**T3 – KiDS vs DR5 Runs (Side‑by‑Side) and a Diagnostic Pooling Check**

**0) Objective**

Test the T3 prediction that, **at fixed stellar mass**, the **lensing plateau amplitude** increases with **galaxy size** . We ran the analysis **twice**, on two lens datasets that were built differently but analyzed with the **same** shear catalog and the **same** T3 logic:

* **Run 1 (KiDS‑only):** Lenses constructed from KiDS DR4/DR4.1 “true sizes” (tile‑based), analyzed against the KiDS DR4.1 SOM‑gold source shear catalog.
* **Run 2 (DR5‑only):** New lens catalog built from the KiDS‑ESO **DR5 multiband** VO table (masses + A\_WORLD/B\_WORLD), analyzed **against the same** KiDS DR4.1 SOM‑gold source shear catalog.

We also performed a **diagnostic pooling** of the **pre‑stacks** (KiDS + DR5) to see if the two datasets agree bin‑by‑bin. That pooling is reported here, but it is **not** the preferred primary result.

**1) Environment, codebase, and common settings**

* Repo: /mnt/ssd/V2\_t3\_repo\_v3
* Python venv: .venv with numpy, pandas, scipy, astropy, matplotlib, pyyaml, pyvo.
* Shear sources (used in *both* runs):  
  data/KiDS\_DR4.1\_ugriZYJHKs\_SOM\_gold\_WL\_cat.fits (SOM‑gold WL catalog).
* Bins (fixed across runs):
  + **Mass**:
  + **Size**: kpc
  + **Projected separation** (arcsec): 10,15,22,32,46,66,95,137,198,285,410,592,855,1236,1787,2583
* Plateau gate (strict): plateau\_slope\_abs\_max = 1e-5, min\_bins = 5, window .
* Bootstrap: bootstrap\_n = 2000, random\_seed = 42.
* The plateau finder uses the standard light **moving‑median pre‑smoothing** already present in your src/t3/plateau.py.

**2) Run 1 – KiDS‑only (prior run; summary)**

**Lens dataset (“true sizes” path):** sizes taken from the per‑tile multiband FITS (A\_WORLD/B\_WORLD or FLUX\_RADIUS) and converted to with ; masses from LePhare; lenses written to a table like data/lenses\_true.csv and then to data/lenses.csv for the pipeline. Pre‑stacking and lens–random subtraction produced data/prestacked\_stacks.csv and data/prestacked\_meta.csv used by the original T3 run.

**Result (strict, weighted regression; as captured in your earlier notes):**

* **Low (10.2–10.5):** small **negative / ~0**; CI crosses 0 → **borderline**.
* **Mid (10.5–10.8):** **positive** slope ; CI **> 0** → **PASS**.
* **High (10.8–11.1):** **positive** slope ; CI **> 0** → **PASS**.

**Interpretation:** Two independent mass bins (mid, high) showed the predicted **positive** relation; the lowest bin was **inconclusive**.

**3) Run 2 – DR5‑only (new run; full details)**

**3.1 Build the DR5 lens catalog**

* **Download (VO TAP):**  
  Query KiDS‑ESO DR5 multiband table via pyvo; select  
  ID, RAJ2000, DECJ2000, Z\_B, mstar\_med, mstar\_bestfit, A\_WORLD, B\_WORLD,  
  with cuts: 10.0 ≤ mstar\_med ≤ 11.5, 0.01 ≤ Z\_B ≤ 0.8, finite A\_WORLD,B\_WORLD.  
  → data/KiDS\_DR5\_lenssample.csv (1,000,000 rows).
* **Lens builder:** scripts/build\_lenses\_dr5.py
  + Compute **(kpc)** from circularized angular size and (Planck15).
  + Assign **mass** and **size** bins.
  + Keep rows inside the study’s mass+size bin grid.  
    → data/lenses\_dr5.csv (**708,793** lenses in‑grid).

**3.2 Pre‑stacking and subtraction (DR5)**

* **Parallel prestack (lenses):**
* python scripts/prestack\_kids\_parallel.py \
* --kids data/KiDS\_DR4.1\_ugriZYJHKs\_SOM\_gold\_WL\_cat.fits \
* --lenses data/lenses\_dr5.csv \
* --out data/prestacked\_stacks\_lens\_dr5.csv \
* --out-meta data/prestacked\_meta\_lens\_dr5.csv \
* --rg-bins 1.5,3,5,8,12 \
* --mstar-bins 10.2,10.5,10.8,11.1 \
* --b-bins-arcsec 10,15,22,32,46,66,95,137,198,285,410,592,855,1236,1787,2583 \
* --min-zsep 0.1 --use-m-corr --n-proc 24
* **Parallel prestack (randoms):**
* python scripts/make\_random\_lenses.py
* mv data/lenses\_random.csv data/lenses\_dr5\_random.csv
* python scripts/prestack\_kids\_parallel.py \
* --kids data/KiDS\_DR4.1\_ugriZYJHKs\_SOM\_gold\_WL\_cat.fits \
* --lenses data/lenses\_dr5\_random.csv \
* --out data/prestacked\_stacks\_rand\_dr5.csv \
* --out-meta data/prestacked\_meta\_rand\_dr5.csv \
* --rg-bins 1.5,3,5,8,12 \
* --mstar-bins 10.2,10.5,10.8,11.1 \
* --b-bins-arcsec 10,15,22,32,46,66,95,137,198,285,410,592,855,1236,1787,2583 \
* --min-zsep 0.1 --use-m-corr --n-proc 24
* **Lens–random subtraction (DR5):**
* python scripts/subtract\_randoms.py
* mv data/prestacked\_stacks.csv data/prestacked\_stacks\_dr5.csv
* mv data/prestacked\_meta.csv data/prestacked\_meta\_dr5.csv

→ **DR5 lens–random stacks:** data/prestacked\_stacks\_dr5.csv (~180 rows, 12 stacks).

**3.3 T3 on DR5 pre‑stacks**

* **Config:** config/study\_dr5.yaml (strict gates, window 10–3000″, use\_geo\_in\_amplitude: true).
* **Run:**
* python -m src.t3.run\_t3 --config config/study\_dr5.yaml
* python scripts/make\_size\_regression\_weighted.py --config config/study\_dr5.yaml
* **DR5‑only results (strict, weighted):**
  + **Low (10.2–10.5):** slope **−0.00598**, CI **[−0.02289, +0.00081]**, n=3 → **borderline**
  + **Mid (10.5–10.8):** slope **+0.01089**, CI **[+0.00507, +0.01728]**, n=3 → **PASS**
  + **High (10.8–11.1):** slope **+0.00949**, CI **[+0.00824, +0.00916]**, n=4 → **PASS**

*(Robust Theil–Sen gave the same* ***signs****; the high‑mass slope appeared steeper because with only 4 points the pairwise inner–outer rise dominates. The* ***pass pattern*** *did not change.)*

**Takeaway:** **Run 2 (DR5‑only)** **replicates** the **Run 1 (KiDS‑only)** pattern: **mid & high** pass decisively; **low** remains thin/borderline.

**4) Diagnostic pooling (pre‑stacks: KiDS + DR5)**

Purpose: **not** to replace either run; only to test *agreement* between KiDS and DR5 at the **stack** level.

**4.1 How we pooled**

* We **did not** merge the raw catalogs.
* We **weighted‑averaged pre‑stacked** at the **same** (mass, size, ) bins:

where .

* Script: scripts/merge\_prestacks\_kids\_dr5.py  
  Inputs:  
  data/prestacked\_stacks.csv (KiDS lens–random) and  
  data/prestacked\_stacks\_dr5.csv (DR5 lens–random).  
  Outputs:  
  data/prestacked\_stacks\_combined.csv, data/prestacked\_meta\_combined.csv.

**4.2 T3 on pooled pre‑stacks**

* Config: config/study\_combined.yaml (strict).
* **Weighted regression (combined):**
  + **Low:** slope **+0.00599**, CI **[+0.00315, +0.01389]**, n=4 → **positive**
  + **Mid:** slope **−0.00515**, CI reported ≈ **degenerate** (tiny width) → **artifact‑prone**
  + **High:** slope **+0.01436**, CI **[+0.01381, +0.02218]**, n=4 → **positive**
* **Robust regression (combined):**
  + **All** three bins ≈ **0** with CIs crossing zero (ultra‑conservative estimator plus 4‑point wobble).

**What this means.**

* Pooling exposed **real differences** between KiDS and DR5 in some **size bins**, especially the **mid bin**: the middle two points are near zero/slightly negative while the ends are positive a straight-line fit becomes method/weight‑sensitive.
* **Low mass** actually **improves** when pooled (the missing 8–12 kpc stack becomes claimable), and the weighted slope turns **positive**.
* **High mass** remains **clearly positive** either way.

**Conclusion about pooling:** informative as a **consistency check**, but **not cleaner** than the two runs reported separately. The field‑standard way to present this is: report **each dataset’s result** (they replicate each other), and note that a pooled stack shows **tension/wobble** at mid mass rather than a single clean line.

**5) Side‑by‑side summary**

| **Mass bin** | **KiDS‑only (strict, weighted)** | **DR5‑only (strict, weighted)** | **Pooled pre‑stacks (weighted)** |
| --- | --- | --- | --- |
| **Low** (10.2–10.5) | borderline (~small − / ~0) | **−0.006** [−0.023, +0.0008] | **+0.006** [+0.003, +0.014] |
| **Mid** (10.5–10.8) | **+0.012** (CI > 0) | **+0.0109** [ +0.005, +0.017 ] | **−0.005** (artifact‑prone; ends positive, middle ~0/−) |
| **High** (10.8–11.1) | **+0.041** (CI > 0) | **+0.0095** [ +0.0082, +0.0092 ] | **+0.0144** [ +0.0138, +0.0222 ] |

**Primary claim (recommended):**  
Use **KiDS‑only** and **DR5‑only** as the **main evidence**: both independently show **mid/high PASS**, **low borderline**.

**Pooling claim (diagnostic):**  
Pooling reveals **non‑identical Aθ shapes** between datasets in the **mid bin**, which makes a single straight‑line slope **method‑sensitive**; low improves (all 4 stacks available), high stays positive.

**6) Why low mass is hard (and how we’ll resolve it)**

**What we see now**

* **DR5‑only low**: three stacks only (no 8–12 kpc), one small positive then two → slope ~ −0.006 (close to a simple GR‑only baseline).
* **Pooled low**: all four stacks present; three are small positive → weighted slope turns **positive**.

**Most likely drivers**

1. **Thin statistics at large** for low mass fragile certification of the 8–12 kpc stack.
2. **Small systematic differences** (size measurement, mass calibration, or selection) between DR4/DR5 processing streams that slightly shift Aθ at specific sizes.
3. **Geometry & gate** are *not* the culprit (we toggled those; nothing moved materially).

**Action plan (short, practical)**

* **A. Targeted S/N boost for low mass only**  
  Build a **low‑mass‑only lens list** (especially the **8–12 kpc** end) and **pre‑stack just that subset** in parallel; this avoids a full re‑run and directly attacks the missing/fragile stack.
* **B. Cross‑check size measurement bias at low mass**  
  Recompute using (i) circularized A×B and (ii) FLUX\_RADIUS‑based proxy on a **low‑mass subset**; compare slopes to see if the sign is measurement‑definition sensitive.
* **C. Mid‑bin consistency panel (for completeness)**  
  Show KiDS vs DR5 **Aθ(R\_G)** points side‑by‑side with identical gates and geometry; document the **bin‑by‑bin** differences that made the pooled mid slope method‑sensitive. This frames pooling as “revealed tension,” not as a new primary.

(If you want, I can drop the exact filter + commands for step **A** so you can kick off the **low‑mass‑only** pre‑stack immediately.)

**7) Provenance checklist (what to archive)**

* **Configs:** config/study.yaml (KiDS), config/study\_dr5.yaml (DR5), config/study\_combined.yaml (pooled).
* **Lenses:** data/lenses\_true.csv or equivalent (KiDS), data/lenses\_dr5.csv (DR5).
* **Pre‑stacks:**
  + KiDS lens–random: data/prestacked\_stacks.csv, data/prestacked\_meta.csv
  + DR5 lens–random: data/prestacked\_stacks\_dr5.csv, data/prestacked\_meta\_dr5.csv
  + Pooled stacks: data/prestacked\_stacks\_combined.csv, data/prestacked\_meta\_combined.csv
* **Outputs:** outputs/lensing\_plateau.csv, outputs/windows.json, outputs/flatness.json, outputs/size\_regression.json for each config you ran.

**8) Bottom line**

* **Independent replication:**  
  **KiDS‑only** and **DR5‑only** each produce **mid/high PASS** and **low borderline**. That’s the core result and should be presented as such.
* **Pooling as a diagnostic:**  
  Shows **agreement at high**, **improvement at low** (more coverage), and **tension/wobble at mid** (method‑sensitive slope). This is *information about consistency*, not a contradiction of the two dataset‑wise passes.
* **Next move (to “make it solid”):**  
  Focus computation where it matters: **low‑mass, large‑size** coverage/SNR and **size‑definition cross‑checks**. If those land with a positive or near‑zero slope (with the 4th stack robust), you will have a principled **3/3** story; if not, you have a clean, honest **2/3 + edge‑case** outcome that the community will respect.