

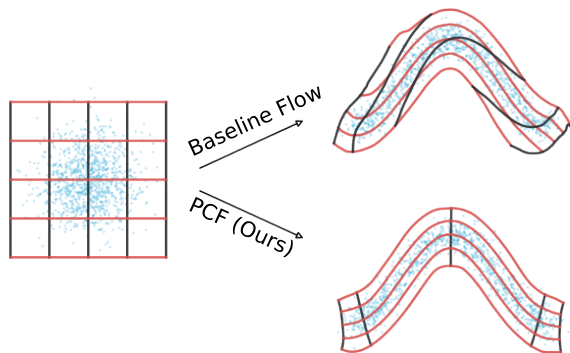
Principal Component Flows

Eddie Cunningham¹, Adam Cobb², Susmit Jha²

¹UMass Amherst

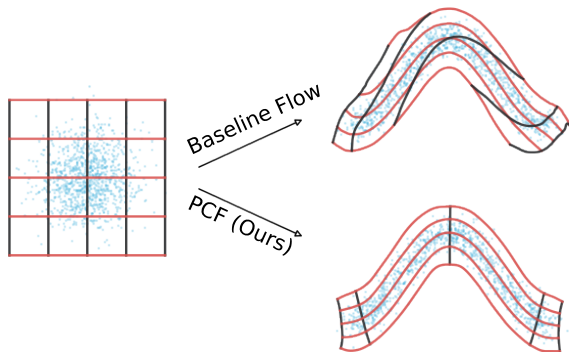
²SRI International

Contributions

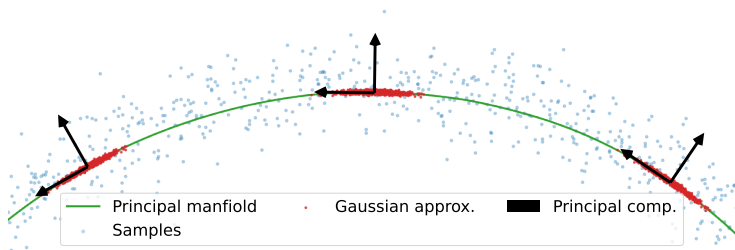


- 1 Novel class of flows, principal component flows (PCF), that can generate their principal manifolds.
- 2 Novel lower bound on injective change of variables formula.
- 3 Perform density estimation on variable dimensional manifolds at test time.

Motivation

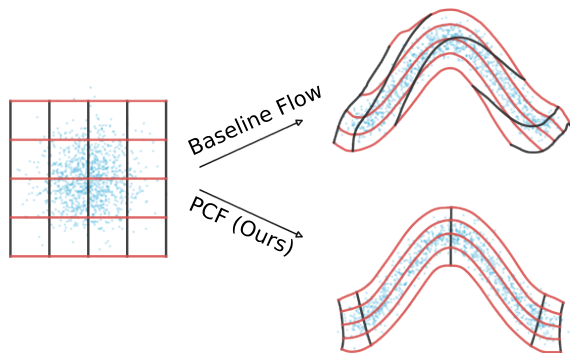


Principal components



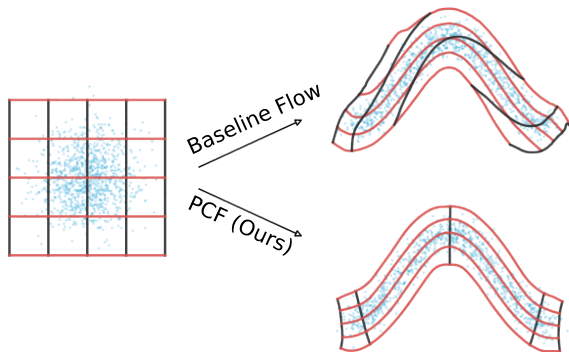
- Defined locally by passing noise around latent variable to data space and performing PCA on resulting distribution.
- Integrating over principal components yields principal manifolds.

Contours of a flow



- Use contours that emerge from fixing one latent variable and varying the rest
- Each contour has a valid probability distribution
- Want contours to capture structure

Principal component flows



Pointwise mutual information

- Pointwise mutual information:

$$\mathcal{I}_{s,t} = \log \frac{p(f_{s+t}(z_s, z_t))}{p(f_s(z_s))p(f_t(z_t))} \geq 0 \quad (1)$$

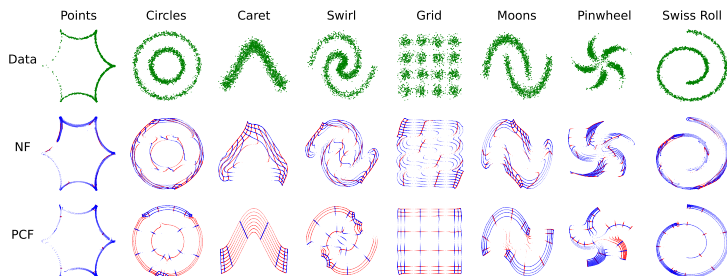
- $\mathcal{I}_{s,t} = 0$ when contours are orthogonal
- Have related term, $\hat{\mathcal{I}}$, that can be computed using any flow that is also 0 when contours are orthogonal.
- Incorporate either term as regularizer to train PCF.

- For flows with low dimensional latent space, sum of change of variables formula and \mathcal{I} cancels Jacobian determinant term.

$$\underbrace{\left(-\log p_z(z) + \frac{1}{2} \log |J^T J|\right)}_{-\log p_x(x)} + \underbrace{\left(-\frac{1}{2} \log |J^T J| + \sum_{i=1}^{\dim(z)} \|J_i\|_2^2\right)}_{\mathcal{I}} \quad (2)$$

$$= -\log p_z(z) + \sum_{i=1}^{\dim(z)} \|J_i\|_2^2 \quad (3)$$

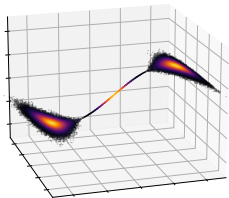
Experiments



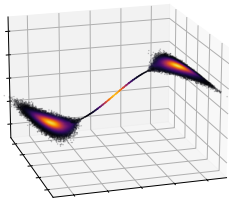
		Points	Circles	Caret	Swirl	Grid	Moons	Pinwheel	Swiss Roll
$\log p(x)(\uparrow)$	NF	-1.60	-3.10	-1.89	-0.19	-6.02	-0.64	-3.28	-4.67
	PCF	-1.62	-3.12	-1.89	-0.20	-6.02	-0.66	-3.29	-4.68
$\mathcal{I}_{\mathcal{P}}(\downarrow)$	NF	1.60	1.18	0.61	0.71	0.39	0.64	0.77	1.38
	PCF	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Experiments

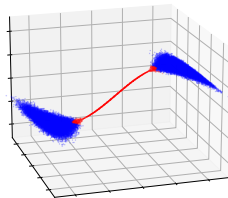
True density



Learned density

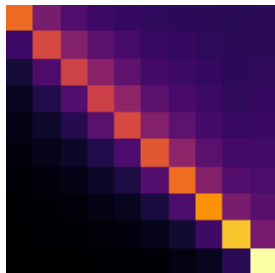


Learned rank

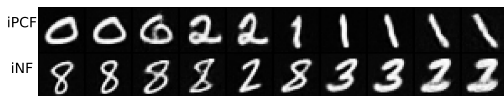
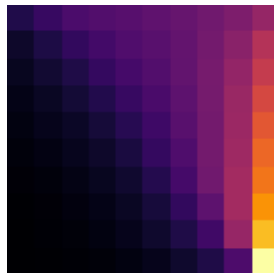


Experiments

iPCF



iNF



(a) Largest contour



(b) 5th largest contour