two labels. Guided by this constraint, we propose the minimal ansatz:

$$\ln\left(\frac{m_1}{m_2}\right) = d_1^{\zeta_1} d_2^{\zeta_2} \kappa^{[(d_1|q_1|)^{\gamma(\mu)} - (d_2|q_2|)^{\gamma(\mu)}]}, \quad (1)$$

where $m_1 > m_2$,² all masses are $\overline{\text{MS}}$ —six reference scales and uncertainties (across ten orders of magnitude) taken from [5]³—and with a mildly running exponent,⁴

¹ While a rare intersecting-brane Pati–Salam compactification has been shown to reproduce all Standard Model masses and mixings [4], that construction requires 12 adjoint Higgs fields and significant flux tuning, placing it well outside the simple texture, Koide-type, GUT, or seesaw ansatz considered here.

² Here the subscripts 1/2 denote the heavier/lighter mass *within that pair*; they are not the global indices used later in the basis scan.

³ All masses up to 10^{12} GeV are taken from [5], which are derived from PDG world averages and evolved using five-loop quantum chromodynamics (QCD) renormalization-group equations; they coincide with the latest PDG inputs at their respective renormalization scales within quoted uncertainties. For comparison with independent lattice determinations, see the FLAG 24 averages [6].

⁴ This running exponent plays a role similar to an anomalous dimension in quantum field theory: it governs the scale dependence of effective mass ratios, even though the formula itself is not derived from a renormalizable Lagrangian.

^a e-mail: samir@vsasset.com (corresponding author)

$$\gamma(\mu) = a + b \ln(\mu/M_Z) \text{ [5 param.], or}$$

$$\gamma(\mu) = a + b \ln(\mu/M_Z) + c \ln^2(\mu/M_Z) \text{ [6 param.].} \quad (2)$$

Equation (1) reproduces

1. all 30 independent quark mass ratios between M_Z and 10^{12} GeV,
2. both charged-lepton ratios over the same scales,⁵ and
3. the two neutrino ratios implied by NuFIT-6.0 [7] oscillation data.

No structural modification is required as we move from one sector to the next, and all fitted parameters remain $\mathcal{O}(1)$.

The empirical strength of Eq. (1) varies with data richness. For quarks, 30 measurements confront at most five or six parameters,⁶ yielding $\chi^2/\text{dof} \sim 10^{-5}$ and residuals below the quoted experimental uncertainties. Charged leptons offer a consistency test with under-determined parameters; nonetheless, the fitted values stay natural in size. For neutrinos, where $q = 0$ removes the κ term entirely, the formula reduces to d^z scaling and predicts absolute masses once the lightest eigenvalue is anchored. In each section, we contrast these results with the best-known alternatives and show explicitly why those alternatives cannot describe all three sectors at once.

The decision to parametrize mass ratios using only generation index and electric charge—the two labels that universally distinguish Standard Model fermions—was motivated by seeking maximum empirical coverage with minimal theoretical bias. The surprising success of this ansatz, particularly the cross-sector universality and the permutation analysis, suggests these labels may play a more fundamental role in flavor physics than previously appreciated [8]. Section 7 translates these empirical findings into concrete guidance for model builders.

Our main findings are summarized here and detailed in the following sections:

- Section 3 demonstrates the quark fits across multiple approaches, establishes parameter stability when fitting with or without experimental uncertainties, shows that even constraining κ to be scale-independent preserves the empirical pattern, and proves through comprehensive permutation analysis that only the physical assignment of quantum numbers to mass ratios yields acceptable fits.
- Section 4 applies the same structure to charged leptons, highlighting natural parameter sizes despite under-determination.

⁵ Charged-fermion ratios drift by $(1.0 \pm 1.4) \times 10^{-3}$ from M_Z to 10^{12} GeV, consistent with zero, so a constant γ suffices.

⁶ We test both a five-parameter $\gamma(\mu) = a + b \ln(\mu/\mu_0)$ and six-parameter $\gamma(\mu) = a + b \ln(\mu/\mu_0) + c \ln^2(\mu/\mu_0)$, with $\mu_0 = M_Z$.

- Section 5 treats the $q = 0$ neutrino limit, derives absolute-mass predictions, and explains why the ratios are expected not to run.
- Section 6 collects falsifiable consequences: quark ratios at 10^{14} GeV via deviations from RunDec v3.1 [9] at five-loop QCD and two-loop electroweak orders, and values for Σm_ν , m_β , and $m_{\beta\beta}$ relevant to upcoming surveys.
- Section 7 identifies specific theoretical modifications required for flavor models to reproduce the empirical pattern, including constraints on Froggatt–Nielsen, warped extra dimensions, and discrete symmetry frameworks.
- Section 8 discusses the hierarchy of evidence strength across sectors, compares Eq. (1) with texture and grand unified theory (GUT) relations, and outlines possible theoretical origins.
- Change-of-basis scan (information criteria). Among all 720 independent ordered chained five-ratio bases (adjacent basis plus 719 alternatives), the adjacent basis is uniquely preferred: it is the *only* basis whose five-parameter global fit attains a *negative* Akaike information criterion (AIC; same conclusion with the Bayesian information criterion [BIC]). Of the 719 non-adjacent bases, only 343 even admit a finite-likelihood fit; the remaining 376 fail to converge (effectively “ $+\infty$ ” AIC). See Fig. 2 and Table 1.

2 A puzzle in plain sight

Figure 1 plots the five quark-mass ratios at the six reference scales used in this study. Neither a linear (shown) nor the logarithmic view (not shown) reveals a monotonic or polynomial trend; the points resemble uncorrelated scatter. What follows shows that the “scatter” is in fact threaded by a simple expression whose four shape parameters are surprisingly stable under every refitting scheme we could devise.

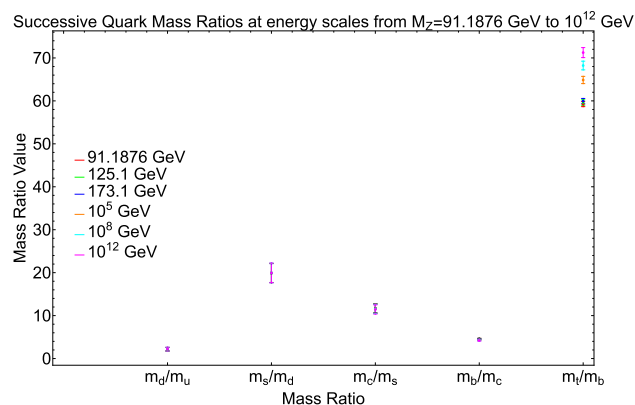


Fig. 1 Quark-mass ratios at six reference scales. No simple trend is apparent, motivating the search for an empirical organizing principle

Throughout the paper, we underscore three empirical facts: (i) there is no a priori reason why the same functional form should succeed across all three sectors, each with different levels of empirical constraint; (ii) alternative relations fail to do so without sector-specific amendments; and (iii) all fitted parameters remain natural in magnitude. Taken together, these points argue that Eq. (1) captures a genuine organizing principle behind the fermion mass spectrum, independent of any ultraviolet model.

3 Quark sector

Six reference scales—91.1876 GeV (M_Z), 125.1 GeV, 173.1 GeV, 10^5 GeV, 10^8 GeV, and 10^{12} GeV—are provided by the running-mass compilation of Huang [5]. At each scale, we construct the five successive ratios,

$$\{m_d/m_u, m_s/m_d, m_c/m_s, m_b/m_c, m_t/m_b\},$$

so that all 30 independent quark ratios enter the analysis.

Why consecutive ratios suffice The five adjacent-generation ratios form an algebraically independent basis for all quark mass ratios. Writing $r_k \equiv \ln(m_{k+1}/m_k)$ after fixing a *global* ascending enumeration $(u, d, s, c, b, t) \equiv (1, \dots, 6)$ (indices increase with mass; this enumeration is independent of the pairwise convention $m_1 > m_2$ used in Eq. (1)),

$$\ln\left(\frac{m_i}{m_j}\right) = \sum_{k=j}^{i-1} r_k \quad (i > j).$$

Thus, in log space the vector of adjacent ratios $\mathbf{r} = (r_1, \dots, r_5)^T$ spans the $\mathbb{Z}$ -module of all ratio logs; any other independent five-tuple of ratios is obtained by a left multiplication $\mathbf{r}' = A\mathbf{r}$ with an integer unimodular matrix $A \in \text{GL}(5, \mathbb{Z})$. In a properly propagated (co)variance treatment this change of basis leaves the weighted least-squares optimum invariant up to numerical rounding, so the choice of adjacent ratios is a *convenience*, not a bias (see Appendix B for a compact proof).

Practical basis choice: information content and numerical stability Appendix C proves that, with full covariance propagation, changing the basis of ratios is a unimodular transformation that leaves the weighted least-squares optimum invariant. In practice, however, one often works with central values or with incomplete covariance, and then the basis choice can affect both convergence and model selection metrics. To test robustness we evaluated the *same* five-parameter global fit (Sect. 3.3) over *all* integer, full-rank five-tuple ordered chained “ratio bases” built from pairwise quark-mass ratios at each reference scale (720 ordered bases

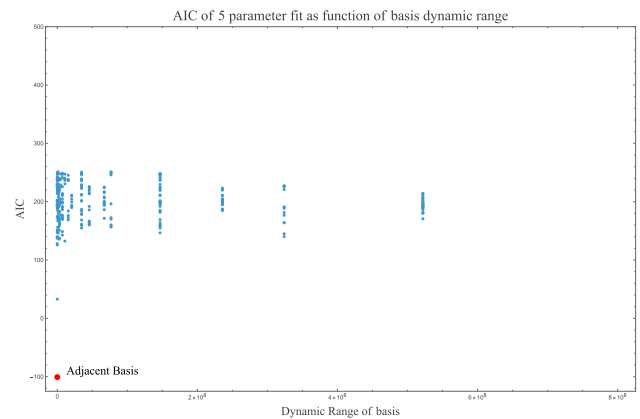


Fig. 2 Information-criterion scan over all five-ratio bases. Each point is one basis; ordinate: AIC from the five-parameter global fit; abscissa: dynamic range of that basis (max/min across its five ratios pooled over the six reference scales). Red dot: adjacent basis, the only one with *negative* AIC. Blue points: other bases that converge (all with large, positive AIC). Non-convergent bases (376 of the 719 non-adjacent) are not plotted and correspond to effectively infinite AIC. BIC shows the same qualitative separation. AICs > 500 are not shown

in total, including the adjacent basis). For each basis we computed the Akaike information criterion (AIC) under a common Gaussian least-squares likelihood with the same N and k across bases.⁷ Figure 2 shows AIC versus a proxy for numerical conditioning (“dynamic range”; see Appendix C).

Specifically, we enumerated all 720 ordered *chained* five-ratio bases (Hamiltonian paths), of the form $m_{i_1}/m_{i_2}, m_{i_2}/m_{i_3}, m_{i_3}/m_{i_4}, m_{i_4}/m_{i_5}, m_{i_5}/m_{i_6}$ over all 6! permutations $(i_1, \dots, i_6)$. These are a distinguished subset of the admissible spanning-tree bases (any five ratios form a basis iff their numerator–denominator graph is a connected tree). We focus on chained trees because—while not well conditioned—they are *less badly pathological*: they compress the dynamic range of the five basis ratios (typically smaller $|\ln r_k|$ than non-chained choices). With only diagonal uncertainties propagated, the larger dynamic ranges of non-chained trees make the normal equations more ill-conditioned, leading to frequent non-convergence and decisively worse information criteria.⁸

⁷ Because N and k are fixed across the basis scan, BIC shifts all AIC values by a constant and yields the same ordering; the adjacent basis is also the only one with negative BIC. We therefore show AIC only. R^2 is not informative for this non-linear, cross-basis comparison because it does not penalize model complexity and is not invariant under basis rescalings.

⁸ There are $6^{6-2} = 1296$ labeled spanning trees on six masses (Cayley). Accounting for orientation of each ratio (numerator/denominator) gives $1296 \times 2^5 = 41,472$ unordered, oriented bases; allowing arbitrary ordering of the five ratios yields $1296 \times 2^5 \times 5! = 4,976,640$ ordered, oriented bases. We restrict attention to the $6! = 720$ chained (Hamiltonian-path) bases because, under diagonal-only uncertainties, they are *less badly pathological* (smaller dynamic-range spreads) than generic trees, which typically suffer markedly worse conditioning.

Table 1 Outcome of the change-of-basis scan for the five-parameter global fit

Category	Count	AIC sign	Notes
Adjacent basis	1	Negative	Unique best (also negative BIC)
Other bases: converged	343	Positive	All decisively worse
Other bases: no fit	376	$+\infty$	Optimizer fails to find a finite likelihood
Total non-adjacent	719		<i>More bases fail than succeed</i>

Table 2 Four-parameter fit to the five quark ratios at each scale. Errors are statistical 1σ values propagated from Huang [5]; current lattice-QCD systematic uncertainties, estimated at $\sim 0.5\%$ for bottom-quark masses—based on $N_f = 2 + 1 + 1$ lattice determinations [6]—would

rescale all errors uniformly and do not affect the central values. The same convention is used in Tables 3 and 4. All quoted masses are $\overline{MS}$ values at the scales indicated. Corresponding p-values range from ~ 0.01 to 0.03 for individual parameters

μ [GeV]	Fit type	κ	ζ_1	ζ_2	γ	χ^2/dof
91.2	With uncertainties	2.367(5)	1.1564(12)	$-0.7671(17)$	1.0949(24)	2.40×10^{-6}
125	With uncertainties	2.368(5)	1.1562(12)	$-0.7691(18)$	1.0975(25)	2.80×10^{-6}
173	With uncertainties	2.364(5)	1.1558(11)	$-0.7699(17)$	1.1011(23)	2.42×10^{-6}
10^5	With uncertainties	2.314(5)	1.1634(13)	$-0.8019(19)$	1.1565(26)	4.54×10^{-6}
10^8	With uncertainties	2.278(6)	1.1697(14)	$-0.8241(21)$	1.1945(29)	7.40×10^{-6}
10^{12}	With uncertainties	2.239(5)	1.1766(14)	$-0.8448(22)$	1.2310(29)	9.04×10^{-6}
91.2	Central values	2.340	1.1609	-0.7719	1.1061	8.31×10^{-4}
125	Central values	2.339	1.1611	-0.7743	1.1096	9.59×10^{-4}
173	Central values	2.338	1.1602	-0.7745	1.1119	7.99×10^{-4}
10^5	Central values	2.287	1.1681	-0.8067	1.1682	1.05×10^{-3}
10^8	Central values	2.249	1.1749	-0.8292	1.2074	1.38×10^{-3}
10^{12}	Central values	2.212	1.1819	-0.8499	1.2443	1.50×10^{-3}

Two empirical facts stand out. (i) The adjacent basis is the *only* basis with negative AIC (same for BIC), and it lies at the extreme left-minimal dynamic range. (ii) Many bases simply *do not fit*: excluding the adjacent basis, only 343 out of the remaining 719 bases yielded any convergent fit at all; 376 failed to converge or returned ill-posed parameters, i.e., effectively infinite AIC/BIC. Thus, more bases fail than succeed. These observations support our use of adjacent ratios in the main analysis: beyond the formal change-of-basis invariance, the adjacent basis is also the information-optimal and numerically stable choice in practice. See Fig. 2 where the abscissa uses the observed ratio dynamic range across scales; Appendix C also defines a purely algebraic proxy $D(A)$ for basis conditioning. The two are highly correlated across the 720 bases and lead to the same separation.

3.1 $(4 \times 6) = 24$ -parameter per-scale fit

Equation (1) was fitted independently at each scale with four free parameters (κ , ζ_1 , ζ_2 , γ). Table 2 lists the best-fit values, one-sigma errors, and the resulting χ^2/dof . Residuals never exceed $\sim 4\%$ and $\chi^2/\text{dof} \leq 9 \times 10^{-6}$. All parameters remain in the natural range $2.2 \lesssim \kappa \lesssim 2.4$, $1.15 \lesssim \zeta_1 \lesssim 1.18$, $-0.85 \lesssim \zeta_2 \lesssim -0.77$, and $1.09 \lesssim \gamma \lesssim 1.23$. With residuals $\lesssim 4\%$ confronting experimental errors (which dominate the

lattice uncertainties) from 0.25% (top/bottom at 10^{12} GeV) to 24% (down/up at M_Z), the resulting $\chi^2/\text{dof} \sim 10^{-5}$ reflects the formula's accuracy in reproducing central values rather than over-fitting.⁹

Crucially, fitting to central values alone (ignoring experimental uncertainties entirely) yields virtually identical parameter values and only modest increases in χ^2/dof to $\sim 10^{-3}$. This demonstrates that the empirical pattern is not an artifact of the uncertainty weighting scheme but reflects genuine structure in the mass ratios themselves. The parameter stability across both fitting approaches, both with and without uncertainties, as well as the quality of the fit as shown in Fig. 3, is remarkable.

3.2 $(3 \times 6) = 18$ -parameter per-scale fit (fixed κ)

Freezing κ at its M_Z value, $\kappa = 2.36707$, reduces the fit to three parameters per scale. Table 3 shows that χ^2/dof *decreases* for every scale up to 10^5 GeV and then increases only slightly up to 10^{12} GeV, a strong indication that κ is a

⁹ Because experimental uncertainties in quark mass evolution are dominated by correlated renormalization scale uncertainties that affect all ratios similarly, the effective degrees of freedom are reduced, leading to $\chi^2/\text{dof} \ll 1$; see Appendix A for details.

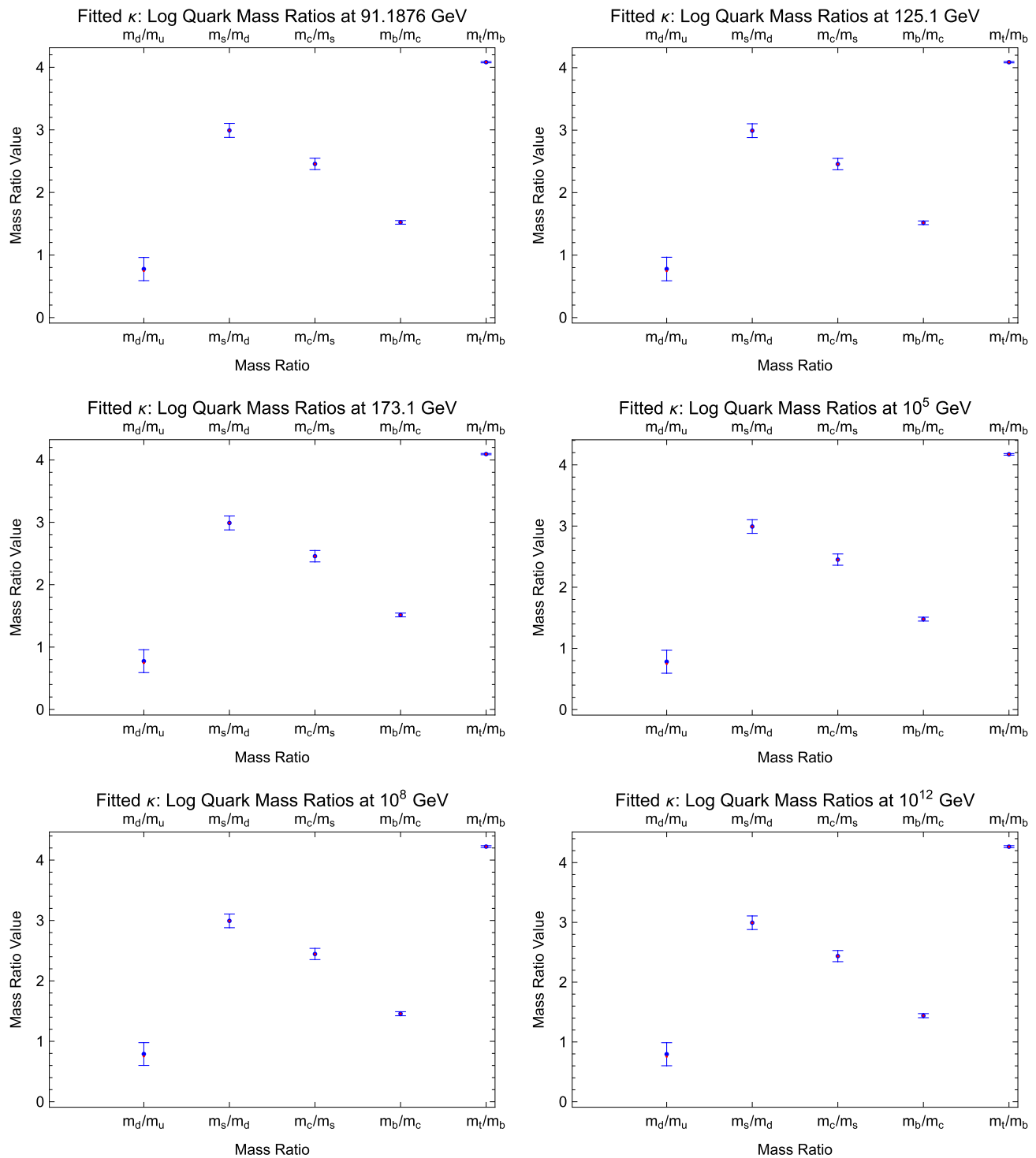


Fig. 3 Comparison of all 24-parameter fits (κ free) across the six energy scales. Even when κ is allowed to vary, the parameter values remain remarkably stable

genuine scale-independent factor rather than a tuning knob. Overall parameter shifts are $\lesssim 5.5\%$ and residuals increase slightly with the maximum residual being $\sim 5\%$. Figure 4 illustrates the high accuracy of the fit.

Again, as Table 3 shows, fitting to central values alone produces nearly identical parameters with χ^2/dof rising only modestly to $\sim 10^{-3}$. The parameter stability across both uncertainty-weighted and central-value-only fits confirms

Table 3 Three-parameter refit with κ fixed to its M_Z value. Corresponding p-values range from $\sim 10^{-6}$ to $\sim 10^{-3}$

μ [GeV]	Fit type	ζ_1	ζ_2	γ	χ^2/dof
91.2	With uncertainties	1.15643(6)	−0.76707(9)	1.09492(3)	1.20×10^{-6}
125	With uncertainties	1.15645(6)	−0.76944(9)	1.09811(3)	1.40×10^{-6}
173	With uncertainties	1.15539(6)	−0.76930(9)	1.09991(3)	1.22×10^{-6}
10^5	With uncertainties	1.15513(9)	−0.78898(14)	1.13158(5)	4.51×10^{-6}
10^8	With uncertainties	1.15542(14)	−0.80187(21)	1.15121(7)	1.25×10^{-5}
10^{12}	With uncertainties	1.15571(19)	−0.81221(28)	1.16709(10)	2.76×10^{-5}
91.2	Central values	1.1565	−0.7663	1.0943	4.45×10^{-4}
125	Central values	1.1565	−0.7686	1.0974	5.10×10^{-4}
173	Central values	1.1554	−0.7687	1.0994	4.33×10^{-4}
10^5	Central values	1.1546	−0.7904	1.1326	8.18×10^{-4}
10^8	Central values	1.1543	−0.8050	1.1537	1.39×10^{-3}
10^{12}	Central values	1.1540	−0.8177	1.1714	2.08×10^{-3}

that the fixed- κ constraint captures genuine physical structure rather than statistical fluctuations.

3.3 Global fits

Treating all 30 ratios simultaneously, we test both linear and quadratic parametrizations of the running exponent:

$$\gamma(\mu) = a + b \ln(\mu/M_Z) \quad [5 \text{ param.}]$$

$$\gamma(\mu) = a + b \ln(\mu/M_Z) + c \ln^2(\mu/M_Z) \quad [6 \text{ param.}].$$

Table 4 presents both fits. The five-parameter form achieves $\chi^2/\text{dof} = 9.29 \times 10^{-6}$, while adding the very mild curvature term reduces this to 3.86×10^{-6} . For the fits with uncertainties, there is a substantial improvement in the information theoretic criteria when moving from the five-parameter fit to the six-parameter fit. Specifically, the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) go from -100.237 and -91.8299 to -126.26 and -116.45 , respectively (more negative is an improvement). However, the central value fits show only insignificant changes in AIC and BIC when moving from a five-parameter fit to a six-parameter fit. This suggests that the information-criterion pay-off is marginal. Thus, parsimony favors the five-parameter form; the six-parameter version merely confirms that a small quadratic bend in $\gamma(\mu)$ is numerically *allowed* but not *required*.

Figure 5 displays both uncertainty-weighted global fits, showing the exceptional agreement between the formula and all 30 data points across six energy scales. Every residual lies well within the quoted experimental uncertainties, with the formula tracking both the central values and their error bars with remarkable precision. The $\chi^2/\text{dof} \sim 10^{-6}$ reflects the

formula's ability to reproduce not just the pattern of mass ratios, but their experimental precision as well.

Most remarkably, both global fits show parameter stability when fitted to central values alone, completely ignoring experimental uncertainties. The quality of the central value fit is apparent from Fig. 6: almost every predicted value is coincident with the central value being fitted. The five-parameter central-value fit yields $\kappa = 2.287$ (vs. 2.316 with uncertainties), $\zeta_1 = 1.169$ (vs. 1.164), $\zeta_2 = -0.802$ (vs. -0.795), and $a = 1.149$ (vs. 1.134)—variations of less than 2%. The six-parameter fit shows similar stability. This consistency across fitting methodologies demonstrates that our empirical law captures genuine structure in the fermion mass spectrum rather than artifacts of the uncertainty weighting scheme. Both central-value fits reproduce every data point with residuals $|\Delta \ln(m_i/m_j)| \leq 0.04$, approximately 4% in the ratios themselves (see Fig. 6), so their $\chi^2/\text{dof} \ll 1$ reflects the formula's intrinsic accuracy, not over-fitting to experimental errors.

3.4 Comprehensive stability analysis

The parameter stability demonstrated in each individual fit above represents significant robustness. Figure 7 displays the residuals of the 18-parameter fit; every point lies within $|\Delta \ln(m_i/m_j)| \leq 0.04$ (4% in the ratio itself). Similarly, Fig. 8 displays the residuals of the two global fits. Table 5 collects the χ^2/dof values for all eight fits (with uncertainties and central values). Table 6 lists the corresponding parameters, demonstrating their stability at the $\lesssim 5\%$ level across all fitting strategies.

The key empirical finding is that the same four $\mathcal{O}(1)$ numbers emerge regardless of fitting methodology: whether we use 24 parameters (4×6), 18 parameters (3×6), or global fits with 5–6 parameters; whether we weight by experimental

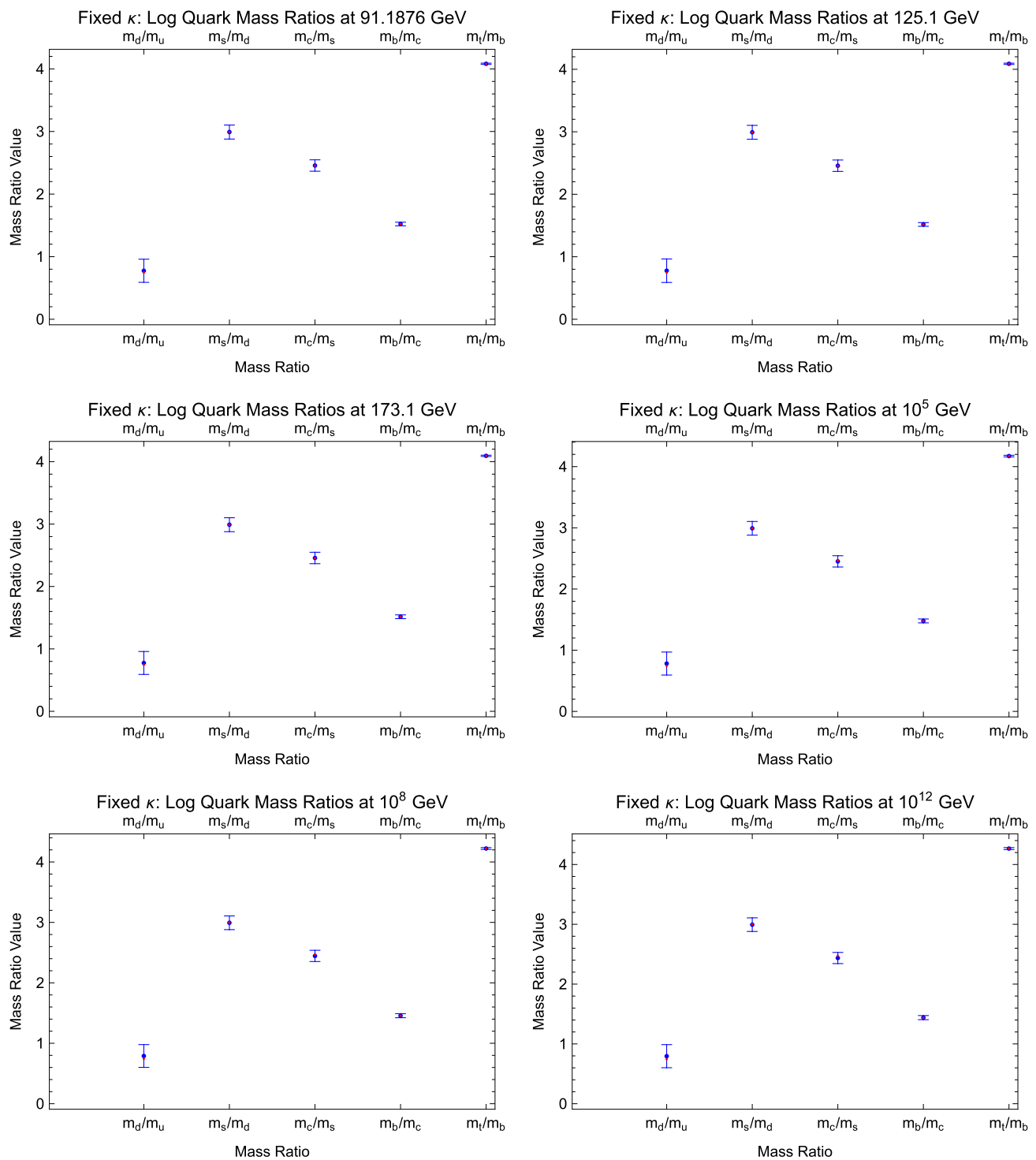


Fig. 4 Comparison of all 18-parameter fits (κ fixed) across the six energy scales. The remarkable parameter stability across scales demonstrates the robustness of the empirical pattern

uncertainties or fit central values only; whether we allow all parameters to float or constrain κ to be scale-independent. This consistency suggests we have identified a fundamen-

tal organizing principle in the fermion mass spectrum, not a statistical artifact.

The same four numbers reappear with consistent $\mathcal{O}(1)$ values in the lepton and neutrino sectors.

Table 4 Five- and six-parameter global fits with $\gamma(\mu) = a + b \ln(\mu/M_Z)$ and $\gamma(\mu) = a + b \ln(\mu/M_Z) + c \ln^2(\mu/M_Z)$, respectively. Corresponding p-values are extremely small, ranging from $\sim 10^{-30}$ to $\sim 10^{-6}$ for individual parameters, indicating highly significant fits

Fit Type	κ	ζ_1	ζ_2	a	$b [10^{-3}]$	$c [10^{-5}]$	χ^2/dof
5-param w/unc.	2.316(8)	1.1638(19)	-0.7949(29)	1.1340(40)	1.587(9)	–	9.29×10^{-6}
5-param central	2.287	1.1686	-0.8019	1.1494	1.440	–	6.26×10^{-4}
6-param w/unc.	2.315(5)	1.1639(13)	-0.7950(19)	1.1201(25)	3.130(26)	-5.20(9)	3.86×10^{-6}
6-param central	2.287	1.1686	-0.8019	1.1353	2.971	-4.948	5.98×10^{-4}

3.5 Specificity test: permutation analysis

To address potential concerns about mathematical flexibility versus physical content, we performed a comprehensive test of Eq. (1)'s specificity. Given that, in our empirical setup, each Standard Model mass eigenstate is uniquely *labeled* by the pair $\{d, q\}$, we ask: does our functional form work *only* when mass ratios are correctly assigned to their physical (d, q) labels, or could it fit any arbitrary assignment equally well?

We systematically tested all $5! = 120$ possible permutations of the five quark mass ratios at each energy scale, keeping the $\{d, q\}$ assignments fixed but reassigning which measured mass ratio corresponds to which quantum number pair. For each permutation, we performed the most unconstrained fit (four free parameters per scale) and computed the root-mean-square error (RMSE) relative to the unpermuted, physical assignment.

Figure 9 displays the results as a function of the normalized Euclidean distance from the physical assignment.¹⁰ The pattern is striking and consistent across all six energy scales: *only* the physical assignment (distance = 0) yields acceptable fits with $\text{RMSE} \lesssim 0.01$. Every alternative permutation produces dramatically worse results, with RMSE values typically very substantially larger. No other assignment among the 119 alternatives approaches the quality of the physical mapping.

This result demonstrates that Eq. (1) is not a flexible curve-fitting expression but rather encodes specific physical structure. The functional form requires the *correct* correspondence between quantum numbers and mass ratios to achieve its remarkable precision, providing strong evidence that it captures genuine organizing principles in the fermion mass spectrum rather than mathematical coincidence.

¹⁰ The normalized Euclidean distance between physical mass ratios m_i and permuted ratios m'_i is defined as $\sqrt{\sum_i \left(\frac{m_i - m'_i}{(m_i + m'_i)/2} \right)^2}$, where each difference is normalized by the average of the corresponding ratios to ensure scale invariance.

3.6 Change-of-basis and information criteria

Equation (1) is written in terms of *adjacent* generation ratios, but any five ratios whose numerator–denominator graph is a spanning tree (i.e., a connected set of five ratios linking all six flavors) form a valid basis. In an ideal analysis with the full 5×5 covariance propagated for each basis, weighted least squares would be invariant under unimodular changes of basis (Appendix B). In practice, published inputs provide diagonal (or effectively diagonal) uncertainties for the five ratios at each scale; with that pragmatic weighting the invariance is not guaranteed. We therefore performed an exhaustive, information-theoretic scan over all independent, chained, five-ratio bases:

- for each basis we repeated the *five-parameter global fit* (Sect. 3.3);
- we recorded the Akaike information criterion, $\text{AIC} = 2k - 2 \ln \hat{\mathcal{L}}$, and the Bayesian information criterion, $\text{BIC} = k \ln N - 2 \ln \hat{\mathcal{L}}$, with $k = 5$ parameters and $N = 30$ data points;
- to characterize conditioning, we computed the *dynamic range* of a basis as the ratio of the largest to the smallest of its five ratios pooled over the six reference scales (larger values indicate a more ill-conditioned basis).

Results Out of the 720 bases (adjacent + 719 alternatives), only the adjacent basis achieves a *negative* AIC (and BIC); every other basis that converges yields large, *positive* AIC/BIC values, i.e., decisively worse fits. Moreover, more bases fail than succeed: among the 719 non-adjacent bases, only 343 admit any finite-likelihood optimum, while 376 do not converge at all under the same fitting setup (effectively $\text{AIC} \rightarrow +\infty$). Figure 2 visualizes AIC versus basis dynamic range. The adjacent basis (red dot) sits alone in the negative-AIC quadrant; all other fits (blue points) lie far away. Because AIC/BIC compare penalized likelihoods, these gaps correspond to $\Delta\text{AIC} = \mathcal{O}(10^2)$ or larger—decisive on the usual Jeffreys/Burnham–Anderson scale—and make separate R^2 discussion unnecessary (Table 1).¹¹

¹¹ AIC and BIC are computed for the same data and parameter count across bases; “more negative” indicates better support. The exact neg-

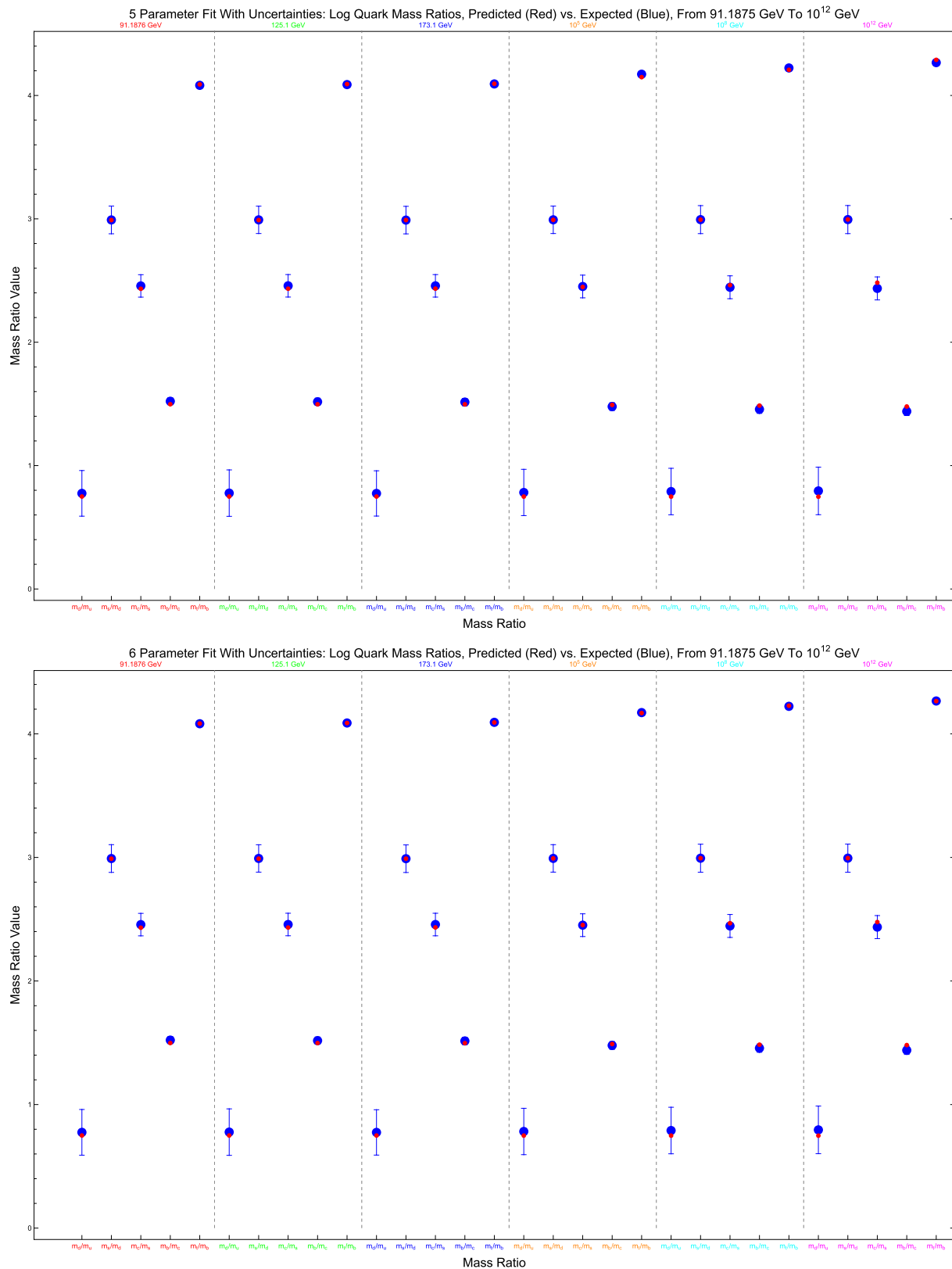


Fig. 5 Global fits with experimental uncertainties included. Top: five-parameter fit with $\chi^2/\text{dof} = 9.29 \times 10^{-6}$. Bottom: six-parameter fit with $\chi^2/\text{dof} = 3.86 \times 10^{-6}$. The empirical value is in blue and the

fitted value is in red. The excellent agreement between theory and data demonstrates the robustness of the empirical pattern



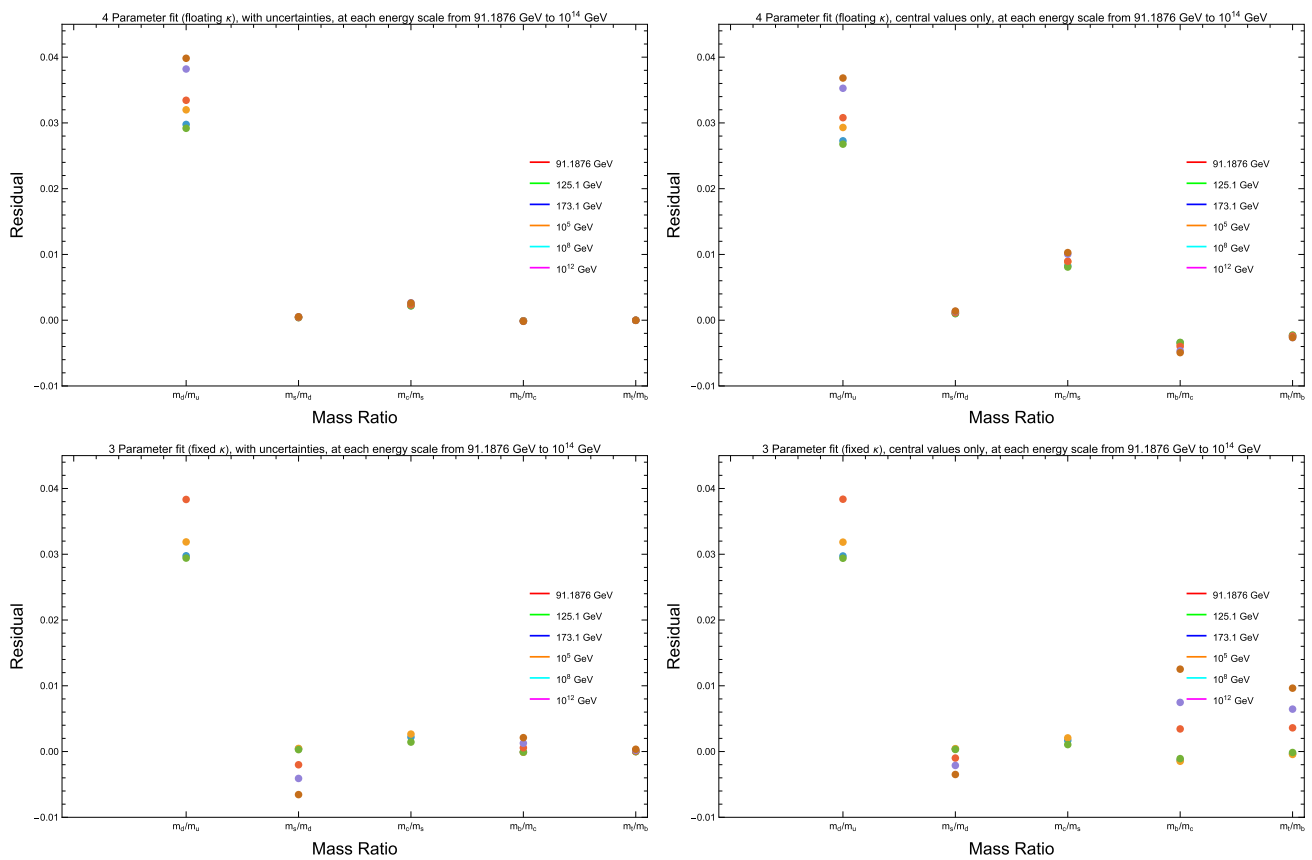


Fig. 7 Residuals of the 18-parameter (3×6) fit

This scan complements the permutation test of Sect. 3.5. The permutation test fixes the *quantum-number assignment* and shows that only the physical mapping of (d, q) to measured ratios is viable. The present scan fixes the mapping and varies the *ratio basis*; it shows that, under realistic weighting, the adjacent basis is uniquely well posed and information-theoretically preferred.

3.7 Alternative functional forms

To check whether simpler or differently factorized expressions can perform as well as Eq. (1), we repeated the *global quark fit with uncertainties* for a representative set of alternatives. In all fits we evaluate $(d|q|)^{\gamma(\mu)}$ with $q \rightarrow |q|$ so that non-integer $\gamma(\mu)$ yields real powers; the signed-power variant (q without $|\cdot|$) fails numerically. Table 7 summarizes the results.

Two robust conclusions emerge. First, the generation prefactor $d_1^{\zeta_1} d_2^{\zeta_2}$ is *required*: removing it (“No- d prefactor”)

active value for the adjacent basis depends slightly on whether one uses the five- or six-parameter global form; both yield the same qualitative conclusion.

degrades the information criteria by $\Delta\text{AIC} \sim \mathcal{O}(10^2)$ relative to the five-parameter baseline in Table 4. Second, electric charge must enter through the *combined* structure $(d|q|)^{\gamma(\mu)}$: using $|q|^{\gamma(\mu)}$ alone, or forcing $\zeta_2 = -\zeta_1$ (“Symmetric ζ ”), similarly worsens AIC/BIC by large margins. Allowing a trivial overall scale ρ inside the κ exponent does not help and effectively reparameterizes the baseline. The only variant that improves the information criteria is a very mild quadratic bend in $\gamma(\mu)$, consistent with the global six-parameter fit already reported in Table 4. Overall, the five-parameter baseline is the most parsimonious form that succeeds. For the basis scan in Sect. 3, we therefore report AIC only. Because k and N are fixed across that scan, BIC differs from AIC by an additive constant and yields the *same* ordering; the adjacent basis is also the only one with negative BIC.

These functional-form checks are orthogonal to the change-of-basis scan in Sect. 3.6: the former holds the basis fixed and varies the ansatz, while the latter fixes the ansatz and varies the basis; both favor the baseline expression in the *adjacent* basis.

These checks, together with the permutation analysis in Sect. 3.5, show that the multiplicative structure linking gener-

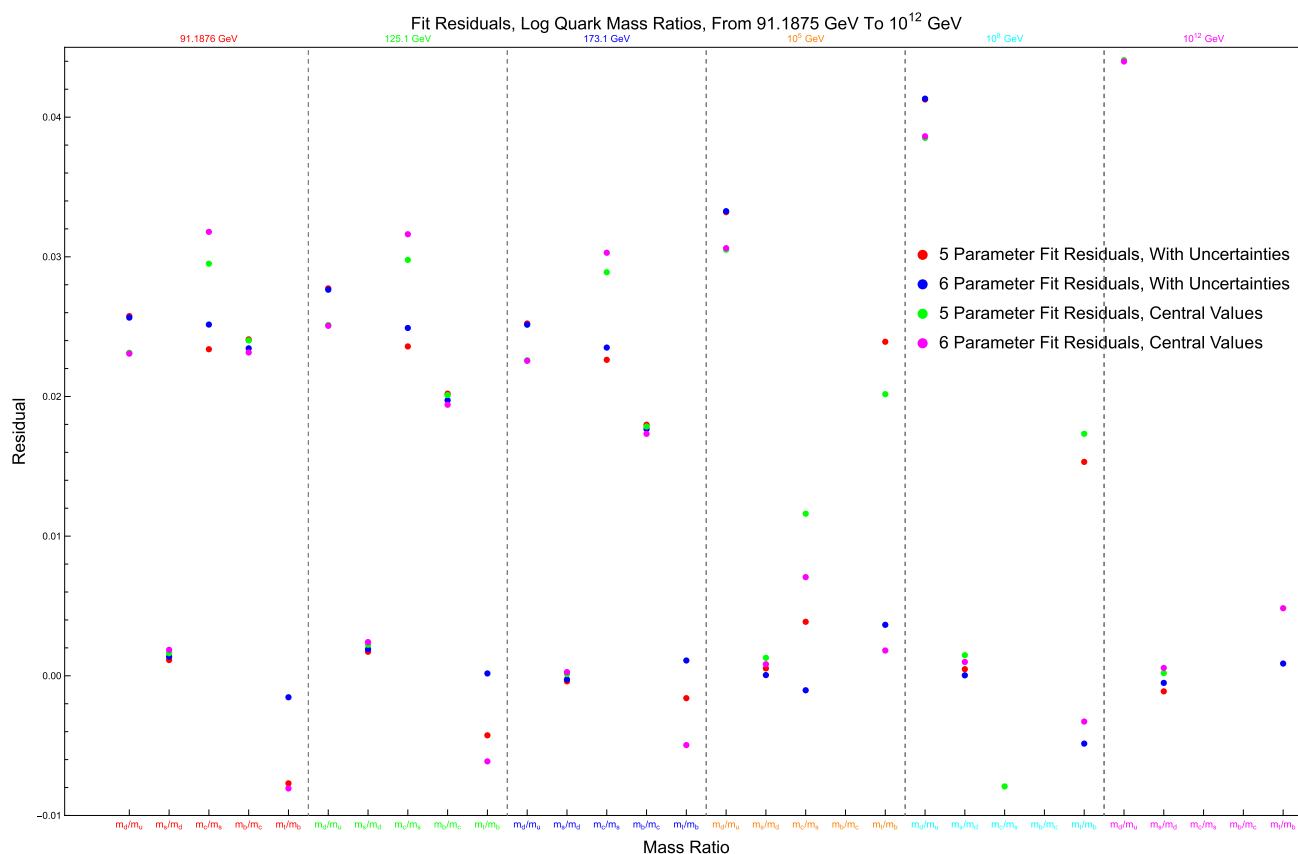


Fig. 8 Residuals from all global fits, showing consistency between uncertainty-weighted and central-value-only approaches

ation and charge is genuinely specific. Simpler or factorized alternatives are either decisively disfavored by the information criteria (typically by $\mathcal{O}(10^2)$) or not mathematically well defined for the data. Apart from a very mild quadratic running in $\gamma(\mu)$, the five-parameter baseline is the most economical successful form.

3.8 Comparison with texture approaches

Standard texture models, such as two-charge Froggatt–Nielsen schemes [3], fit at most one ratio per generation and typically require ten or more parameters. Koide’s [1] relation does not extend to quarks; the Georgi–Jarlskog relation [2] matches a single ratio at a unification scale but fails elsewhere. In contrast, Eq. (1) reproduces *all* 30 quark ratios with at most six parameters and $\chi^2/\text{dof} \ll 10^{-5}$.

The same analytic form is now applied, unchanged, to the charged lepton and neutrino sectors.

4 Charged lepton sector

At each reference scale, we work with the two ascending ratios m_μ/m_e and m_τ/m_μ .¹² Because these ratios vary by less than 0.0001 ± 0.0014 (i.e., consistent with zero) between M_Z and 10^{12} GeV, the 12 data points provide only a modest lever arm against the four parameters in Eq. (1). The fit is therefore a *consistency* test: success is not guaranteed even though the system is formally under-determined.¹³

4.1 Four-parameter fit

Table 8 shows the best-fit values. Residuals are below 2.5×10^{-4} , giving $\chi^2/\text{dof} = 9.6 \times 10^{-11}$. All parameters remain $\mathcal{O}(1)$ and of the same magnitude as in the quark sector. That a single analytic form reproduces both charged leptons and quarks with natural coefficients is, by itself, non-trivial.

¹² For charged leptons, we use $|q| = 1$; the sign of q does not enter Eq. (1).

¹³ Over the same ten-decade range as the quarks, the ratios drift by $(0.0 \pm 0.8) \times 10^{-3}$ for m_μ/m_e and $(0.1 \pm 1.4) \times 10^{-3}$ for m_τ/m_μ , completely consistent with scale-independence.

Table 5 χ^2/ν for the quark sector. The first four fits are performed scale by scale; the five- and six-parameter rows are global fits to all 30 ratios

Fit and weighting	91 GeV	125 GeV	173 GeV	10^5 GeV	10^8 GeV	10^{12} GeV
(4×6) w/unc.	2.40×10^{-6}	2.80×10^{-6}	2.42×10^{-6}	4.54×10^{-6}	7.40×10^{-6}	9.04×10^{-6}
(4×6) <i>central</i>	8.31×10^{-4}	9.59×10^{-4}	7.99×10^{-4}	1.05×10^{-3}	1.38×10^{-3}	1.50×10^{-3}
(3×6) w/unc.	1.20×10^{-6}	1.40×10^{-6}	1.22×10^{-6}	4.51×10^{-6}	1.25×10^{-5}	2.76×10^{-5}
(3×6) <i>central</i>	4.45×10^{-4}	5.10×10^{-4}	4.33×10^{-4}	8.18×10^{-4}	1.39×10^{-3}	2.08×10^{-3}
6-par w/unc.	3.86×10^{-6} (global)					
6-par <i>central</i>	5.98×10^{-4} (global)					
5-par w/unc.	9.29×10^{-6} (global)					
5-par <i>central</i>	6.26×10^{-4} (global)					

Table 6 Parameter values for all eight quark-sector fits. For the per-scale fits, the quoted $\pm$ is the scatter of the six independent determinations; for global fits, it is the fit uncertainty

Fit and weighting	κ	ζ_1	ζ_2	$\gamma(M_Z)$
(4×6) w/unc.	2.30 ± 0.06	1.163 ± 0.009	-0.797 ± 0.033	1.147 ± 0.059
(4×6) <i>central</i>	2.29 ± 0.06	1.168 ± 0.009	-0.802 ± 0.033	1.158 ± 0.059
(3×6) w/unc.	2.367 (fixed)	1.156 ± 0.001	-0.785 ± 0.020	1.124 ± 0.031
(3×6) <i>central</i>	2.367 (fixed)	1.155 ± 0.001	-0.786 ± 0.022	1.125 ± 0.033
6-par w/unc.	2.315 ± 0.015	1.164 ± 0.004	-0.795 ± 0.004	1.120 ± 0.002
6-par <i>central</i>	2.287	1.169	-0.802	1.135
5-par w/unc.	2.316 ± 0.015	1.164 ± 0.004	-0.795 ± 0.004	1.134 ± 0.002
5-par <i>central</i>	2.287	1.170	-0.802	1.149

Running $\gamma(\mu)$

Allowing $\gamma(\mu) = a + b \ln(\mu/M_Z)$ or $a + b \ln(\mu/M_Z) + c \ln^2(\mu/M_Z)$ produces $\chi^2/\text{dof} \sim 10^{-9}$ but returns order-unity or larger *standard errors* for some parameters. The degeneracy simply reflects the near scale-independence of the lepton ratios: current data cannot meaningfully distinguish a constant from a slowly varying exponent. For leptons, we therefore keep the constant- γ fit for the rest of the paper.

4.2 Context

Koide's [1] relation provides a single constraint on the three charged-lepton masses using a different functional form that fails for quarks and neutrinos. Texture models fit the lepton ratios only at the cost of abandoning the quark hierarchy or of introducing sector-specific matrices. Equation (1) achieves simultaneous agreement with the quark fits (Sect. 3) *without* any additional sector-dependent factor and with $\mathcal{O}(1)$ coefficients. This cross-sector coherence strengthens the empirical claim that Eq. (1) captures a genuine organizing principle in the fermion spectrum.

The neutrino sector, where $q = 0$ eliminates the κ term entirely, provides an even more severe test and is treated next.

5 Neutrino sector

In the neutrino sector, we normally discuss mass eigenstates rather than flavor eigenstates. On the other hand, our label, d , represents the generation number. In many flavor constructions, the labels that diagonalize the mass matrix need not coincide with the weak-interaction basis, so one often works with “effective” generation indices. A well-known example is in A_4 models of lepton flavor (e.g., [10]) where the mass eigenstates arise from non-trivial mixtures of the symmetry triplet, and one effectively re-labels them by descending mass rather than gauge quantum numbers. In our ansatz, we similarly take $d_i \in \{1, 2, 3\}$ to index the mass eigenstates in order of magnitude, understanding that these are shorthand for the actual mixed states in the weak basis. Therefore, for mass eigenstates, d represents an effective generation index reflecting the average generation content of each mass eigenstate rather than integer labels.

In normal ordering for neutrinos, $m_1 < m_2 < m_3$. However, our convention for quarks and leptons has been $m_1 > m_2$ in the ansatz. For neutrinos, therefore, we write the master ansatz as:

$$\ln \left(\frac{m_{i+1}}{m_i} \right) = d_{i+1}^{\zeta_1} d_i^{\zeta_2} \kappa^{[(d_{i+1} q_{i+1})^\gamma - (d_i q_i)^\gamma]},$$

where we label the heavier mass as $i + 1$ and the lighter mass as i . However, the ansatz simplifies radically for neu-

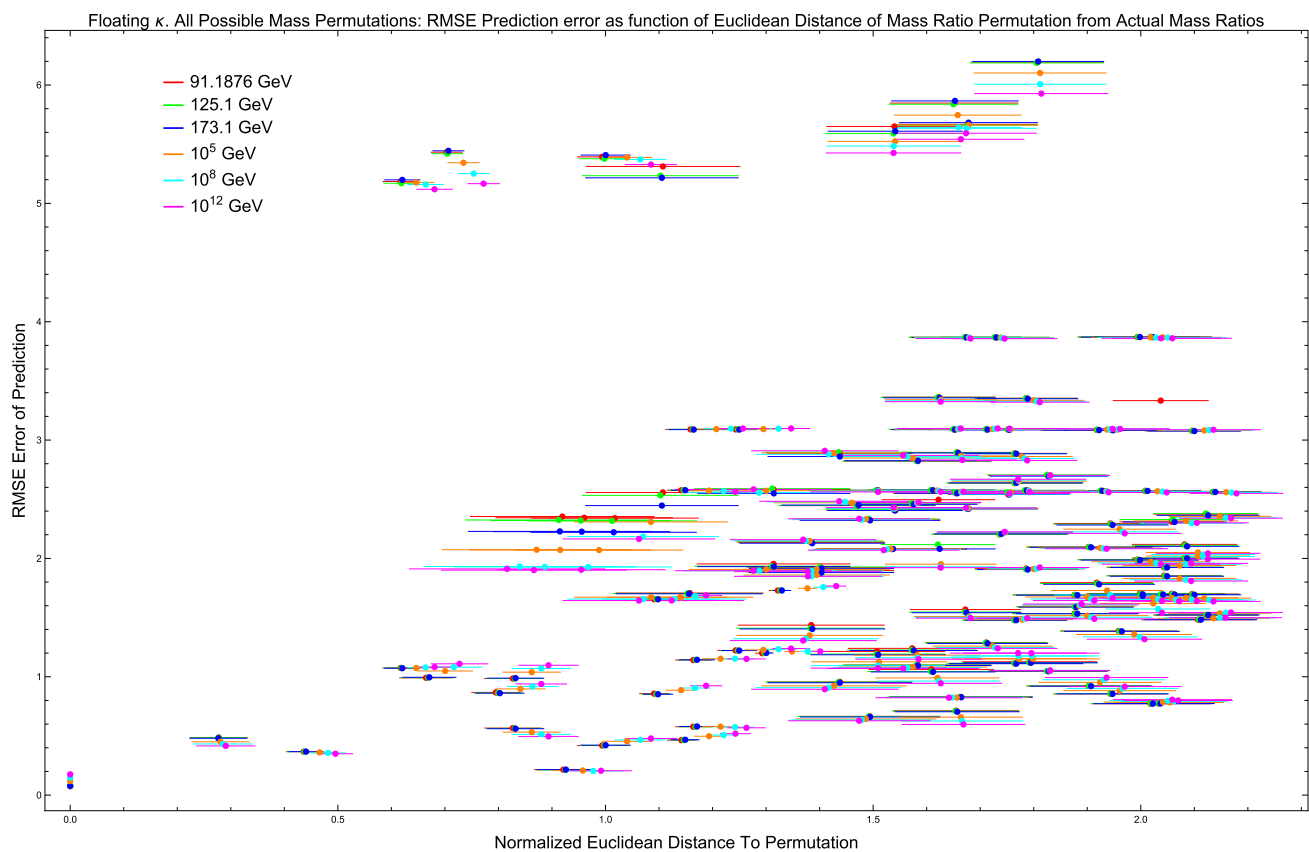


Fig. 9 RMSE as a function of normalized Euclidean distance for all 120 possible permutations of quark mass ratio assignments. Only the physical assignment (distance = 0) yields acceptable fits across

all energy scales. Alternative assignments produce dramatically larger residuals, demonstrating the specificity of the empirical pattern

Table 7 Functional-form checks (global quark fit with uncertainties; same data as Table 4). The five-parameter baseline in Table 4 has AIC ≈ -100.24 and BIC ≈ -91.83 . Smaller (more negative) AIC/BIC indi-

cate better fits. “Ill-posed” means the optimizer fails or returns complex values due to non-integer powers of signed charges

ID	Functional form	Params	AIC	BIC
Baseline	$\ln \frac{m_1}{m_2} = d_1^{\zeta_1} d_2^{\zeta_2} \kappa (d_1 q_1)^{\gamma(\mu)} - (d_2 q_2)^{\gamma(\mu)}$	5	-100.24	-91.83
No- d prefactor	$\ln \frac{m_1}{m_2} = \kappa (d_1 q_1)^{\gamma(\mu)} - (d_2 q_2)^{\gamma(\mu)}$	3	80.64	86.25
$ q $ only	$\ln \frac{m_1}{m_2} = d_1^{\zeta_1} d_2^{\zeta_2} \kappa q_1 ^{\gamma(\mu)} - q_2 ^{\gamma(\mu)}$	5	3.16	11.57
Symmetric ζ	$\ln \frac{m_1}{m_2} = d_1^{\zeta} d_2^{-\zeta} \kappa (d_1 q_1)^{\gamma(\mu)} - (d_2 q_2)^{\gamma(\mu)}$	4	12.73	19.73
Log-additive split	$\ln \frac{m_1}{m_2} = c_0 + \alpha_1 \ln d_1 + \alpha_2 \ln d_2 + \rho (q_1 ^{\gamma(\mu)} - q_2 ^{\gamma(\mu)})$	6	95.61	105.42
Poly. in $(d q)$	$\ln \frac{m_1}{m_2} = d_1^{\zeta_1} d_2^{\zeta_2} \kappa c_1 [(d_1 q_1) - (d_2 q_2)] + c_2 [(d_1 q_1)^2 - (d_2 q_2)^2]$	6	Ill-posed	Ill-posed
No $ \cdot $ on q	Same as baseline with $q \rightarrow q$ (signed power)	5	Ill-posed	Ill-posed
Scale in exponent	$\ln \frac{m_1}{m_2} = d_1^{\zeta_1} d_2^{\zeta_2} \kappa \rho [(d_1 q_1)^{\gamma(\mu)} - (d_2 q_2)^{\gamma(\mu)}]$	6	-95.97	-87.56
$\gamma(\mu)$ quadratic	Baseline with $\gamma(\mu) = a + b \ln(\mu/M_Z) + c \ln^2(\mu/M_Z)$	6	-126.26	-116.45

trinos because $q_i = 0$. As we said earlier, however, we do not know the precise values for d_{i+1} and d_i for the neutrino sector since we are talking about mass eigenstates. For illustration, we take the d values to be 1, 2, 3 as representative

generation indices, though the actual effective values may differ depending on the flavor-mass eigenstate mixing pattern. The key point is not the specific ζ values, which depend on the choice of effective generation indices, but rather that

Table 8 Charged-lepton fit with four free parameters. Quoted errors are 1σ statistical uncertainties propagated from Ref. [5]

Parameter	Estimate	Std. error [10^{-6}]	t -stat. [10^5]	χ^2/dof
κ	2.91835	2.94	9.9	9.6×10^{-11}
γ	0.885797	5.33	1.7	
ζ_1	1.10965	1.76	6.3	
ζ_2	-1.49050	13.3	-1.1	

natural $\mathcal{O}(1)$ values emerge for any reasonable assignment of d_1, d_2, d_3 . Because d_1, d_2, d_3 are each constrained to lie between 1 and 3, other reasonable assignments, e.g., taking d_i weighted by Pontecorvo–Maki–Nakagawa–Sakata (PMNS) mixing angles, would yield qualitatively similar $\mathcal{O}(1)$ natural parameter values, confirming the robustness of the cross-sector consistency (for a full renormalization-group analysis including mixing effects, see [11]).

Furthermore, *because* $q = 0$, the κ term vanishes entirely. Using 1, 2, 3 for the generation index, we see that the ansatz simplifies radically¹⁴:

$$\ln\left(\frac{m_2}{m_1}\right) = 2\zeta_1, \quad \ln\left(\frac{m_3}{m_2}\right) = 3\zeta_1 + 2\zeta_2. \quad (3)$$

There are only two unknown exponents and exactly two measured mass ratios, so *any* choice of a lightest mass m_1 fixes (ζ_1, ζ_2) algebraically. The neutrino sector therefore serves as a self-consistency check rather than a genuine fit.¹⁵ No matter what the actual values of d_1, d_2, d_3 are, the key point is that the ζ values would remain reasonably $\mathcal{O}(1)$ regardless of what non-integer values of d_i we used since those values are constrained to be between 1 and 3. We continue therefore with $d_1, d_2, d_3 = 1, 2, 3$ and proceed as follows.

Log-log form and anchor scan With $q_i = 0$, the κ -term vanishes, so we work entirely in double logs. Denoting $L_{21} \equiv \ln(\ln(m_2/m_1))$ and $L_{32} \equiv \ln(\ln(m_3/m_2))$, Eq. (1) reduces to

$$L_{21} = \zeta_1 \ln 2, \quad L_{32} = \zeta_1 \ln 3 + \zeta_2 \ln 2. \quad (4)$$

¹⁴ Any reasonable assignment of effective generation indices $d_i \in [1, 3]$ for neutrino mass eigenstates yields $\mathcal{O}(1)$ values for $\zeta_{1,2}$, confirming the robustness of cross-sector consistency. The specific choice $d_i = i$ is illustrative; other assignments (e.g., weighted by PMNS mixing angles) produce qualitatively similar natural parameter values. A complete treatment would require solving the full mixing problem, which is beyond our empirical scope.

¹⁵ If the lightest mass were exactly zero, the successive ratio m_2/m_1 (or m_3/m_2 in inverted ordering) would be undefined, and Eq. (3) would simply not apply rather than be falsified; present limits still allow that possibility, so throughout this paper we assume a small but non-zero smallest mass.

Table 9 Neutrino observables at two illustrative anchors that satisfy $\zeta_1 > 0$, $\zeta_2 < 0$ and $|\zeta_i| \sim 1$. Errors are 1σ uncertainties from NuFIT-6.0 [7]

Quantity	$m_1 = 1.0 \text{ meV}$	$m_1 = 1.5 \text{ meV}$
$m_2 \text{ (meV)}$	8.71 ± 0.11	8.78 ± 0.11
$m_3 \text{ (meV)}$	50.35 ± 0.25	50.36 ± 0.25
$\Sigma m_\nu \text{ (eV)}$	0.0601 ± 0.0003	0.0606 ± 0.0003
$m_\beta \text{ (meV)}$	8.95 ± 0.11	9.02 ± 0.11
ζ_1	1.114 ± 0.008	0.822 ± 0.010
ζ_2	-0.955 ± 0.017	-0.498 ± 0.019

Hence,

$$\zeta_1 = \frac{L_{21}}{\ln 2}, \quad \zeta_2 = \frac{L_{32} - \zeta_1 \ln 3}{\ln 2}.$$

Setting a minimum mass for $m_1 = 0.0001 \text{ eV}$ and scanning the physically allowed window $m_1 \leq 0.03 \text{ eV}$ (NuFIT 6.0 splittings and 1σ errors [7]) gives the results summarized in Table 9.

Cross-sector consistency Quarks and charged leptons favor $\zeta_1 \approx +1$ and $\zeta_2 \approx -1$. Demanding the *same sign pattern* and $\mathcal{O}(1)$ size restricts the anchor to¹⁶

$$m_1 \lesssim 1.5 \text{ meV}.$$

Anchors near 1 meV yield $\zeta_1 = 1.11 \pm 0.01$, $\zeta_2 = -0.96 \pm 0.02$, remarkably close to the quark and lepton values.

Numerical predictions We adopt the standard observables $\Sigma m_\nu \equiv m_1 + m_2 + m_3$, $m_\beta^2 \equiv \sum_i |U_{ei}|^2 m_i^2$, and $m_{\beta\beta} \equiv |\sum_i U_{ei}^2 m_i|$, with U_{ei} the PMNS matrix elements. For the preferred meV-scale anchors, we obtain

$$\begin{aligned} \Sigma m_\nu &\simeq 0.061 \text{ eV}, \quad m_\beta \simeq 9 \text{ meV}, \\ m_{\beta\beta}^{\min} &\simeq 0.46 \text{ meV}, \quad m_{\beta\beta}^{\max} \simeq 4.76 \text{ meV}. \end{aligned} \quad (5)$$

These lie just below the current Planck+BAO limit ($\Sigma m_\nu < 0.09 \text{ eV}$) and beneath KATRIN's 0.2 eV reach, but sit

¹⁶ The sign of ζ_1 flips from positive to negative at $m_1 \simeq 3.426 \text{ meV}$ and the sign of ζ_2 flips from negative to positive at $m_1 \simeq 2.17 \text{ meV}$.

Table 10 Predicted high-scale log quark mass ratios with one- σ uncertainties propagated from low-scale inputs via RunDec v3.1 [9]. Numbers in parentheses denote one standard deviation

Ratio	d/u	s/d	c/s	b/c	t/b
5-param. fit ln ratio pred.	0.748 (8)	2.996 (31)	2.493 (27)	1.475 (9)	4.333 (11)
6-param. fit ln ratio pred.	0.748 (5)	2.996 (20)	2.496 (10)	1.474 (6)	4.345 (8)
RunDec v3.1	0.77 (18)	2.99 (11)	2.46 (8)	1.515 (30)	4.094 (11)
5-param. fit, dev. from RunDec	-3(23)%	0.0(4)%	2.0(4)%	-2.6(2.0)%	5.8(0.4)%
6-param. fit, dev. from RunDec	-3(23)%	0.0(4)%	2.0(4)%	-2.7(2.0)%	6.1(0.4)%

squaring in the target ranges of next-generation CMB-S4 surveys and β -decay experiments such as the Project 8 [12] and HOLMES [13].

Outlook If a future measurement pins down any one of Σm_ν , m_β , or $m_{\beta\beta}$ and the effective generation indices are established, the log-log relations in Eq. (4) lock $\zeta_{1,2}$ and make the neutrino sector as predictive—and falsifiable—as the quark and charged-lepton sectors.

6 Predictions and tests

The empirical pattern established above yields specific, falsifiable predictions across all three fermion sectors that can be tested by near-term experimental and theoretical advances:

1. **Quark sector** Using the $a + b \ln \mu$ and $a + b \ln(\mu/M_Z) + c \ln^2(\mu/M_Z)$ fits of Sect. 3, we evolve the five successive ratios from M_Z up to 10^{14} GeV. The relative departure from a straight RunDec v3.1 prediction is shown in Table 10.¹⁷ High-precision lattice QCD could, with continued algorithmic progress, be able to test the (t/b) prediction at the $\sim 5\%$ level, provided bottom-quark mass uncertainties improve from the current $\sim 0.5\%$, as summarized in FLAG's latest review [6], toward sub-percent accuracy. Since m_t is already tightly constrained by collider data, a lattice determination of m_b at this level would enable a direct, sharp test of our model.
2. **Charged leptons** The fixed- κ fit implies that the two observed ratios are essentially scale-independent; any future improvement in the tau mass immediately tightens the allowed (ζ_1, ζ_2) window.

¹⁷ To produce a definitive high-scale prediction, we evolved all six quark masses to $\mu_* = 10^{14}$ GeV with RunDec v3.1 [9], using its five-loop QCD β -function and four-loop heavy-quark decoupling in a six-flavor scheme. We took $\alpha_s^{(5)}(M_Z) = 0.1179 \pm 0.0009$, included two-loop electroweak corrections, and applied standard threshold matching at each heavy-quark mass. At μ_* (in GeV), we find $m_u = 0.00049(8)$, $m_d = 0.00107(8)$, $m_s = 0.0213(17)$, $m_c = 0.248(7)$, $m_b = 1.127(11)$, $m_t = 67.6(4)$.

3. **Neutrinos** Assuming $d_1, d_2, d_3 = 1, 2, 3$, cross-sector naturalness ($\zeta_1 \approx +1$, $\zeta_2 \approx -1$) confines the lightest mass to $m_1 \lesssim 1.5$ meV. In that window, the model predicts $\Sigma m_\nu \simeq 0.06$ eV and $m_\beta \simeq 9$ meV. Stage-IV CMB surveys and sub-40 meV β -decay experiments (Project 8 [12], HOLMES [13]) will probe this range directly.

- No unconditional prediction can be made until both the effective generation indices and an anchor observable are determined. Once a β -decay experiment fixes m_β , Eq. (4) determines (ζ_1, ζ_2) ; the model can then be cross-checked against Σm_ν or $m_{\beta\beta}$.
- *Tritium β decay.* KATRIN reaches only $m_\beta \simeq 0.2$ eV, well above our 9 meV target, but the Project 8 [12] and HOLMES [13] aim for sub-40 meV sensitivity and could directly test the prediction.

7 Theoretical implications and model-building guidance

The empirical pattern uncovered in Sects. 3–5 imposes stringent constraints on flavor model construction. Rather than broadly categorizing models as “favored” or “disfavored,” we outline specific modifications required for representative theoretical frameworks to reproduce our ansatz.

Because Eq. (1) is multiplicative in $X_{ij} \equiv \ln(m_i/m_j)$, taking a logarithm renders it additive in $\{\ln d, (d|q|)^\gamma\}$. In what follows we therefore work at the exponent level, i.e., directly with $X_{ij} = \ln(m_i/m_j)$, which the mechanisms below naturally generate as sums.

7.1 Universal constraints from the empirical pattern

Any viable flavor model must explain:

1. The multiplicative structure for the exponent: $X_{ij} \equiv \ln(m_i/m_j)$ scales as $d_i^{\zeta_1} d_j^{\zeta_2}$ (equivalently, $m_i/m_j \propto \exp[d_i^{\zeta_1} d_j^{\zeta_2}]$ when $q = 0$).
2. The electric-charge dependence enters through $\kappa(d|q|)^\gamma$ with $\gamma \approx 1.1$, where q is the electric charge in units of $|e|$.

3. Cross-sector universality: similar $\mathcal{O}(1)$ κ , ζ_1 , ζ_2 values for quarks and charged leptons
4. The vanishing of the κ term for electrically neutral fermions ($q = 0$)
5. Permutation specificity: among 120 possible quantum number assignments, only the physical one yields acceptable fits.

These features suggest a deep connection between flavor structure and gauge quantum numbers that goes beyond conventional model constructions.

7.2 Illustrative modifications for model classes

We present three minimal worked examples that reproduce the same additive building blocks underlying Eq. (1): (i) a two-spurion Froggatt–Nielsen (FN) map (Sect. 7.2.1); (ii) a partial-compositeness (PC) parametrization (Sect. 7.2.2); and (iii) a warped-overlap construction in Randall–Sundrum (RS) geometry (Sect. 7.2.3).

7.2.1 Minimal example: two-spurion Froggatt–Nielsen

A horizontal $U(1)_F$ with two spurions reproduces the additive (exponent-level) dependence on $\{\ln d, (d|q|)^\gamma\}$ without committing to a UV completion. Let

$$\epsilon_d \equiv \frac{\langle \theta_d \rangle}{\Lambda}, \quad \epsilon_q \equiv \frac{\langle \theta_q \rangle}{\Lambda}, \quad 0 < \epsilon_{d,q} < 1,$$

and collect $\mathcal{O}(1)$ coefficients into a structured prefactor $C_{ij} \equiv d_i^{\zeta_1/2} d_j^{\zeta_2/2}$, reflecting the fitted generation dependence. Assign effective exponents

$$N_i \equiv \alpha_1 \ln d_i + \alpha_0, \quad M_i \equiv \beta (d_i |q_i|)^{\gamma(\mu)},$$

with $d_i \in \{1, 2, 3\}$ and $|q_i| \in \{2/3, 1/3, 1, 0\}$ ($q_i = 0$ for neutrinos). Then diagonal Yukawas take the schematic form

$$Y_{ii} = C_{ii} \epsilon_d^{N_i} \epsilon_q^{M_i}, \quad m_i \propto Y_{ii}.$$

Taking ratios and logarithms gives

$$\begin{aligned} \ln \frac{m_i}{m_j} &= \zeta_1 \ln d_i + \zeta_2 \ln d_j \\ &+ (\ln \epsilon_q^{-1}) \left[(d_j |q_j|)^{\gamma(\mu)} - (d_i |q_i|)^{\gamma(\mu)} \right] \\ &+ (\ln \epsilon_d^{-1}) \Delta N_{ij}. \end{aligned} \quad (6)$$

Choosing the basis so that ΔN_{ij} is absorbed into the generation term (fixing $\alpha_{1,0}$ accordingly) and defining $\kappa \equiv \epsilon_q^{-1}$ yields the canonical form

$$\ln \frac{m_i}{m_j} = \zeta_1 \ln d_i + \zeta_2 \ln d_j + (\ln \kappa) \left[(d_j |q_j|)^{\gamma(\mu)} - (d_i |q_i|)^{\gamma(\mu)} \right], \quad (7)$$

which matches the RS result in Eq. (15) up to notation. The $q=0$ neutrino limit is automatic since the charge-generation term vanishes. Our empirical template Eq. (1) encodes the same combination via a simple non-linearity; we do not claim a UV-complete derivation here.

7.2.2 Minimal example: partial compositeness

In composite Higgs models with partial compositeness, fermion masses arise from linear mixings $y_{L_i} \bar{q}_{L_i} \mathcal{O}_{R_i}$ and $y_{R_i} \bar{\psi}_{L_i} \psi_{R_i}$ with a strong sector, giving $m_i \sim Y_* y_{L_i} y_{R_i} v$ where Y_* is an $\mathcal{O}(1)$ strong-sector Yukawa. Parametrize the mixings as

$$\begin{aligned} y_{L_i} &\propto d_i^{\zeta_1/2} \exp \left[-\frac{\rho}{2} (d_i |q_i|)^{\gamma(\mu)} \right], \\ y_{R_i} &\propto d_i^{\zeta_2/2} \exp \left[-\frac{\rho}{2} (d_i |q_i|)^{\gamma(\mu)} \right], \end{aligned} \quad (8)$$

encoding mild generation dependence and a small charge-generation deformation. Then

$$\ln \frac{m_i}{m_j} = \zeta_1 \ln d_i + \zeta_2 \ln d_j + \rho \left[(d_j |q_j|)^{\gamma(\mu)} - (d_i |q_i|)^{\gamma(\mu)} \right], \quad (9)$$

matching the same additive structure as Eqs. (15) and (7); for $q=0$ the charge term vanishes. This is a minimal parametrization (no UV claim), included to illustrate that the required $d-|q|$ dependence can also arise in PC-type settings.

7.2.3 Minimal example: warped extra dimensions

In Randall–Sundrum models [14–16], fermion masses arise from wavefunction overlaps. Standard implementations use constant bulk masses. Our pattern suggests a more complex structure. We show that a standard Randall–Sundrum (RS) setup produces the two key ingredients of Eq. (1)—a multiplicative generation factor and a charge-weighted $(d|q|)^\gamma$ term—under mild, explicit assumptions.

Setup Consider RS₁ with metric $ds^2 = e^{-2ky} \eta_{\mu\nu} dx^\mu dx^\nu + dy^2$, $y \in [0, \pi R]$, and bulk Dirac masses $m_{5,L_i} = c_{L_i} k$, $m_{5,R_j} = c_{R_j} k$ for the i th left- and j th right-handed zero modes. The canonically normalized zero-mode wavefunctions scale as $f_{L_i}(y) \propto e^{(1/2-c_{L_i})ky}$, $f_{R_j}(y) \propto e^{(1/2-c_{R_j})ky}$. The Yukawa entry is the overlap

$$Y_{ij} \propto \int_0^{\pi R} dy e^{-4ky} f_{L_i}(y) f_{R_j}(y) \simeq \mathcal{N}_{L_i} \mathcal{N}_{R_j} e^{[1-(c_{L_i}+c_{R_j})]k\pi R}, \quad (10)$$

where $\mathcal{N}_{L_i}$ and $\mathcal{N}_{R_j}$ are the normalization factors (explicit forms not needed below). Hence the diagonal masses scale

as

$$m_i \propto \mathcal{N}_{L_i} \mathcal{N}_{R_i} e^{[1-(c_{L_i}+c_{R_i})]k\pi R}. \quad (11)$$

Assumption 1 (generation prefactor) Empirically the fit prefers multiplicative factors $d_i^{\zeta_1}$ and $d_j^{\zeta_2}$ in Eq. (1). In the RS overlap, such factors arise from the smooth c -dependence of the zero-mode normalizations when the bulk masses vary mildly with a generation label d (mass ordering); cf. Eq. (11). We parametrize this by

$$\mathcal{N}_{L_i} \mathcal{N}_{R_j} \propto d_i^{\zeta_1/2} d_j^{\zeta_2/2}, \quad (12)$$

with signs $\zeta_1 > 0$, $\zeta_2 < 0$ matching the quark and charged-lepton fits (Sects. 3, 4).

Assumption 2 (charge-generation coupling) To capture the observed $(d|q|)^\gamma$ dependence (and the $q \rightarrow 0$ neutrino limit), take the *sum* of bulk masses to acquire a small, universal deformation that couples electric charge to generation:

$$c_{L_i} + c_{R_i} = c_0 + \lambda(d_i |q_i|)^\gamma, \quad (13)$$

with λ dimensionless and γ the same exponent that appears empirically. This preserves standard RS localization while linking flavor and gauge structure in the minimal way suggested by the data.

Result Using (11), the ratio of diagonal masses is

$$\begin{aligned} \ln \frac{m_i}{m_j} &= \frac{\zeta_1}{2} \ln d_i + \frac{\zeta_2}{2} \ln d_j + k\pi R [(c_{L_j} + c_{R_j}) - (c_{L_i} + c_{R_i})] \\ &= \underbrace{\zeta_1 \ln d_i + \zeta_2 \ln d_j}_{\text{generation factor}} + \underbrace{(2\lambda k\pi R) [(d_j |q_j|)^\gamma - (d_i |q_i|)^\gamma]}_{\text{charge-generation term}}. \end{aligned} \quad (14)$$

after absorbing the factor $1/2$ into $\zeta_{1,2}$ (pure convention).

Defining $\rho \equiv 2\lambda k\pi R$ and recalling $|q| = 0$ for neutrinos, Eq. (14) implies

$$\ln \frac{m_i}{m_j} = \zeta_1 \ln d_i + \zeta_2 \ln d_j + \rho [(d_j |q_j|)^\gamma - (d_i |q_i|)^\gamma], \quad (15)$$

i.e., the same *additive* dependence on $\{\ln d, (d|q|)^\gamma\}$ that underlies our empirical pattern. Our template Eq. (1) uses a non-linear (exponentiated) encoding of this linear combination at the level of $\ln(m_i/m_j)$; empirically this captures the data most parsimoniously (Sect. 3.7). The $q = 0$ neutrino limit is immediate since the charge-generation term vanishes. We therefore view Eq. (15) as a minimal RS mechanism that yields the required d – $|q|$ dependence, without claiming identity to Eq. (1).

Comments (i) Nothing in Eq. (15) is tuned: a mild deformation of the bulk masses yields $\rho = \mathcal{O}(1)$, while $k\pi R$

takes its usual RS value. (ii) Allowing a small left–right split $c_{L_i} \neq c_{R_i}$ only shifts $\zeta_{1,2}$ without affecting the $(d|q|)^\gamma$ structure. (iii) Because $q \rightarrow 0$ kills the second term, the $q = 0$ constraint from the neutrino sector is automatic. A more elaborate UV origin for the deformation can be provided (e.g., brane kinetic terms or gauged horizontal symmetries), but is not required for this minimal example. We do not claim a UV-complete derivation here; the FN and RS constructions are explicit, economical frameworks that reproduce the observed multiplicative d – $|q|$ structure with $\mathcal{O}(1)$ parameters. Building a fully UV-complete realization is beyond the scope of this empirical paper.

We note that both FN-type $U(1)_F$ settings and RS geometries admit UV-complete embeddings in explicit compactifications; our goal here is only to show that the required multiplicative $d - |q|$ structure arises from minimal, standard ingredients.

7.2.4 Discrete flavor symmetries

Models based on A_4 [10,17], S_4 [18], etc., typically yield fixed Clebsch–Gordan coefficients. Reproducing our pattern requires:

- Multiple flavon fields with hierarchical VEVs: $\langle \phi_1 \rangle / \Lambda_1 \ll \langle \phi_2 \rangle / \Lambda_2 \ll 1$.
- Effective operators at different orders: $(L_i \phi_1^{n_1} \phi_2^{n_2} E_j H) / \Lambda^{n_1+n_2}$.
- The powers n_1, n_2 must be generation- and charge-dependent to yield the observed structure.
- Note: The generation-dependent power-law $d_i^{\zeta_1} d_j^{\zeta_2}$ requires going beyond minimal discrete implementations, possibly through sequential symmetry breaking or additional continuous symmetries.

These modifications enable discrete symmetry models to approximate the ansatz’s generation- and charge-dependent structure, though achieving the exact form may require additional theoretical mechanisms.

7.3 The neutrino constraint: implications of $q = 0$

A particularly stringent constraint comes from the neutrino sector where $q = 0$ causes the κ term to vanish entirely. This implies:

- Any UV completion must explain why electric charge enters through $(d|q|)^\gamma$ rather than q^γ alone.
- Models attempting to unify quark and lepton sectors must naturally produce this vanishing.
- The residual pattern $\frac{m_{i+1}}{m_i} = \exp \left[d_{i+1}^{\zeta_1} d_i^{\zeta_2} \right]$ for neutrinos provides an independent test.

- See Eq. (3) for the explicit $q = 0$ limit.

This feature severely constrains grand unified theories where quarks and leptons sit in common multiplets—the mechanism generating the $\kappa^{(d|q)\gamma}$ term must respect the $q \rightarrow 0$ limit.

7.4 Common themes and future directions

All successful modifications share key features:

1. Flavor and gauge structures must be intrinsically linked, not independent, as required by the $(d|q)^\gamma$ term and the neutrino sector's $q = 0$ constraint.
2. The generation index must enter through approximate continuous symmetries or geometric factors.
3. The electric charge dependence suggests electroweak symmetry breaking may play a role in flavor.

These modifications are illustrative rather than exhaustive. The empirical pattern provides a clear target [10, 19] for model builders: any complete theory of flavor should explain not just the existence of hierarchies, but their specific functional form across all fermion sectors.

8 Discussion and outlook

8.1 Hierarchy of empirical support

1. **Quarks: over-constrained success** Thirty independent ratios across six scales confront four parameters per scale (or five/six in the global fit). The resulting $\chi^2/\text{dof} \ll 1$ and the modest, $\lesssim 6\%$, deviations from RunDec v3.1 evolution (Sect. 6) make the quark sector the most compelling evidence for Eq. (1).
2. **Charged leptons: under-constrained but natural** Effectively two ratios versus four parameters render the fit tautological, yet the extracted $(\kappa, \zeta_1, \zeta_2, \gamma)$ are all $\mathcal{O}(1)$ and the ratios stay flat across six decades in energy. Any future detection of significant running not captured by $\gamma(\mu)$ would immediately falsify the ansatz.
3. **Neutrinos: consistency only (for now)** With $q = 0$, the formula collapses to Eq. (3); two exponents are fixed by the two measured mass splittings. The test is merely that *some* (ζ_1, ζ_2) of natural size exist for the allowed anchor window. Once an external observable pins m_1 and the effective generation indices are established, the sector will become predictive and directly comparable to quarks and leptons.

8.2 Relation to other empirical mass relations

- *Koide-type relations.* Koide's [1] celebrated lepton formula provides a striking mass relation that remains very stable under renormalization-group evolution, but cannot be extended to quarks where RG running is significant, nor to neutrinos without sector-specific modifications. More fundamentally, such relations work beautifully in the charged-lepton sector but cannot be applied across all three fermion sectors with a single functional form.
- *Texture and GUT relations* (e.g., Georgi–Jarlskog [2], Froggatt–Nielsen [3]). These link selected Yukawa eigenvalues at the unification scale, typically $m_d = m_e$ or $m_b = m_\tau$, but fail when run to low energy or across all generations simultaneously. By contrast, Eq. (1) fits *successive* ratios at every scale and works unchanged in all three sectors.
- *Democratic/seesaw models.* Approaches based on random $\mathcal{O}(1)$ Yukawa entries reproduce hierarchies statistically, not algebraically, and make no precise ratio predictions. The purely multiplicative structure of Eq. (1) suggests a broken horizontal $U(1)$ or discrete flavor symmetry acting on the ordered pair (d, q) [8]. Whether realized through Froggatt–Nielsen charges or wavefunction localization in warped space (see Sect. 7), any viable theory should now reproduce *both* the running exponent $\gamma(\mu)$ and the $\mathcal{O}(1)$ coefficients inferred here.

8.3 Near-term tests

- *High-precision lattice QCD* can confront the (t/b) prediction at the $\sim 5\%$ level, provided bottom-quark mass uncertainties improve from the current $\sim 0.5\%$ —as reported in recent $N_f = 2 + 1 + 1$ heavy-quark simulations [6, 20]—toward sub-percent accuracy.
- *Improved m_τ measurements* at Belle II will shrink the allowed charged-lepton parameter space and test the scale-independence implied by a constant γ .
- *Cosmology or $0\nu\beta\beta$* that pins any neutrino anchor observable will fix (ζ_1, ζ_2) ; the ansatz then predicts the remaining observables and becomes falsifiable in the neutrino sector.

9 Conclusions

The pattern revealed in this work is empirically robust. A single functional form, constructed from the only two labels that distinguish Standard Model fermions, fits all measured fermion mass ratios with stable parameters.

For the quark sector, whether fitted scale-by-scale (24 or 18 parameters), globally (five or six parameters), with or

without experimental uncertainties, the same four $\mathcal{O}(1)$ numbers emerge: $\kappa \sim 2.3$, $\zeta_1 \sim 1.16$, $\zeta_2 \sim -0.8$, $\gamma \sim 1.1$.

This stability is remarkable: parameter variations for quarks remain below 5% across eight distinct fitting strategies spanning ten orders of magnitude in energy scale. The pattern extends beyond quarks—charged leptons and neutrinos respect the identical functional form with equally natural parameter values. Thus, the quark sector provides strong evidence, the charged-lepton sector supplies a non-trivial consistency check, and the neutrino sector offers a clear path to future falsification. This cross-sector universality, combined with the minimal input requirements (generation index and electric charge only), suggests a fundamental organizing principle in flavor physics.

The specificity of this empirical pattern is further demonstrated by comprehensive permutation analysis: among all 120 possible assignments of mass ratios to quantum number pairs, only the physical assignment yields acceptable fits (see Fig. 9). Alternative mappings produce residuals from 3 to 100 times larger, indicating that Eq. (1) encodes specific physical structure rather than mathematical flexibility. This combination of parameter stability and assignment specificity provides compelling evidence for genuine organizing principles in the fermion mass spectrum. A complementary change-of-basis scan over all 720 chained (Hamiltonian-path) independent five-ratio bases (Sect. 3.6) shows that the adjacent basis is the only one with negative AIC/BIC, and that more alternative bases fail to fit than succeed (376 vs. 343), underscoring both specificity and well-posedness.

Previous empirical relations have generally not achieved both a comprehensive scope and parameter stability. Koide-type relations work beautifully within single sectors but cannot be extended. Texture and GUT models require sector-specific modifications and many more parameters. Our four-parameter expression simultaneously describes all three fermion sectors with residuals below 4%.

The theoretical implications outlined in Sect. 7 demonstrate that our empirical pattern places non-trivial constraints on model building. The required modifications—linking flavor structure to electric charge—point toward a more unified understanding of these supposedly independent parameters. Whether this reflects a fundamental symmetry principle or emerges from dynamics remains an open question for future investigation. For completeness, Sect. 7.2 provides three minimal worked examples that reproduce the multiplicative $d - |q|$ structure.

Our analysis provides empirical targets rather than theoretical explanations. The remarkable fit quality and cross-sector universality suggest these patterns reflect genuine organizing principles, but identifying the underlying physics remains an open challenge.

A complete theoretical explanation of this empirical pattern would need to address several key questions: Why

do generation index and electric charge enter multiplicatively? What dynamics produce the specific values $\kappa \sim 2.3$, $\zeta_1 \sim 1.16$, $\zeta_2 \sim -0.8$? Why does the exponent γ run logarithmically for quarks but remain constant for leptons? And most fundamentally, what is the origin of the $(d|q|)^\gamma$ structure that respects the $q \rightarrow 0$ limit for neutrinos? The theoretical modifications outlined in Sect. 7 provide necessary conditions, but a fully satisfactory theory must derive these features from first principles.

We therefore propose this empirical relation as an important constraint that realistic flavor models may wish to consider. Any viable theory of fermion masses should not only explain individual sector patterns, but also account for why the same four natural numbers appear to govern all fermion mass ratios across the entire Standard Model spectrum.

Whether this pattern is a clue to flavor dynamics or a numerical accident will be decided by the next generation of precision mass measurements.

Data Availability Statement This manuscript has no associated data. [Author's comment: Data sharing not applicable to this article as no new datasets were generated or analyzed; all inputs are taken from published sources cited in the manuscript. Derived numbers that underlie the figures are reproducible from those sources with the code described below.]

Code Availability Statement Code/software will be made available on reasonable request. [Author's comment: The code/software generated during and/or analysed during the current study is available from the corresponding author on reasonable request.]

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. Funded by SCOAP³.

A Expected χ^2 for correlated uncertainties

When experimental errors are highly correlated, the expected χ^2/dof for a correct model can be significantly less than 1. This arises because correlated uncertainties reduce the effective number of independent measurements, leading to a suppressed χ^2 value. Mass running studies often exhibit such correlations since systematic uncertainties in renormalization scale choices affect all energy scales similarly.

For N observables with a single fully correlated scale error σ , the covariance matrix is $C_{ij} = \sigma^2$. Its Moore–Penrose

inverse is $C_{ij}^+ = 1/(N\sigma^2)$, so for residuals δ_i we have

$$\chi^2 = \delta^T C^+ \delta = \frac{(\sum_i \delta_i)^2}{N\sigma^2} \sim \chi_1^2,$$

whose expectation value is $\langle \chi^2 \rangle = 1$, i.e., $\langle \chi^2/\text{dof} \rangle = 1/(N-k) \ll 1$ for $N \gtrsim 10$ and $k \leq 5$. This general effect explains why excellent fits to correlated data can yield $\chi^2/\text{dof} \ll 1$ without indicating overfitting.

B Adjacent-ratio basis and change-of-basis invariance

Let $(m_1, \dots, m_6) \equiv (m_u, m_d, m_s, m_c, m_b, m_t)$ at a fixed scale and define the five adjacent log-ratios $r_k \equiv \ln(m_{k+1}/m_k)$, $k = 1, \dots, 5$. For any $i > j$,

$$\ln\left(\frac{m_i}{m_j}\right) = \sum_{k=j}^{i-1} r_k,$$

so the mapping from adjacent ratios to arbitrary pairwise ratios is linear with integer coefficients. Any other independent choice of five ratios corresponds to $\mathbf{r}' = A\mathbf{r}$ with $A \in \text{GL}(5, \mathbb{Z})$. If the full covariance C of $\mathbf{r}$ is propagated to $\mathbf{r}'$ as $C' = ACA^T$, then the weighted least-squares objective

$$\chi^2(\theta) = [\mathbf{r} - \mathbf{f}(\theta)]^T C^{-1} [\mathbf{r} - \mathbf{f}(\theta)]$$

is unchanged under the simultaneous transformation $(\mathbf{r}, \mathbf{f}, C) \mapsto (\mathbf{r}', A\mathbf{f}, C')$, and the normal equations for θ are identical up to floating-point rounding. Hence, using adjacent-generation ratios is a basis choice that does not bias parameter estimation.

Practical note The invariance proven above requires carrying through the full 5×5 covariance for each basis. Public inputs for quark running provide diagonal (or effectively diagonal) errors for the five ratios at each scale; with that pragmatic weighting, the log-likelihood is *not* strictly invariant under $A \in \text{GL}(5, \mathbb{Z})$, and numerical conditioning becomes basis dependent. This is precisely what the information-criterion scan in Sect. 3.6 quantifies: the adjacent basis is uniquely well posed (negative AIC/BIC), many alternative bases converge but are decisively worse (large positive AIC/BIC), and a majority fail to admit a finite-likelihood optimum under identical settings.

C Basis enumeration, dynamic range and AIC/BIC

What we call a “basis” At a fixed scale, let

$$\mathbf{r} = (\ln(m_d/m_u), \ln(m_s/m_d), \ln(m_c/m_s), \ln(m_b/m_c),$$

$$\ln(m_t/m_b))^T$$

denote the adjacent-ratio log vector. Any other independent five-tuple of pairwise ratio logs can be written as $\mathbf{r}' = A\mathbf{r}$ with an integer unimodular matrix $A \in \text{GL}(5, \mathbb{Z})$ whose rows each represent a difference of two log masses (a single pairwise ratio).¹⁸ We enumerate all 720 ordered chained (Hamiltonian-path) integer bases of this type used in Fig. 2; the adjacent basis is the identity $A = \mathbb{I}$.

Conditioning and dynamic range. Although every spanning tree yields a valid $\mathbb{Z}$ -basis, their numerical behaviour differs when only diagonal uncertainties are used. Chained (adjacent) trees keep the five basis ratios closer to unity in magnitude (smaller $|\ln r_k|$), whereas non-chained trees—especially those linking the lightest to the heaviest masses—produce basis ratios with widely varying $|\ln r_k|$. In that setting the effective weighting becomes strongly anisotropic and the normal equations are more ill-conditioned. Empirically this manifests as higher failure rates and worse AIC/BIC for non-chained trees. Thus chained bases, though not well conditioned, are *less badly pathological* and comparatively robust in our scans.

Remark (when do five ratios form a basis?) Let $r_k = \ln(m_{a_k}/m_{b_k})$ and let G be the undirected graph on $\{1, \dots, 6\}$ with edges $\{a_k, b_k\}$. With five edges on six vertices, the rows $\{e_{a_k} - e_{b_k}\}$ span the five-dimensional log-ratio space iff G is connected, i.e., a spanning tree. In that case, for any u, v the unique path $P(u \rightarrow v)$ in G yields

$$\ln \frac{m_u}{m_v} = \sum_{(a,b) \in P(u \rightarrow v)} \sigma_{ab} \ln \frac{m_a}{m_b}, \quad \sigma_{ab} \in \{\pm 1\},$$

so all 30 pairwise ratios are reconstructible. Our scan considers the $6! = 720$ Hamiltonian-path (“chained”) spanning-tree bases; other spanning trees are related by unimodular transformations in $\text{GL}(5, \mathbb{Z})$.

Dynamic range (conditioning proxy) For each basis we compute a proxy for numerical conditioning,

$$\mathcal{D}(A) \equiv \frac{\max_{i,j} |A_{ij}|}{\min_{i,j: A_{ij} \neq 0} |A_{ij}|}. \quad (16)$$

This scales like a condition number for the linear map $\mathbf{r} \mapsto \mathbf{r}'$ and correlates with estimator instability. Any monotone proxy for conditioning gives the same qualitative picture; the adjacent basis minimizes $\mathcal{D}$ by construction.¹⁹

¹⁸ Appendix B gives the change-of-basis invariance when the full covariance is propagated.

¹⁹ Equivalently, one may multiply by $\text{cond}_2(A)$ to penalize near-singular maps; both choices pick out the adjacent basis in practice.

Likelihood and criteria For each basis we refit the *same* five-parameter global model to the 30 data points expressed in that basis and evaluate the Gaussian least-squares likelihood

$$\mathcal{L}(\theta) = \frac{1}{(2\pi)^{N/2} |C|^{1/2}} \exp \left[-\frac{1}{2} \Delta(\theta)^\top C^{-1} \Delta(\theta) \right], \quad (17)$$

$$\Delta(\theta) \equiv r - f(\theta),$$

where r collects the five ratio logs at each of the six reference scales (see Appendix B), $f(\theta)$ are the model predictions, and C is the (block-)diagonal covariance built from the quoted per-ratio uncertainties at those scales. The information criteria are then

$$\text{AIC} = 2k - 2 \ln \hat{\mathcal{L}}, \quad \text{BIC} = k \ln N - 2 \ln \hat{\mathcal{L}}, \quad (18)$$

with $k = 5$ and $N = 30$ fixed across the scan and $\hat{\mathcal{L}} = \mathcal{L}(\hat{\theta})$ the maximized Gaussian likelihood evaluated at the best-fit parameters for that basis. Because k and N are constant, BIC differs from AIC by an additive constant and yields the same ranking; hence only AIC is shown in Fig. 2. Bases for which the optimizer fails to converge or return well-posed parameters correspond to effectively infinite AIC/BIC; we report these as “no fit” in the main text (376 of 719 non-adjacent bases).

References

1. Y. Koide, Phys. Lett. B **120**(1–3), 161 (1983). [https://doi.org/10.1016/0370-2693\(83\)90644-5](https://doi.org/10.1016/0370-2693(83)90644-5)
2. H. Georgi, C. Jarlskog, Phys. Lett. B **86**, 297 (1979). [https://doi.org/10.1016/0370-2693\(79\)90842-6](https://doi.org/10.1016/0370-2693(79)90842-6)
3. C. Froggatt, H. Nielsen, Nucl. Phys. B **147**, 277 (1979). [https://doi.org/10.1016/0550-3213\(79\)90316-X](https://doi.org/10.1016/0550-3213(79)90316-X)
4. M. Sabir, A. Mansha, T. Li, Z.W. Wang, J. High Energy Phys. **2025**(1) (2025). [https://doi.org/10.1007/jhep01\(2025\)201](https://doi.org/10.1007/jhep01(2025)201)
5. G.Y. Huang, S. Zhou, Phys. Rev. D **103**(1), 016010 (2021). <https://doi.org/10.1103/PhysRevD.103.016010>
6. Y. Aoki, T. Blum, S. Collins, L.D. Debbio, M.D. Morte, P. Dimopoulos, X. Feng, M. Golterman, S. Gottlieb, R. Gupta, G. Herdoiza, P. Hernandez, A. Jüttner, T. Kaneko, E. Lunghi, S. Meinel, C. Monahan, A. Nicholson, T. Onogi, P. Petreczky, A. Portelli, A. Ramos, S.R. Sharpe, J.N. Simone, S. Sint, R. Sommer, N. Tantalo, R.V. de Water, A. Vaquero, U. Wenger, H. Wittig (2025). <https://arxiv.org/abs/2411.04268>
7. I. Esteban, M.C. Gonzalez-Garcia, M. Maltoni, I. Martinez-Soler, J.P. Pinheiro, T. Schwetz, J. High Energy Phys. **2024**(12) (2024). [https://doi.org/10.1007/jhep12\(2024\)216](https://doi.org/10.1007/jhep12(2024)216)
8. F. Feruglio, A. Romanino, Rev. Mod. Phys. **93**, 015007 (2021). <https://doi.org/10.1103/RevModPhys.93.015007>
9. F. Herren, M. Steinhauser, Comput. Phys. Commun. **224**, 333 (2018). <https://doi.org/10.1016/j.cpc.2017.11.014>
10. G. Altarelli, F. Feruglio, Nucl. Phys. B **741**(1–2), 215–235 (2006). <https://doi.org/10.1016/j.nuclphysb.2006.02.015>
11. T. Ohlsson, S. Zhou, Nat. Commun. **5**, 5153 (2014). <https://doi.org/10.1038/ncomms6153>
12. Project 8 Collaboration, A.A. Esfahani, S. Böser, N. Buzinsky, M.C. Carmona-Benitez, C. Claessens, L. de Viveiros, P.J. Doe, S. Enomoto, M. Ferl, J.A. Formaggio, J.K. Gaisson, M. Grando, K.M. Heeger, X. Huan, A.M. Jones, K. Kazkaz, M. Li, A. Lindman, C. Matthé, R. Mohiuddin, B. Monreal, R. Mueller, J.A. Nikkel, E. Novitski, N.S. Oblath, J.I. Peña, W. Pettus, R. Reimann, R.G.H. Robertson, G. Rybka, L. Saldaña, M. Schram, P.L. Slocum, J. Stachurska, Y.H. Sun, P.T. Surukuchi, J.R. Tedeschi, A.B. Telles, F. Thomas, M. Thomas, L.A. Thorne, T. Thümmel, W.V.D. Pontseele, B.A. VanDevender, T.E. Weiss, T. Wendler, A. Ziegler, The project 8 neutrino mass experiment (2022). [arXiv:2203.07349](https://arxiv.org/abs/2203.07349)
13. M. Borghesi, B. Alpert, M. Balata, D. Becker, D. Bennet, E. Celasco, N. Cerboni, M. De Gerone, R. Dressler, M. Faverzani, M. Fedkevych, E. Ferri, J. Fowler, G. Gallucci, J. Gard, F. Gatti, A. Giachero, G. Hilton, U. Koster, D. Labranca, M. Lusignoli, J. Mates, E. Maugeri, S. Nisi, A. Nucciotti, L. Origo, G. Pessina, S. Ragazzi, C. Reintsema, D. Schmidt, D. Schumann, D. Swetz, J. Ullom, L. Vale, Nucl. Instrum. Methods Phys. Res. Sect. A Accel. Spectrom. Detect. Assoc. Equip. **1051** (2023). <https://doi.org/10.1016/j.nima.2023.168205>
14. L. Randall, R. Sundrum, Phys. Rev. Lett. **83**, 3370 (1999). <https://doi.org/10.1103/PhysRevLett.83.3370>
15. L. Randall, R. Sundrum, Phys. Rev. Lett. **83**, 4690 (1999). <https://doi.org/10.1103/PhysRevLett.83.4690>
16. T. Gherghetta, A. Pomarol, Nucl. Phys. B **586**, 141 (2000). [https://doi.org/10.1016/S0550-3213\(00\)00392-8](https://doi.org/10.1016/S0550-3213(00)00392-8)
17. E. Ma, Phys. Rev. D **70**, 031901 (2004). <https://doi.org/10.1103/PhysRevD.70.031901>
18. C.S. Lam, Phys. Rev. D **78**, 073015 (2008). <https://doi.org/10.1103/PhysRevD.78.073015>
19. C.S. Lam, Phys. Rev. D **87**, 013001 (2013). <https://doi.org/10.1103/PhysRevD.87.013001>
20. J.T. Tsang, Status and outlook of quark flavour physics (2025). <https://arxiv.org/abs/2503.05554>