Simphony: GPU Optical Photon Monte Carlo for noble liquid TPC Simulations & ML Ready Optical Datasets

Validated against Geant4 in a 14.7 kt liquid-argon benchmark with two-stage wavelength shifting

G. Galgoczi, X. Ning, D. Smirnov, B. Viren, C. Zhang

Brookhaven National Laboratory, Upton, NY 11973, USA · Paper: arXiv:2606.05385 (v2) · Code: github.com/BNLNPPS/simphony · Contact: ggalgoczi1@bnl.gov

BROOKHAVEN
NATIONAL LABORATORY

Neutrino 2026 · UC Irvine · June 21–26, 2026

Code · github.com/BNLNPPS/simphony Paper · arXiv:2606.05385 (v2)

14.7 kt
LArTPC optical benchmark
60 × 13.5 × 13 m active volume

80.2 ± 1.6
M photons s^{-1} sustained
4 stacked e⁻ events, RTX 4090

1053 ± 55×
optical-transport speedup
vs single-thread Geant4

< 0.25%
Hit count disagreement with Geant4
e⁻, μ⁻, p source topologies

1 Why Simphony

Large noble-liquid TPCs make O(10 M) optical photons per GeV-scale interaction, full CPU optical MC is accurate but too costly for repeated detector scans and ML dataset generation.

Simphony is up to a 1000x faster on a consumer GPU than Geant4.

Light in LArTPCs enables:

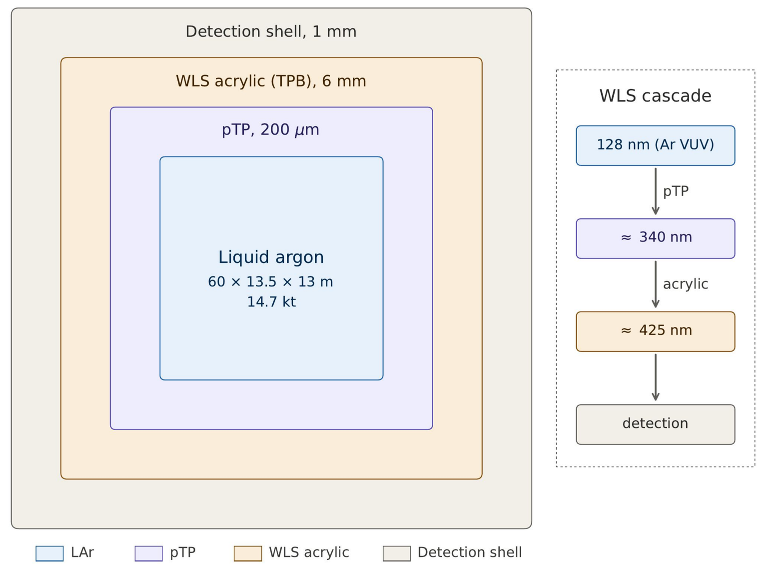
- **t_0 / triggering** — fast timing anchor
- **charge–light match** — event association
- **light calorimetry** — complementary energy
- **MeV physics** — solar & supernova
- **detector design** — coverage · WLS · surfaces

Easy install

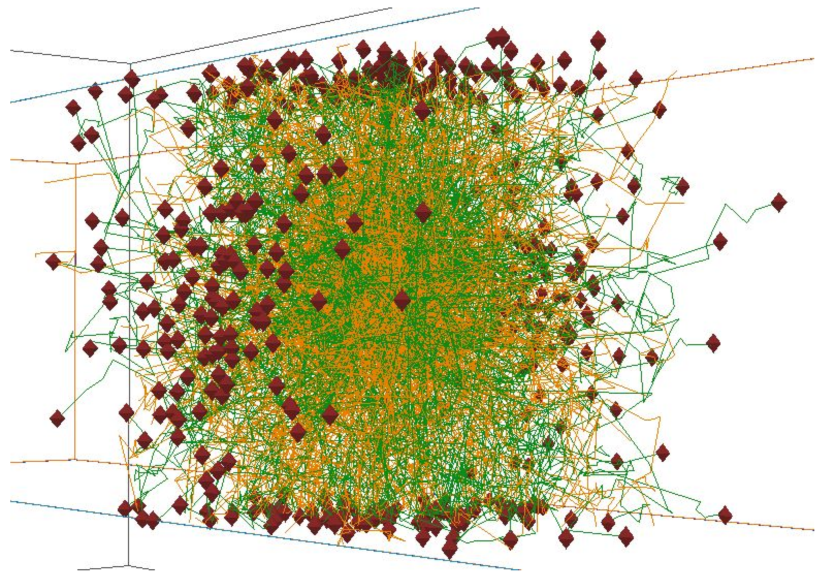
Simphony can be installed with a few spack commands. We also provide docker images for different CUDA capabilities that can be downloaded from our GitHub repository. We have tested Simphony on NERSC. Runs on any NVIDIA GPU Turing, 2018, or newer.

2 Method: benchmark & GPU offload

Controlled optical-transport benchmark — identical photon source for CPU and GPU.



14.7 kt LAr in nested pTP / WLS-acrylic / detector layers; 128→340→425 nm WLS cascade



GPU event display: detected on first entry to WLS (green), detected on second entry (orange), detection points (red)

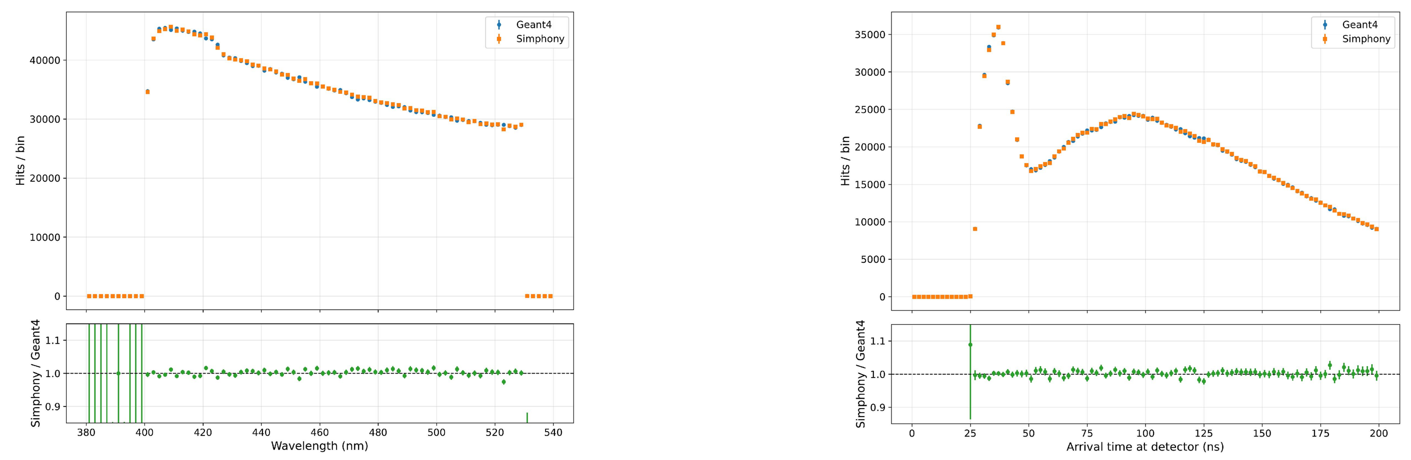
Geant4 — energy deposition → gensteps ➡ **Simphony GPU** — photon generation · transport + WLS ➡ **Detected hits** — compacted records + genstep truth

Identical genstep input isolates optical transport, WLS, and boundary handling from the upstream particle shower.

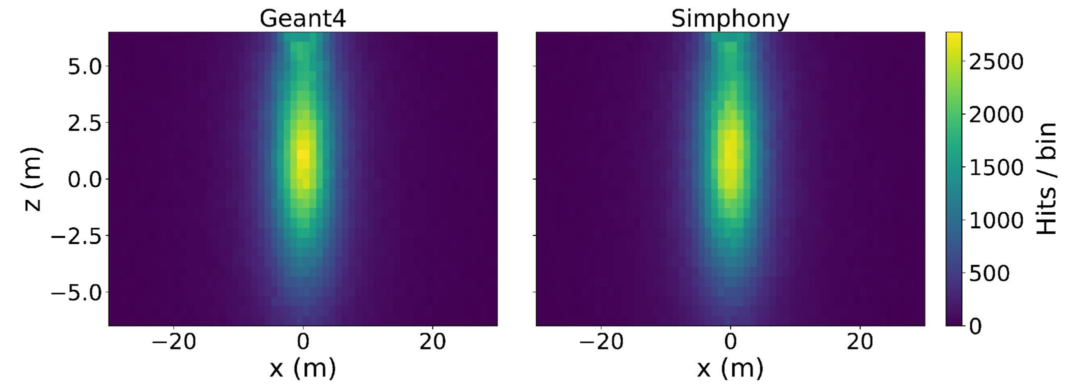
- **LAr scintillation** — 24 000 γ / MeV
- **Fast / slow** — 7 / 1400 ns
- **pTP WLS** — → 340 nm
- **TPB acrylic WLS** — → 425 nm

3 Validation against Geant4

Same photon distribution input → differences test optical transport, WLS, boundaries, and hit formation, not the upstream shower.



Detected wavelength and timing spectra agree with Geant4 under common genstep input.



Hit maps reproduce the spatial response on the +Y detector face (Geant4 | Simphony).

Primary	E kin	Geant4 hits (10^6)	Simphony hits (10^6)	R□ = S/G	χ^2/ndf t / λ
e⁻	2.5 GeV	2.378 ± 0.055	2.383 ± 0.055	**1.0022 ± 0.0002**	0.98 / 1.08
μ⁻	1.0 GeV	1.421 ± 0.034	1.423 ± 0.034	**1.0017 ± 0.0008**	1.01 / 0.90
p	400 MeV	0.259 ± 0.025	0.259 ± 0.024	**1.0005 ± 0.0014**	1.07 / 1.03

Sub-0.25% integrated agreement across three source topologies (mean ± sd, 3 paired seeds; χ^2/ndf on seed-42 events).

4 ML data product: dataset schema

gensteps (input)

```
x0, t0
direction
parent_track_id
n_photons
```

➡ **GPU photon transport**: generation, propagation, WLS × 2 stages, boundaries, detection ➡

hits (point cloud)

```
x, y, z, t
wavelength
direction,
polarization
boundary_flags
material_flags
process_mask
photon_index
```

➡ **labels (per hit)**

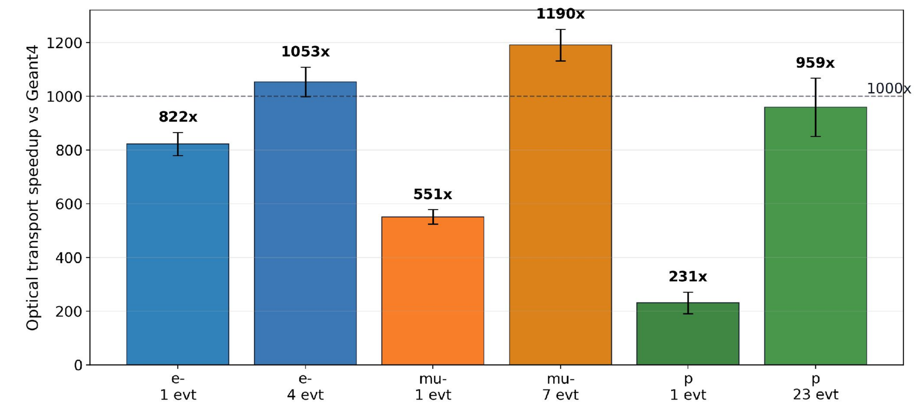
```
event_id
source_pos,
source_t
parent_track_id
→ ML sample
```

Detected hits can include the Monte Carlo truth. Simphony stores compact truth for detected photons rather than full optical histories, making the output directly usable for surrogate training, validation, and domain shift studies.

Truth lookup runs only for detected hits (~4%) — labeling cost amortized over all launched photons.

5 Performance

243 M photons in 3.03 ± 0.06 s on one NVIDIA RTX 4090 → 80.2 ± 1.6 M photons s^{-1}.



Event stacking amortizes fixed GPU launch overhead; all speedups vs single-thread Geant4.

Host — Geant4 event gen · setup · I/O — **136.3 ns / photon (~92% of wall)** ➡ **GPU** — OptiX propagation + compaction — **10.1 ns / photon** ➡ **Return** — compacted hit download — **1.2 ns / photon**

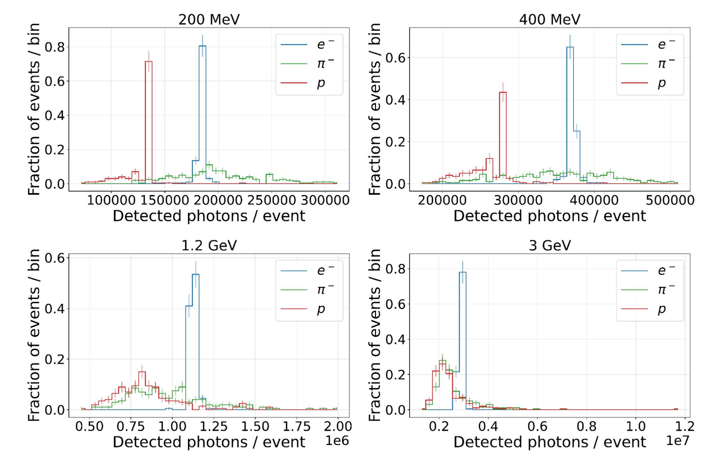
After GPU optical acceleration, host-side event generation dominates end-to-end wall time.

T_wall ≈ 0.31 s + 147 ns × N_photons — reproduces measurements to < 5% (61–243 M photons).

1053 ± 55× — optical-transport speedup · vs 1-thread Geant4

89 ± 5× — end-to-end speedup · vs 1-thread Geant4

6 ML benchmark tasks enabled by Simphony



Light-response scan: 3000 events (e⁻, π⁻, p × 5 energies × 200) with explicit transport in 3.83 h on one RTX 4090 (≈10 CPU-thread-days). Species-dependent structure — an optical-response benchmark dataset, not a detector-level PID claim.

A · Optical surrogate learning
gensteps + geometry → hit time / λ / position / count. Metrics: Poisson NLL, KS / Wasserstein on timing, hit-count ratio.

B · Detector-design response model
material / WLS / layout variations → response maps & timing PDFs — no CPU optical re-runs per design point.

C · Light-only toy PID / calorimetry
photon count, prompt fraction, timing quantiles → species / energy. Benchmark task only — idealized detector shell.

D · Domain-shift robustness
train on nominal optical model → test on shifted Rayleigh length, WLS spectrum / timing, reflectivity, PDE / coverage.

7 Validation & Plans

Validated now

- ✓ Common-genstep Geant4 parity
- ✓ WLS timing & wavelength spectra
- ✓ Hit maps: e⁻, μ⁻, p topologies
- ✓ Truth-labeled detected hits
- ✓ Single-GPU throughput (RTX 4090)

Next validation

- → Realistic photon detector modules
- → Experiment specific geometries
- → Calibrated recombination & quenching source
- → Calibration / test-beam / detector-data comparison

Code support today: automatic GDML import · ~12 CSG primitives + Booleans · GPU WLS matched to G4OpWLS · optical surfaces & material tables.

Talk to me about: optical ML benchmark dataset releases · surrogate-model targets & formats · geometry primitives your detector needs · experiment integration · multi-GPU scaling · multi-thread CPU baselines

Acknowledgments
U.S. DOE Office of Science, Office of High Energy Physics, Contract No. DE-SC0012704; BNL LDRD 26-055 and BNL LDRD 26-794.

Benchmark stack (reproducibility)
Geant4 11.3.2, single CPU thread (Intel Xeon w7-3445, 3.4 GHz) vs Simphony on Opticks · CUDA 13.0 · OptiX 9.0 · driver 580.105.08 · RTX 4090 (24 GiB).
Per-thread algorithmic reference, not a production-node comparison.

Selected references
[1] S. Blyth, EPJ Web Conf. 214 (2019) 02027. [2] G. Galgoczi et al., arXiv:2512.06061. [3] X. Ning et al., PRD 111 (2025) 032007.
[4] M. Lei et al., arXiv:2211.01505 (SIREN). [5] This work: arXiv:2606.05385 (v2) · github.com/BNLNPPS/simphony