

Tier 2 (Sidecar Integration) Checklist

1) DeFi → Residual Traces

- Build ``defi_mapper.emit_traces(state, oracle)`` per primitive.
- Decide features (swap deltas, HF/LTV margins, balances).
- Lock T (trace length) and preprocessing (`z_score`, `common_mode`).

2) WDD Parameters

- Tune warp, matched filter detect, and denoise configs.
- Set per-primitive gates (relative, absolute).
- Establish production defaults for EMA/median hybrid.

3) Null Calibration

- Implement per-primitive null threshold with permutations.
- Decide K (e.g., 40–120) and quantile q (e.g., 0.93–0.97).
- Serialize thresholds into artifacts.

4) Verifier Wiring

- Implement ``verify(plan, state, oracle, policy)``.
- Cover HF/LTV guards, limits, fees, oracle freshness.
- Return abstain reasons cleanly on fail.

5) Sidecar API/Contract

- Expose function/HTTP endpoint: ``plan(prompt, state, oracle, policy)``.
- Golden contract tests with schema lock.

6) (Optional) LLM Tap Priors

- Implement pooled latent probe → priors π per primitive.
- Reweight slopes after smoothing with small λ (0.1–0.2).
- Default OFF for MVP.

7) Data & Calibration

- Run calibration suite on historical state/oracle snapshots.
- Dump JSONL of {state, oracle, prompt, truth_plan}.

- Fit thresholds and denoise params.

8) Observability & Safeguards

- Log per request: z margins, phantom index, SNR, verify deltas.
- Classify abstains (conf vs exec) with reason strings.

9) Benchmark Harness

- Extend `defi_benchmark_consolidated.py` to real traces.
- Modes: Tier0, Tier2 core (WDD), optional priors.
- Metrics: exact, F1, hallucination, omission, abstains.

10) Rollout & KillSwitch

- Rails config flag to enable/disable WDD.
- Default planner hybrid/rails, priors OFF.
- Any verify fail → abstain with reason.

Definition of Done (Tier2)

- Deterministic mapper traces with tests.
- WDD tuned on real signals; null thresholds locked.
- Verifier enforces HF/LTV/limits; abstains with reasons.
- Sidecar API stable with golden tests.
- Benchmark run shows: hallucination \leq Tier0, omission ↓, exact ↑.
- Telemetry stable on shadow traffic (healthy z margins, low phantom index).