HPS 2021 MC Production: pass5_v9 and pass5_v9 Bugfix

Status and Comparison

M. Graham

October 20, 2026

Overview

- **pass5_v9** (nominal): original 2021 MC production, partially produced, located in `mcprod_2021/all_prod`.
- **pass5_v9 bugfix** (`mcprod_2021_v5.3.1-lowp-bugfix`): reprocessing campaign to fix a beam-spread bug and rerun recon with updated low-p Kalman tracking.
- Species common to both: tritrig, wab, rad, ap (A'), beam, simp (dark photon signal, 13+6 mass points).
- A third, distinct effort also appears below: a **v10-dev detector pilot** for tritrig/wab/rad/ap, serving as an early test run for the future **pass6_v10** production – not part of either pass5_v9 production proper.
- This talk: what was produced, where it stands, and known gaps/caveats.

The Bug and the Fix

- **Bug**: `stdhep_beam_coords` applied the beam-spread smearing sigma **twice**, making the simulated beam spot far too narrow relative to data.
- **Fix**: `mgignac-BS-width-bugfix` branch → corrected binary, invoked via a dedicated `hps-mc-env-bsfix.sh` environment.
- All new bugfix-production samples carry a `bsfix` tag and write to a separate output tree (`mc_v5.3.1-lowp-bugfix/`) so they never collide with the original (buggy) pass5_v9 output.
- Recon steering also updated: `pass6dev_lowp` low-momentum Kalman tracking (HitSmear/HitKill), vs. nominal pass5_v9's `pass5_physics` Kalman steering.
- Some species (tritrig/wab/rad/ap, non-bsfix tag) were **reprocessed with the new recon steering only**, reusing existing nominal pass5_v9 (v9-detector) generation/SLIC output – these rows are not new end-to-end samples.
- A separate, later effort reran gen→slic→recon for tritrig/wab/rad/ap on a **new v10-dev detector geometry**, as an early pilot for the future **pass6_v10** production – distinct from both the nominal and bsfix campaigns (own frame below).

Pipeline Stages

- **gen**: event generation (MadGraph5/EGS5/etc.) → raw stdhep.
- **gen_to_slic / prep_and_slic**: beam-coords smearing + `SLIC` detector simulation → SLCIO.
- **overlay_to_recon**: readout (optionally merged with pulser background), then reconstruction → recon SLCIO.
- **hpstr**: flat-tuple ntuple production (ROOT) from recon output.
- Merge factors (signal files combined per readout job): tritrig $\times 8$, wab $\times 20$, rad $\times 5$, ap $\times 5$, simp $\times 5$; beam has no merge (no pulser overlay).

Nominal pass5_v9 (all_prod) — Status (old-Kalman recon only)

Species (tag)	gen	slic	hpstr	Total events
tritrig_pulser, 14272	20000 (lhe)	7999	1000	∼1.77M
tritrig_pulser, 14596	8000 (lhe)	8000	1000	∼1.78M
wab_pulser, 14272	80000 (lhe)	80000	1000	∼0.11M
wab_pulser, 14596	39911 (lhe)	39929	4000	∼0.46M
rad_pulser, 14272	2000	2000	0*	–
rad_pulser, 14596	2000	2000	0*	–
beam, 14272	1000	1000	847	∼2.28M
ap_pulser, 14272 (13 masses)	1000/mass	1000/mass	0/mass*	–
ap_pulser, 14596 (4 masses)	965–1000/mass	matches gen	0/mass*	–
simp_pulser (19 masses)	1000/mass (lhe)	1000/mass	200/mass	∼0.55M/mass

gen column: lhe counts for tritrig/wab (retained on disk; stdhep cleaned up post-SLIC) and simp; stdhep counts for rad/ap/beam. * rad and ap never had the original (old-Kalman) recon/hpstr stage run in this nominal campaign on disk – only the newer low-p Kalman reprocess exists (see later frames); reported as 0/incomplete here rather than substituting those numbers. Total-events columns use sampled hpstr entries/file $\times$ hpstr file count; omitted where hpstr count is 0.

Nominal pass5_v9 — Key Gaps

- **rad and AP (A' signal) never got an old-Kalman recon/hpstr pass at all** in this nominal campaign – only the newer low-p Kalman reprocess exists on disk for these two species (see the "v9 reprocess" frame below). Treated as 0/incomplete for the nominal old-Kalman table rather than borrowing those numbers.
- **tritrig/wab/beam** have genuine old-Kalman recon/hpstr output and are at or near their original targets (beam ≈85% at 847/1000*).
- **SIMP is fully complete** (200/mass, all 19 masses) under the original recon steering.
- SIMP recon/hpstr output in nominal prod lives under a differently-named directory (`simp_pulser_new`) than readout (`simp_pulser`) – a naming quirk not present in the bugfix production, which uses one consistent path throughout.

* excludes a small legacy `smeared/merged` aggregate subtree.

Bugfix Production — SIMP Signal (dark photon)

- 13 nominal mass points (30–210 MeV, 15 MeV steps) + 6 extended masses (240–440 MeV).
- 1000 gen_to_slic jobs, 200 overlay_to_recon jobs per mass point; gen (lhe) pool shared with nominal.

Mass range	gen (lhe)	slic	hpstr	Total events
30–210 MeV (13 pts)	1000/mass	1000/mass	200/mass	∼0.46M/mass
280–440 MeV (5 pts)	1000/mass	1000/mass	200/mass	∼0.46M/mass
240 MeV	1000	1000	**1**	Stalled – needs rerun

Total: 2600/2600 jobs clean across the 13 core masses (confirmed via disk-based output audit, not sacct exit codes – SLURM silently reports `COMPLETED` even on fatal conditions-DB errors that leave zero-byte output). Total-events uses sampled hpstr entries/file (∼2313, simp30MeV) × 200.

Bugfix Production — Other Species (bsfix, new end-to-end)

Species (tag)	gen (stdhep)	slic	hpstr	Total events
tritrig_bsfix_14272_v10dev	1000	1000	125	∼0.23M
wab_bsfix_14272_v10dev	5000	5000	250	∼0.03M
rad_bsfix_14272_v10dev	1000	1000	200	∼1.06M
ap_bsfix_14272_v10dev (180 MeV)	880	880	180	∼2.71M
beam_bsfix_14272_v10dev	1000*	903	594	∼1.62M

All early-stage / partial – these are genuinely new (gen→slic→recon) bugfix samples on the v10-dev detector, well short of eventual target job counts (2500/4000/2000 for tritrig/wab/rad). SIMP is the only species fully finished in this production so far. * beam gen reuses the shared 1000-file raw EGS5 pool.

Bugfix Production — v9 Reprocess (true nominal gen/SLIC reuse, new recon only)

Species (tag)	gen	slic	hpstr	Total events
tritrig_14272 (v9)	20000 (lhe)†	7999†	1000	∼1.77M
tritrig_14596 (v9)	8000 (lhe)†	8000†	1000	∼1.78M
wab_14272 (v9)	80000 (lhe)†	80000†	4000	∼0.45M
wab_14596 (v9)	39911 (lhe)†	39929†	2000	∼0.23M
rad_14272 (v9)	2000†	2000†	400	∼2.12M
rad_14596 (v9)	2000†	2000†	400	∼2.12M
ap_14272 (v9, 13 masses)	1000/mass†	1000/mass†	200/mass	∼3.03M/mass
ap_14596 (v9, 4 masses)	965–1000/mass†	matches gen†	68–200/mass	varies

Genuine reuse of nominal pass5_v9 (v9-detector) gen/SLIC output; only readout+recon rerun with the new `pass6dev_lowp` steering. †gen/slic counts shared with the nominal pass5_v9 table (same physical files; not double-produced). This is distinct from both the bsfix species above and the v10-dev pilot below.

pass6_v10 Pilot (v10-dev detector, new end-to-end)

- Early test run on a **new v10-dev detector geometry**, using this production's `pass6dev_lowp` recon steering – a preview of the future **pass6_v10** production, not part of either pass5_v9 campaign.
- Originally filed under the nominal pass5_v9 species directories by mistake; genuinely new gen→slic→recon samples.

Species (tag)	gen (stdhep)	slic	hpstr	Total events
tritrig_14272 (v10-dev)	5000	5000	625	∼1.13M
wab_14272 (v10-dev)	10000	10000	500	∼0.06M
rad_14272 (v10-dev)	2403	2403	480	∼2.56M
ap_14272 (v10-dev, 120 MeV)	1000	1000	200	∼3.05M
ap_14272 (v10-dev, 160 MeV)	1000	1000	200	∼3.03M
ap_14272 (v10-dev, 180 MeV)	967	967	0	–

hpstr-stage output for tritrig/wab/rad lives under the bugfix tree (`mc_v5.3.1-lowp-bugfix/hpstr/...`) despite gen/slic being filed under the nominal `mc/gen/...` tree – a path split left over from the mislabeling. `ap_oldkalman_14272` (180 MeV) is a small 5-job timing-comparison stub on the v9 detector, not a physics sample – not included above.

Operational Notes / Lessons

- SLURM/sacct report `COMPLETED, ExitCode 0:0` even when a job's conditions-DB connection fails fatally mid-job – **disk-based output auditing is the only trustworthy completion signal** in this framework.
- Two distinct DB failure signatures seen during bugfix SIMP production: `Too many connections` (original, motivated a running-job throttle) and `CommunicationsException: Communications link failure` (transient, cleared up after the first few mass points).
- Failed jobs can leave a **zero-byte placeholder** at the final output path – since job filtering only checks file *existence*, not size, these are silently treated as "done" and never auto-resubmitted. Cleanup requires deleting zero-byte files before rerunning.
- SLURM preemption (`qos=preemptable`, no dedicated allocation on `milano`) is a separate, unrelated cause of missing jobs – also not auto-resubmitted.

Summary

- **Nominal pass5_v9** (old-Kalman recon): tritrig/wab/beam/simp have genuine original-steering output; **rad and AP never got an old-Kalman recon/hpstr pass at all** – only the newer lowp-Kalman reprocess exists for them.
- **v9 reprocess** (true pass5_v9 gen/SLIC reuse, new `pass6dev_lowp` recon only): essentially complete for tritrig/wab/rad/ap at their original nominal job counts.
- **Bugfix (bsfix) production**: SIMP signal (13 core mass points, 2600 jobs) is complete and clean; one extended mass (240 MeV) needs a rerun. All other bsfix species (tritrig/wab/rad/ap/beam) are early-stage/partial, well short of target job counts.
- **pass6_v10 pilot** (v10-dev detector): early, smaller-scale test run for tritrig/wab/rad/ap, run alongside the bugfix campaign – not a target-count production, just a preview of the next pass6_v10 effort.
- Next steps: continue bsfix gen→recon chain for tritrig/wab/rad/ap/beam to reach target job counts; rerun stalled simp 240 MeV; use the pass6_v10 pilot results to inform scope/scheduling of the actual future pass6_v10 production.