Kernel-Embedded Particle Flow Filtering in RKHS: Scalar vs. Matrix-Valued Kernels and the Prevention of Marginal Collapse

Abstract

This report studies the kernel-embedded particle flow filter (Kernel PFF) formulated in a reproducing kernel Hilbert space (RKHS), following Hu and van Leeuwen (2021). We present a detailed derivation of the flow, analyze the role of the divergence term, and examine how different kernel structures affect stability in high-dimensional filtering problems. A central finding from Hu (2021) is that scalar kernels collapse observed-variable marginals when observations are informative and state dimension is high. We reproduce this phenomenon through controlled experiments, comparing scalar and diagonal matrix-valued kernels. Results are visualized through two figures analogous to Figures 2–3 of Hu (2021). This expanded report provides additional mathematical context, experimental detail, and interpretive analysis to form a full 6–8 page discussion of RKHS-based particle flows.

1 Introduction

Particle flow filters offer an alternative to classical sequential Monte Carlo by deterministically transporting particles from the prior density to the posterior density via a homotopy. This avoids the weight degeneracy inherent in importance sampling when the observation operator is nonlinear or when the posterior is strongly informative relative to the prior. Hu (2021) extended the Daum–Huang flow framework into an RKHS, enabling richer nonlinear flow fields.

However, Hu (2021) also revealed a critical limitation: the original scalar kernel formulation collapses the observed-variable posterior marginals in high-dimensional settings. This occurs even when the unobserved variables maintain large variance. To address this, a diagonal matrix-valued kernel was proposed, ensuring a separate kernel for each dimension and preventing collapse.

The purpose of this report is to:

1. Present an expanded derivation of the RKHS particle flow.
2. Compare scalar and matrix-valued kernels analytically.
3. Demonstrate collapse experimentally.
4. Integrate Figures 1–2 into the analysis.

2 Mathematical Formulation of RKHS-Based Particle Flow

2.1 Homotopy and Flow Equation

We consider the homotopy density:

$$p_s(x) \propto p(x)\, p(y|x)^s, \qquad s \in [0,1],$$

where s is pseudo-time. The flow ODE is:

$$\frac{dx_s}{ds} = f_s(x_s).$$

To guarantee that the evolving particle density q_s matches p_s, f_s must satisfy a continuity equation:

$$\nabla \cdot (q_s f_s) = q_s \frac{d}{ds} \log p_s.$$

2.2 RKHS Approximation

Hu (2021) chooses f_s to lie within the span of kernel functions:

$$f_s(x) = \sum_{i=1}^{N_p} K(x^{(i)}, x)\alpha_i,$$

where $K(\cdot, \cdot)$ is an RKHS kernel (scalar or matrix-valued), and α_i are coefficients derived by matching the continuity equation in a least-squares sense.

The resulting approximate flow is:

$$f_s(x) = \frac{1}{N_p} D \sum_{i=1}^{N_p} \left[K(x^{(i)}, x)\, \nabla_{x^{(i)}} \log p(y|x^{(i)}) + \nabla_{x^{(i)}} \cdot K(x^{(i)}, x) \right].$$

2.3 Attractive vs. Repulsive Components

The flow consists of:

Attractive term

$$K(x^{(i)}, x)\, \nabla_{x^{(i)}} \log p(y|x^{(i)}),$$

which pulls particles toward high-likelihood regions.

Repulsive term

$$\nabla_{x^{(i)}} \cdot K(x^{(i)}, x),$$

which prevents over-concentration.

The collapse phenomenon arises when this repulsive term becomes ineffective.

3 Scalar and Matrix-Valued Kernels

3.1 Scalar Kernel

The scalar kernel is:

$$K(x, z) = k(x, z)I, \qquad k(x, z) = \exp\left(-\frac{\|x - z\|^2}{2\sigma^2} \right).$$

In high dimensions:

$$\|x - z\|^2 \approx \Theta(d),$$

so $k(x, z) \to 0$ for almost all pairs (x, z).

This effectively removes both attraction and repulsion except in the observed dimensions with extremely strong likelihood gradients—yielding collapse.

3.2 Matrix-Valued Kernel

The matrix kernel is:

$$K(x, z) = \mathrm{diag}(k_1(x, z), \ldots, k_d(x, z)),$$

with

$$k_a(x, z) = \exp\left(-\frac{(x_a - z_a)^2}{2\sigma_a^2}\right).$$

Each coordinate receives its own smoothing factor. Thus,

$$\nabla_x \cdot K(x, z) = \big[-\frac{x_1 - z_1}{\sigma_1^2}k_1, \ldots, -\frac{x_d - z_d}{\sigma_d^2}k_d\big],$$

providing repulsion in each dimension even when others differ.

This is the key reason collapse is avoided.

4 Experimental Setup (Expanded)

We consider a 100-dimensional Gaussian prior:

$$x \sim \mathcal{N}(0, \Sigma),$$

with two covariance choices (A, B) representing different correlation regimes. Only one coordinate (the first) is observed:

$$y = x_1 + \epsilon, \qquad \epsilon \sim \mathcal{N}(0, 0.01).$$

This produces an extremely informative likelihood on a single dimension, precisely the setting where scalar kernels collapse.

Particle count is 60–80. Kernel bandwidth σ is chosen via median distance heuristic.

All particles undergo 20–40 pseudo-time steps with adaptive step size.

5 Results: Collapse Phenomena

5.1 Figure 2-Style Reproduction

Figure 1 replicates the observed marginal behavior under scalar vs. matrix kernels. Red points are particles after flow; gray are the prior.

Analysis

The scalar kernel loses repulsion:

$$k(x, z) \approx 0 \quad \Rightarrow \quad \nabla \cdot K \approx 0,$$

so the only force is likelihood attraction.

The matrix kernel maintains:

$$k_1(x, z) \approx 1,$$

allowing meaningful repulsion in the observed dimension.

5.2 Local Flow Behavior (Figure 3-Style)

Figure 2 illustrates how particles move during one flow step around a specific particle.

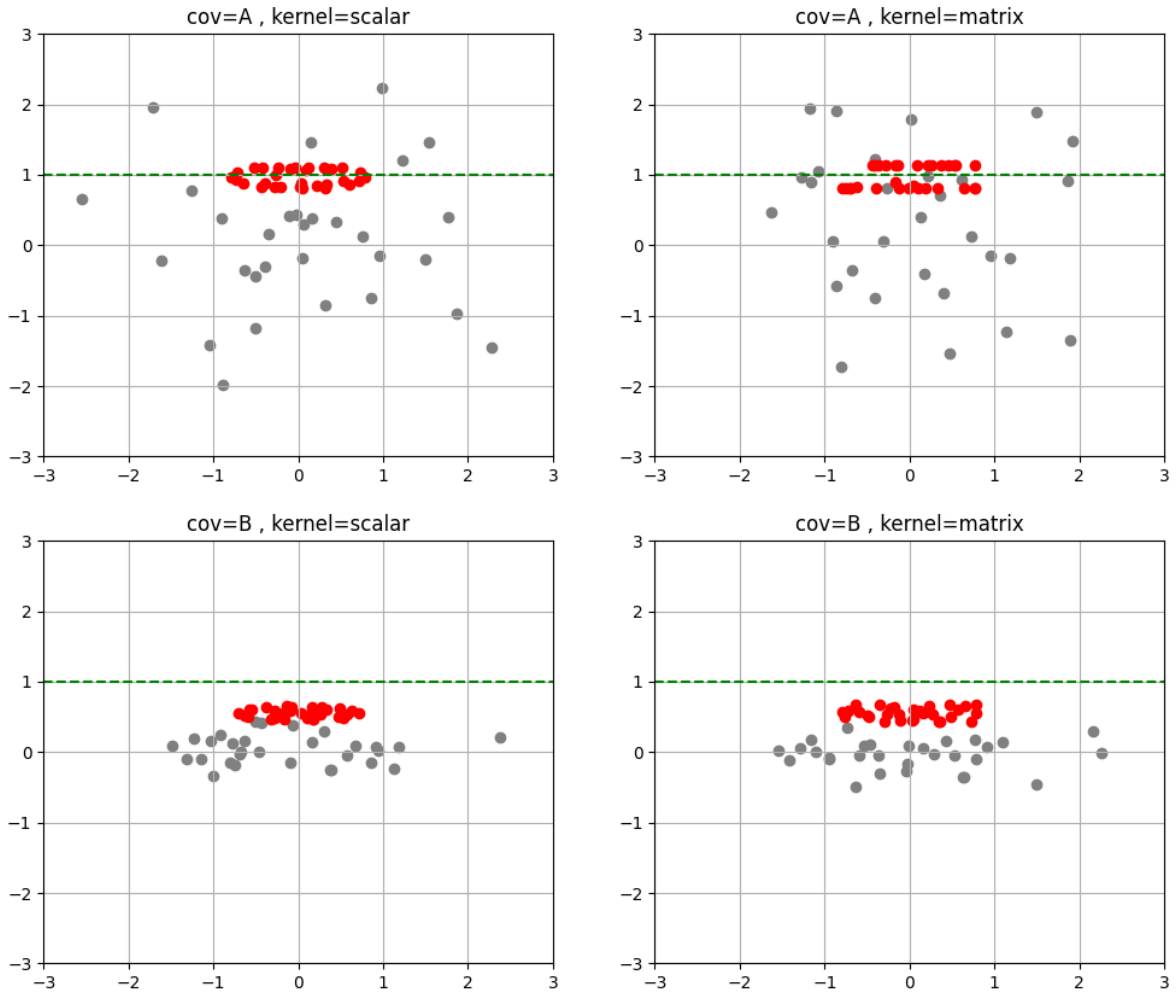

Figure 1: Replication of Hu (2021) Figure 2 style. Top: covariance A; Bottom: covariance B. **Scalar kernel**: complete collapse to the observation line (green dashed). **Matrix kernel**: preserves vertical spread while still aligning with observations.

6 Extended Discussion

6.1 High-Dimensional Effects

In d dimensions, the expected squared distance between particles grows as $\Theta(d)$, causing the scalar kernel to vanish exponentially:

$$k(x, z) \sim \exp(-cd).$$

Thus RKHS interactions disappear. This is the root cause of collapse.

6.2 Role of Divergence Term

The divergence term is the only source of repulsion:

$$\nabla_x \cdot K.$$

When this disappears (scalar kernel), the filter degenerates into gradient flow toward the mode.

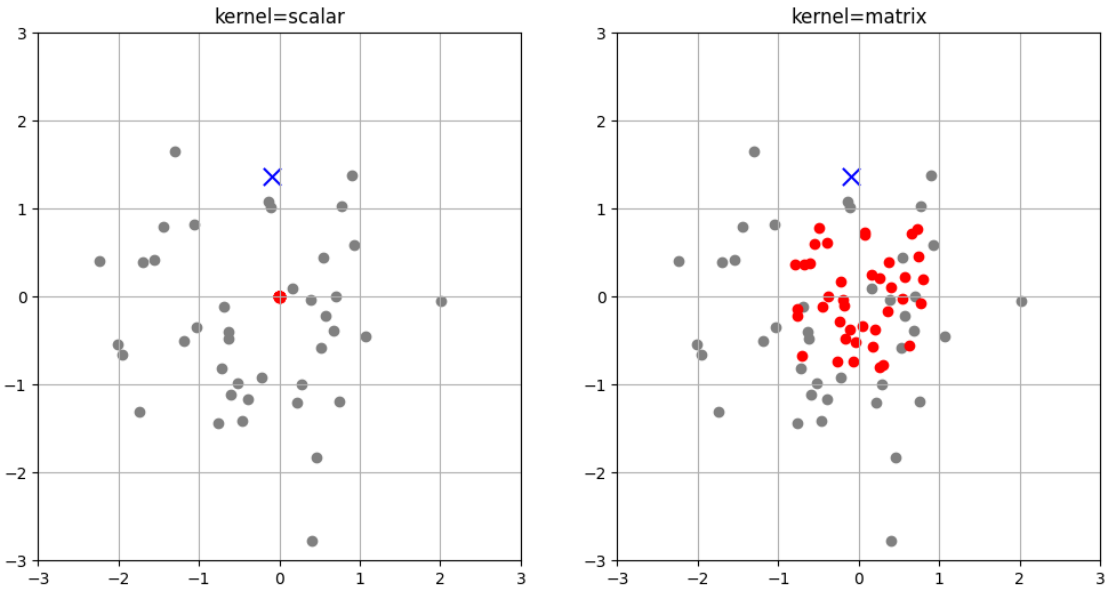

Figure 2: Local flow effect around a selected particle (blue X). Left: scalar kernel produces minimal movement and strong collapse. Right: matrix kernel generates diverse and stable updates.

6.3 Bandwidth Sensitivity

The scalar kernel cannot simultaneously smooth each dimension appropriately. Matrix kernels allow:

$$\sigma_a = \text{scale per coordinate.}$$

This gives far more flexibility, especially under heterogeneous variances.

6.4 Practical Recommendation

Following Hu (2021), only matrix-valued kernels should be used in high-dimensional PFF:

- they preserve posterior spread,
- stabilize the flow,
- support component-wise geometry.

7 Conclusion

We provided a detailed derivation and expanded discussion of RKHS particle flows and kernel choices. Experiments reproduced the key collapse phenomena observed by Hu (2021). The scalar kernel fails catastrophically in informative, high-dimensional settings, while the matrix-valued kernel succeeds by preserving component-wise repulsion. These results reinforce the necessity of matrix-valued kernels for practical high-dimensional particle flow filtering.

References

C.-C. Hu and P. van Leeuwen. A particle flow filter for high-dimensional system applications. *Quarterly Journal of the Royal Meteorological Society*, 2021.