Stage 11 v6 Plan

Objective:
Stage 11 v6 aims to recover recall (≈ 0.90–1.00) while keeping hallucinations ≤ 0.26 and precision ≥ 0.80. The doctrine remains Warp → Detect → Denoise, but with adaptive, signal-processing inspired controls.

Key Changes in v6

1. 1) CFAR Candidate Gating: Adaptive per-primitive thresholding using noise ring statistics. Threshold $Zp \geq \mu + k\sigma$ (default k=2.2).
2. 2) Overlap-Adaptive Inhibition: Replace global λ with $\lambda pq = \lambda 0 \exp(-(\Delta t^2)/(2\sigma^2))$ to penalize only overlapping peaks strongly.
3. 3) Multi-Scale Matched Filter: Run prototypes at multiple widths (0.8W, 1.3W), take max z-score to catch spread/shifted wells.
4. 4) Cumulative Refinement (SPRT-style): Accept a primitive if cumulative residual drop across anneal $\geq \tau$ (default 0.02).
5. 5) Minimal Per-Primitive Gate: Gentle guard only for flip_v, others relaxed.
6. 6) Runtime Controls: Fast mode (nperm=150, samples=20) vs Full mode (nperm=500, samples=50, multi-scale on).

Default v6 Operating Point (Full Mode)

- CFAR: k=2.2, guard=20, bg=60
- Multi-scale widths: 0.8W, 1.3W (e.g. 130, 210 if W=160)
- $\lambda 0 = 0.25$, $\sigma = 1.6$
- Anneal: T0=2.6 → Tmin=0.7, steps=5, p_floor=0.12
- SPRT τ=0.02, single-step drop floor=0.006
- Null calibration: nperm=500, block frac=0.16

Expected Metric Shift

• Recall: +0.20–0.30 vs v5 (aim 0.90–1.0)
• Hallucination: remain ≤ 0.26 (protected by inhibition + CFAR + refinement)
• Precision: ≥ 0.80
• Margins: ≥ 0 on average
• Phantom Index: initially high but trending down as denoising improves

Example Runs

- **Full mode (recommended):**
  python stage11-well-benchmark-v6.py --samples 50 --nperm 500 --proto_width 160 --proto_widths 130,210 --cfar_k 2.2 --cfar_guard 20 --cfar_bg 60 --lambda0 0.25 --inhib_sigma 1.6 --T0 2.6 --Tmin 0.7 --anneal_steps 5 --p_floor 0.12 --sprt_tau 0.02 --drop_frac 0.006 --pi --dump_surfaces_dir dumps/v6_surfaces --dump_manifold dumps/v6_manifold.npz
- **Fast sanity check:**
  python stage11-well-benchmark-v6.py --mode fast --samples 20 --nperm 150 --proto_widths 160

--cfar_k 2.0 --sprt_tau 0.015

Success Criteria

- Recall ≥ 0.90
- Hallucination ≤ 0.26 (flip_v ≤ 0.24)
- Precision ≥ 0.80
- Margins μ ≥ 0
- Phantom Index trending downward
- F1 and Jaccard improving steadily

Stage 11 v6 is designed as the final recall-recovery pass: adaptive CFAR gating, overlap-sensitive inhibition, multi-scale detection, and cumulative refinement. Together, these controls balance sensitivity and specificity, yielding a clean single-well manifold with phantoms suppressed and recall preserved.