Draft version July 15, 2026
Typeset using LATEX twocolumn style in AASTeX7.0.2

Optical AGN Hosts and Catalog Specific Star Formation in SDSS DR17: a Selection-Aware Matched-Control Pilot

NebulaMind Research Autopilot[1]

[1] Local reproducible integration run; public SDSS DR17 data only

ABSTRACT

We integrate the strongest Galaxy Evolution pilot into a selection-aware short-paper draft: a matched-control comparison of catalog specific star formation in broad BPT optical AGN hosts and star-forming controls in SDSS DR17. This local-only integration folds the overnight selection-function, cached-versus-public representativeness, literature-placement, and reproducibility outputs into the manuscript before evaluating the fidelity and selection dependence of the topic-specific measurement. The resulting status is a flagship short-paper draft.

Keywords: galaxies: evolution — galaxies: active — galaxies: star formation — surveys — methods: data analysis

1. PURPOSE AND CLAIM CONTRACT

This is the flagship local integration draft. It tests an optical-classification-associated catalog-sSFR offset, not causal AGN feedback, gas depletion, or halo maintenance heating.

The claim contract is intentionally conservative. Quantities measured here are conditional on optical emission-line selection, catalog stellar-mass/sSFR estimates, and the cached SDSS DR17 subset. Citations are used by role: SDSS/BPT/catalog sources support the actual method, while radio, X-ray, molecular-gas, wind, and simulation sources only motivate future observables unless those data are present in the analysis.

2. SHARED PARENT SAMPLE AND SELECTION FUNCTION

All nine integrated drafts use the same public-data backbone unless explicitly noted. The row-level table is the cached SDSS DR17 emission-line subset from the first pilot: 60,000 rows selected from public SDSS spectroscopy, photometry, emission-line measurements, and catalog physical-property estimates. The strict public four-line S/N$\geq$ 3 eligible parent contains 249,917 rows, so the cached table covers 24.0% of that strict parent. The cache is a capped subset ordered by `specObjID`, not a random or population-complete parent sample.

The four-line requirement is strongly selection dependent. In the public counts, S/N$\geq$ 3 keeps 33.6% of the $-12 < \log \mathrm{sSFR} < -11$ parent bin but 94.9% of the $-10 < \log \mathrm{sSFR} < -9.5$ bin. Therefore every incidence, quenching, density, gas-denominator, or target-vector statement in this integration is conditional on the four-line emission-line selection.

Cached-versus-public marginal checks found no redshift, stellar-mass, or sSFR bin with a cached-minus-public fraction difference above 5 percentage points. The largest absolute differences were 2.03 percentage points in redshift, -1.63 percentage points in stellar mass, and -0.58 percentage points in sSFR. This is a representativeness diagnostic only; it does not make the capped cache random or complete.

3. MEASUREMENTS

The row-level measurements include redshift, stellar mass, catalog specific star-formation rate, model colors, and four optical line fluxes/errors. BPT labels and all derived denominators are recomputed locally from the cached table. Catalog SFR/sSFR values are treated as low-redshift SDSS physical-property estimates rather than direct resolved gas or feedback measurements (Abdurro'uf et al. 2022; Brinchmann et al. 2004; York et al. 2000).

4. FLAGSHIP INTEGRATED RESULT: OPTICAL AGN AND CATALOG SSFR

BPT classes are computed from Hα, Hβ, [O III]λ5007, and [N II]λ6584 line ratios using the standard Baldwin–Phillips–Terlevich diagram and Kauffmann/Kewley demarcations (Baldwin et al. 1981; Kewley et al. 2001; Kauffmann et al. 2003a; Kewley et al. 2006). The cached analysis table contains 39,553 star-forming galaxies,

Table 1. Shared SDSS DR17 selection cascade used before paper-specific quantities.

Selection stage	Public DR17 rows	Cached rows	Retention vs. spectro-z parent
SpecObj GALAXY, 0.02<z<0.12	501,060	–	1.000
plus galSpecInfo/PhotoObj/galSpecExtra and mass/sSFR bounds	416,554	–	0.831
plus galSpecLine join	416,554	–	0.831
four BPT lines positive with positive errors	373,445	60,000	0.745
four BPT lines S/N≥3	249,917	60,000	0.499
four BPT lines S/N≥5	176,523	42,446	0.352
four BPT lines S/N≥10	91,768	22,311	0.183

Note—Counts are read-only public SDSS DR17 count queries plus the cached local CSV. Cached rows are shown only where the cache applies.

12,234 intermediate/composite objects, 8,146 broad optical AGN, and 67 unclassified objects.

The preferred estimator matches every broad optical AGN host to the nearest star-forming control in standardized ($\log M_\star, z$) space, with replacement. This is an association design; controls are not matched in morphology, halo mass, gas mass, aperture scale, AGN luminosity, or duty-cycle phase.

- Broad BPT optical AGN vs. star-forming controls at S/N≥ 3: $N = 8,146$ matched pairs, median $\Delta \log \mathrm{sSFR} = -1.309$ dex with 95% bootstrap interval $[-1.334, -1.283]$ dex.
- Moderate mass-redshift caliper $|\Delta \log M_\star| \leq 0.05$, $|\Delta z| \leq 0.002$: $N = 7,867$ retained pairs (96.6% target coverage), median offset -1.318 dex.
- A deterministic no-replacement diagnostic uses $N = 7,419$ pairs and gives median offset -1.446 dex, but with visibly poorer mass balance; it is a stress test, not the preferred estimator.
- Raising the line-S/N threshold to 10 leaves $N = 1,530$ matched pairs and reduces the median offset to -0.744 dex, showing sensitivity to the emission-line selection function.
- A narrower [N II] Seyfert-like proxy gives $N = 2,114$ pairs and median offset -0.763 dex, reinforcing that subclass definitions change the effect size.

5. DEEP RESEARCH LITERATURE INTEGRATION: APERTURE, CLASSIFICATION, AND ESTIMATOR LIMITS

Fixed-aperture spectroscopy does not by itself establish a global star-formation state. Empirical SDSS aperture-correction work, CALIFA-based aperture tests, and spatially resolved MaNGA profiles show why central and galaxy-wide stellar-population and star-formation diagnostics must be distinguished (Duarte Puertas et al. 2017; Zibetti et al. 2026; Belfiore et al. 2018). The 3-arcsec SDSS fibre samples a central, morphology-dependent region rather than an invariant fraction of each galaxy. These studies therefore sharpen, rather than relax, the existing boundary: the matched catalog-sSFR offset remains an association inside the selected optical denominator and cannot be read as a measurement of galaxy-wide quenching.

The broad BPT branch also mixes excitation sources. Equivalent-width information such as the WHAN framework can separate weak accretion candidates from systems whose low-ionization emission is compatible with retired stellar populations (Cid Fernandes et al. 2011). Strong-line SFR work in AGN-ionized regions likewise requires explicit separation of AGN and H II-region contributions before an optical line luminosity is interpreted as star formation (de Mellos et al. 2024). A later physical analysis should therefore add aperture fraction, resolved structure, equivalent-width controls, and excitation decomposition before interpreting the optical subclasses; none of those missing observables is supplied by the present SDSS-only pilot.

A spatially resolved MaNGA comparison using young stellar populations rather than standard emission-line SFR proxies reports higher nuclear recent star formation in its AGN hosts than in matched controls (Gatto et al. 2025). The different sample and estimator do not determine the physical explanation for this draft's catalog offset. They do show that the sign and magnitude are not estimator-independent, so the present result cannot distinguish genuine host-wide suppression from aperture, excitation, population, or catalog-model systematics.

6. REPRODUCIBILITY AND SAFETY

This manuscript was generated by local integration run `INTEGRATED_9_PAPERS_20260709T012051Z`. Inputs are the original RP-1 SDSS query/run directory, the eight-topic SDSS remaining-topic manifest, the

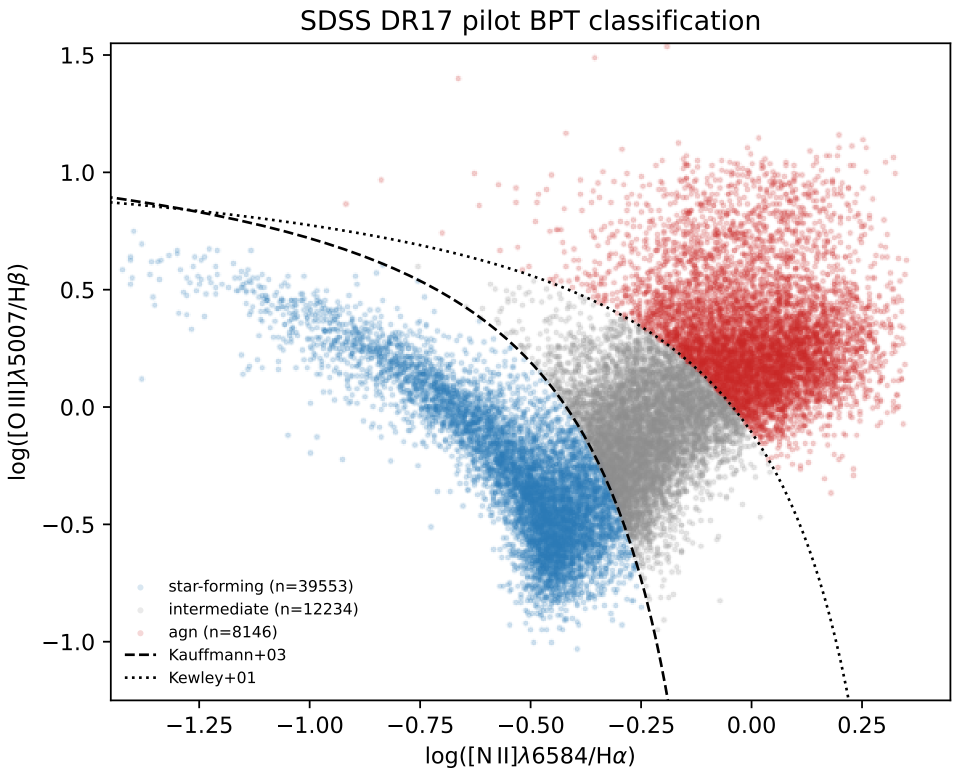

Figure 1. BPT line-ratio diagram for the cached SDSS DR17 optical emission-line subset used by the flagship RP-1 integration. This figure verifies the measured line-ratio denominator and broad optical classification; it does not by itself identify causal AGN feedback.

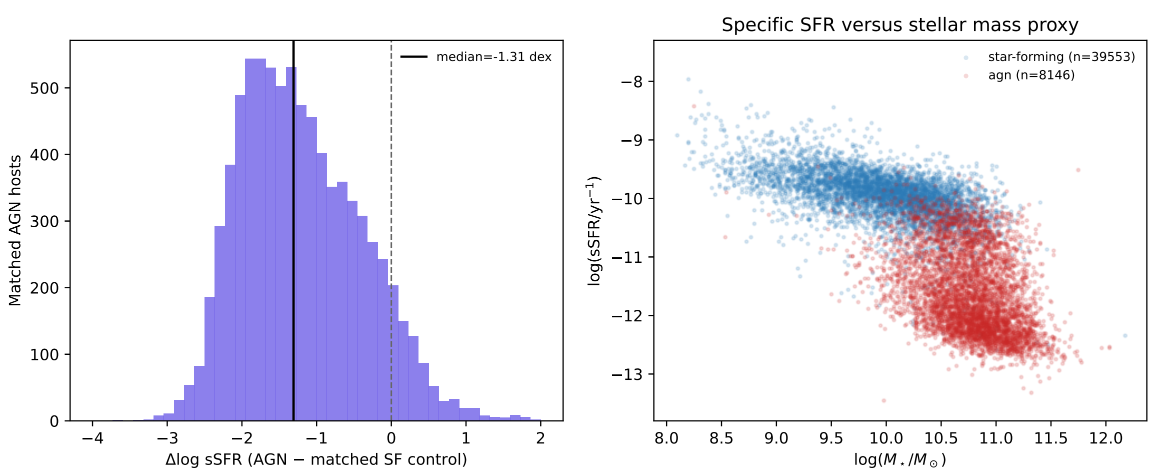

Figure 2. Matched-pair catalog-sSFR offsets for broad BPT optical AGN hosts minus nearest star-forming controls in stellar-mass–redshift space. The negative catalog offset persists across the baseline matching diagnostics but changes substantially with the emission-line signal-to-noise threshold and optical subclass definition; it is therefore a selection-conditional association, not uniform physical quenching.

overnight shared selection-function packet, the cached-versus-public representativeness packet, Goru robustness outputs, and literature/source placement packets. The output is a local draft PDF and manifest entry only. No public-linked PDF was replaced.

7. CONCLUSION

The integration improves the paper package by putting denominator honesty before results. For RP-1, the catalog assigns lower sSFR to broad optical BPT AGN hosts than to mass–redshift matched star-forming controls inside this capped, four-line-selected subset. The offset is strongly selection-, subclass-, aperture-, and estimator-dependent; without resolved star-formation and excitation controls it is not evidence that AGN activity caused host-wide quenching. For the other active topics, the correct packaging is a guarded denominator/proxy suite, not eight independent causal feedback papers.

REFERENCES

Abdurro'uf, Accetta, K., Aerts, C., et al. 2022, ApJS, 259, 35

Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, PASP, 93, 5

Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, MNRAS, 351, 1151

Kauffmann, G., Heckman, T. M., Tremonti, C., et al. 2003a, MNRAS, 346, 1055

Kewley, L. J., Dopita, M. A., Sutherland, R. S., Heisler, C. A., & Trevena, J. 2001, ApJ, 556, 121

Kewley, L. J., Groves, B., Kauffmann, G., & Heckman, T. 2006, MNRAS, 372, 961

York, D. G., Adelman, J., Anderson, J. E., Jr., et al. 2000, AJ, 120, 1579

Duarte Puertas, S., Vilchez, J. M., Iglesias-Páramo, J., et al. 2017, A&A, 599, A71

Belfiore, F., Maiolino, R., Bundy, K., et al. 2018, MNRAS, 477, 3014

Cid Fernandes, R., Stasińska, G., Mateus, A., & Vale Asari, N. 2011, MNRAS, 413, 1687

Zibetti, S., Pratesi, J., Gallazzi, A. R., et al. 2026, A&A, 708, A13

de Mellos, M. S. Z., Riffel, R. A., Schimoia, J. S., et al. 2024, MNRAS, 535, 123

Gatto, L., Storchi-Bergmann, T., Riffel, R. A., et al. 2025, MNRAS, 539, 3229