Flow-based Particle Filtering in the Stochastic Volatility SSM

Abstract

This report evaluates three deterministic flow-based particle filtering variants—Energy-Diffusion Homotopy (EDH), Local EDH (LEDH), and Kernel flow—against a bootstrap baseline on a stochastic volatility (SV) state-space model. We compare state estimation error (RMSE), effective sample size (ESS), and flow diagnostics (average flow magnitude and average Jacobian magnitude along pseudo-time). Experiments are conducted under three observation regimes: (1) standard observations (`obs_gap` $= 1$, $\beta = 0.65$), (2) sparse observations (`obs_gap` $= 5$, $\beta = 0.65$), and (3) strong observations (`obs_gap` $= 1$, $\beta = 1.0$). Overall, all flow methods track the baseline closely in RMSE, while providing modest ESS improvements in typical steps but failing to eliminate rare degeneracy events. Sparse observations require larger flows per pseudo-time step, whereas strong observations produce behavior nearly identical to the standard regime.

1 Model and Methods (brief)

We consider the stochastic volatility SSM used in the previous 2c report. The latent state x_t evolves nonlinearly with process noise, and observations y_t depend on x_t through a heavy-tailed/nonlinear likelihood controlled by a concentration parameter β. We run a particle filter with N particles and compare:

- **Baseline:** bootstrap PF with standard importance weights.
- **EDH:** global homotopy flow using a linearized transport field.
- **LEDH:** local (per-particle) linearization to generate a particle-specific flow.
- **Kernel:** RKHS-smoothed flow field with conservative updates.

For the flow methods, we integrate the transport ODE over L pseudo-time steps $\lambda \in [0, 1]$ (here $L = 8$), and track: (i) average flow magnitude $\mathbb{E}[\|\dot{x}\|]$ per pseudo-step and (ii) average Jacobian magnitude $\mathbb{E}[\|J\|]$ as a proxy for contraction/conditioning.

2 Experimental Cases

We keep the SV parameters and particle count the same as in the previous report, and vary only observation frequency and strength:

1. **Case 1 (standard):** `obs_gap` $= 1$, $\beta = 0.65$.
2. **Case 2 (sparse):** `obs_gap` $= 5$, $\beta = 0.65$.
3. **Case 3 (strong observations):** `obs_gap` $= 1$, $\beta = 1.0$.

3 Results

3.1 Case 1: Standard observations (obs_gap = 1, $\beta = 0.65$)

RMSE. Figure 1 shows the time-resolved RMSE for all four methods. The curves overlap heavily throughout the sequence, indicating that none of the deterministic flows yields a systematic accuracy gain over the bootstrap baseline in this SV setting. Small differences appear at individual times, but these are not persistent: each method alternates between slightly better or worse RMSE depending on the local observation realization, consistent with the strong stochasticity and nonlinearity of the SV dynamics.

ESS. Figure 2 indicates that ESS remains close to the full particle count ($\sim$ 590–600) for most times for all methods. The flow-based approaches (EDH/LEDH/Kernel) sit marginally above the baseline in typical steps, suggesting slightly better weight uniformity after assimilation. However, rare but severe ESS collapses occur—most prominently around $t \approx 235$—where all methods, including the flows, drop dramatically. The Kernel method is conservative on average but does not prevent these worst-case collapses, implying that degeneracy is dominated by the observation shock rather than the transport smoothness.

Flow diagnostics. Figure 3 summarizes pseudo-time behavior. Average flow magnitude decreases gently over λ; LEDH consistently applies the largest transport, EDH is intermediate, and Kernel is smallest. This ordering matches their design: LEDH adapts locally and can move particles more aggressively, while Kernel regularizes and thus moves more cautiously. The Jacobian magnitude is very close to 1 for all methods and increases slightly with λ, meaning the flows remain near volume-preserving with mild expansion as pseudo-time progresses. EDH and LEDH show slightly higher Jacobian norms than Kernel, consistent with their stronger transports.

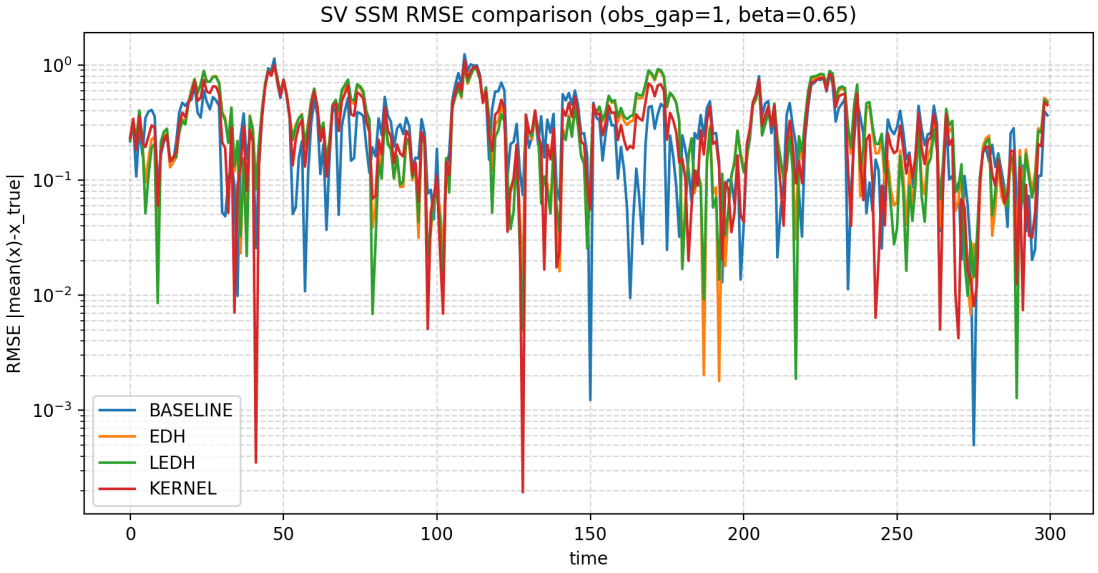


Figure 1: Case 1 RMSE over time. All methods track closely with no consistent winner.

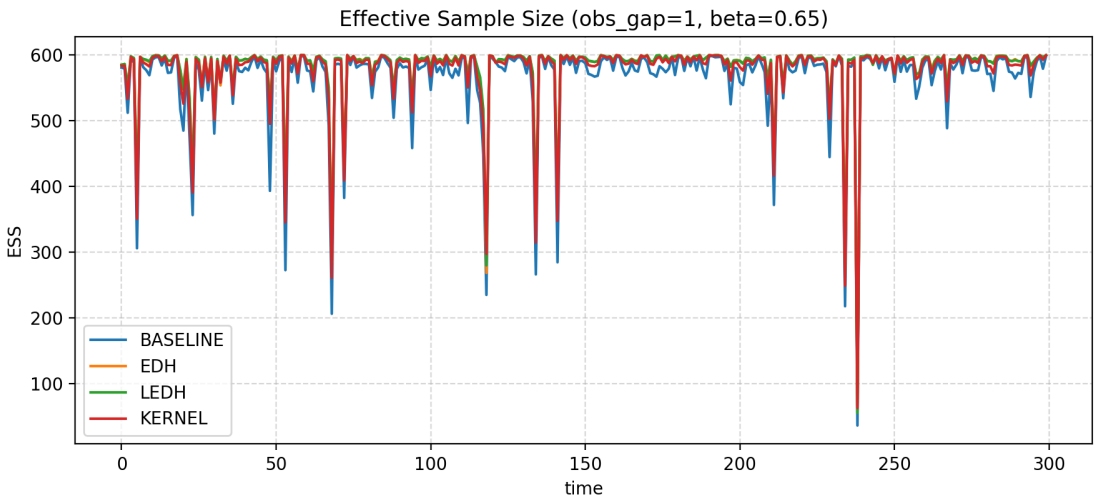
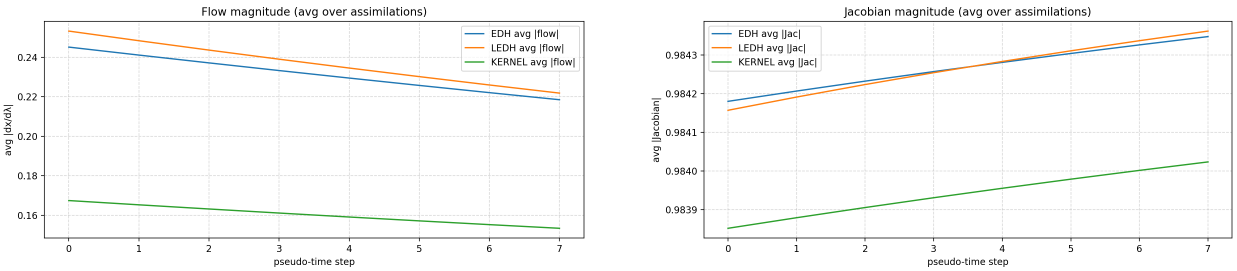


Figure 2: Case 1 ESS over time. Flow methods yield small typical ESS gains but do not eliminate rare degeneracy events.



(a) Average $\|\dot{x}\|$ vs. pseudo-time.

(b) Average $\|J\|$ vs. pseudo-time.

Figure 3: Case 1 flow diagnostics. LEDH uses the strongest transport, Kernel the weakest; Jacobians remain close to 1.

3.2 Case 2: Sparse observations ($\texttt{obs\_gap} = 5$, $\beta = 0.65$)

RMSE. With observations every 5 time steps, Figure 4 shows that RMSE trajectories remain extremely similar across methods, again with no stable ranking. Compared to Case 1, error spikes are slightly more structured between assimilation times (due to longer forecast windows), but the deterministic flows still do not change the overall accuracy envelope. This suggests that in the SV model, forecast error dominates over small differences in assimilation transport.

ESS. Figure 5 indicates ESS behavior nearly mirrors Case 1: high ESS most of the time, with occasional sharp collapses. The flow methods are typically a hair above baseline, but the large degeneracy event around $t \approx 240$ persists for all methods. Notably, sparsity does not substantially worsen ESS overall, implying that the particle set does not drift into extreme weight disparity during the longer forecast intervals, except when a highly informative observation arrives.

Flow diagnostics. Sparse observations require larger corrections when data arrive. Figure 6 confirms this: the average flow magnitudes are noticeably higher than in Case 1 throughout pseudo-

time (LEDH > EDH > Kernel). The magnitudes still decay smoothly with λ, indicating stable ODE integration. Jacobian norms remain near 1 with the same ordering as Case 1, meaning that despite larger corrective moves, conditioning is not significantly degraded.

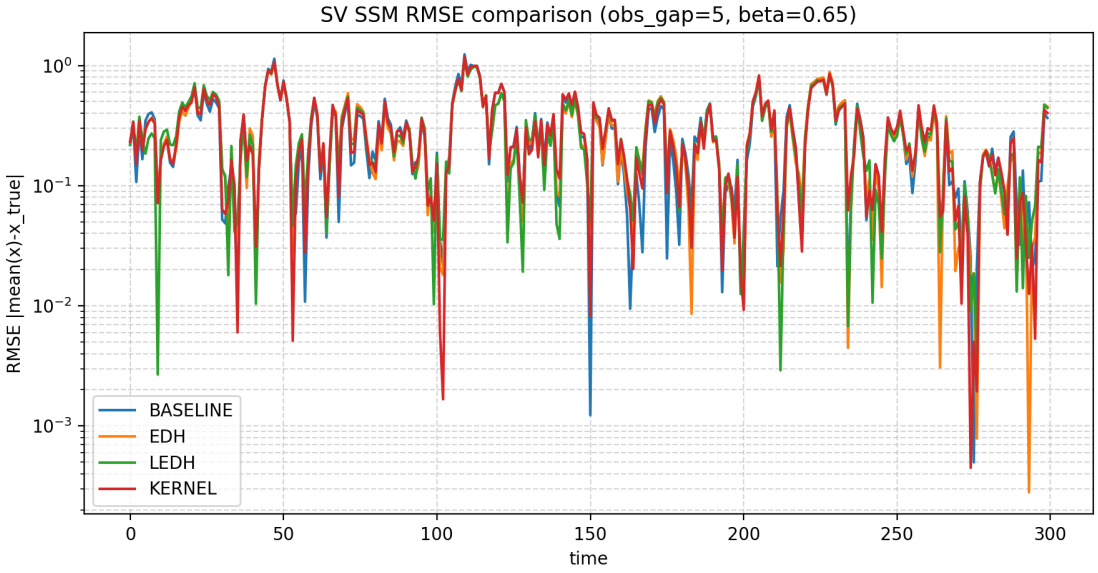


Figure 4: Case 2 RMSE over time with sparse observations (`obs_gap` $= 5$). Methods remain tightly clustered.

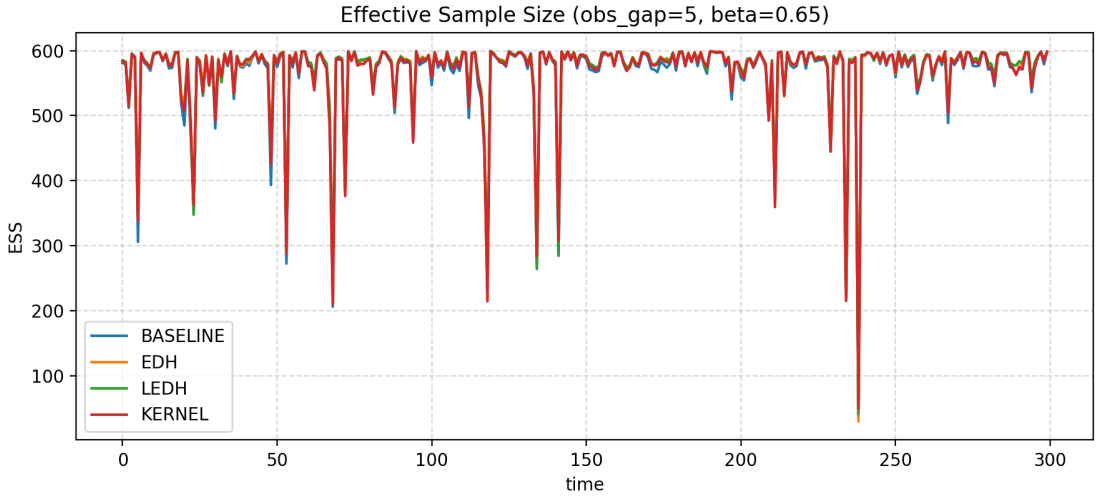
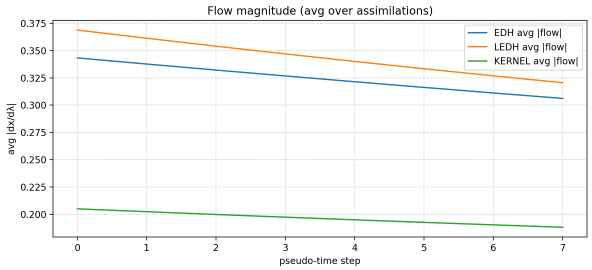
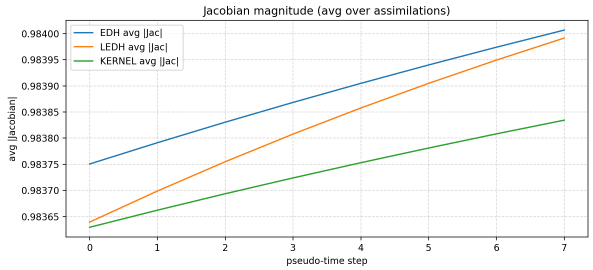


Figure 5: Case 2 ESS over time. Typical ESS is high; rare collapses remain.



(a) Average $\|\dot{x}\|$ vs. pseudo-time.



(b) Average $\|J\|$ vs. pseudo-time.

Figure 6: Case 2 flow diagnostics. Sparse observations induce larger transports, but Jacobians stay close to 1.

3.3 Case 3: Strong observations (`obs_gap` $= 1$, $\beta = 1.0$)

RMSE. Despite a stronger likelihood ($\beta = 1.0$), Figure 7 is virtually indistinguishable from Case 1. RMSE curves again overlap without a consistent ordering. This indicates that for the tested SV parameters and particle count, scaling observation strength within this range does not materially change the posterior transport difficulty.

ESS. Figure 8 likewise matches Case 1. ESS is near full size for most steps, and the same rare collapse around $t \approx 235$ appears for all approaches. Stronger observations therefore do not exacerbate degeneracy in a systematic way here, suggesting that the dominating collapse is tied to specific outlier observations rather than global likelihood sharpness.

Flow diagnostics. Figures 9 show almost the same flow and Jacobian magnitudes as Case 1. Average transport strengths keep the same ordering (LEDH > EDH > Kernel), and Jacobian norms remain stably close to 1. Practically, this means the pseudo-time integration regime is robust to this increase in β.

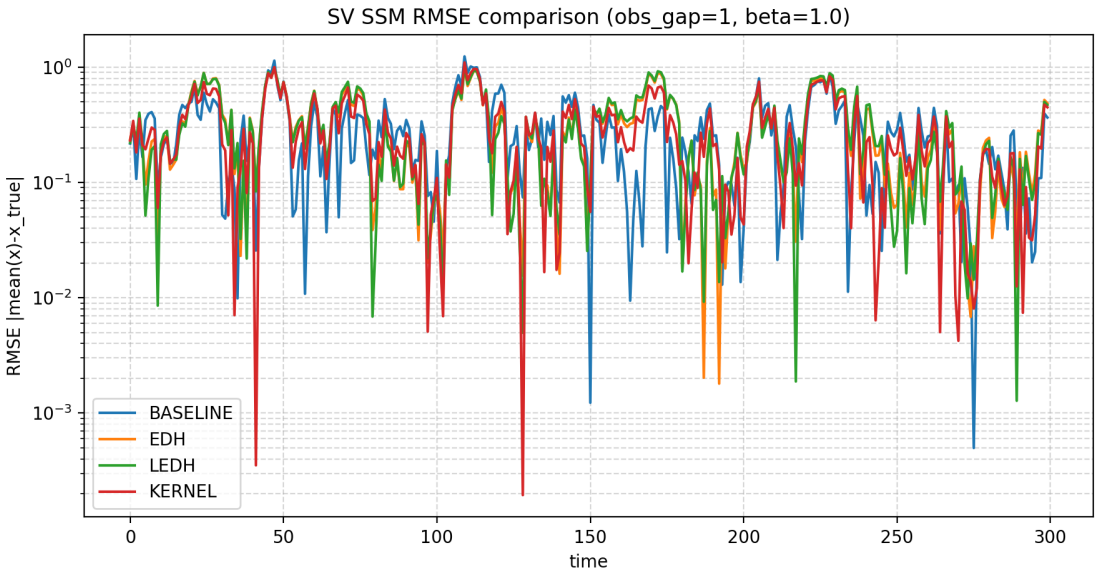


Figure 7: Case 3 RMSE over time with stronger observations ($\beta = 1.0$). Behavior is nearly identical to Case 1.

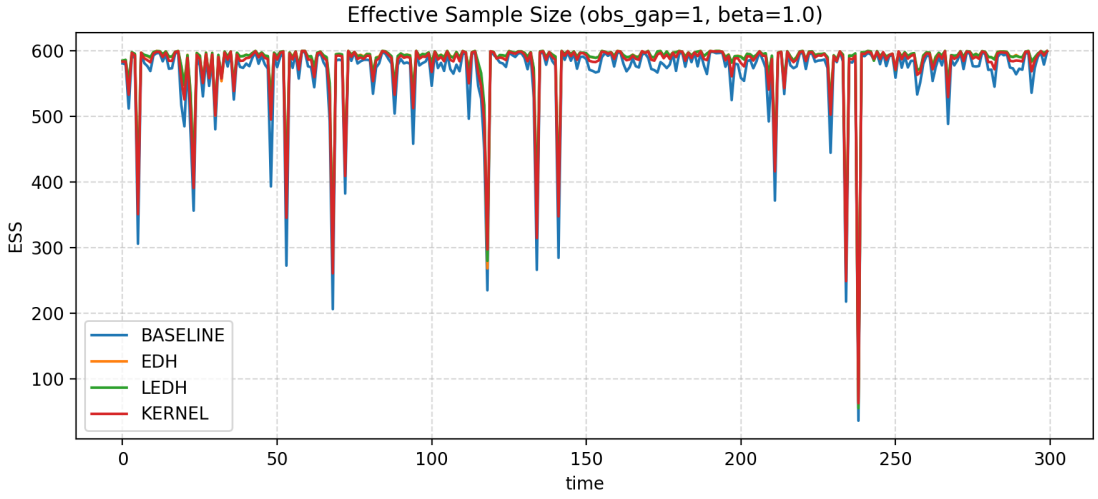
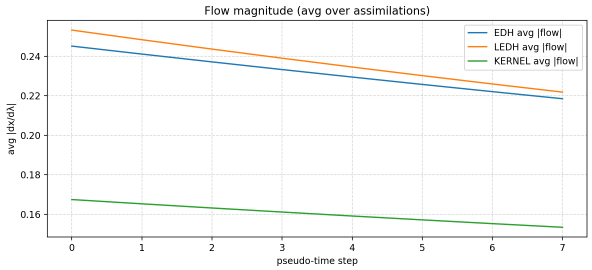
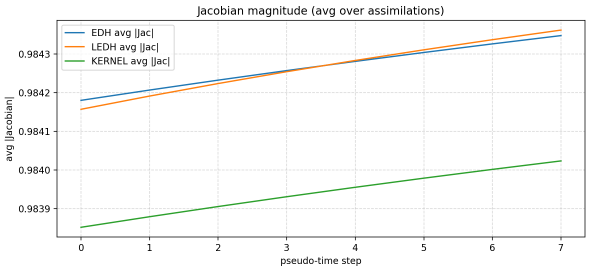


Figure 8: Case 3 ESS over time. Typical ESS is high; the same rare collapse persists.



(a) Average $\|\dot{x}\|$ vs. pseudo-time.



(b) Average $\|J\|$ vs. pseudo-time.

Figure 9: Case 3 flow diagnostics. Flow magnitudes and Jacobians coincide with Case 1.

4 Cross-case Discussion

Across all three regimes, two themes emerge:

- **Accuracy parity.** RMSE traces for EDH, LEDH, Kernel, and baseline are nearly identical in every case. Any advantage from deterministic transport is too small to separate from SV stochasticity at the present particle count.
- **Typical vs. worst-case ESS.** Flow methods slightly improve ESS in typical assimilations (closer to N), but they do *not* prevent rare catastrophic collapses. These failures occur at the same times across methods, indicating that they are driven by extreme observation mismatch rather than by deficiencies in transport linearization.
- **Effect of observation spacing.** Sparse observations (Case 2) increase required transport per assimilation, reflected in higher average flow magnitudes, while leaving Jacobians well-conditioned and RMSE unchanged.
- **Effect of observation strength.** Increasing β to 1.0 (Case 3) has negligible impact on RMSE, ESS, and flow diagnostics for this model and parameter range.

5 Conclusion

In the SV SSM, EDH, LEDH, and Kernel flows behave robustly and yield slightly higher ESS during ordinary steps, but do not provide measurable RMSE gains over a bootstrap PF and cannot eliminate rare degeneracy events. Observation sparsity induces larger but still stable transports, whereas strengthening the likelihood within the tested range leaves behavior essentially unchanged. Future improvements likely require either stronger nonlinearity-aware transport (beyond local linearization) or adaptive resampling/tempering strategies targeted at the rare collapse times.