Active Pointillistic Pattern Search

Carnegie Mellon University

Yifei Ma, Dougal J. Sutherland, Roman Garnett, Jeff Schneider

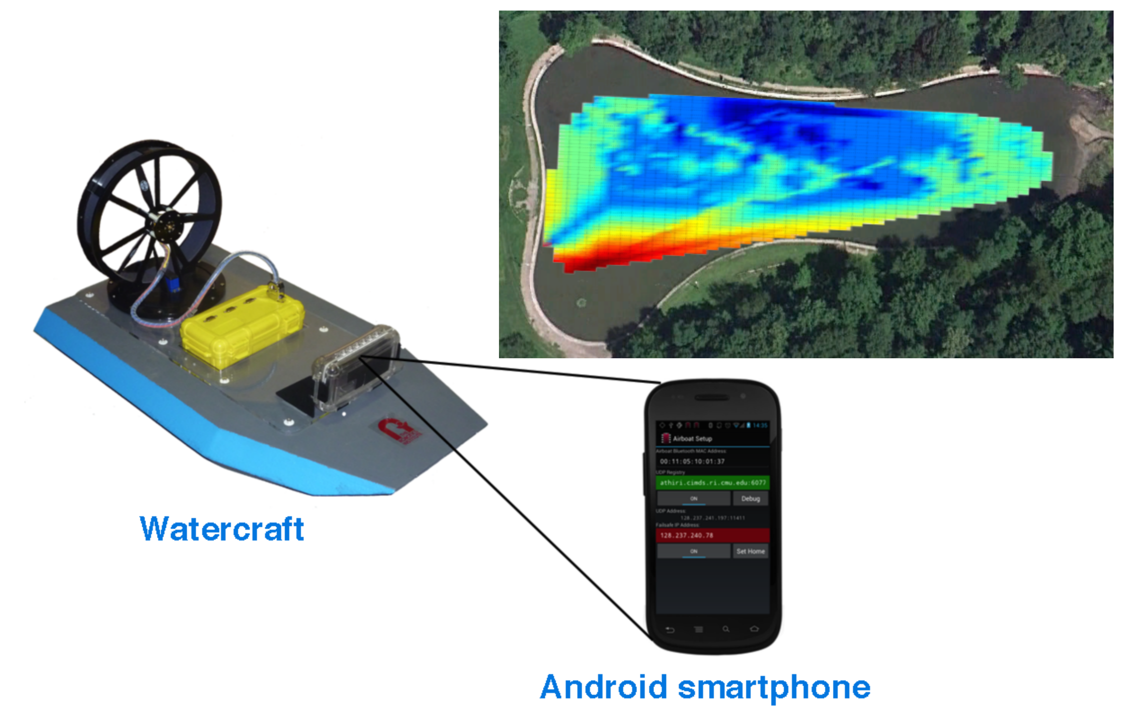
universität bonn

Motivation

- Pointillism: point observations, group patterns.
- Active search: filter as many positives as possible for later processing.
- Flexibility: allow for arbitrary definitions of the desired pattern, in the form a functional classifier.

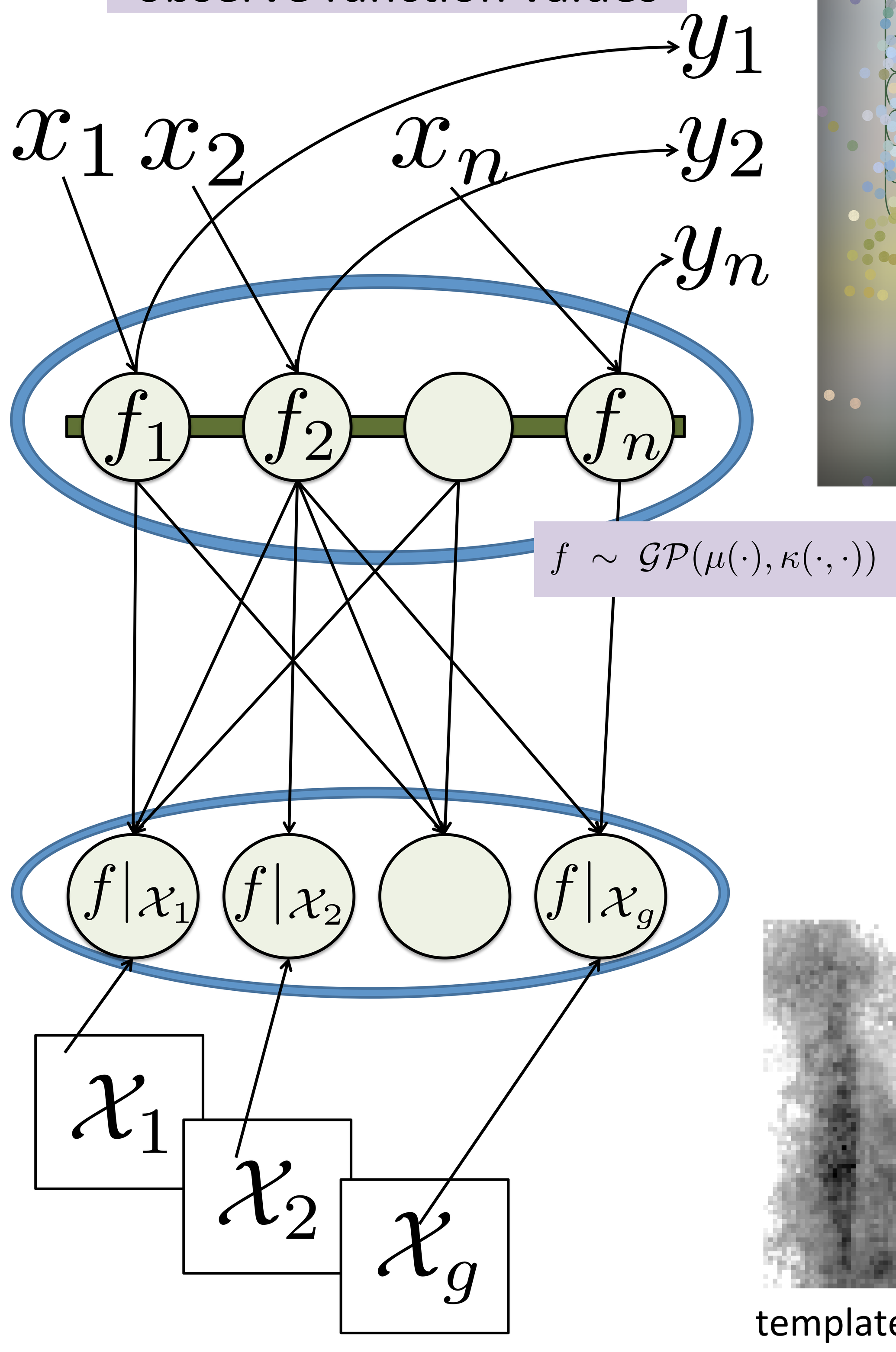
Applications

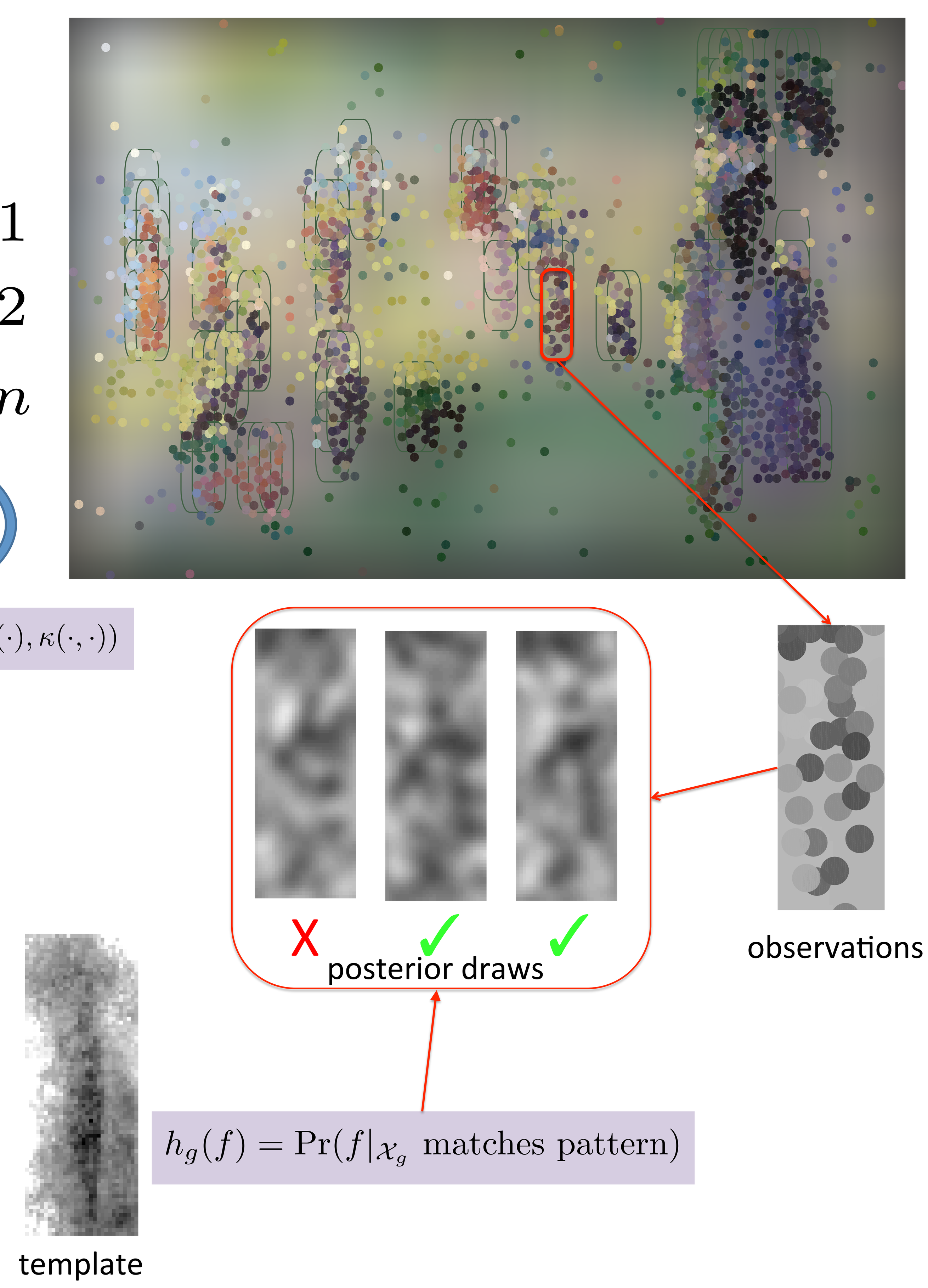
- Environmental monitoring: autonomous boats searching a pond for polluted areas
- Astronomy: choosing where to point a telescope to find interesting objects
- Polling: carefully surveying to find electoral races that need attention.



Problem Setup

1 Select locations to observe function values





2 Collect reward for region matches $r(\mathcal{D}) = \sum_g \mathbb{1}\{\mathbb{E}[h_g(f) \mid \mathcal{D}] > \theta\}$

Algorithm

Choose point to greedily maximize expected reward

$$\max_{x_*} \mathbb{E} \sum_{g \in \mathcal{G}_t} [r_g(\mathcal{D}_*) \mid x_*, \mathcal{D}_t]$$

Estimate expected reward with Monte Carlo:

Sample observation

$$z_* \mid x_*, \mathcal{D}_t \sim \mathcal{N}\Big(\mu_{f|\mathcal{D}_t}(x_*),\ \kappa_{f|\mathcal{D}_t}(x_*, x_*) + \sigma^2\Big)$$

Sample enough of f to get $h_g(f)$

$$r(\mathcal{D}_t \cup \{x_*, z_*\}) = \sum_{g \in \mathcal{G}_t} \mathbb{1}\{\mathbb{E}[h_g(f) | \mathcal{D}_t, x_*, z_*] > \theta\}$$

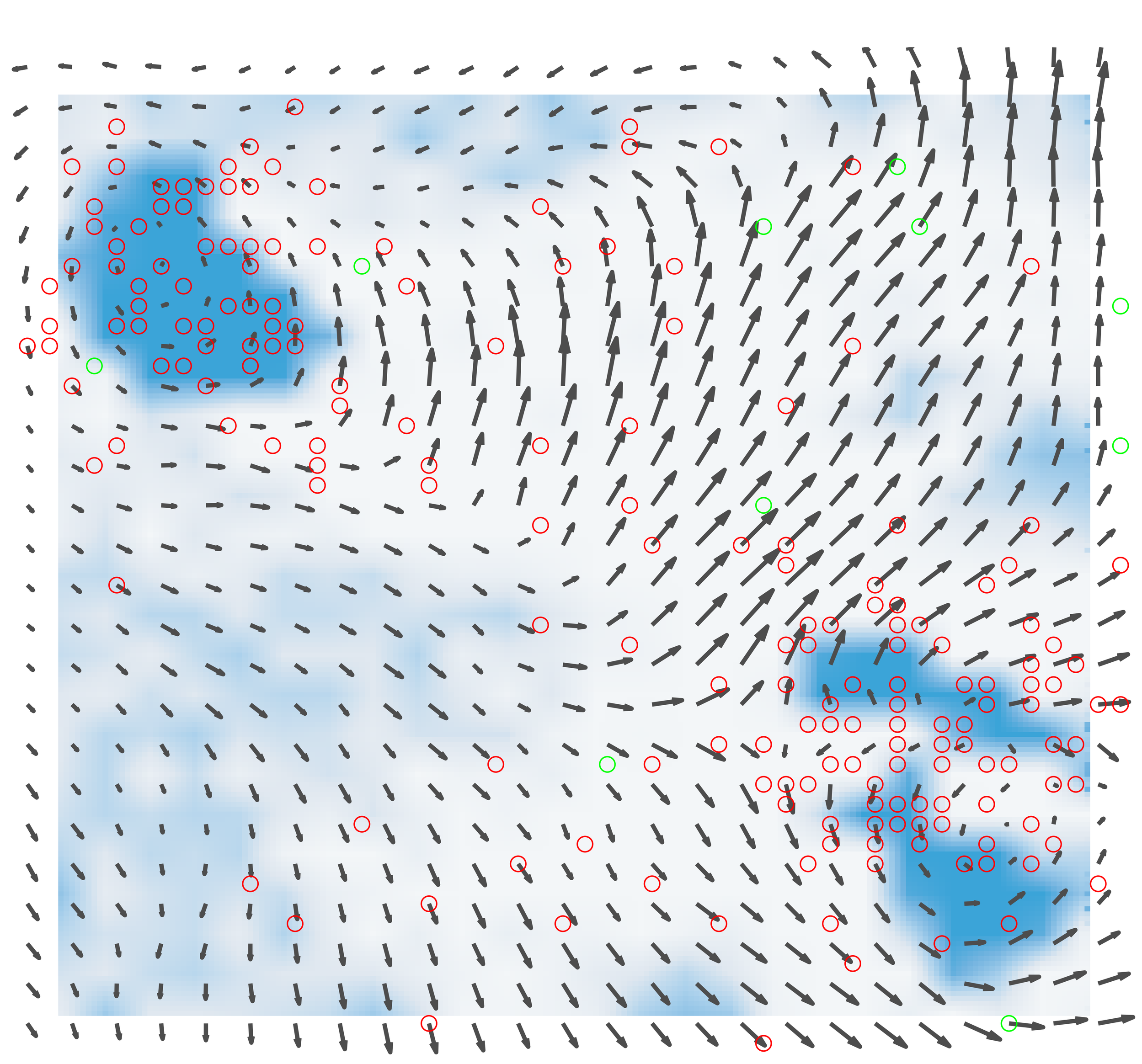
Analytical form:

If $h_g(f) = \Phi\big(L_g f + b\big)$ where $\mathbf{\Phi}$ is normal cdf and L_g is linear, e.g. $L_g f = \int_{x \in \mathcal{X}_g} w(x)^\mathsf{T} f(x)\,\mathrm{d}x$

then $\mathbb{E}\big[r_g(\mathcal{D}_*) \mid x_*, \mathcal{D}\big]$ has a closed form.

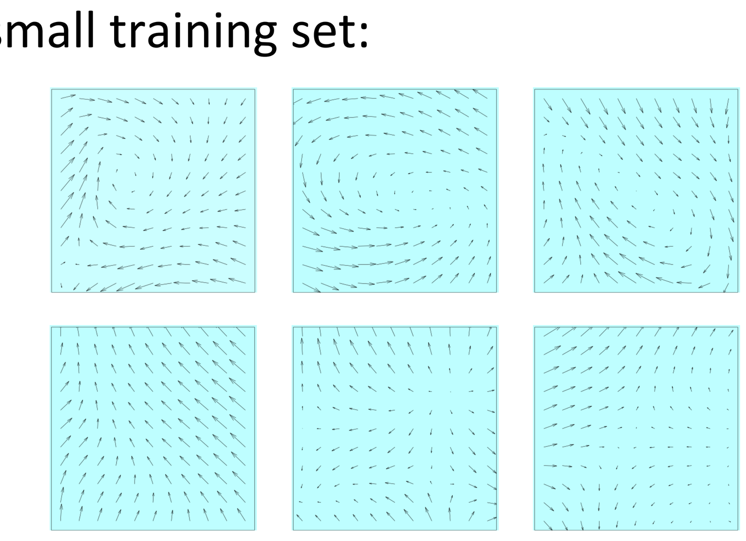
Finding Vortices with a Black-Box Classifier

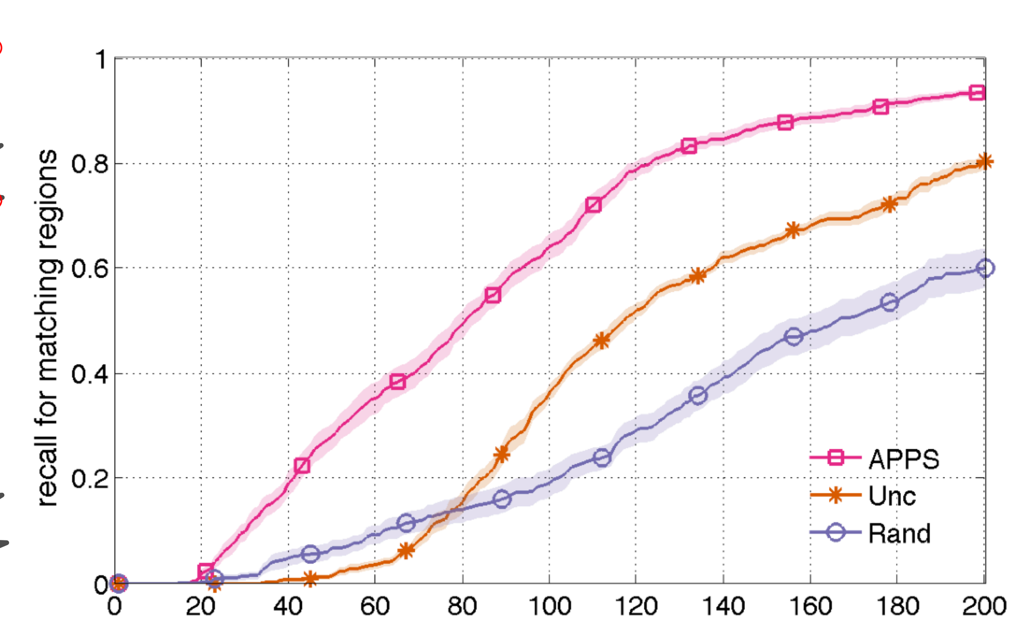
We want to find vortices in a 2d map of fluid flow.



The velocity dataset; each arrow represents the mean of a 2x2 square. This run was initialized with the points at the green circles and selected the ones at the red circles.

Trained a 2-layer neural network on a small training set:





Mean and standard error of recall for matching regions, over 15 runs. True labels are determined by using the classifier on the full velocity dataset.

Intuition

If linear classifiers and regions are independent:

- If two regions have the same marginal probability, picks the one whose variance can be reduced more with one point.
- If two regions' variances can be equally well reduced, picks the one with higher marginal probability.
- In a given region, picks point most correlated to region's label.

Related Work

Most Bayesian optimization:

- Models functions with GPs
- Usually maximization of observable point values

Active search (e.g. Garnett et al., ICML 2012)

- Usually assumes that labels are directly observable and correspond to single points.

Active Area Search (Ma et al., AISTATS 2014)

- Similar setup and algorithm, but can only detect thresholds on mean of a region.
- APPS generalizes to *any* pattern.

Level set estimation (Gotovos et al., IJCAI '13; Low et al., AAMAS '12)

- Actively finds a particular level set in a function.
- Related to AAS; can't model arbitrary patterns.