Longitudinal Λ Polarization in Tagged $Z \to s\bar{s}$ Events with the CEPC AGORA Detector

Native source-dependent full simulation and method closure

Shuo Han Minqia Li

Institute of High Energy Physics, CAS

22 August 2026

One baseline selection and one disjoint enrichment partition

The parity-violating decay is a self-analyser:

$$\frac{1}{\Gamma_s}\frac{d\Gamma_s}{dc} = \frac{1}{2}\left(1 + \alpha_\Lambda P_s c\right).$$

Here $c \equiv \cos\theta^*$. The proton angle is defined in the Λ rest frame; the $\bar{\Lambda}$ angle is charge-flipped. The reconstructed density is the sum of source-dependent responses.

The nominal analysis reports

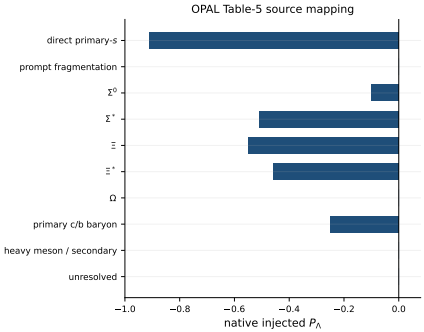
P_{eff} from separate inclusive, enhanced and complement mass–angle fits of all real-Λ peak candidates,
then $P_\Lambda^{\mathrm{direct-s}}$ from the joint, source-resolved enhanced/complement likelihood.

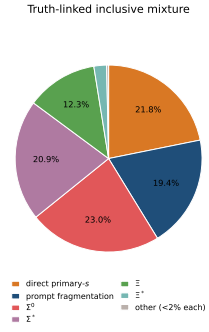
Enhanced and complement are mutually exclusive and add to inclusive. Inclusive has its own reported P_{eff}, but overlaps both categories and is not a third observation in the direct-source fit. Truth labels only build the input benchmark and the independent MC source response; they do not classify fitted entries.

Reporting contract: the inclusive sample is exactly enhanced $\oplus$ complement. All three selected-mixture fits use the same analytic construction; only the conditional direct-source result uses the disjoint pair jointly, under the OPAL non-direct-source assumptions.

Benchmarks: OPAL, Eur. Phys. J. C 2 (1998) 49; PDG 2025 $\alpha_- = 0.749 \pm 0.008$. The charge-flipped $\bar{\Lambda}$ angle uses the same sign convention.

The OPAL source model is injected natively in PYTHIA8





Benchmark assignments
direct primary-s: -0.91
prompt fragmentation: 0
Σ^0: -0.10
Σ^*: -0.51
Ξ: -0.55
Ξ^*: -0.46
primary c/b baryon: -0.25
Ω, heavy meson, unresolved: 0

Nominal source inference
POI: $P_\Lambda^{\mathrm{direct-s}}$
native injection: -0.91
Λ purity: 75.25%
direct-primary-s fraction of all selected: 16.43%

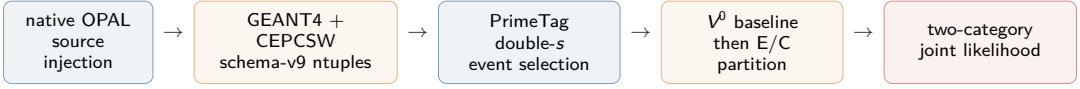
OPAL, $x_E > 0.3$
measured mixture: -0.329 ± 0.076
JETSET default prediction: -0.406
modified-popcorn prediction: -0.244

Current inclusive pseudo-observation
native real Λ: $4\,420 \times (-0.4063)$
Zqq real Λ: 743.6×0
yield-weighted $P_{\mathrm{input}} = \mathbf{-0.3478}$

OPAL Table 5 supplies the source-polarization benchmark. The current input averages truth-linked real Λ candidates only: native PYTHIA8 source labels give the signal term, while legacy-PYTHIA6 Zqq real Λ candidates are assigned the validated unpolarized density. Smooth or truth-unlinked candidates are not part of this polarization average.

Signal and response use independent native 200k samples

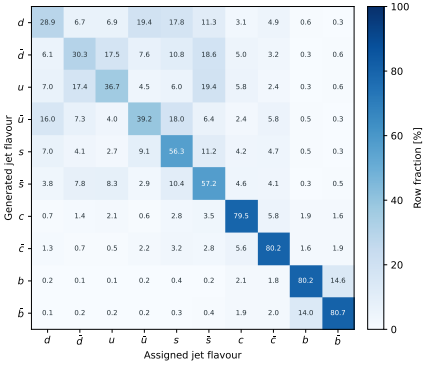
Sample	Events	Generation	Statistical role
native OPAL $Z \to s\bar{s}$	200 000	WHIZARD 3.1.7 + PYTHIA 8.310; source-dependent decay injection; schema-v9 full simulation	pseudo-observation for the disjoint fit categories
independent $Z \to s\bar{s}$, $P = 0$	200 000	same generator, detector and reconstruction chain; independent seeds	source-resolved angular response and DSCB mass auxiliary
explicit $Z \to q\bar{q}$ background	1 800 000	WHIZARD 1.95 + PYTHIA6; TDR25.3.5/25.3.7 reconstruction	all-Λ input dilution; source rate and angle only for the conditional direct-s fit
current five-flavour diagnostic	250 000	WHIZARD 3.1.7 + PYTHIA6; current reconstruction	tag migration only; never enters the likelihood



The source-dependent angular density is imposed before detector simulation, so the analysis applies unit event weights. The independent $P = 0$ sample supplies the angular response and DSCB mass auxiliary. Enhanced and complement counts add exactly to the inclusive count in signal and response.

Source polarizations: OPAL Table 5. Source identity: native PYTHIA8 fragmentation hook propagated through schema-v9; no source component is a fitted POI.

PrimeTag has sizeable signed light-flavour migration



Corrected schema-v8 ntuples; unchanged PrimeTag ONNX model and raw-argmax definition.

Model
ZJet M10, ten signed-flavour outputs **Category**

raw per-jet argmax; one s jet and one $\bar{s}$ jet; no score threshold **Input**

250k events; 499,805 valid jets; rows normalized to 100%

All five samples use re-computed TPC particle-hypothesis χ^2 values.

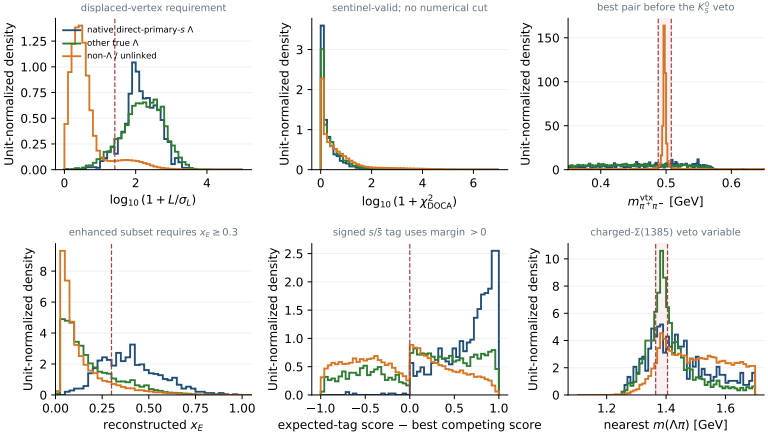
The double-strange category is 35.218% efficient and 79.94% pure

Generated process	N_{gen}	$N_{s\bar{s}\text{-tag}}$	ϵ_f [%]	Selected composition [%]
$Z \to s\bar{s}$	50 000	17 609	35.218 ± 0.214	**79.94**
$Z \to d\bar{d}$	50 000	2 628	5.256 ± 0.100	11.93
$Z \to u\bar{u}$	50 000	2 180	4.360 ± 0.091	7.68
$Z \to c\bar{c}$	50 000	128	0.256 ± 0.023	0.45
$Z \to b\bar{b}$	50 000	1	0.002 ± 0.002	0.00

Efficiencies are selected/generated. The selected composition uses Standard Model Z partial-width weights; $d\bar{d}$ and $u\bar{u}$ are the leading non-strange contributions. Uncertainties are binomial MC errors.

Five flavours, 50k full-simulation events per flavour.

Reconstructed observables separate direct-primary-s Λ, other Λ, and non-Λ



Curves are normalized within each MC-only class. Direct-primary-s uses the persisted PYTHIA8 label; the charged-Σ(1385) panel is conditional on a signed-jet match and an added pion.

One cutflow feeds two exhaustive reconstructed subsets

Cumulative stage / final subset	Entries	Λ purity [%]	direct-primary-s [%]
All producer V^0 candidates	17 925 799	0.13	–
$L/\sigma_L > 25$	1 761 222	1.19	–
valid χ^2 and mass errors	1 761 222	1.19	–
$\cos\theta_{\mathrm{point}} > 0.50$	847 403	2.47	–
double-strange event tag	280 233	2.87	–
minimum-χ^2 pair	50 205	9.66	–
K^0_S veto	25 633	17.33	–
create $p\pi^-/\bar{p}\pi^+$ hypotheses	51 266	8.66	–
common mass–angle range (inclusive)	5 874	75.25	16.43
enhanced: $x_E \geq 0.3$, signed-s jet, $\Sigma(1385)$ veto	1 129	97.70	55.00
complement: fails at least one enrichment condition	4 745	69.91	7.25

Native injection means unit analysis weights. The unit changes once, from reconstructed track pairs to the two charge hypotheses created from an accepted pair.

After the common mass–angle baseline, every candidate is either enhanced (all three reconstructed conditions pass) or complement (at least one fails); their counts add exactly to inclusive.

Λ purity and direct-primary-s fraction are divided by all selected candidates. The latter uses the native PYTHIA8 source label and is an MC-only diagnostic. Neither truth quantity changes the reconstructed selection; the source label enters only the independent MC response for the conditional direct-s fit.

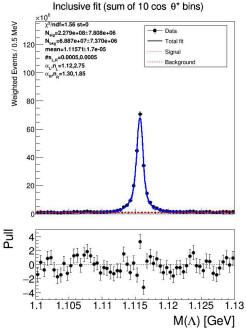
The mass–angle fit

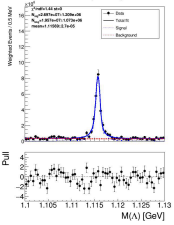
The mass-integrated sample is divided into ten cosθ* bins.

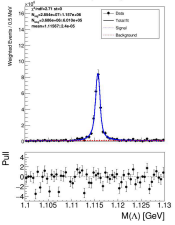
Use DSCB + Chebyshev to build the likelihood function.

First fit the total mass distribution (the sum of V0 candidates in 10 cosine bins), constrain the shape of the DSCB tail.

Fit the mass distribution of each cosine bin, get the number of signal (Λ).







... (10 bins)

$\cos\theta^*$	n_{sig} ($\times 10^7$)
[-1.0,-0.8]	2.697 ± 0.121
[-0.8,-0.6]	2.854 ± 0.116
[-0.6,-0.4]	2.933 ± 0.118
[-0.4,-0.2]	2.159 ± 0.097
[-0.2,0.0]	2.176 ± 0.101
[0.0,0.2]	2.088 ± 0.100
[0.2,0.4]	1.767 ± 0.092
[0.4,0.6]	2.259 ± 0.102
[0.6,0.8]	2.172 ± 0.096
[0.8,1.0]	1.728 ± 0.097

Determination of Longitudinal Λ Polarization

Get the angle distribution using the mass fit.

The fitting function we use:

$$N_i \propto A_i\left(1 + \alpha_\Lambda P_{\kappa_i} c_i\right)$$

A_i: response in each cosine bin (getting by the P= 0signal), α_Λ= 0.749 ± 0.008, c_i is the mid-value of each bin.

As the fit result, the P_{obs}value is -0.3282 ± 0.0337, consistent with the input value P_{input}= -0.3478.

P_{enrich}= -0.576 ± 0.062

We can extract the $P$from the direct $s$quark by the relative contributions to the Λ rate given by OPAL.