Resolving the Light-Flavor Composition of Hadronic Z Decays with DELPHI Open Data

Jianfeng Jiang*, **Shuo Han***, Yuzhi Che, Xinchou Lou, Manqi Ruan

Institute of High Energy Physics, Chinese Academy of Sciences

23 October 2026 · Shanghai

* These authors contributed equally to this work.

A direct strange fraction is the missing piece

$$R_q^{\,o} = \frac{\Gamma(Z \to q\,\bar{q})}{\Gamma_{had}}$$

Flavor-resolved tests of the hadronic Z width

Separate s from u+d

Particle identification separates flavors; response calibration turns tags into a measurement.

Fit R_b, R_c, R_s; derive R_{u+d}.

No imposed $R_d = R_s$.

d and s have the same SM electroweak charges:
$\Gamma_d \simeq \Gamma_s$ for negligible quark masses.
Testing this separately requires resolving u and d.

Previous benchmarks	Result
DELPHI R_b	0.21634 ± 0.00065 ± 0.00060
LEP+SLD R_b	0.21629 ± 0.00066
DELPHI R_c	0.1665 ± 0.0051 ± 0.0081
LEP+SLD R_c	0.1721 ± 0.0030
OPAL $R_{d,s}^{uds}$	0.371 ± 0.016 ± 0.016

OPAL: $R_d = R_s$, normalized to uds.

Fixing R_d instead: $R_s^{uds} = 0.392 \pm 0.043 \pm 0.045$.

> CEPC's large Z sample makes data calibration central to precise flavor measurements.

LEP EWWG, Phys. Rep. 427 (2006); OPAL, Z. Phys. C 76 (1997); PDG, Electroweak Review (2025).

Select hadronic Z decays before flavor tagging

Pre-selection

Hadronic Z candidates passing the preserved DELPHI Team-4 selection.

Charged-particle multiplicity and energy identify multihadronic events.

Require a valid beam / vertex record and reconstructed particles assigned to hemispheres.

Cutflow · collision data

Requirement	1994	1995
Input data before analysis cuts	5,753,343	3,661,917
DELPHI hadronic selection	1,400,252	677,507
+ valid beam / vertex record	1,389,506	674,905

Input: archived DELPHI short-data streams.
Hemisphere-input requirement causes no further loss.

Selected collision data

Year	Round 1 role	All events	Barrel	Endcap
1994	Flavor measurement	1,389,506	906,345	483,161
1995	Response calibration	674,905	438,568	236,337

Preparation through NanoAOD follows the same workflow as Y. Che's DELPHI analysis, this workshop.

Round 1: ParticleNet classifies the full event

≤80 particles · 15 inputs
Kinematics, charge, dE/dx, RICH and impact parameters.

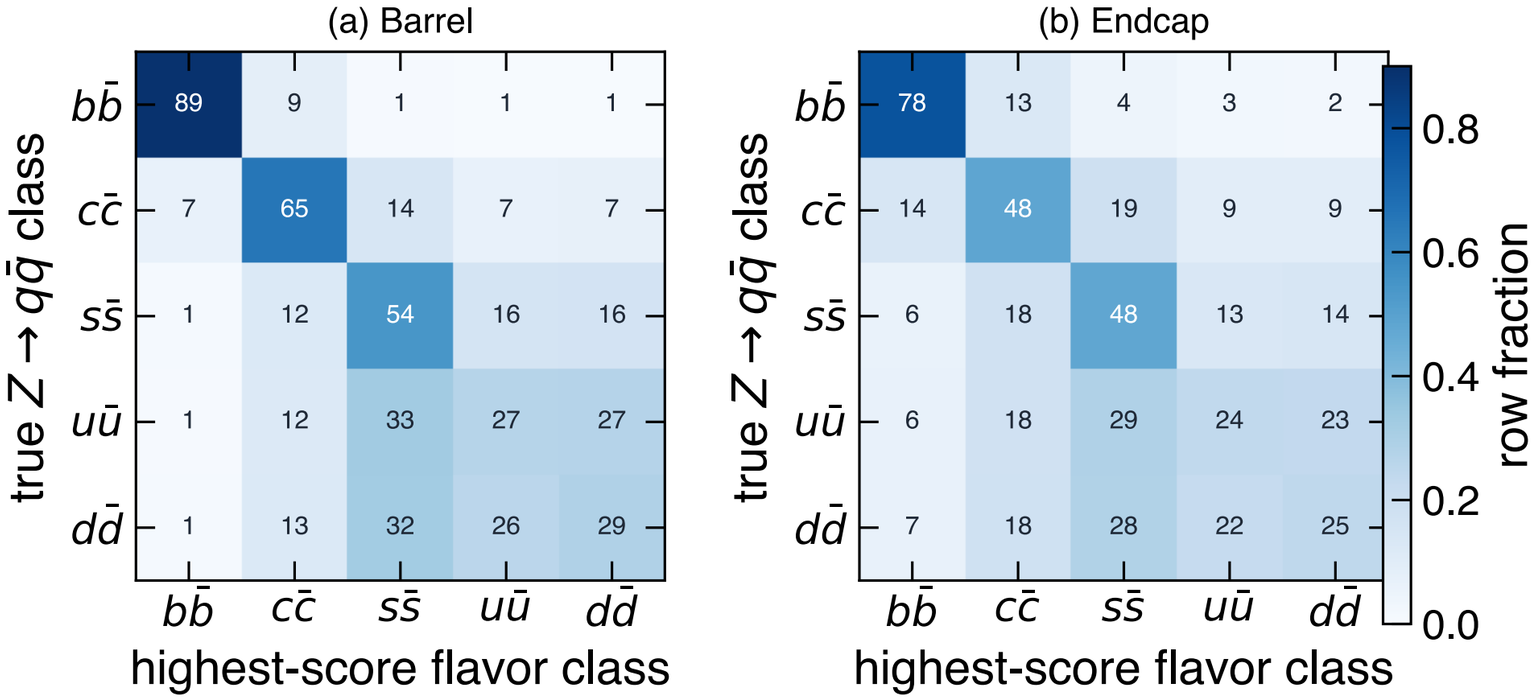
3 EdgeConv blocks · k = 16
64 → 128 → 256 channels

↓

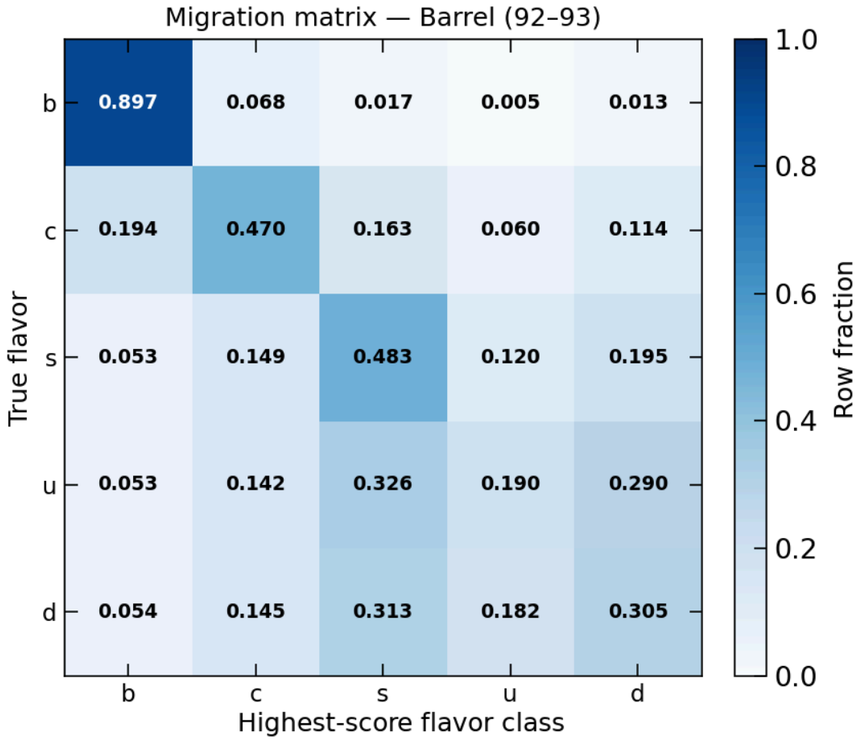
Global pooling → dense 256
Dropout 0.1 → u, d, s, c, b

Training on simulation
500k events + 50k validation
30 epochs · batch 512
Ranger · initial learning rate 0.01
Five-class cross-entropy

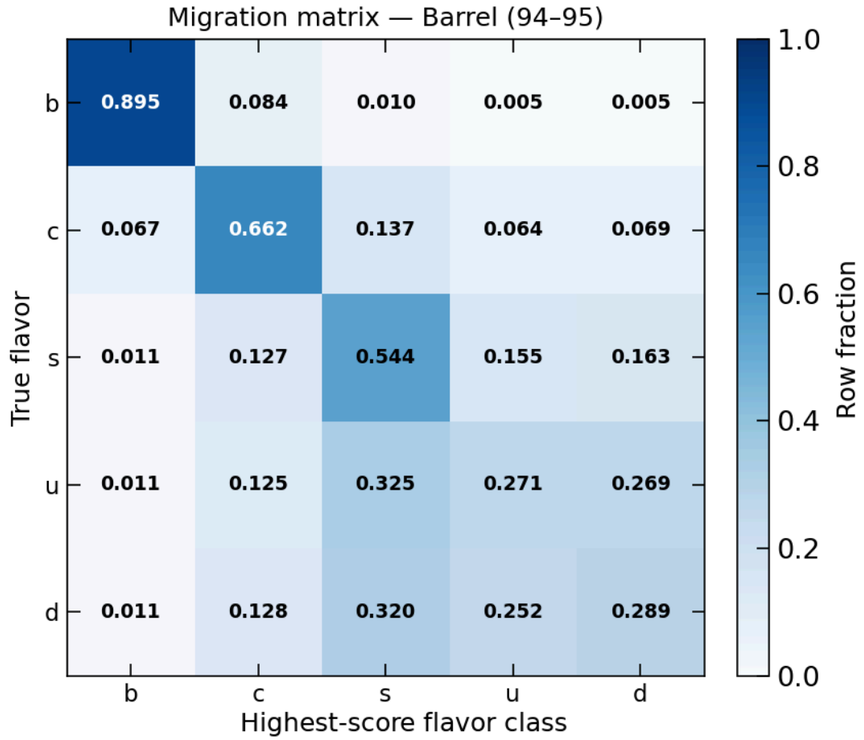
Five-flavor response · highest-score label (argmax)



Why exclude the 1992–1993 data?



1992–1993



1994–1995

> The 1994 vertex-detector upgrade changes flavor separation substantially.

Use 1994–1995; earlier data would require a separate response calibration.

Eight exclusive counts connect tags to flavor fractions

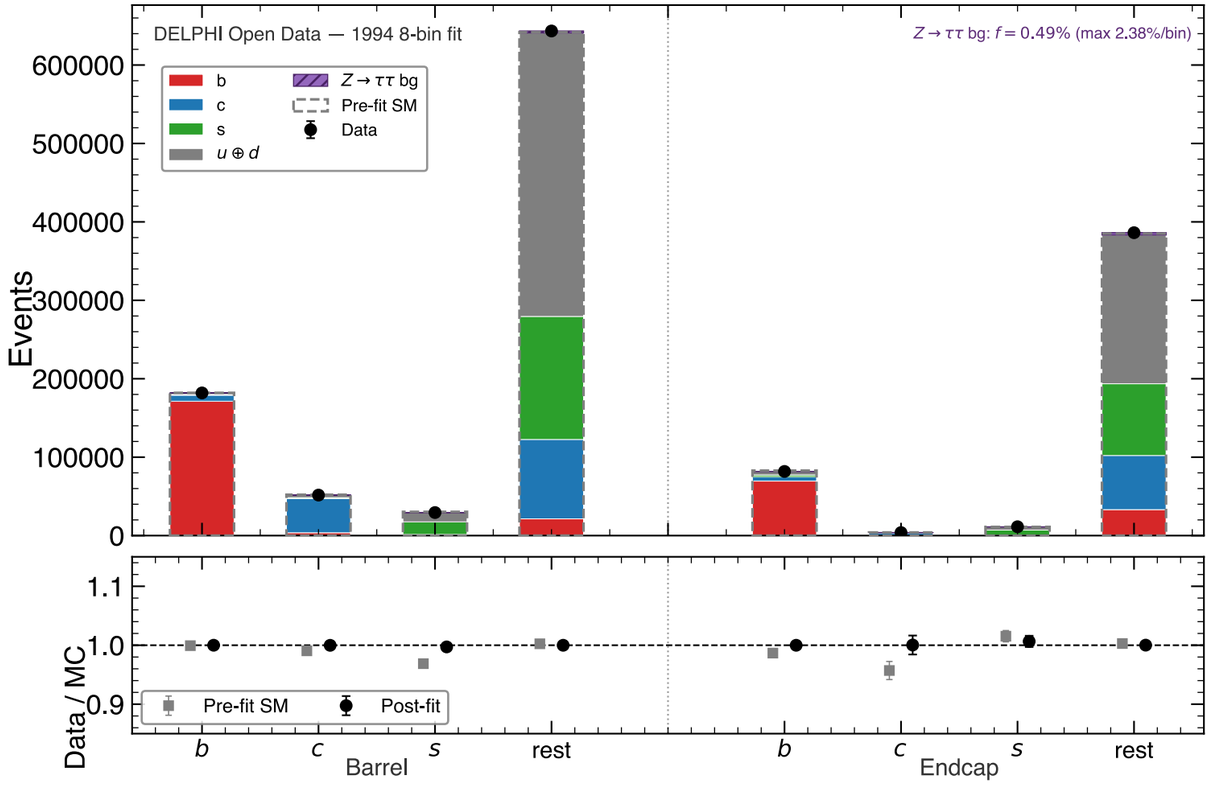
One event, one category

b	highest b score; $s_b \geq \theta_b$
c	otherwise: $s_c \geq \max(s_s, \theta_c)$
s	otherwise: $s_s \geq \max(s_u, s_d, \theta_s)$
remainder	all other events

Barrel + endcap: **4 × 2 = 8 counts**

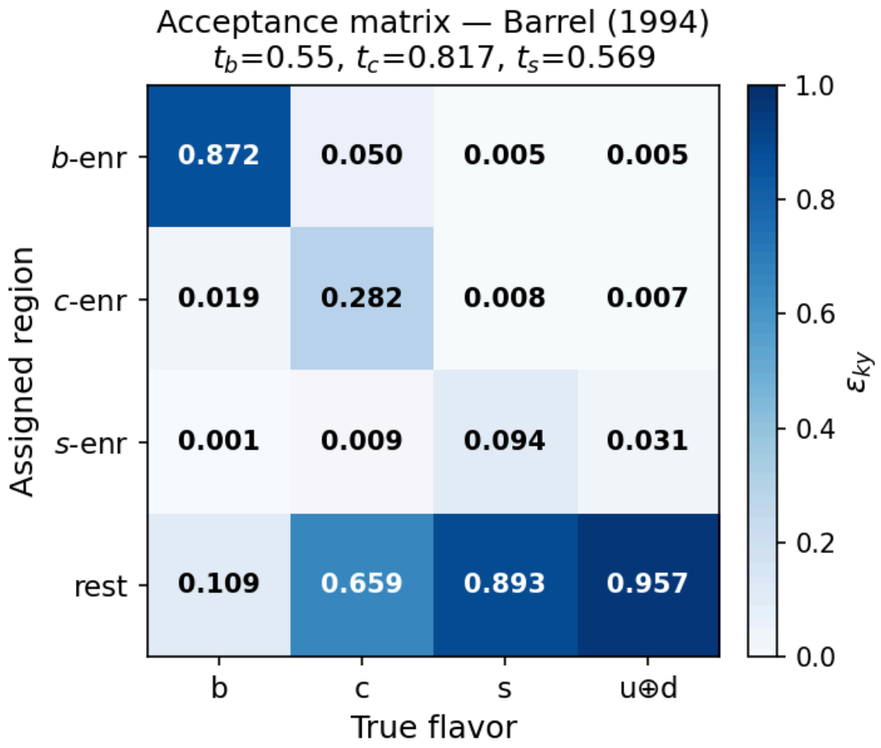
Score cuts	θ_b	θ_c	θ_s
Barrel	0.550	0.817	0.569
Endcap	0.557	0.895	0.587

Fixed in MC with purity targets:
b/c ≥85%; s ≥60%.

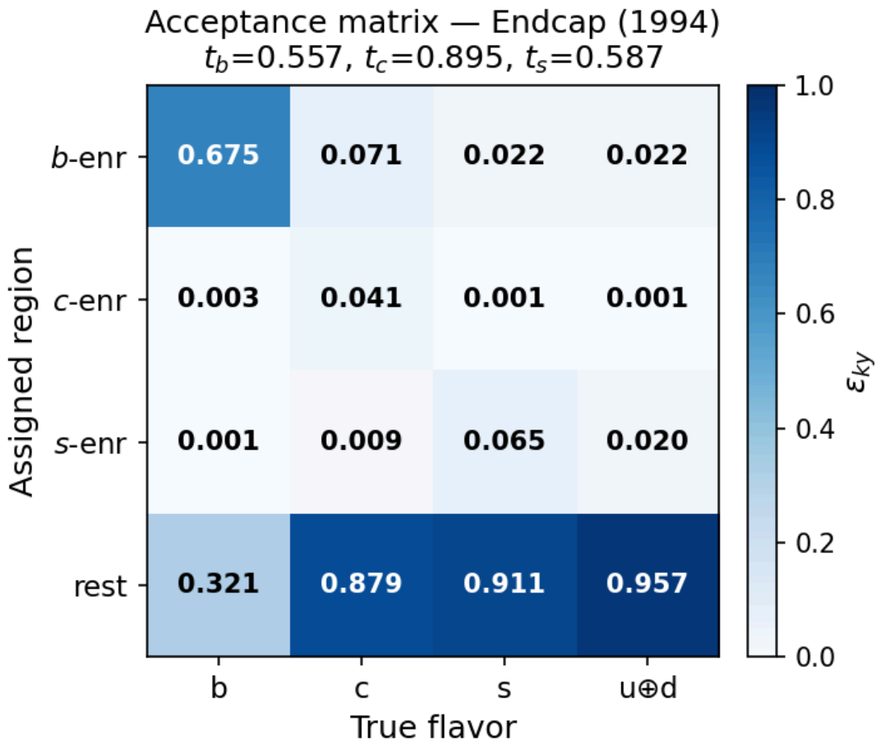


1994 data compared with the five-flavor prediction and ττ background.
Counts trade some score information for fewer response terms to calibrate.

Acceptance maps flavors into the observed counts



Barrel

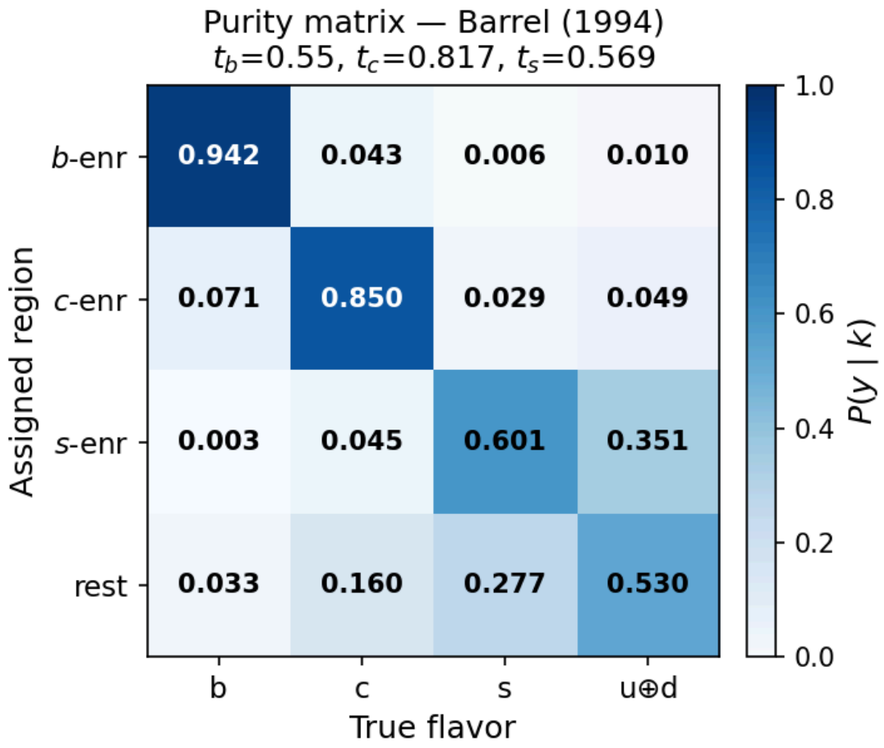


Endcap

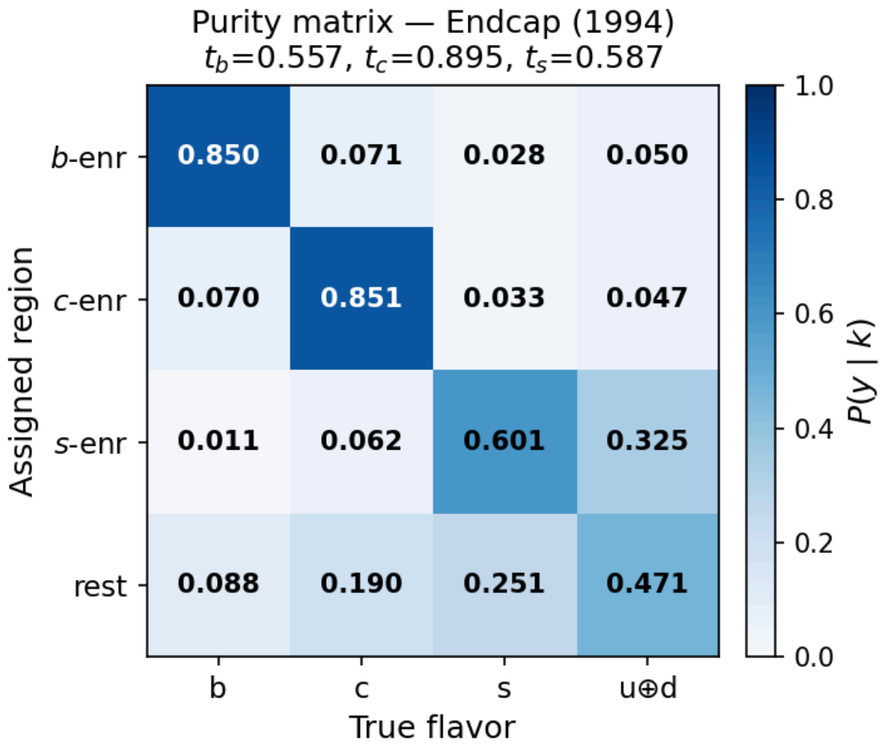
$\mathbf{p_{r,t,f} = P(t \mid f, r)}$ — probability for true flavor f to enter tag category t.

This response, including migration among all flavors, predicts the observed counts.

Purity sets the simulation-selected working points



Barrel



Endcap

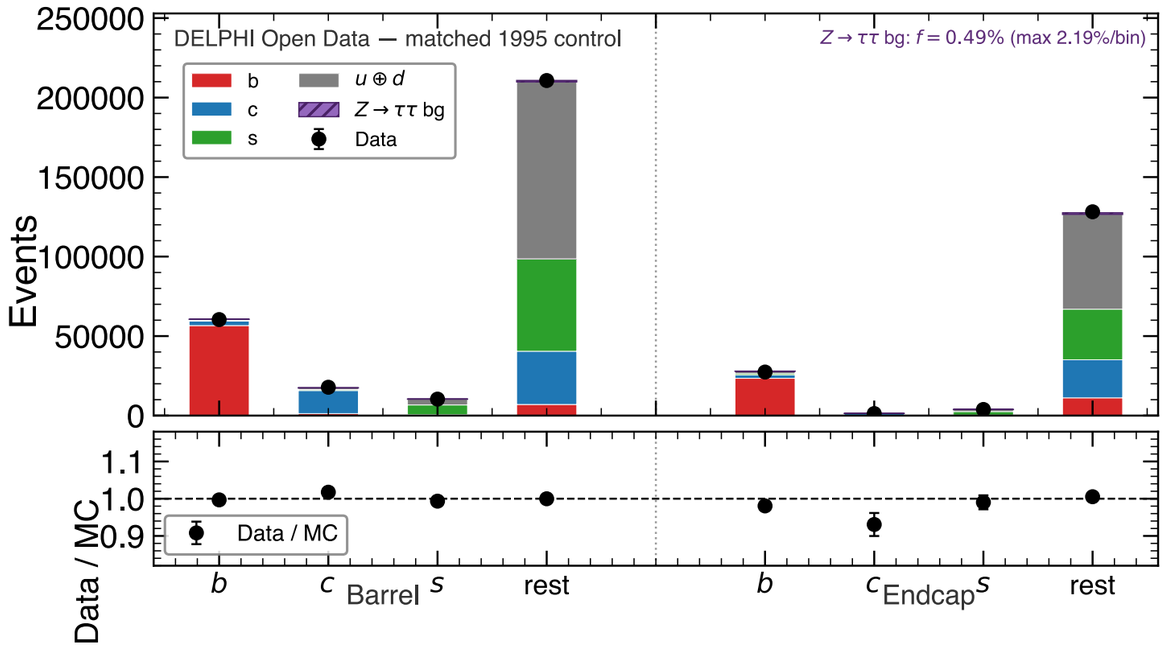
P(f | r, t) — flavor composition of a selected category, at the SM reference.

(b, c, s) purity: barrel (94.2, 85.0, 60.1)%; endcap (85.0, 85.1, 60.1)%.

How 1995 data define the error applied to 1994

1995 control data

Reference-mixture MC; tag efficiencies matched to 1994.



Translate the residual

Flavor-response direction
+ data / MC statistics

Build a covariance

Centered on nominal MC
Residual sets the width

Fit the 1994 fractions

R

Vary the response
within that uncertainty

> **Transfer assumption:** after matching tag efficiencies, the 1995 residual describes a response uncertainty applicable to 1994.

The residual defines an uncertainty; the nominal response is not shifted.

Tag-response calibration dominates the Round 1 uncertainty

Systematic sources: relative impact weights (%)

Source and evaluation	R_b	R_c	R_s
Tag response 1995 data–MC residual + control statistics; kaon correlations retained.	70.9	80.8	93.1
1994 MC statistics Finite response-matrix and selection-efficiency samples.	9.6	16.0	6.2
Detector / RICH PID-validity masking scaled by data–MC mismatch; residual detector / PID terms.	2.3	0.2	0.3
Hadronization MC proxy reweighting for fragmentation / decays; DELPHI-based scale.	17.2	3.0	0.5
Beam energy LEP energy error propagated through the energy dependence of tag counts.	<0.01	<0.01	<0.01



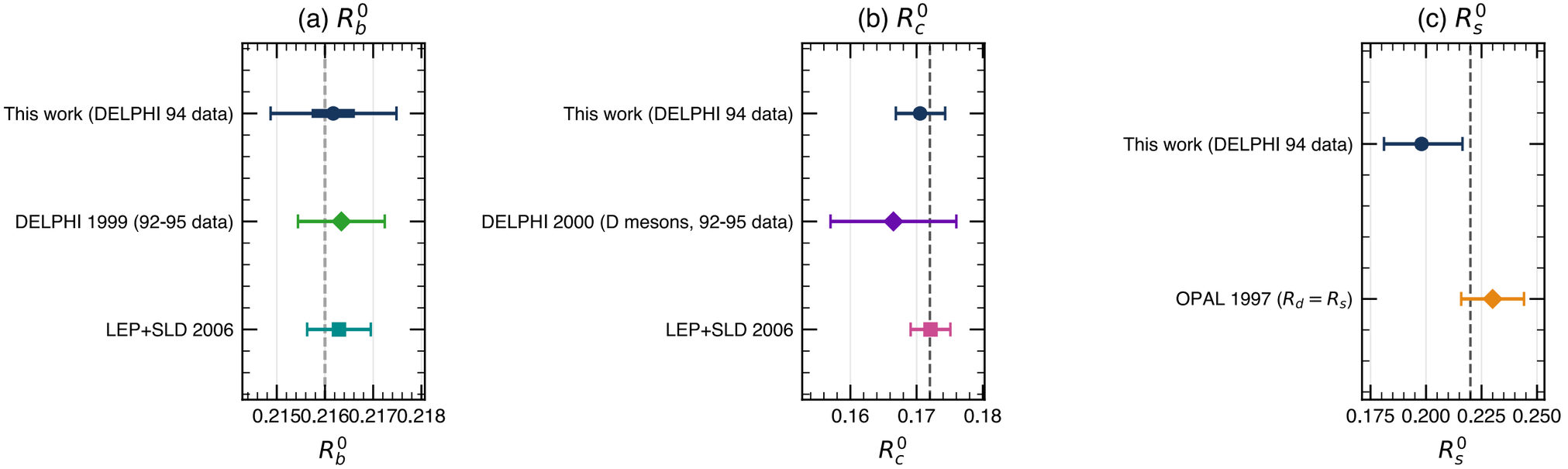
Toy covariance	ΔR_b	ΔR_c	ΔR_s
Counting	0.00045	0.00091	0.00272
Systematic	0.00129	0.00379	0.01740
Total	0.00137	0.00390	0.01761

> Tag-response calibration dominates; counting precision is already strong.

Left table / outer rings: normalized squared source impacts within systematics; not additive ΔR components.
Right table / inner rings: one toy covariance, stat² + syst² = total². Counting: Poisson fluctuations of 1994 data.

Direct flavor fractions without down–strange symmetry

A simultaneous likelihood fit extracts R_b, R_c and R_s from the eight tag counts.



R_b^0	R_c^0	R_s^0	R_{u+d}^0 · derived
$0.21617\,^{+0.00131}_{-0.00130}$	$0.17052\,^{+0.00379}_{-0.00366}$	$0.19801\,^{+0.01842}_{-0.01693}$	$0.41530\,^{+0.01691}_{-0.01798}$

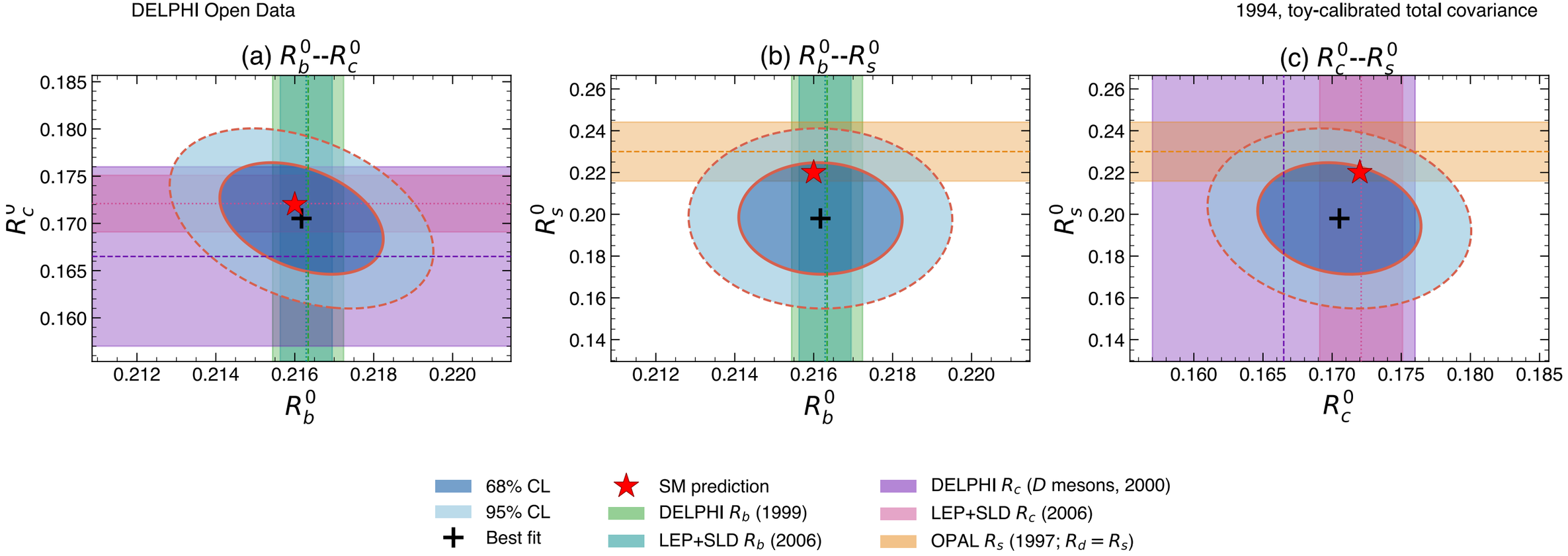
R_b statistical error: 0.00045 vs historical DELPHI 0.00065 — 31% smaller. R_s: 1.4% relative.

$R_{u+d} = 1 - R_b - R_c - R_s$, without $R_d = R_s$. All quoted intervals: 68% profiles.

Historical results share data; OPAL assumes $R_d = R_s$. Inner bars: counting component.

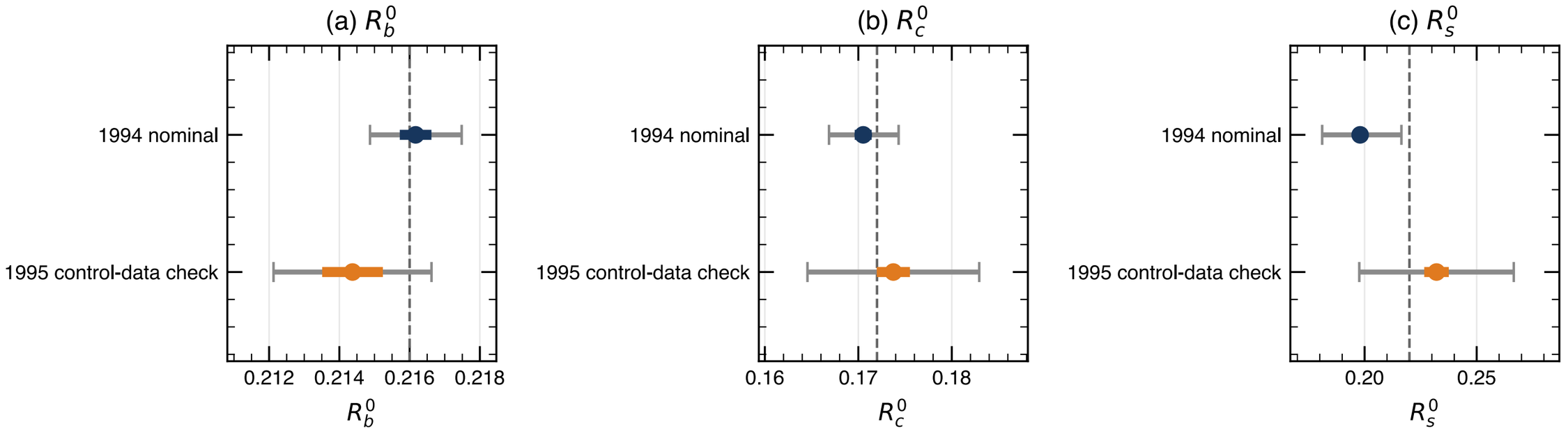
The simultaneous fit retains the joint flavor information

$\rho(R_b, R_c) = -0.368$ $\rho(R_b, R_s) = -0.020$ $\rho(R_c, R_s) = -0.139$



68% / 95% regions from the total covariance; bands show published measurements.

1995 provides a compatible control-data check



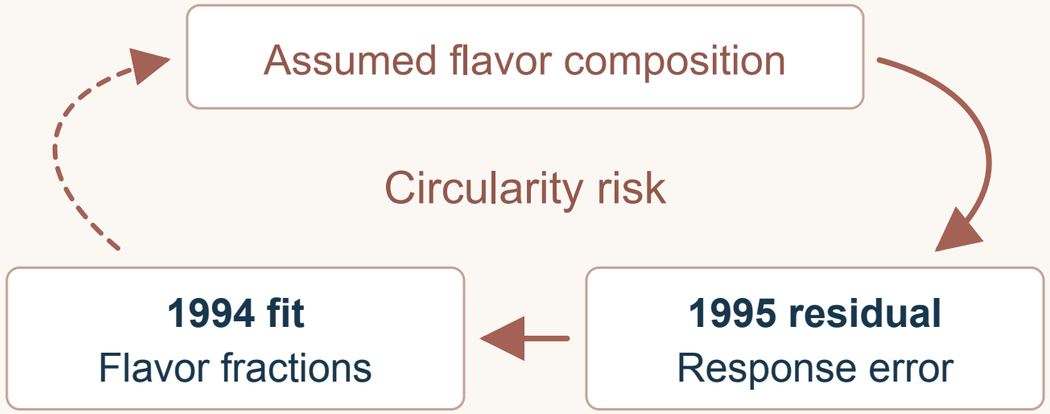
Run-disjoint comparison: **p = 0.708**. The years are not combined.

Thick bars: counting uncertainty. Thin bars: total uncertainty.

Where the circularity risk enters

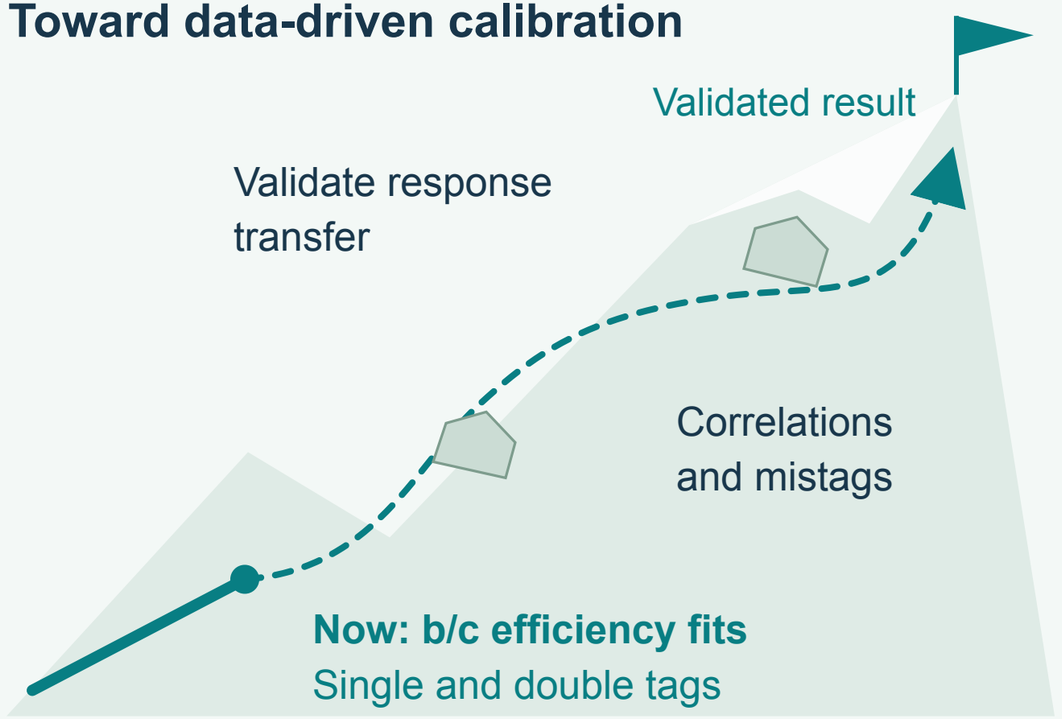
Data − MC can reflect **flavor composition** as well as **tag response**.

Round 1 · where the loop enters



Separate years remove event overlap,
but not the assumed-mixture dependence.

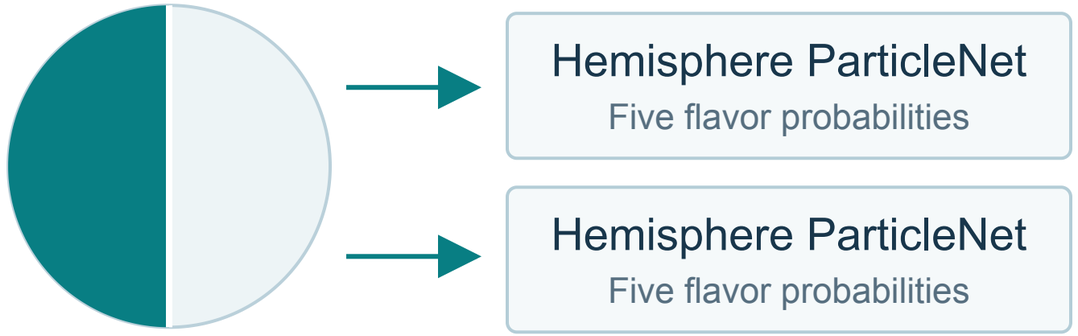
Toward data-driven calibration



A real shift in R_s could be absorbed into the assigned response error.

Toys at SM and best-fit truth test covariance coverage within the assumed response model.
Independent response transfer remains open; bias in the central values has not been established.

Round 2: a shared ParticleNet tags the two hemispheres



61 inputs · ≤80 particles per event

Kinematics, tracking / vertex quality and PID.

Backbone: 3 EdgeConv blocks, k = 16
64 / 128 / 256 channels · dense 256
Dropout 0.1

Training: 500k + 50k validation events
30 epochs · batch 512 · Ranger · LR 0.01

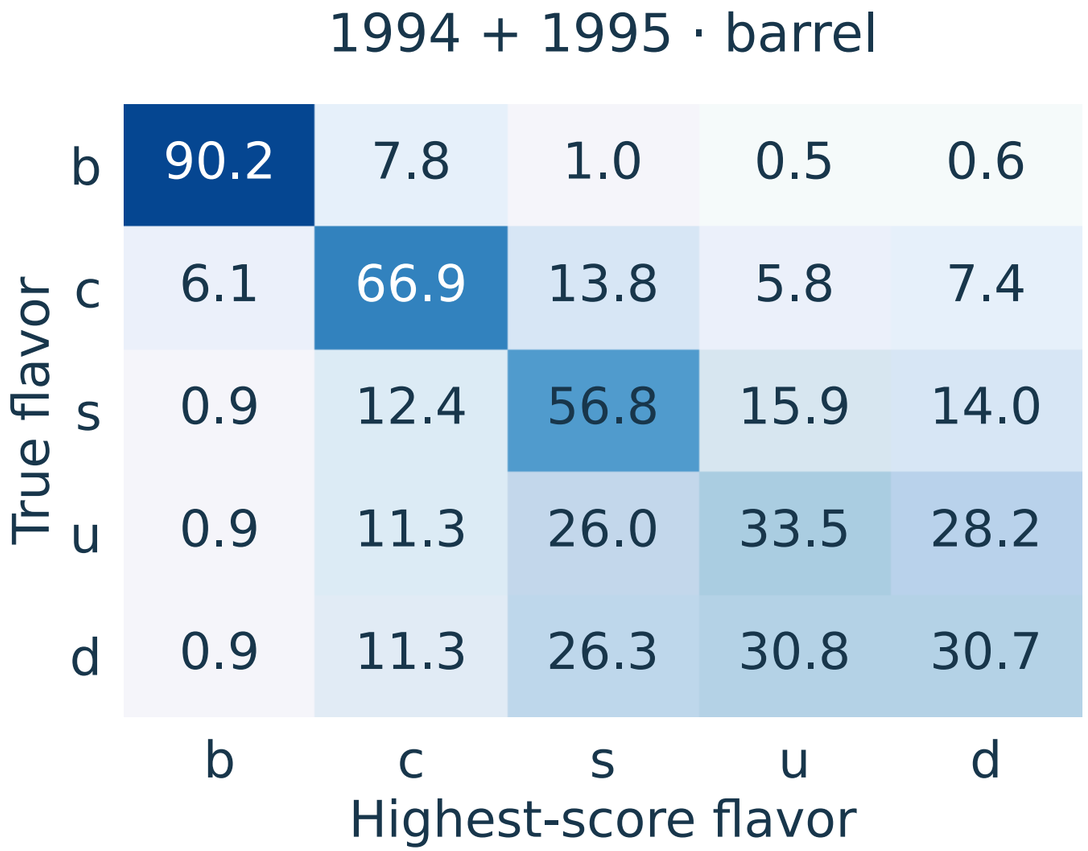
Loss: event cross-entropy + mean hemisphere cross-entropy

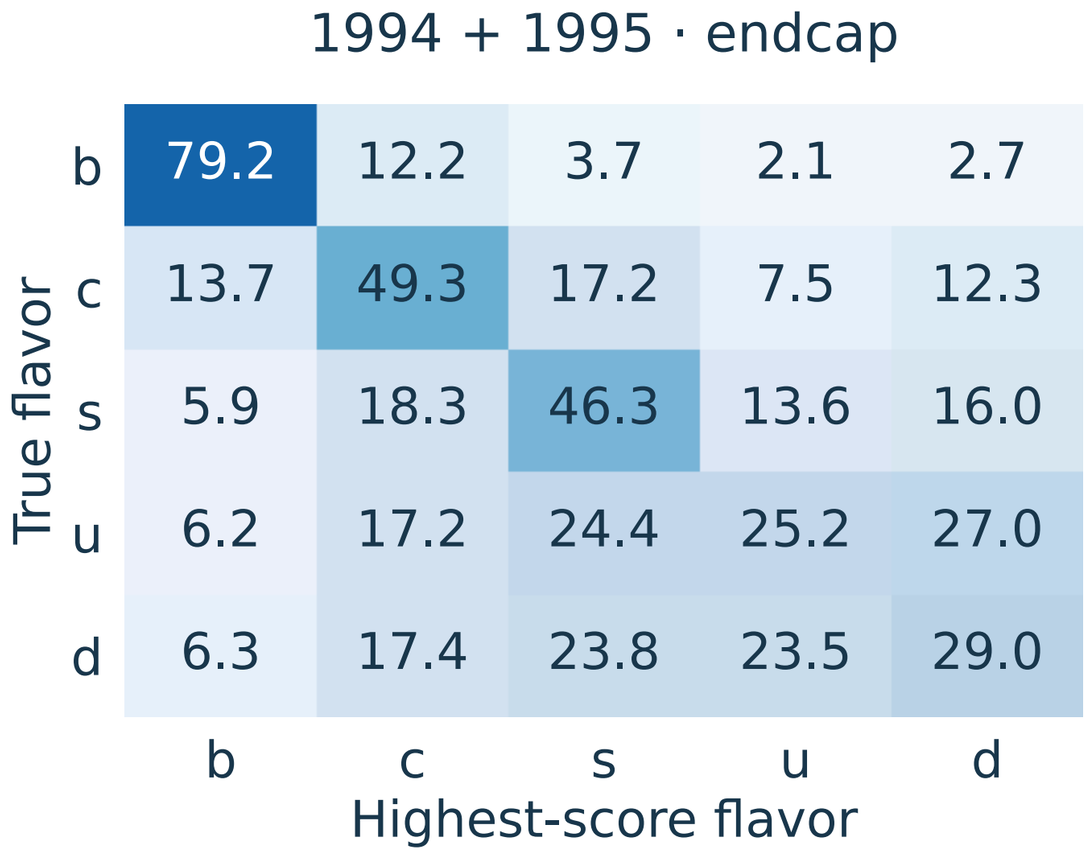
Tag heads: 5 → 32 → 32 → 3, applied to each side separately.
Frozen ParticleNet; heads trained for 40 epochs with Adam, batch 4096, LR 0.003.

> Side labels: **B** if b score ≥ t_B; else **C** if c score ≥ t_C; otherwise **U**.
> Fixed thresholds in MC; the measurement does not use argmax side labels.

Two hemispheres retain event-level flavor separation

Combined 1994–1995 simulation · event label = highest combined score (argmax)





Barrel: $|\cos\theta| < 0.7$; endcap: $|\cos\theta| \geq 0.7$.
The combined event output can be compared with Round 1; the fit also retains the two side tags.

Why use the barrel for the nominal result?

Hemisphere correlations in simulation · common primary vertex

Same-flavor tag correlation

Tag / region	1994	1995
b · barrel	0.992 ± 0.005	0.997 ± 0.003
b · endcap	1.350 ± 0.031	1.374 ± 0.022
c · barrel	1.067 ± 0.020	1.045 ± 0.013
c · endcap	1.693 ± 0.256	1.882 ± 0.190
s · barrel	1.006 ± 0.004	1.011 ± 0.003
s · endcap	1.116 ± 0.013	1.128 ± 0.009

C = P(both tags) / (ε_0 ε_1).
Region averages; errors are MC counting statistics.

Back-to-back jets share |cos θ|



Inner endcap: ε ≈ 24%

Forward: ε ≈ 7%

Both sides efficient — or both sides inefficient

$$C_{\mathrm{mix}} = \langle \varepsilon^2 \rangle \,/\, \langle \varepsilon \rangle^2 \geq 1$$

Even if the two sides are independent at fixed angle.

> Use the barrel and retain two angle bins; C = 1 is a fit assumption.
> The key missing input is a data constraint on the residual correlation in each bin.

Measure the b fraction and tag efficiency together

Ideal b example: no mistags, equal side efficiencies, no acceptance difference; first take $C_b = 1$.

R_b: true $b\bar{b}$ share of hadronic events. **ε_b^{data}:** probability of tagging a true b hemisphere in data.
S_b and D_{bb} below are both measured in data.



Tagged hemispheres

$$S_b = R_b\, \varepsilon_b^{data}$$

S_b: tagged / all hemispheres



Double-tagged events

$$D_{bb} = R_b\, (\varepsilon_b^{data})^2$$

D_{bb}: double tags / all events

$$\varepsilon_b^{data} = D_{bb} / S_b$$

Measure the efficiency from data.
$\varepsilon_b^{data} / \varepsilon_b^{MC}$ is a comparison after the fit.

$$R_b = S_b / \varepsilon_b^{data} = S_b^2 / D_{bb}$$

The same counts determine R_b.
If $C_b \neq 1$: $R_b = C_b\, S_b^2 / D_{bb}$.

Data + finite MC counts
Statistical fluctuations

→

Fit ε_b(data) and R_b
Keep their joint uncertainty

→

Interval on R_b
Profile the efficiency too

> Data determine R_b and ε_b^{data} together; no separate correction to R_b is applied.

Partial intervals: correlation, mistag and acceptance / response transfer incomplete. Multi-flavor fit → Backup 3 (p.24).

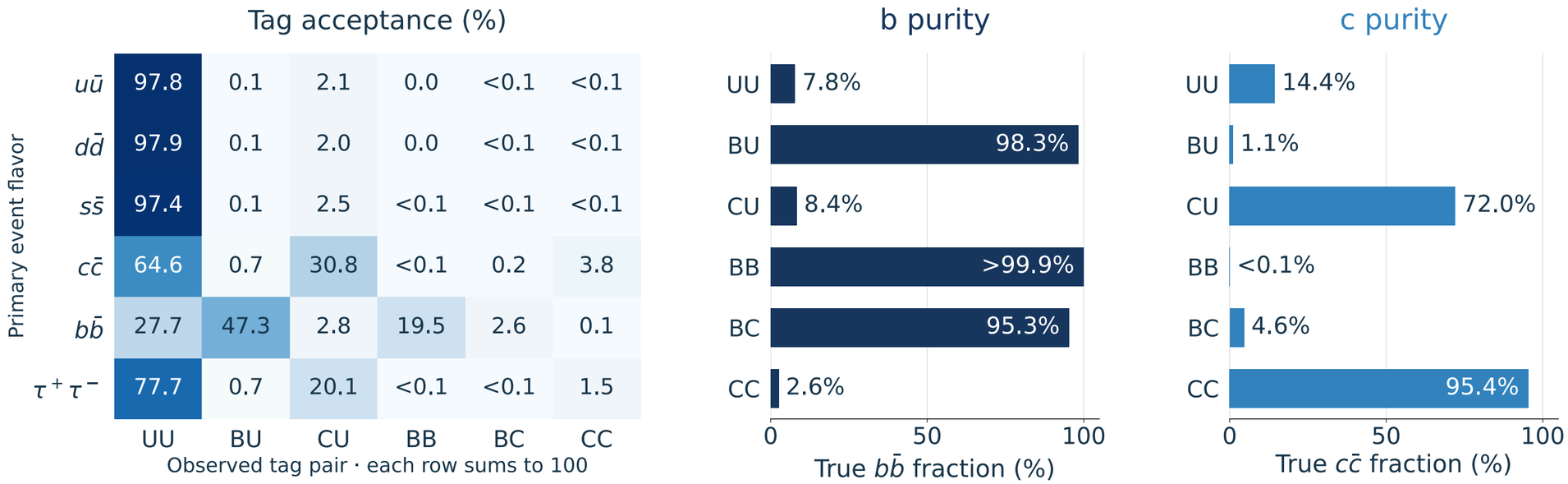
Every tag pair contributes to the same fit

B: ≥99% MC purity; C: tighter MC scan point.
B takes priority, then C; U is neither. Cuts use head scores, not probabilities.

UU: no tags | **BU / CU:** one tag | **BB / CC:** two | **BC:** mixed

No tags: normalization; single/double: efficiency; mixed: cross-flavor response.

Head-score cuts	t_B	t_C
Barrel	9.748	4.209
Endcap	11.448	5.060



Purity = true b b̄ or c c̄ fraction in each category; BC is not a true “bc” decay.
Plots combine UC + CU etc.; the fit retains side order and event category.

Strange extension: also count 0 / 1 / 2 s tags in each event to fit s efficiency.
Independent recovery of R_s is still unresolved; B/C/U pairs alone do not isolate s.

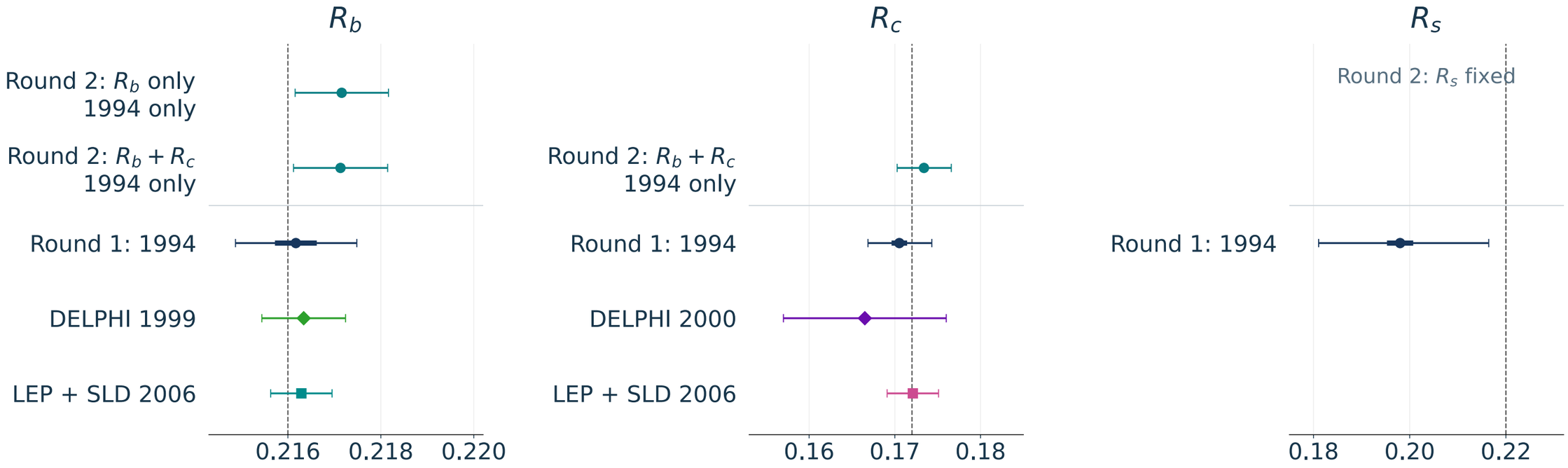
Round 1 and preliminary Round 2 results

Round 2 · partial uncertainties only

1994 barrel · common vertex · C = 1 within year / angle bins

$R_{u+d} = 0.38947$

1994 joint fit · R_s fixed at 0.220 · central only



Round 1: total 68% at the Z pole (thick: statistical). Round 2: partial 68%, before pole conversion. Dashed: SM.
R_b-only / R_b+R_c fix R_s = 0.220; R_b-only also fixes R_c = 0.172. All fits fix internal u:d.

Both fits use 1994 only; R_{u+d} is conditional. The origin of the shift from Round 1 is not yet established.

> Joint fraction / efficiency fits are feasible; a precision gain is not yet established.

Summary and outlook

A workflow for LEP DELPHI open data

We established the workflow and analyzed hadronic $Z \rightarrow q\bar{q}$ events.

Two ParticleNet strategies for flavor separation

We trained holistic event-level and shared hemisphere ParticleNet models to distinguish quark flavors.

Round 1: competitive strange-flavor precision

1995 data set the response uncertainty for the 1994 R_s measurement and derived R_{u+d}.
Tag response dominates; its composition dependence motivates Round 2.

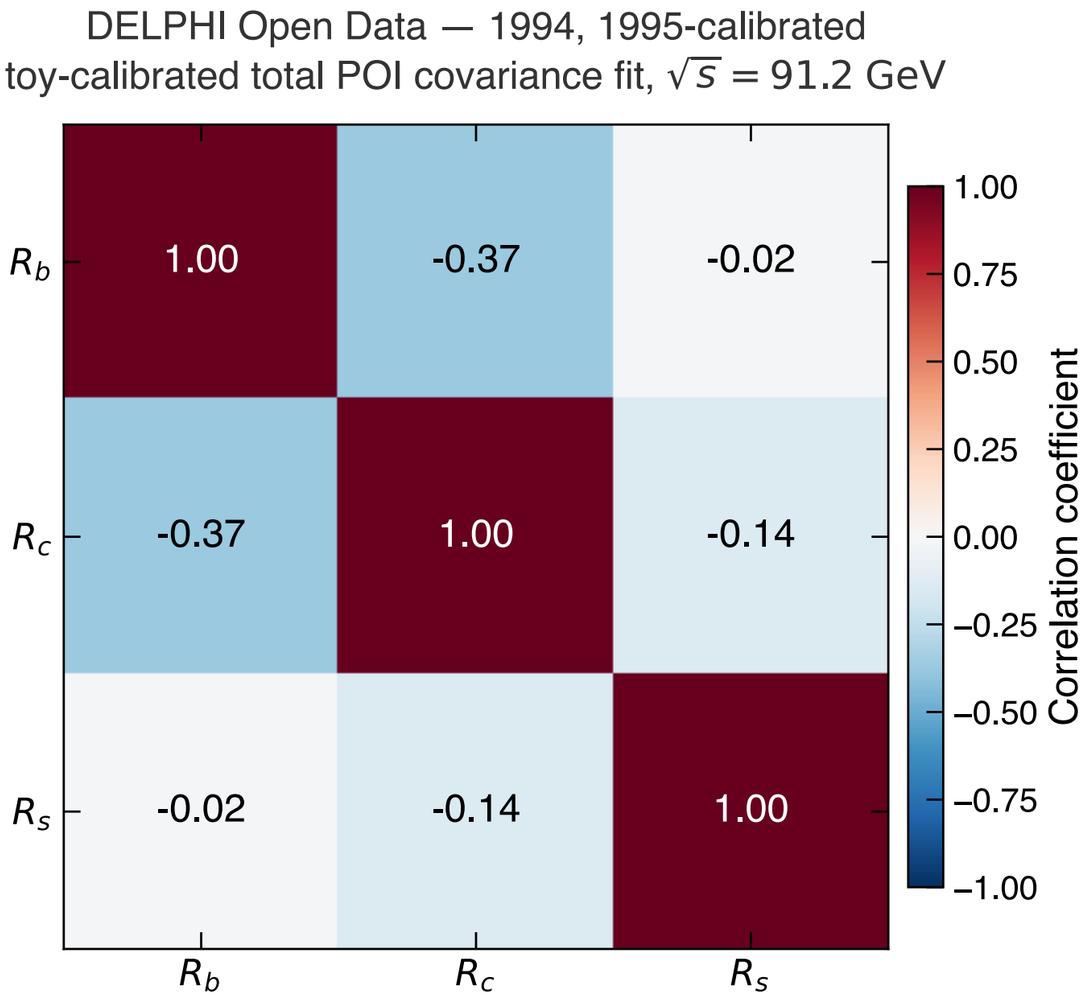
Round 2: fit fractions and efficiencies together

b/c efficiency calibration is demonstrated; the s-efficiency extension is under study.
Correlation constraints and independent strange-response recovery remain open.

A reusable platform for A_{FB} and CEPC

Orient the thrust axis with jet / vertex charge; fit tagged cos θ distributions for A_{FB},
correcting charge dilution and angular acceptance.
LEP data let us test joint flavor / efficiency fits and correlation control for CEPC.

The flavor ratios have a correlated total covariance

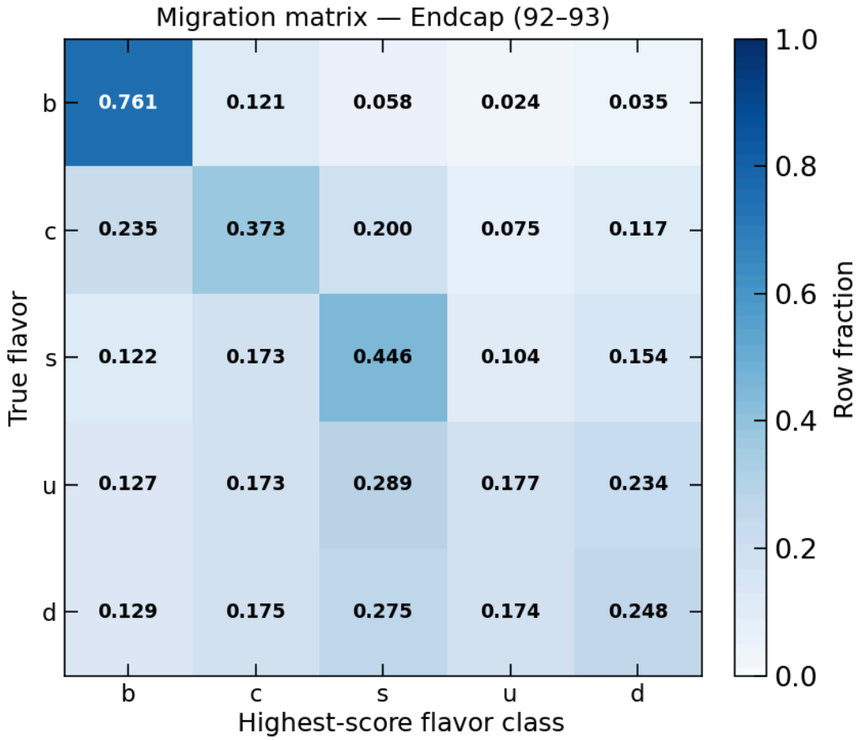


Pair	Correlation
R_b, R_c	-0.368
R_b, R_s	-0.020
R_c, R_s	-0.139

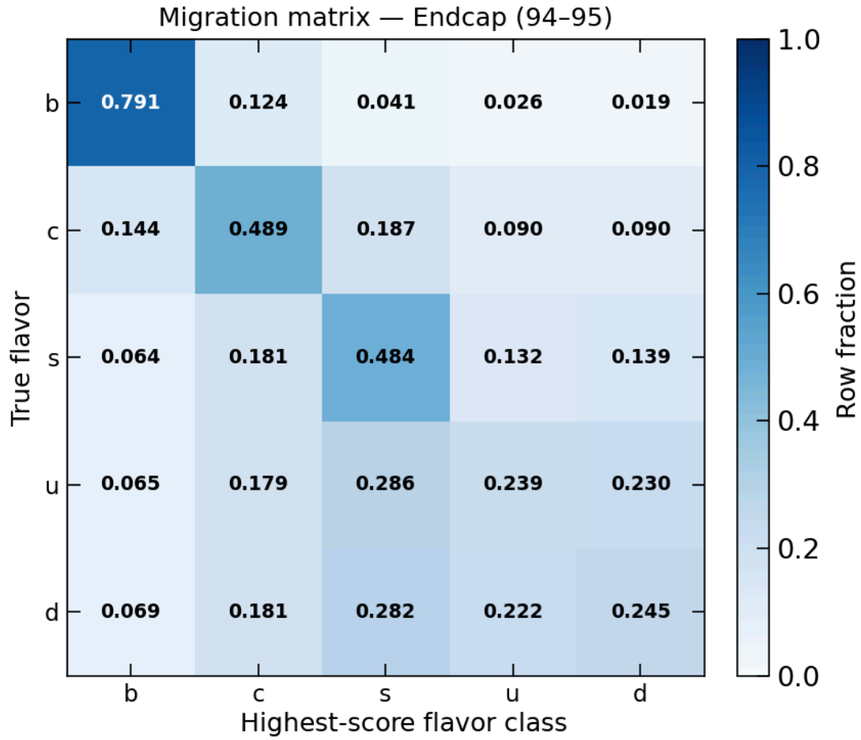
R_{ud} follows from normalization.
$\rho(R_{ud}, R_s) = -0.979$

Toy-calibrated total covariance.
Quoted intervals use 1D profiles.

Detector-era response: endcap



1992–1993



1994–1995

> The endcap and barrel need separate response matrices.

1992–1993 and 1994–1995 simulation · highest-score flavor assignment

From the single-b example to the multi-flavor fit

1 Slide 18: only b makes b tags

$$S_b = R_b\,\varepsilon_b^{\text{data}} \qquad D_{bb} = R_b\,(\varepsilon_b^{\text{data}})^2$$

2 Real b tags include other flavors

$$S_b = \sum_q w_q\,\varepsilon_{q\to b} \qquad D_{bb} = \sum_q w_q\,P_q(BB)$$

Signal-only rates; q = u, d, s, c, b. $\varepsilon_{q\to b}$: b-tag probability for true q (side average); $P_q(BB)$: double-b-tag probability.
No mistags, equal acceptance and independent equal sides → slide 18. Apply the same construction to c tags.

3 Fit the same data as exclusive counts

n_k: ordered B/C/U pairs × event category; each event counted once.
S and D are summaries of this table, not extra independent data.

$$\mu_k = N_{had} \sum_q w_q\,P_{qk}(\varepsilon^{\text{data}}, C, \eta) + B_k$$

$$w_q = R_q A_q \,/\, \sum_f R_f A_f$$

A: acceptance; P: category probability; η: other response terms.
N_{had}: selected hadronic yield; B: background.

Fit fractions and efficiencies

R_b, R_c and b/c side efficiencies;
R_s fixed in the two-flavor fit.

Response inputs still needed

MC acceptance, mistags and conditional response. L_{aux} includes finite MC and other constraints.
$C_b = C_c = 1$ is assumed, not fitted.

To add R_s: count 0 / 1 / 2 s tags; fit s efficiency. s-head cut: 1.205 (barrel), 1.284 (endcap).
$C_s = 1$ and equal s-side efficiencies are assumed. Independent recovery remains unresolved.

$$L = \prod_k \text{Pois}(n_k \mid \mu_k) \times L_{aux}$$

Each R_q interval profiles all other fractions, efficiencies and response terms; $\Delta(-2 \ln L) = 1$ gives partial 68%.