Gaussianization–PCA for One-Class Remote Sensing Image Classification

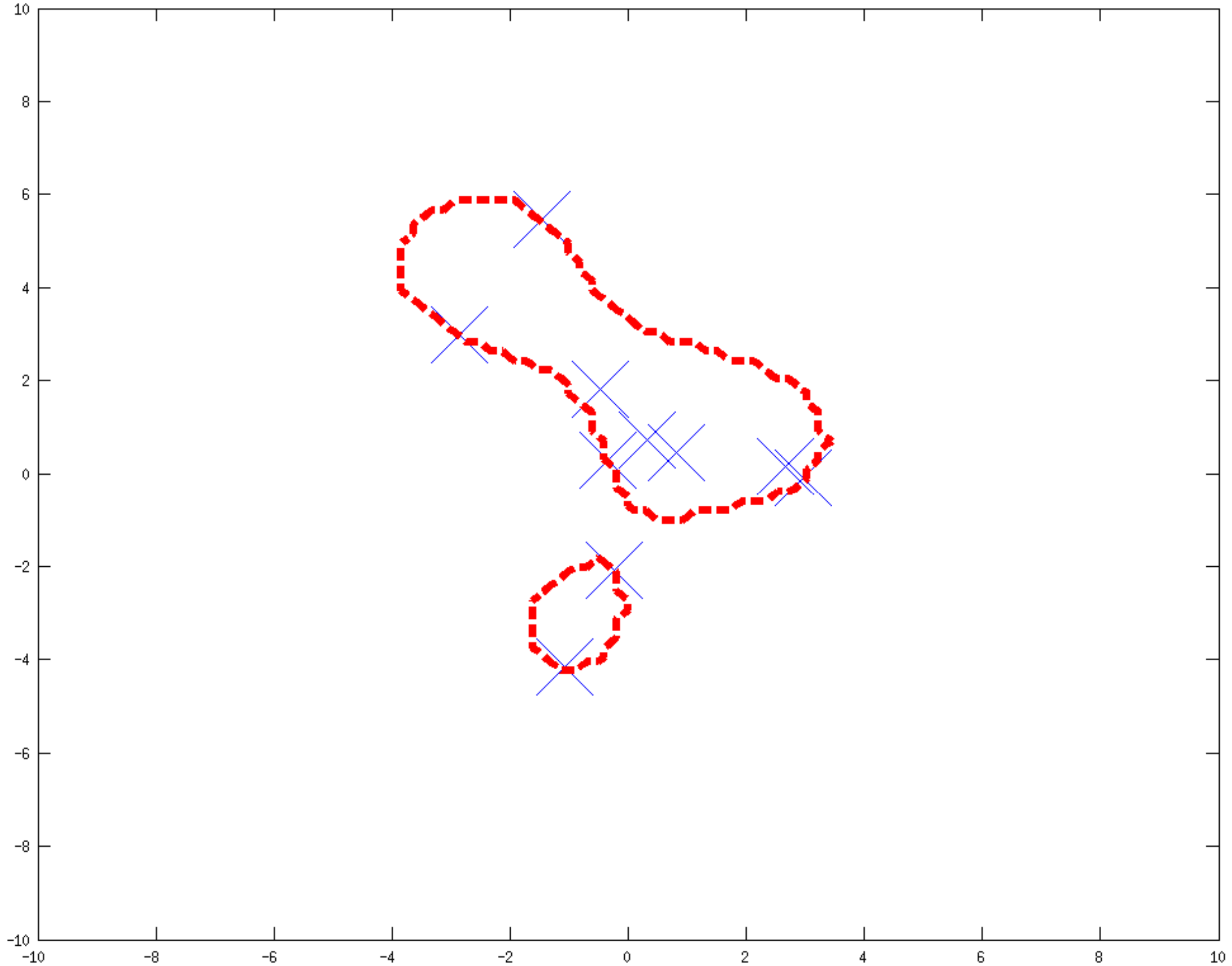
Valero Laparra, Jordi Muñoz-Marí,
Gustavo Camps-Valls and Jesús Malo

Image Processing Laboratory (IPL)
Universitat de València, Spain

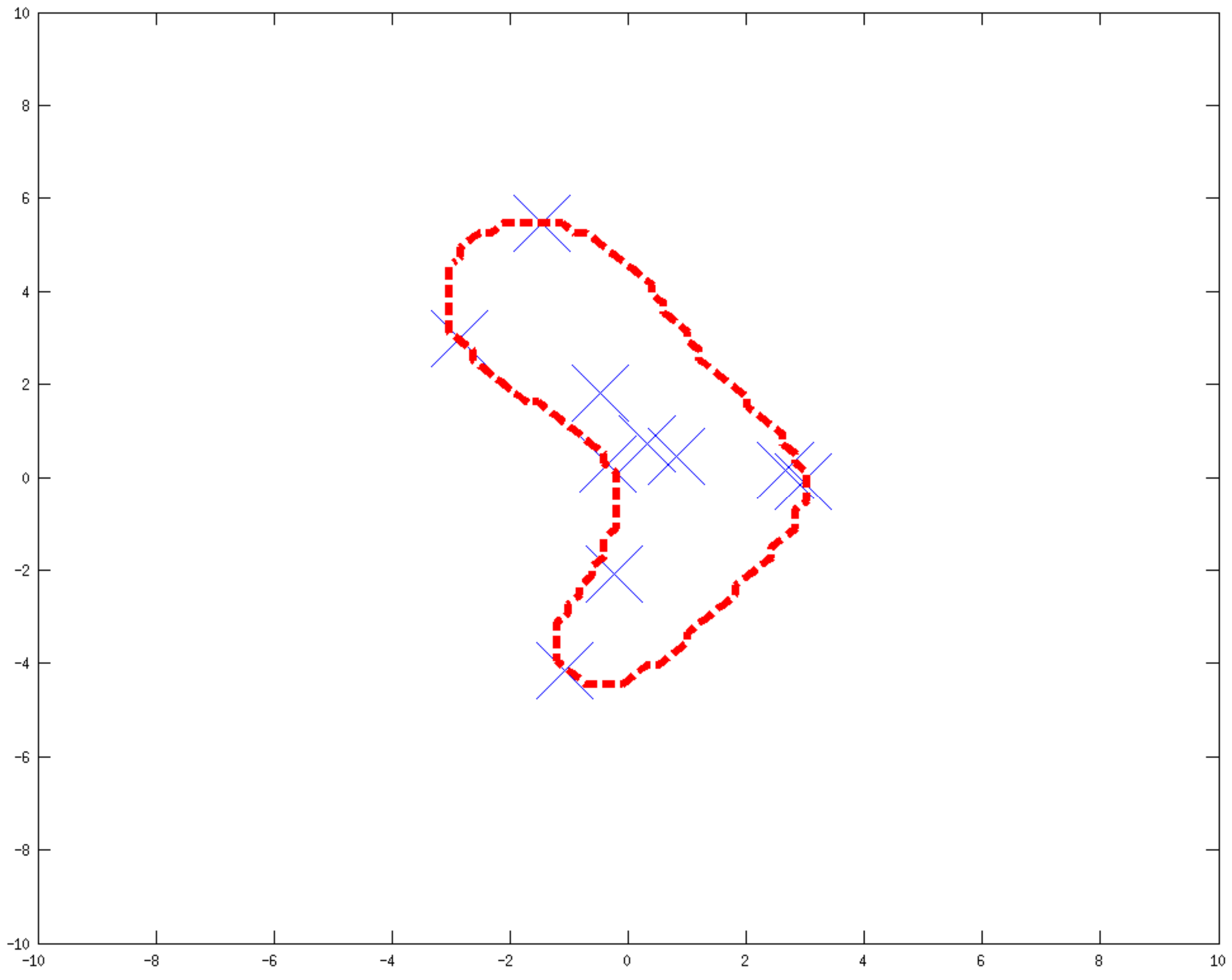
ONE CLASS classification

- **One-class classification** tries to distinguish one class of objects from **all** other possible objects, by learning from a training set containing **only** the objects of that class.

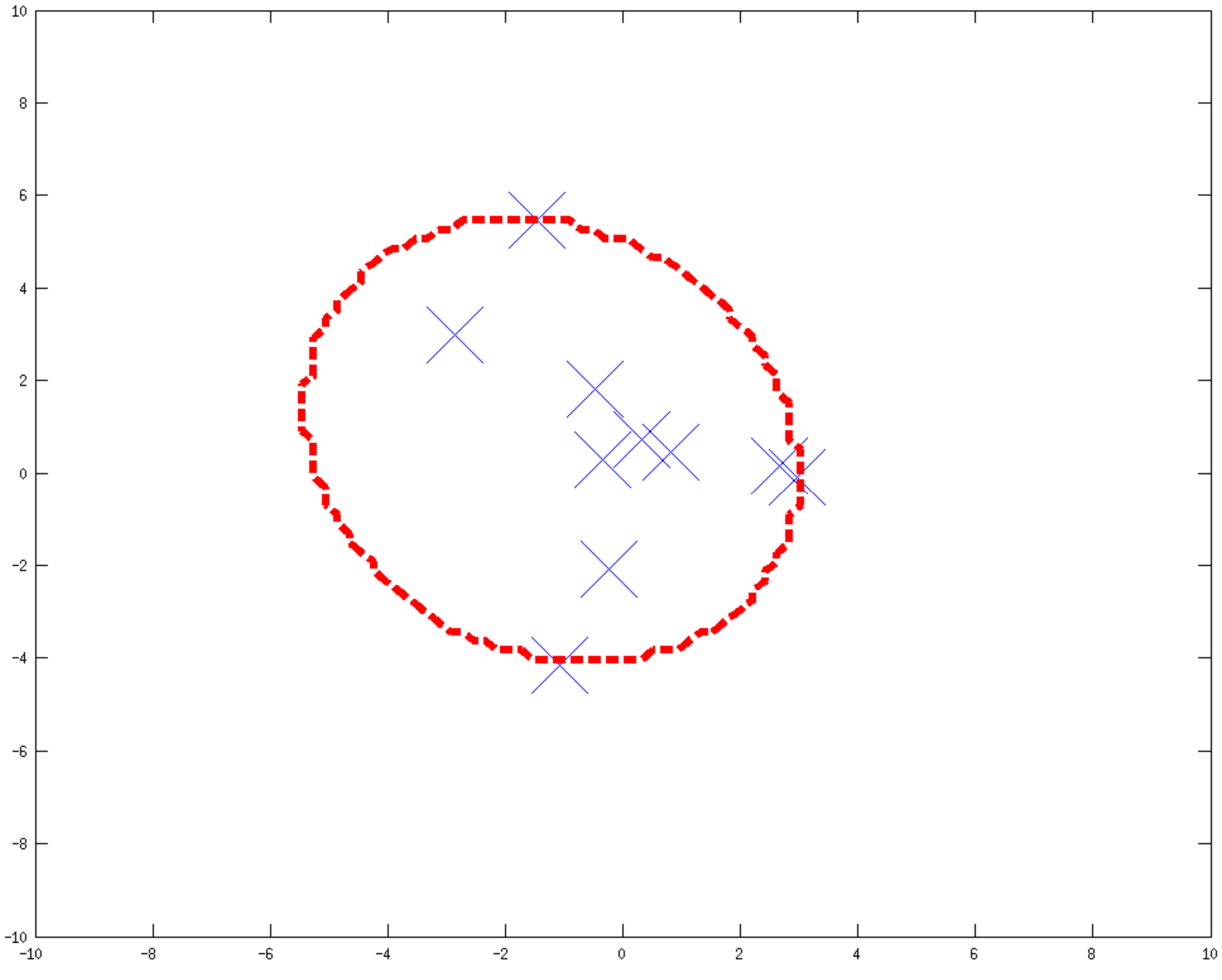
ONE CLASS classification



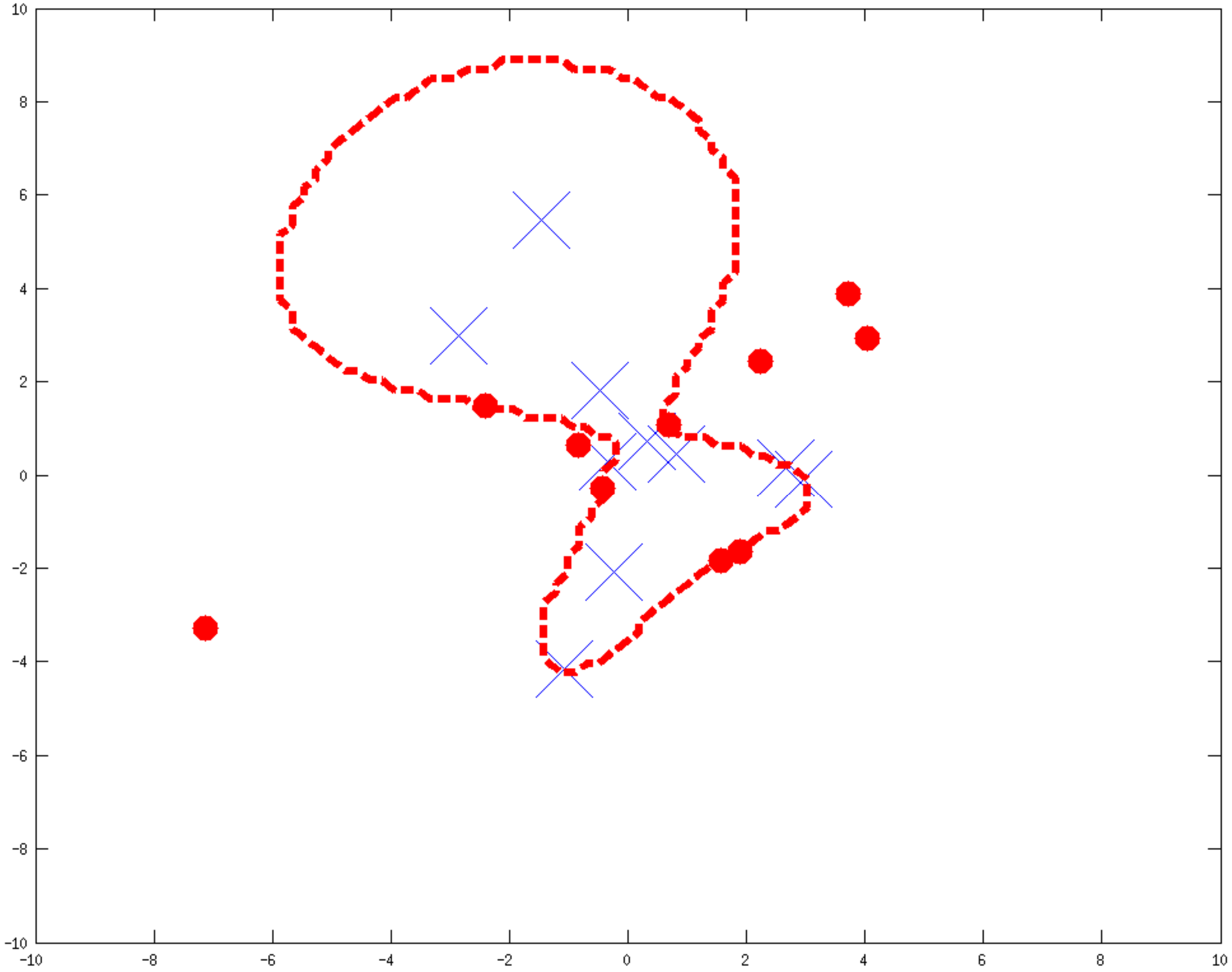
ONE CLASS classification



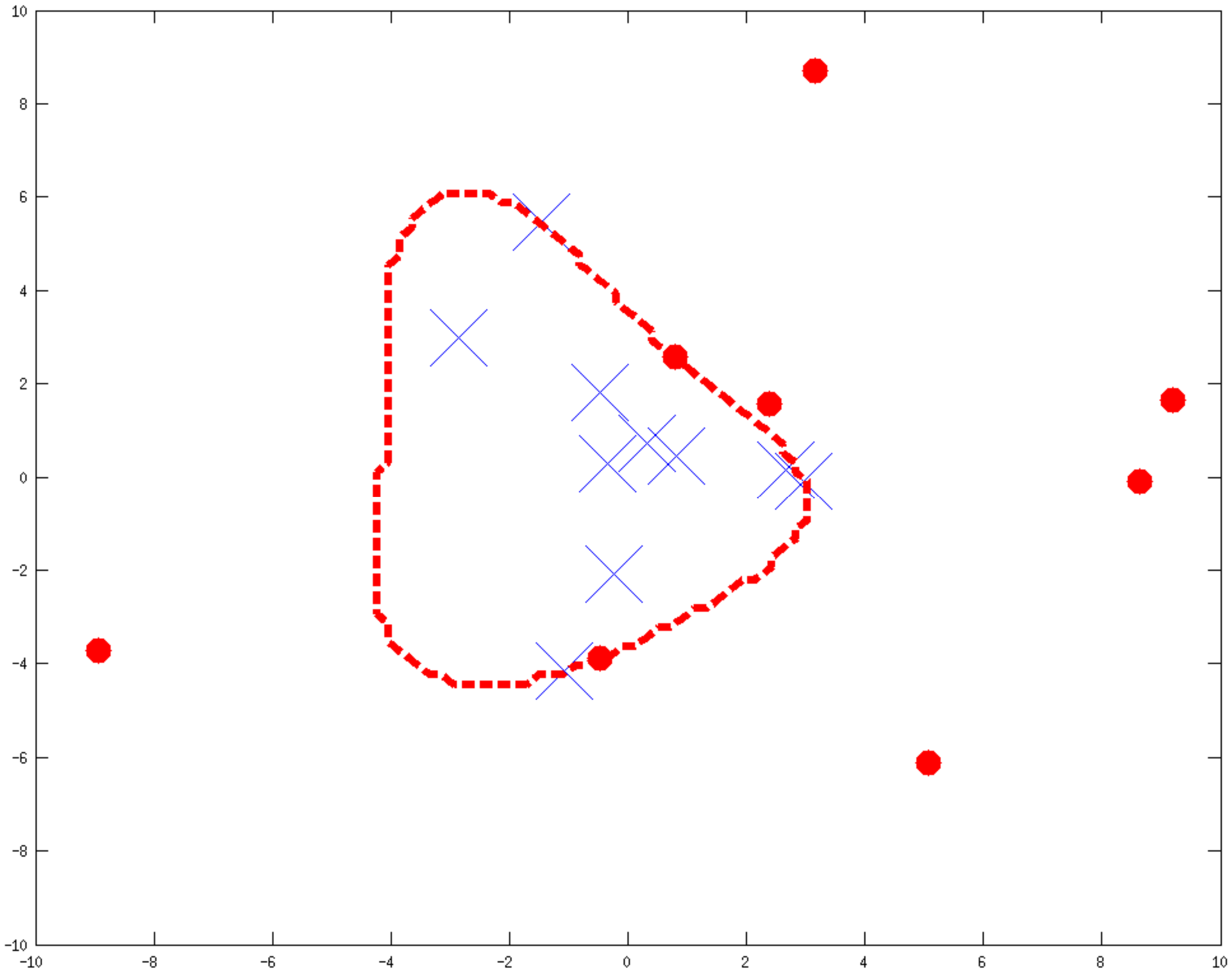
ONE CLASS classification



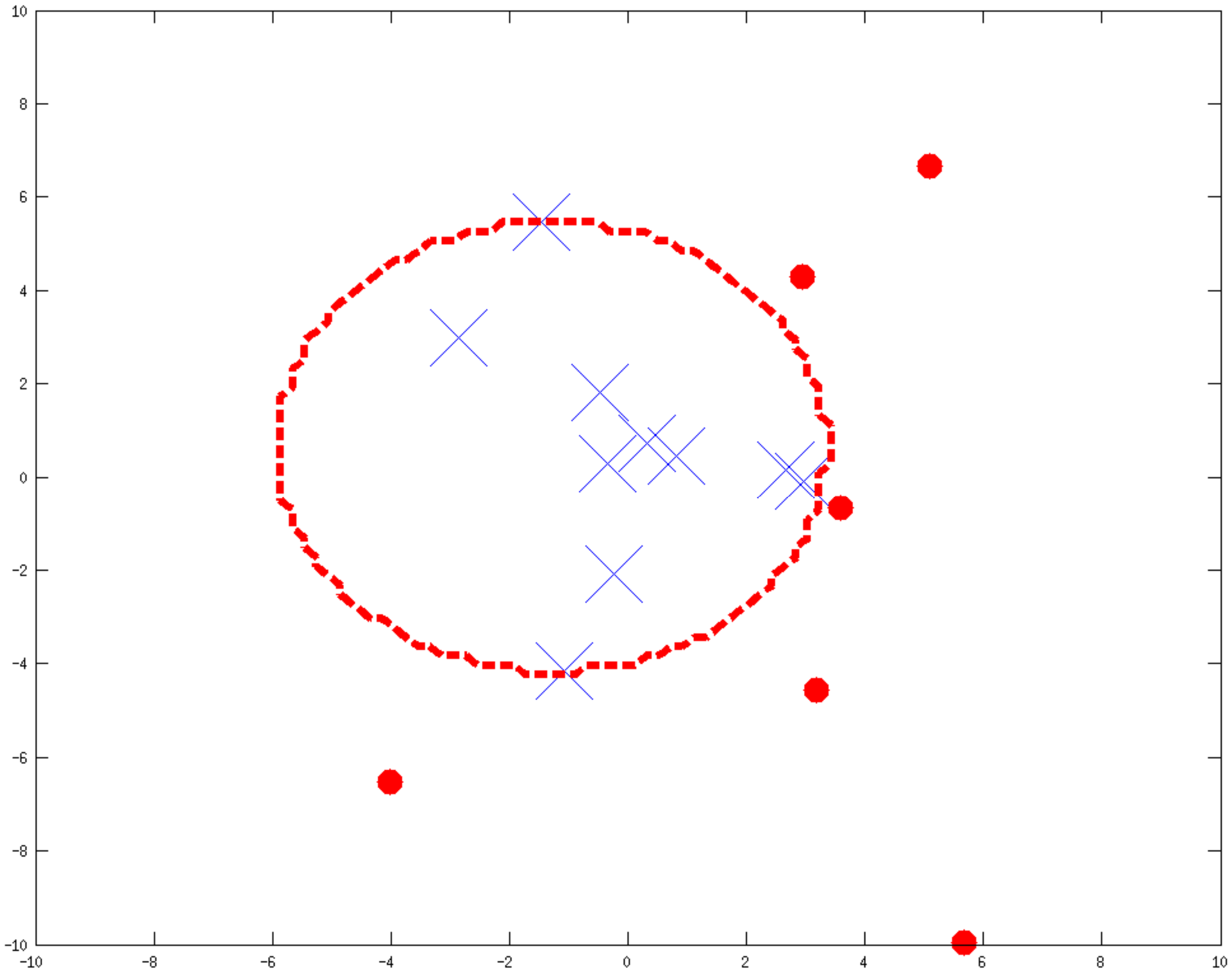
ONE CLASS classification



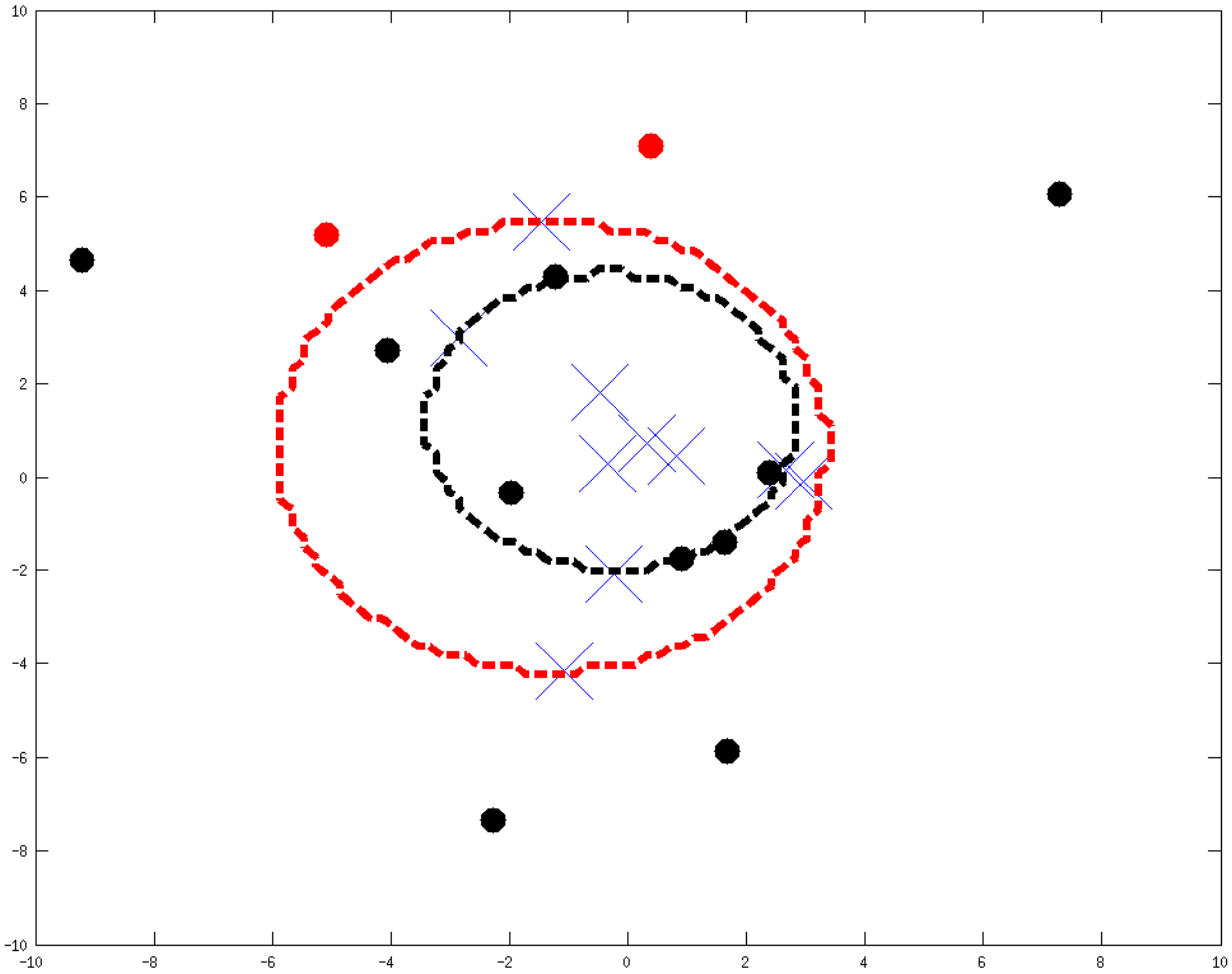
ONE CLASS classification



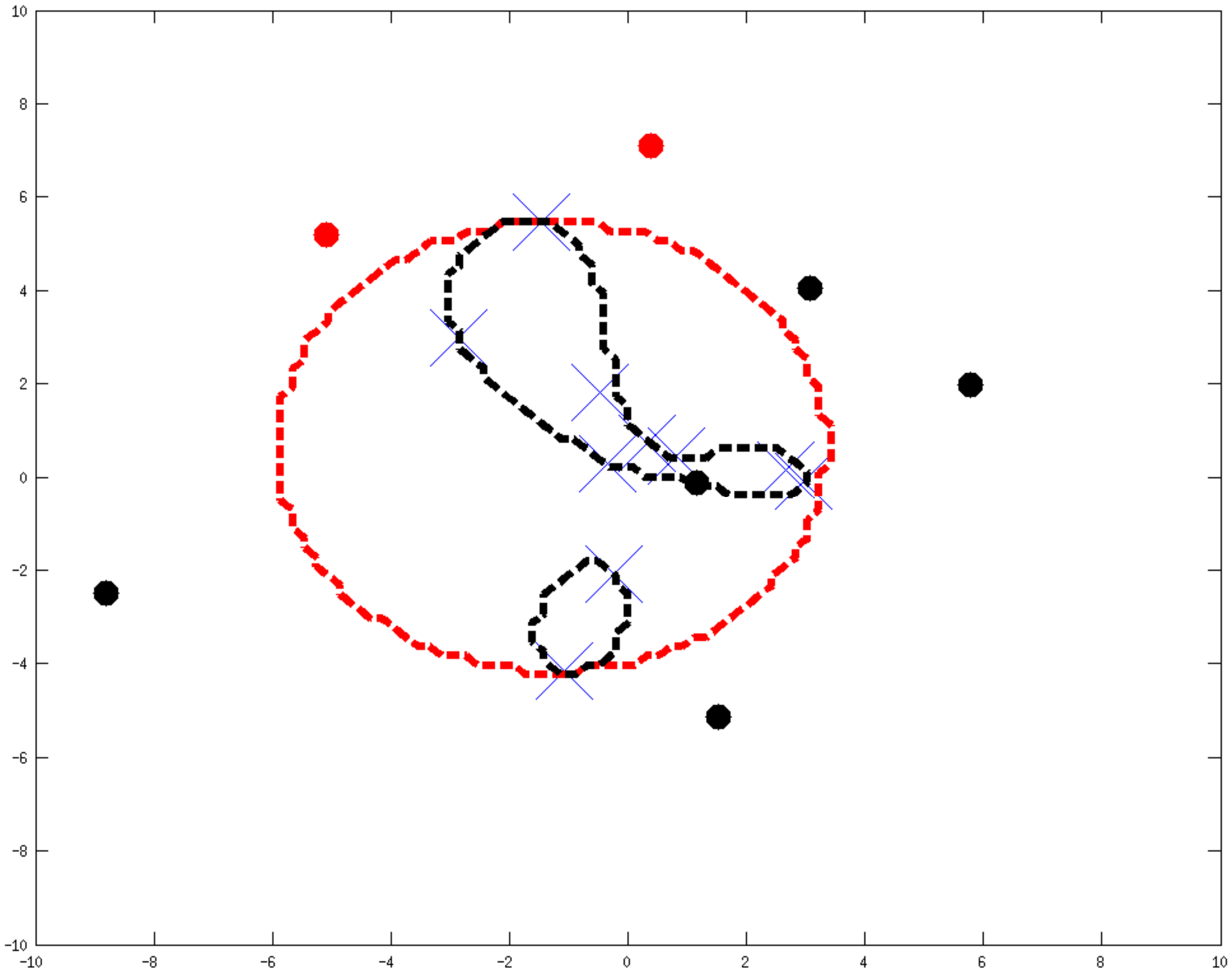
ONE CLASS classification



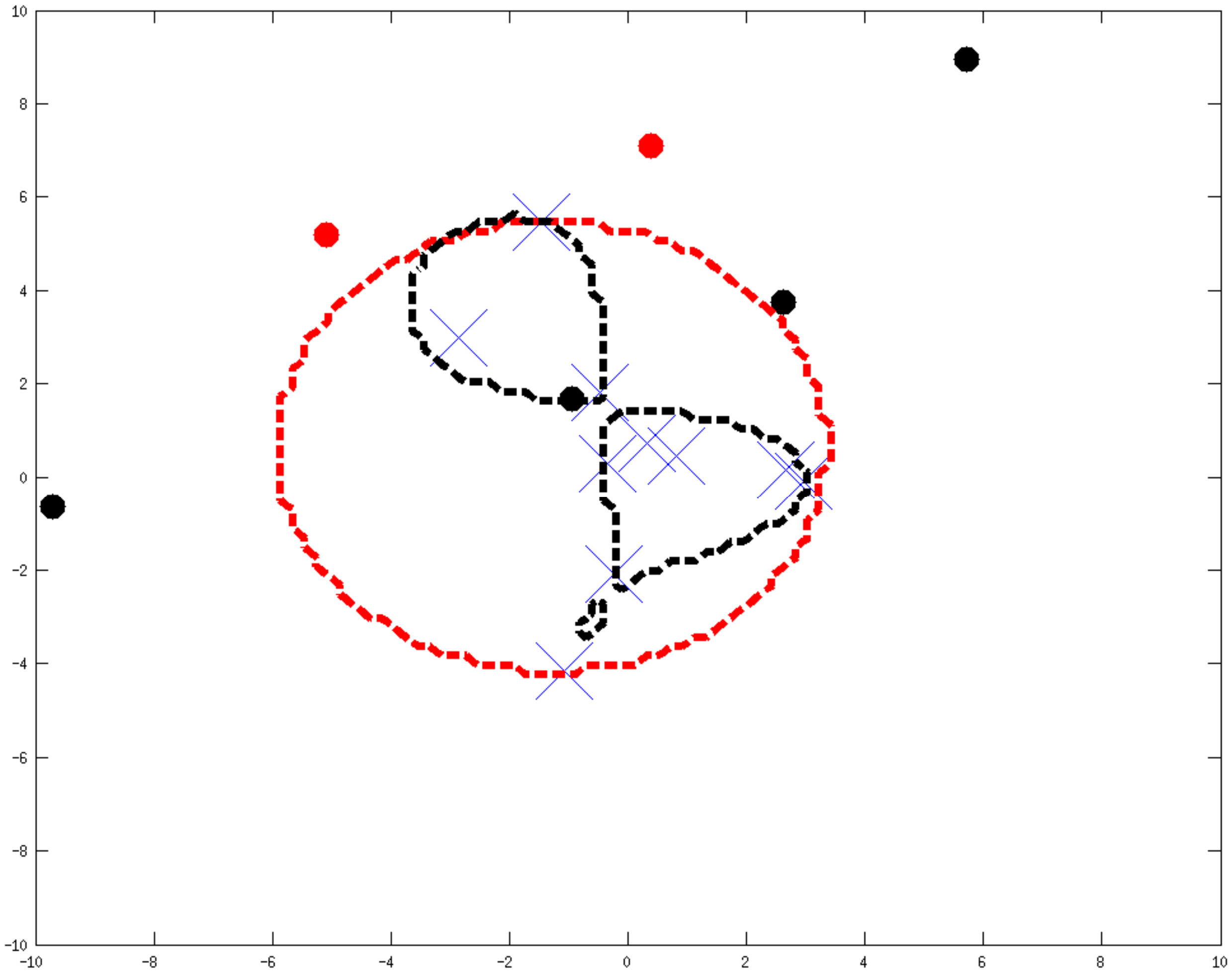
ONE CLASS classification



ONE CLASS classification



ONE CLASS classification



SUMMARY

- One-class classification
- G-PCA
  - Convergence
  - PDF estimation
- Experiments
  - 2D example
  - 10D experiment with real data
- Conclusions & Further work

ONE CLASS classification

Trade-off problem

- General
- Specific

2 approaches

- Boundary

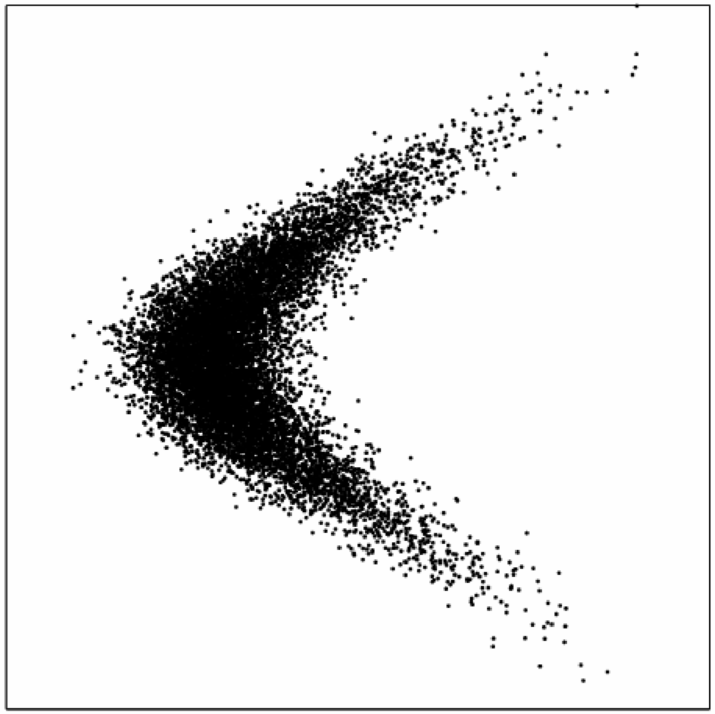
  Support Vector Domain Description SVDD

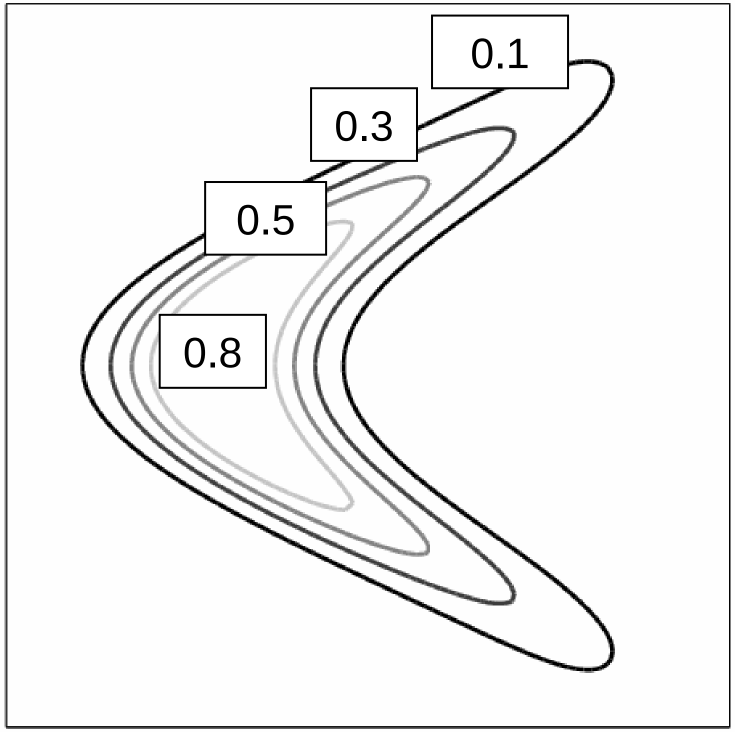
- Probability Density Function (PDF)

  Gaussianization PCA

ONE CLASS - PDF solution

- $P_{cl1}(x_0) >$ THRESHOLD $\rightarrow$ $Class(x_0) = 1$
- $P_{cl1}(x_0) <$ THRESHOLD $\rightarrow$ $Class(x_0) = 0$





ONE CLASS - PDF solution

- PROBLEM: PDF estimation!!
  - Assuming a parametric model (Gaussian, GSM…)
    - Problem: Previous knowledge about the PDF
  - Non assuming a parametric model (histogram…)
    - Problem: number of samples

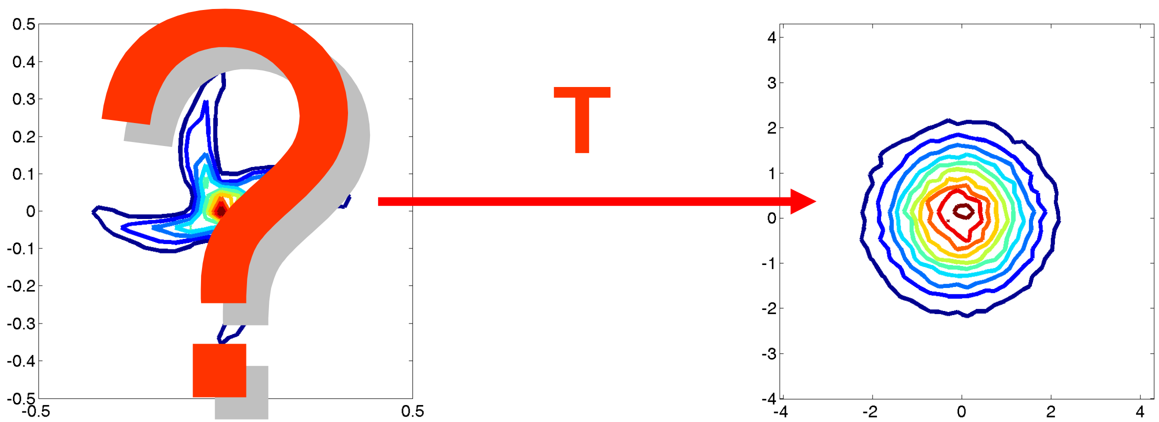
      "curse of dimensionality"

OUR APPROACH

$$Y = T(X)$$

$$P_X = P_Y * |J_T|$$

- From P(x) to P(y) (Gaussian)

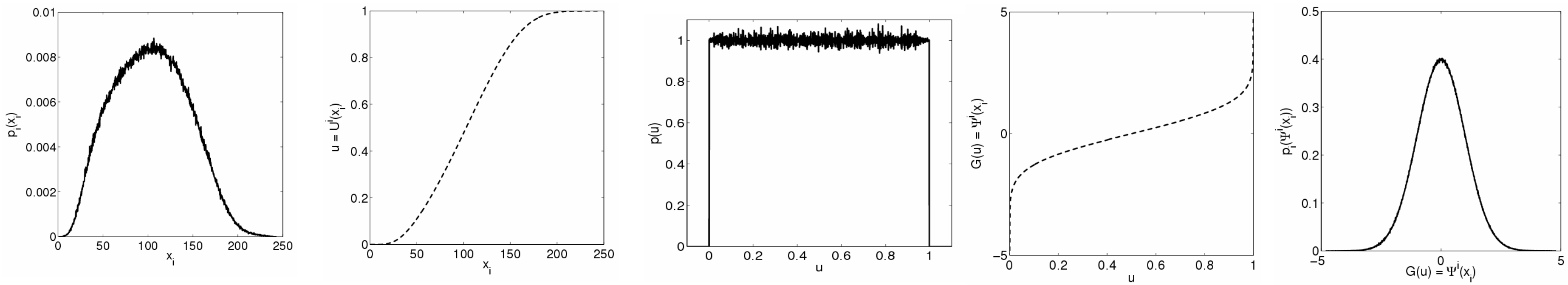


Gaussianization - PCA

$$Y = X^{n+1} = B^T \Psi(X^n)$$

- Gaussianization-PCA 2 steps:

(1) Marginal gaussianization



(2) PCA Rotation

Theoretical convergence Proof

$$Y = X^{n+1} = B^T \Psi(X^n)$$

- Negentropy:

$$J(X) = \int_{-\infty}^{\infty} p(X) log \frac{p(X)}{\mathcal{N}(X)} dX$$

- Difference in Negentropy:

$$\Delta J = J(X^n) - J(X^{n+1}) = \sum_{k=1}^{d} J(X_i^n) + 2_{ord}^{nd}(\Psi(X^n))$$

G-PCA

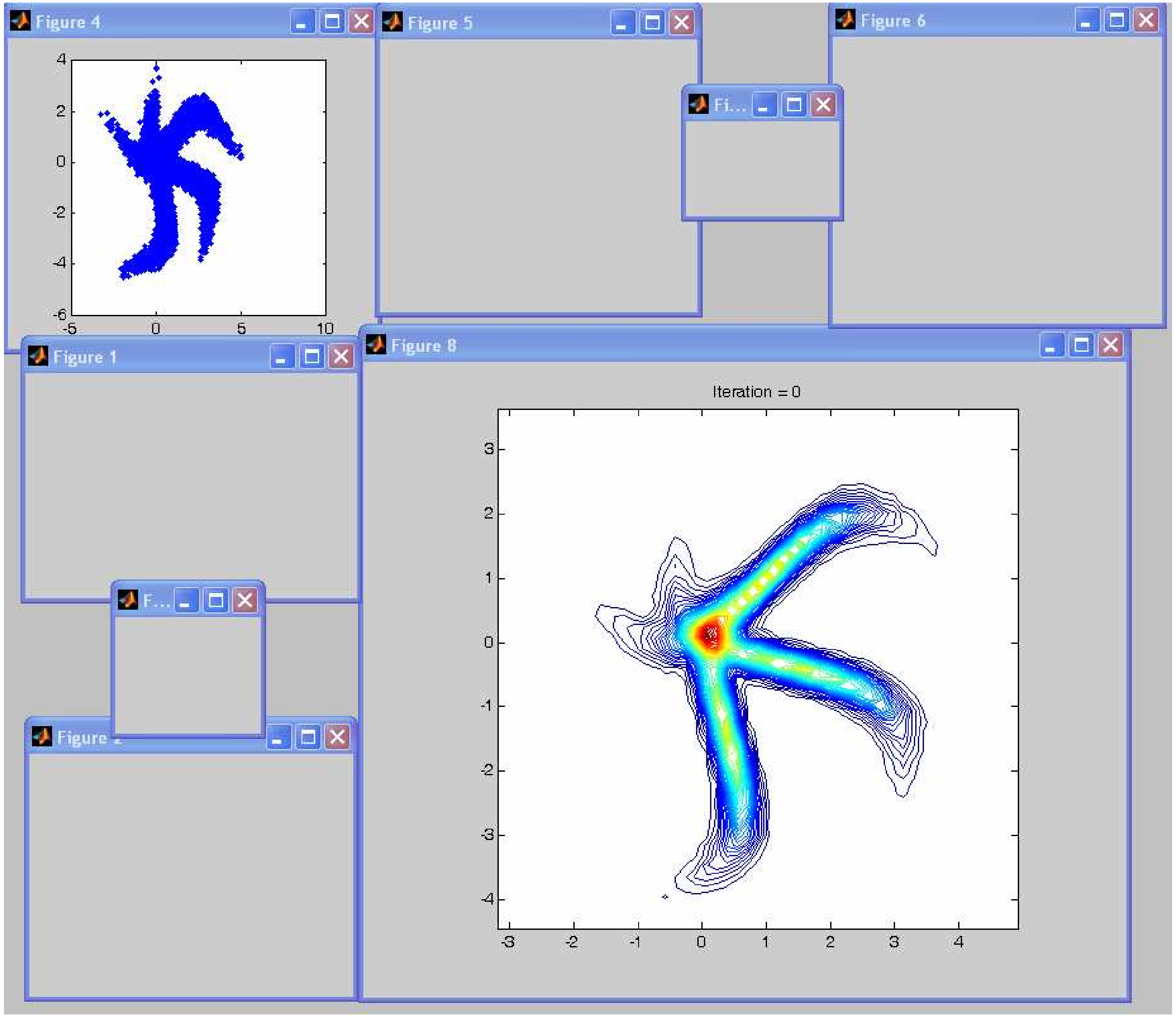


Figure 4
Figure 5
Figure 6
Fi...
Figure 1
F...
Figure 2
Figure 8
Iteration = 0

G-PCA

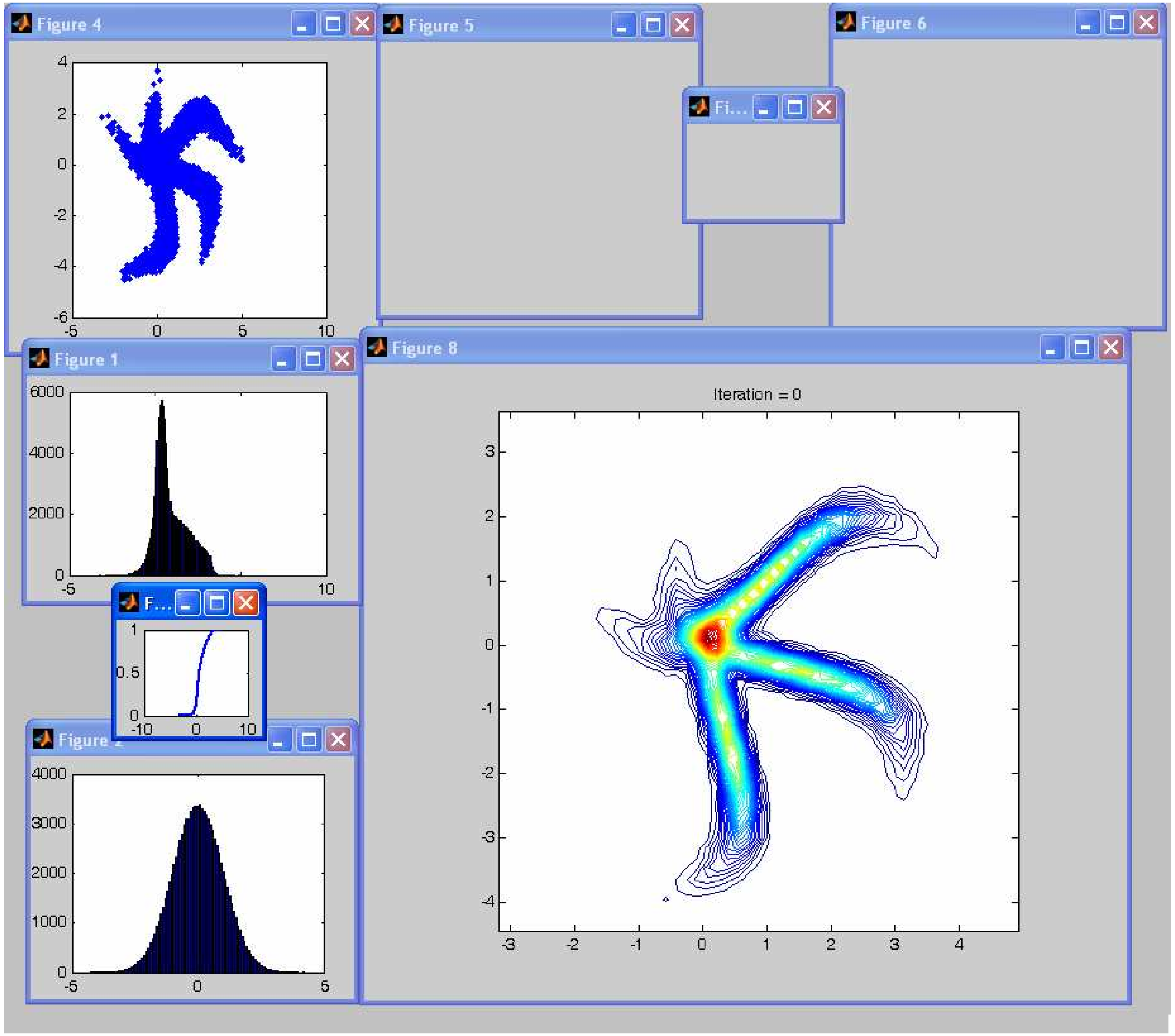


Figure 4
Figure 5
Figure 6
Fi...
Figure 1
Figure 8
Iteration = 0
F...
Figure 2

G-PCA

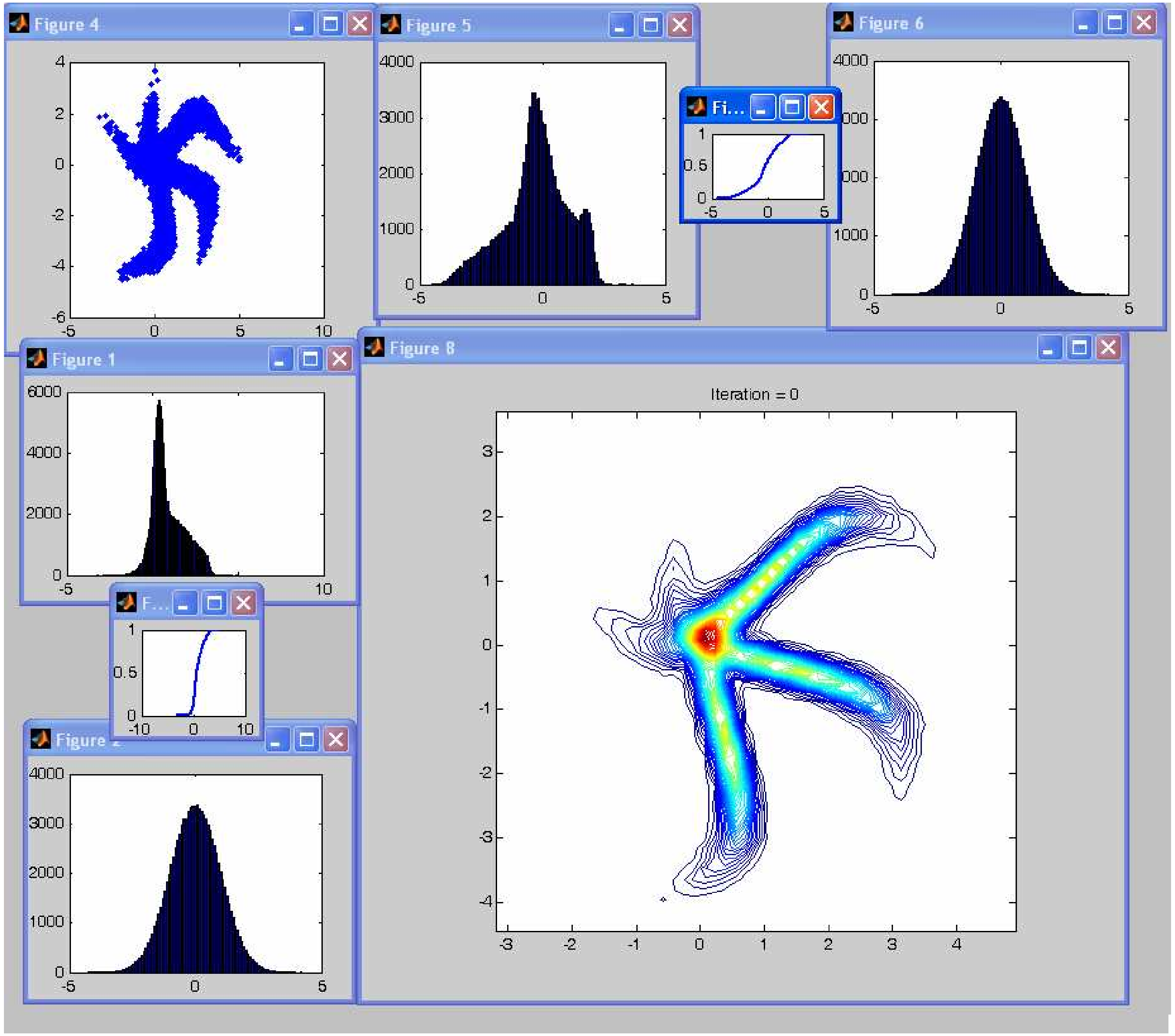


Figure 4
Figure 5
Figure 6
Figure 1
Figure 2
Figure 8
Iteration = 0

G-PCA

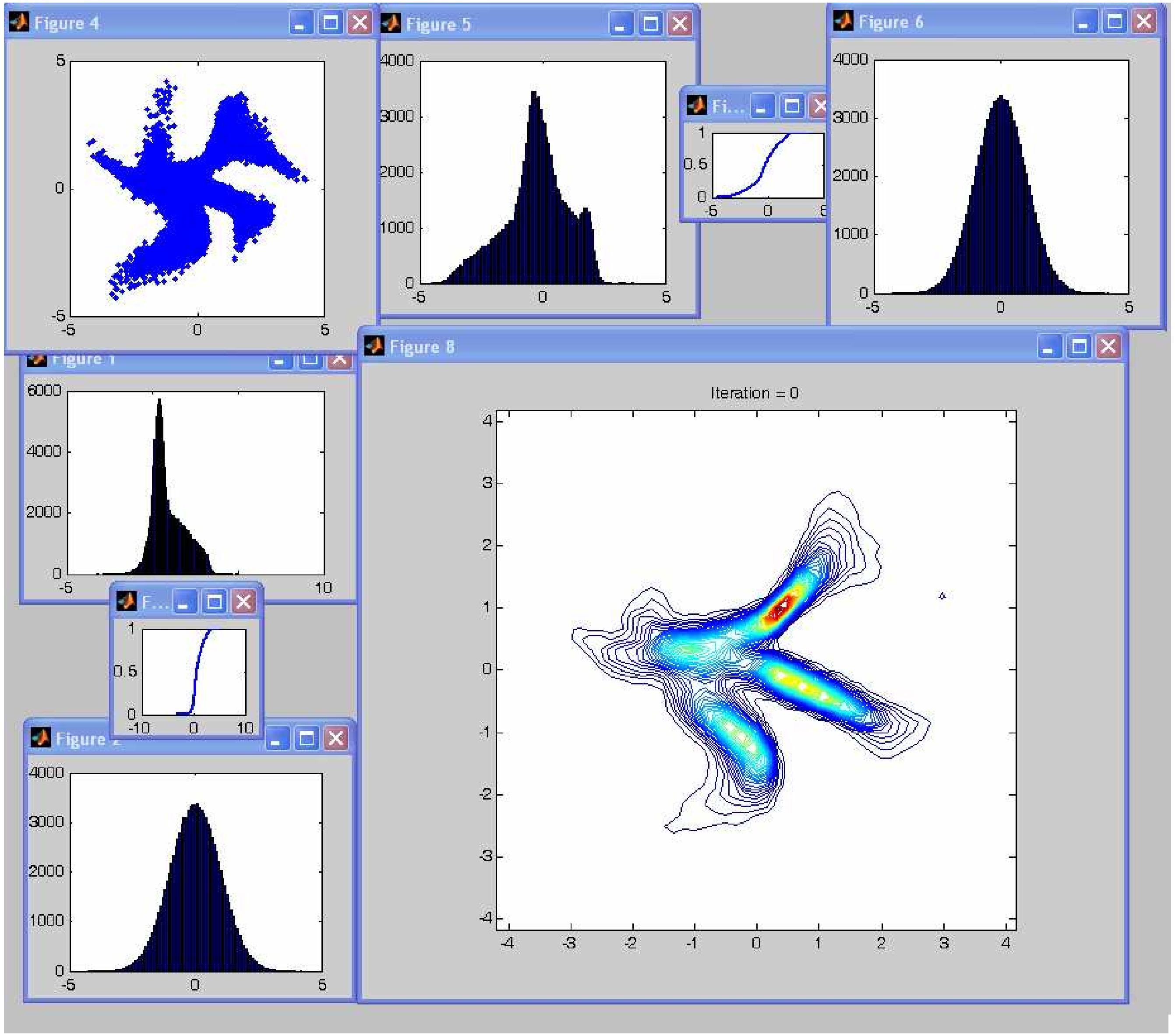


Figure 4
Figure 5
Figure 6
Figure 1
Figure 2
Figure 8
Iteration = 0

G-PCA

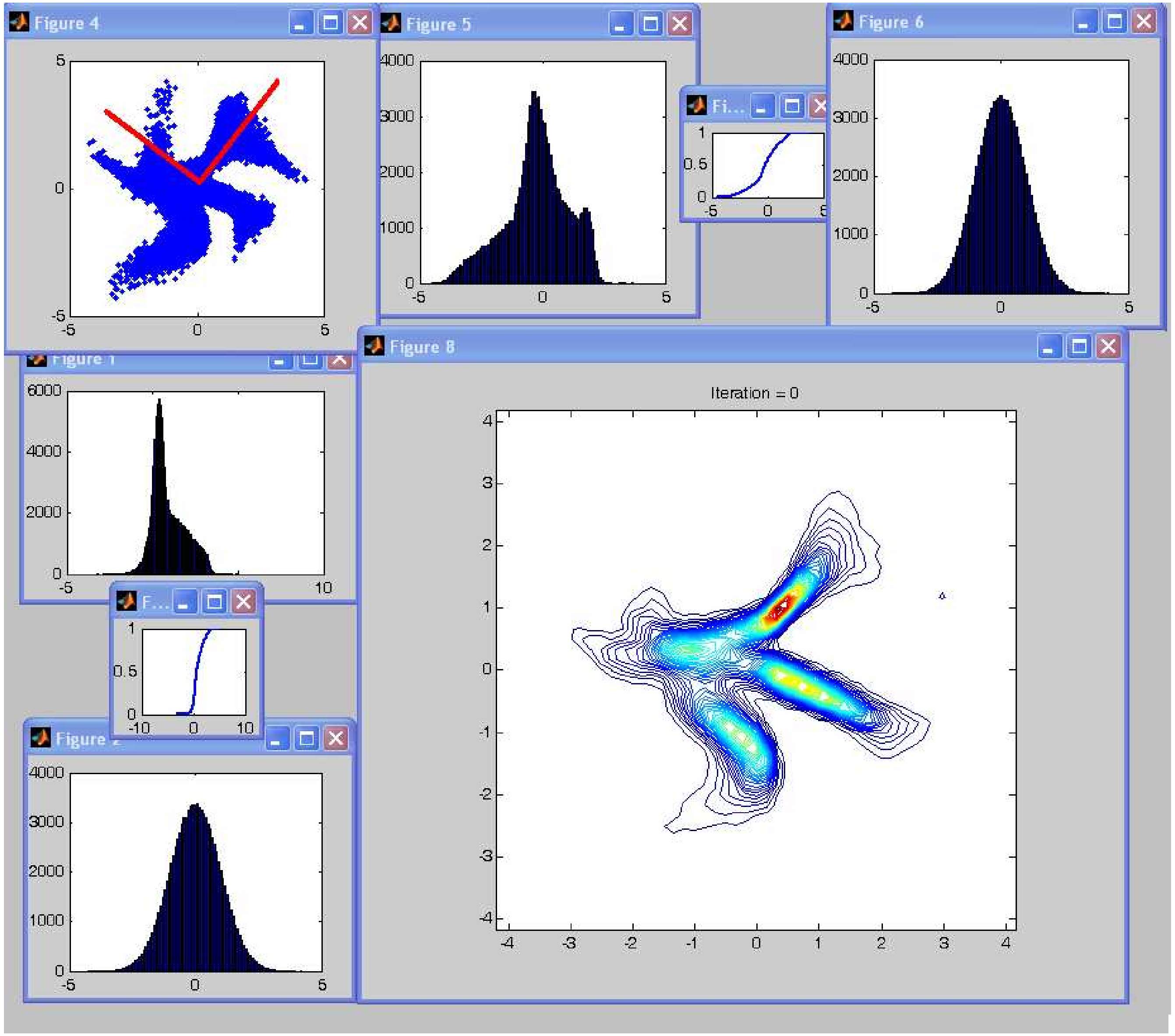
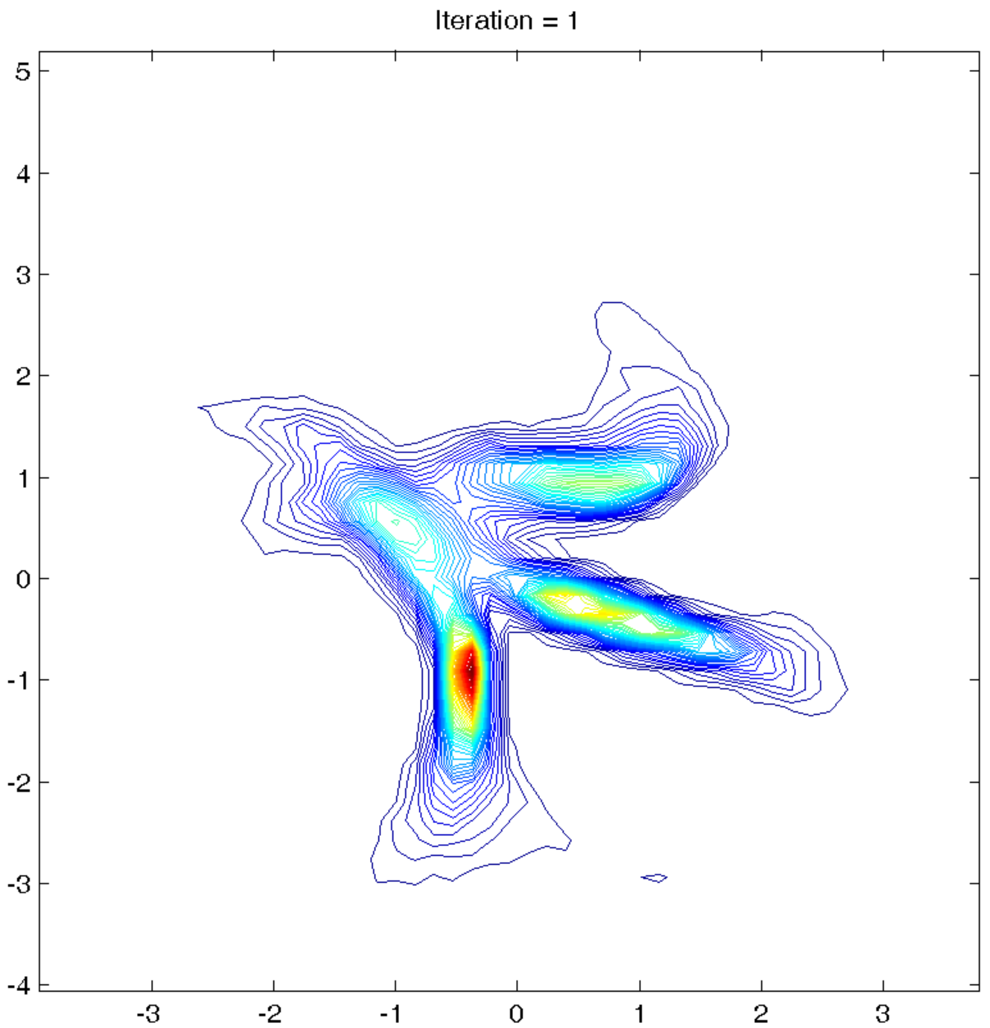
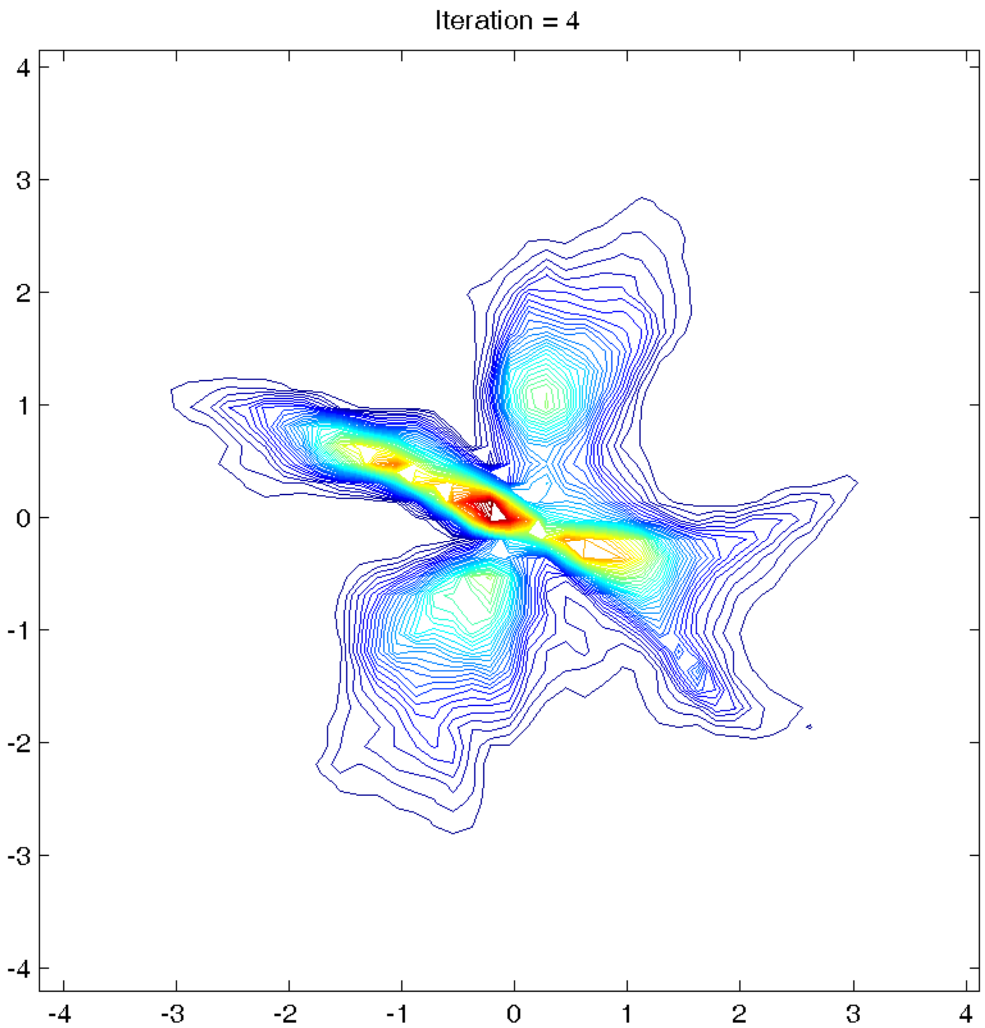


Figure 4
Figure 5
Figure 6
Figure 1
Figure 2
Figure 8
Iteration = 0

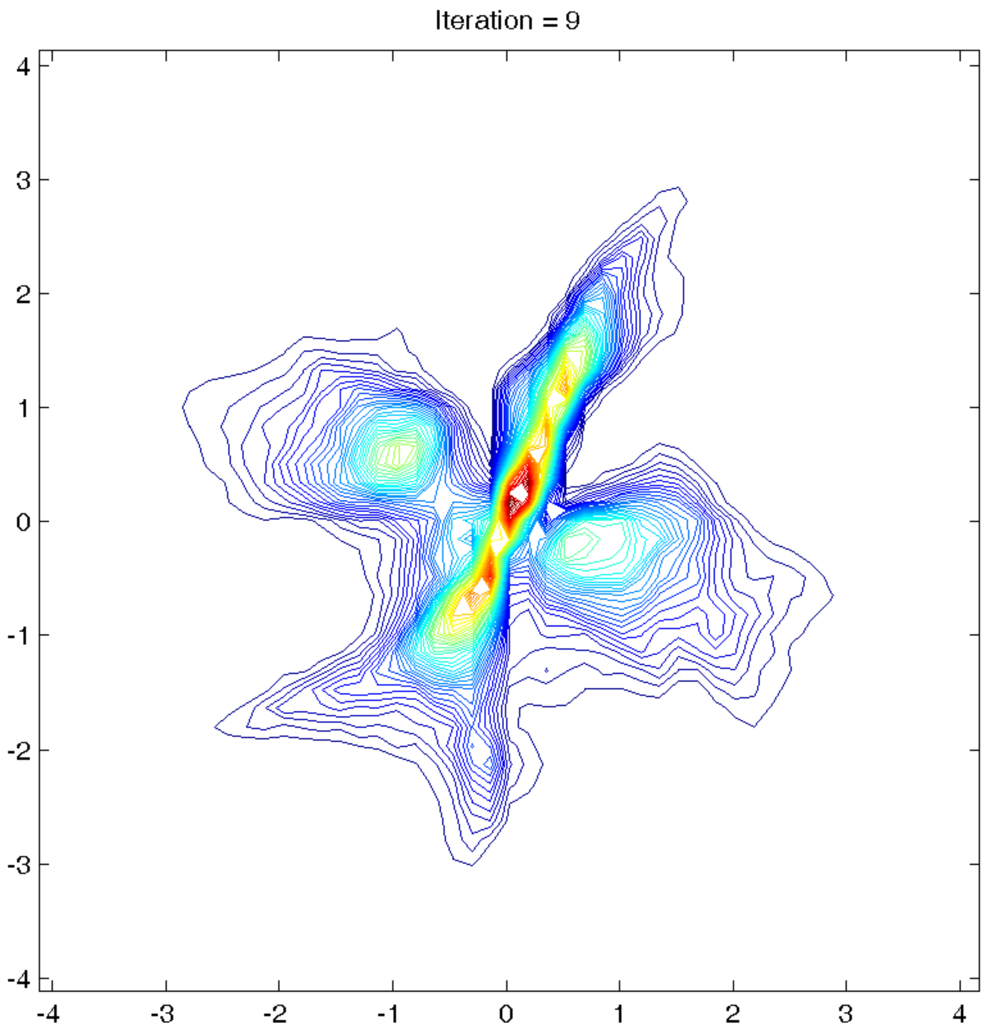
G-PCA



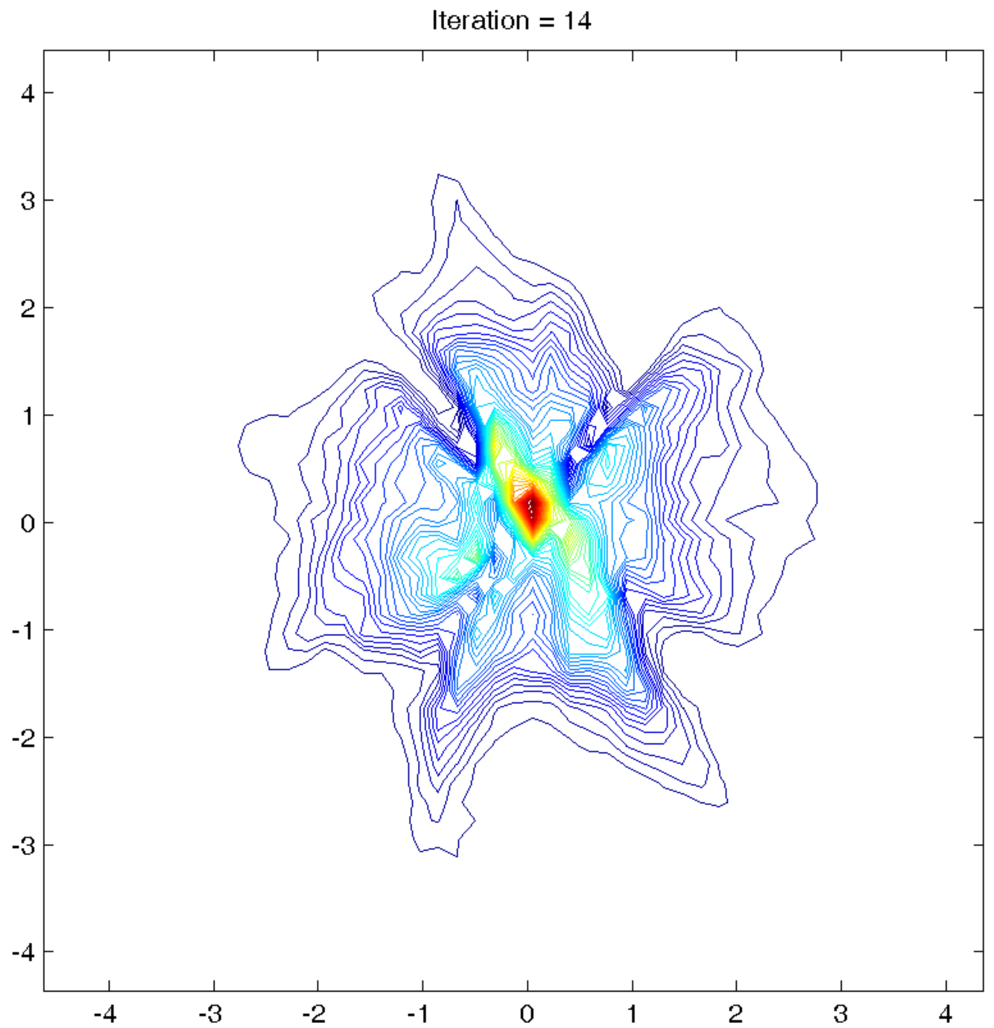
G-PCA



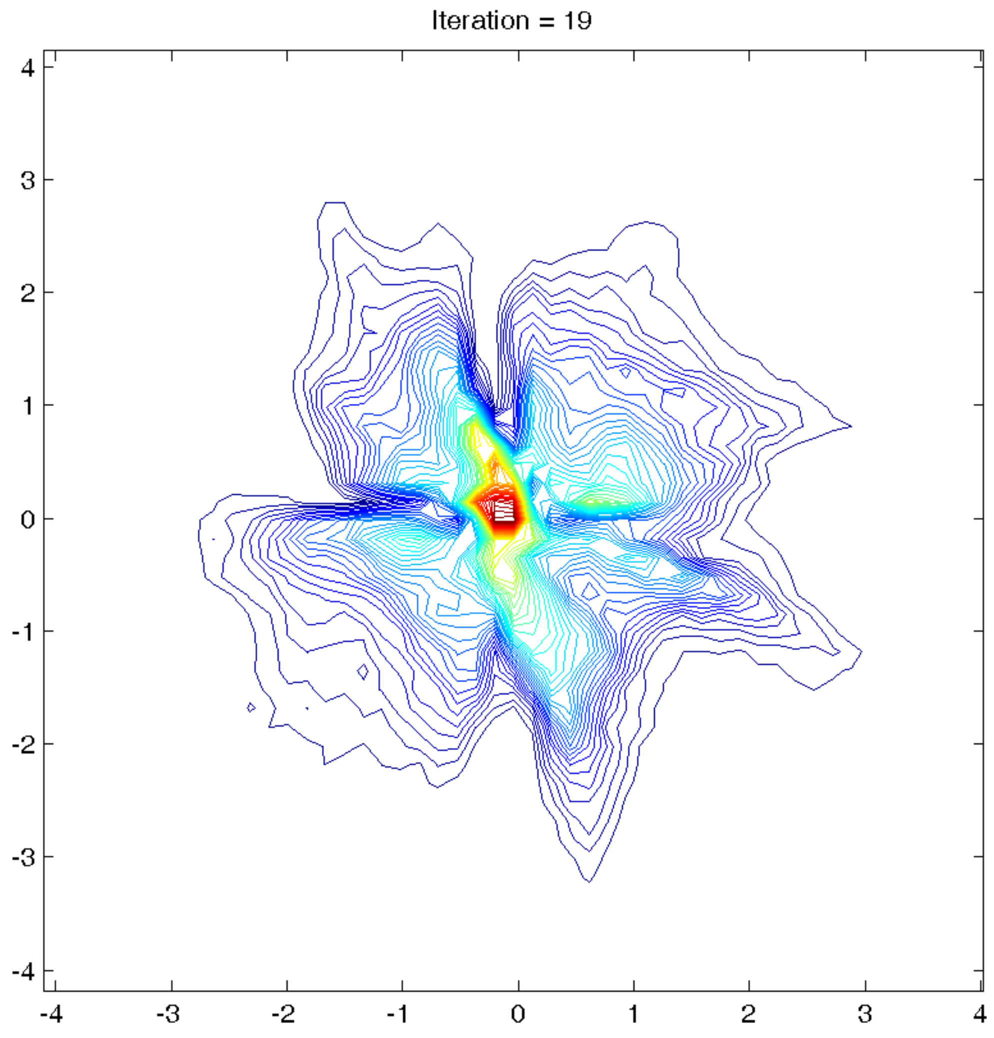
G-PCA



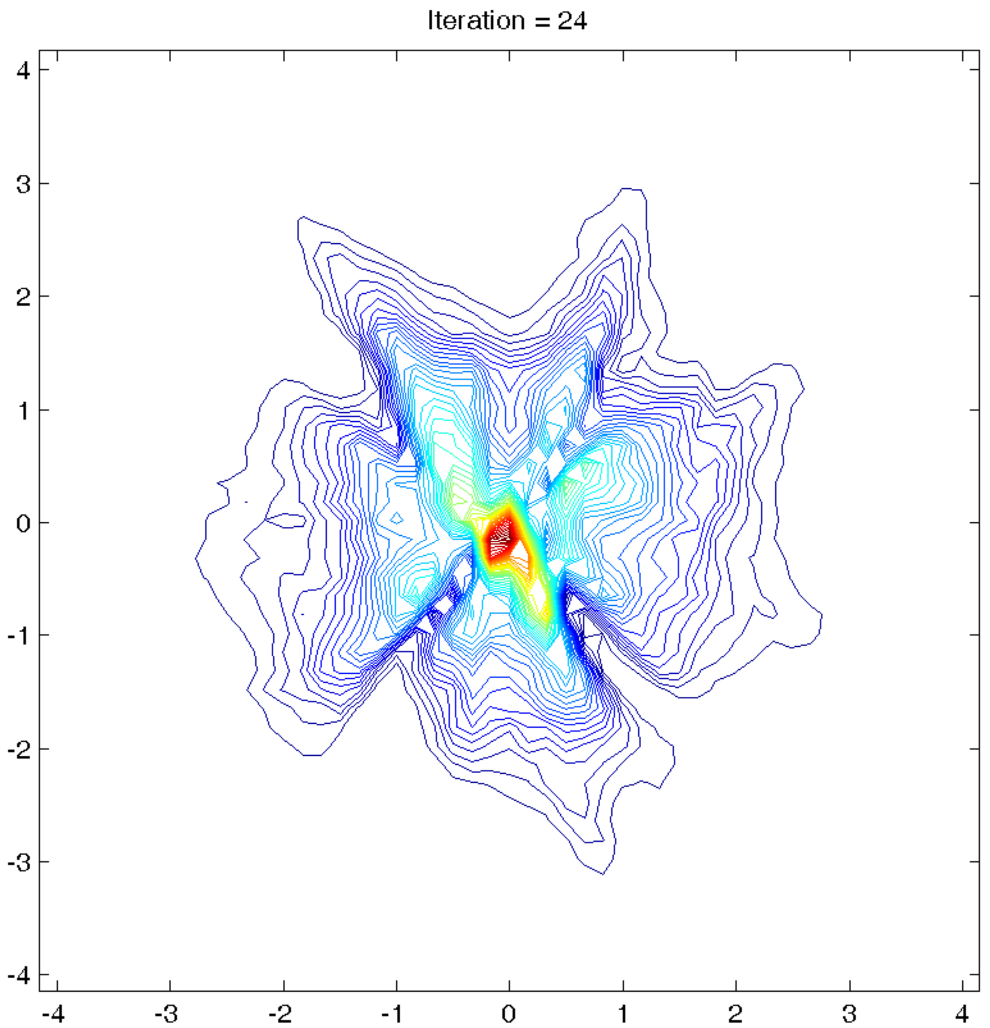
G-PCA



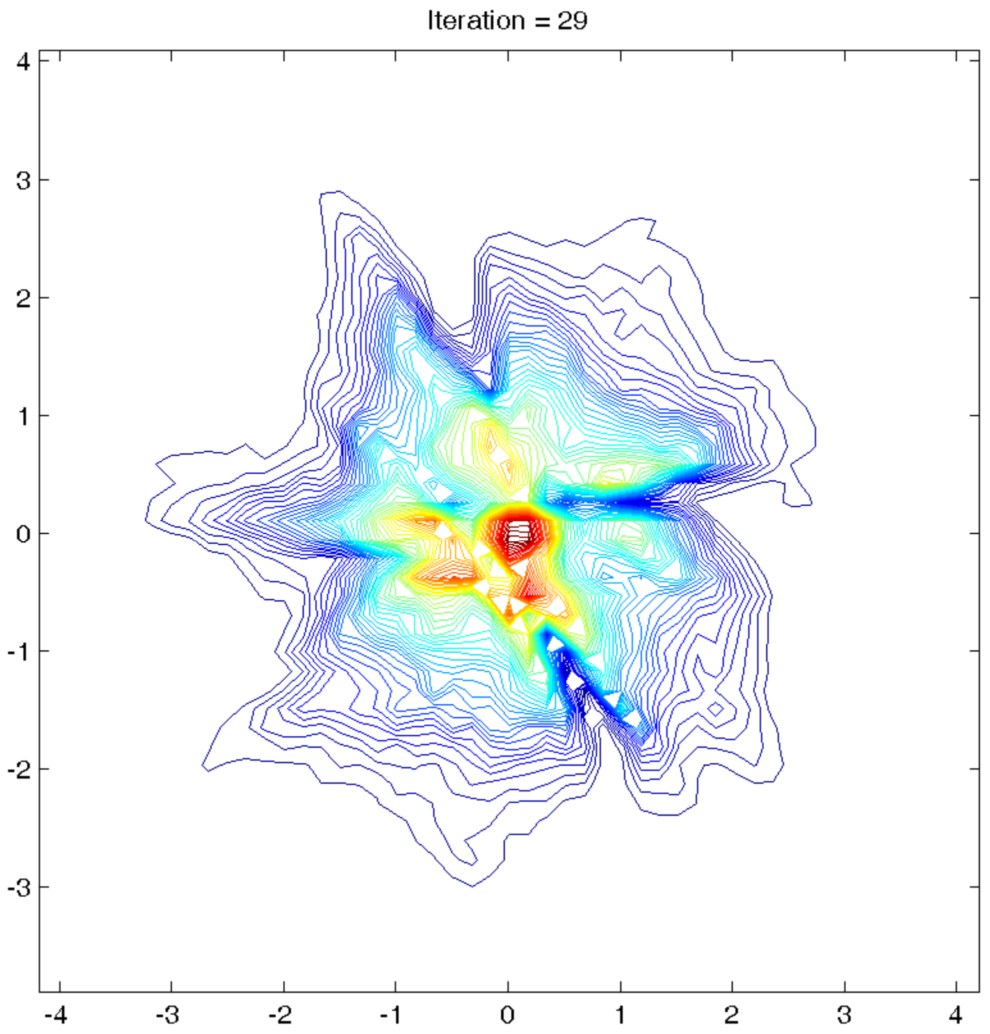
G-PCA



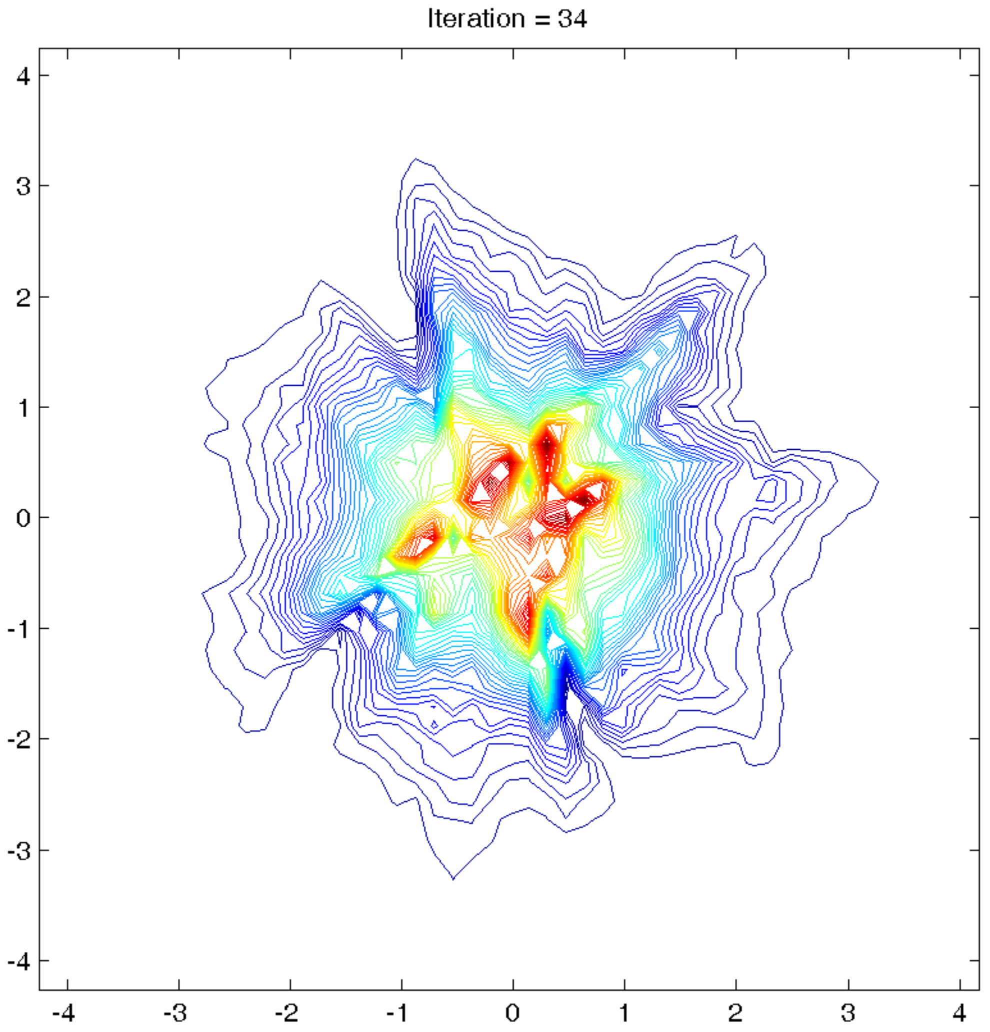
G-PCA



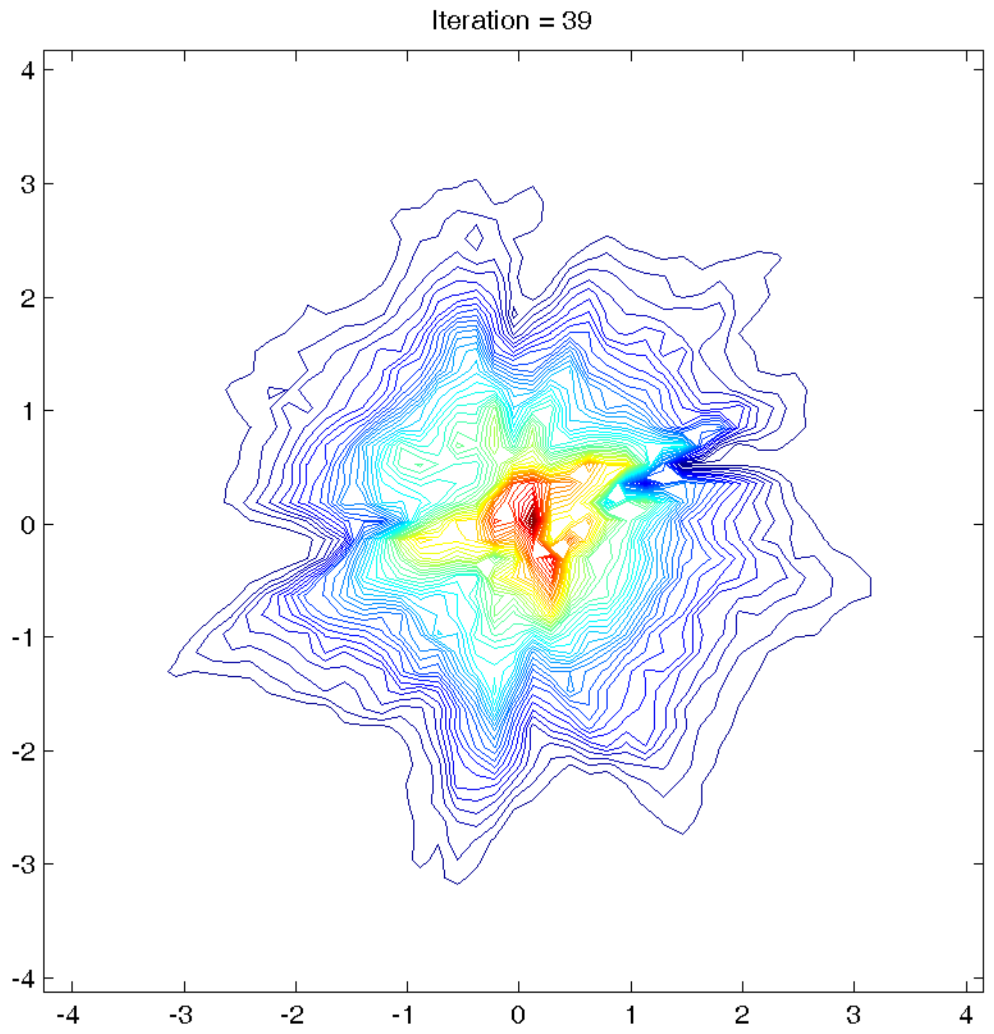
G-PCA



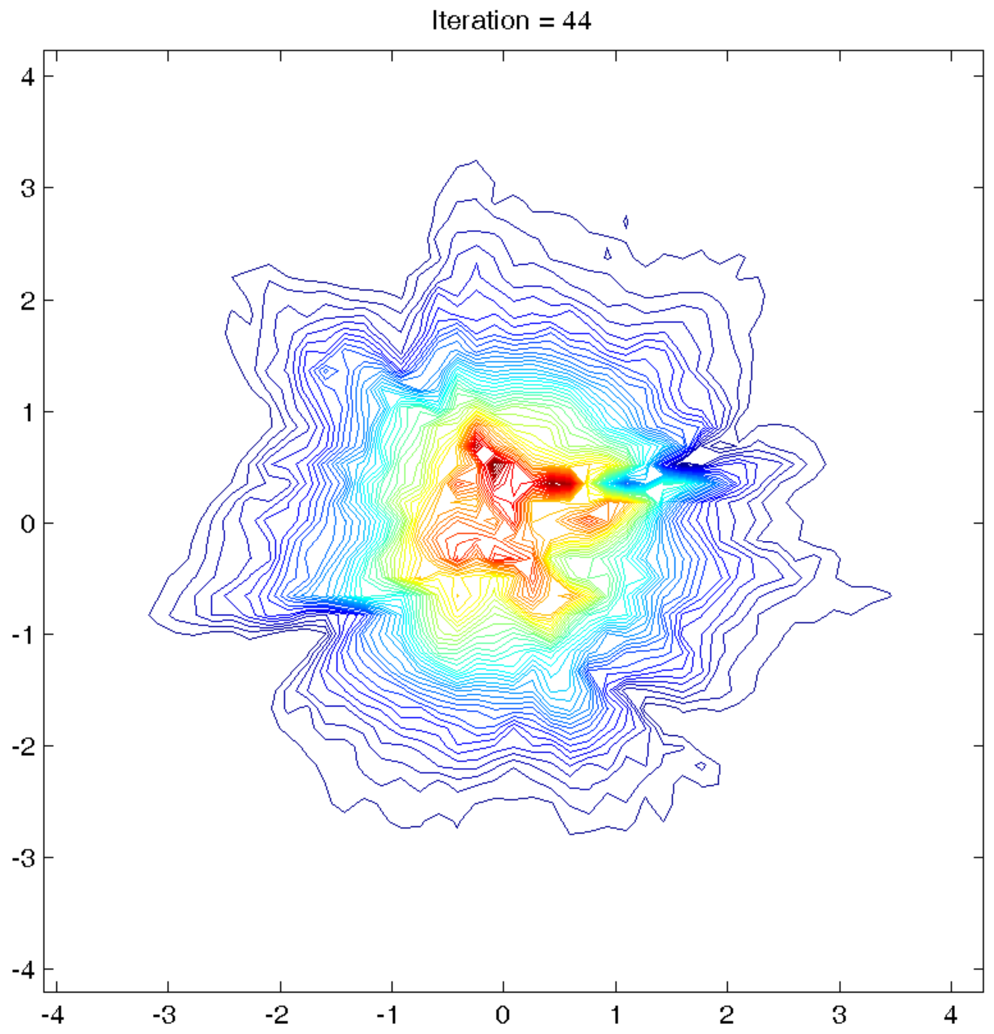
G-PCA



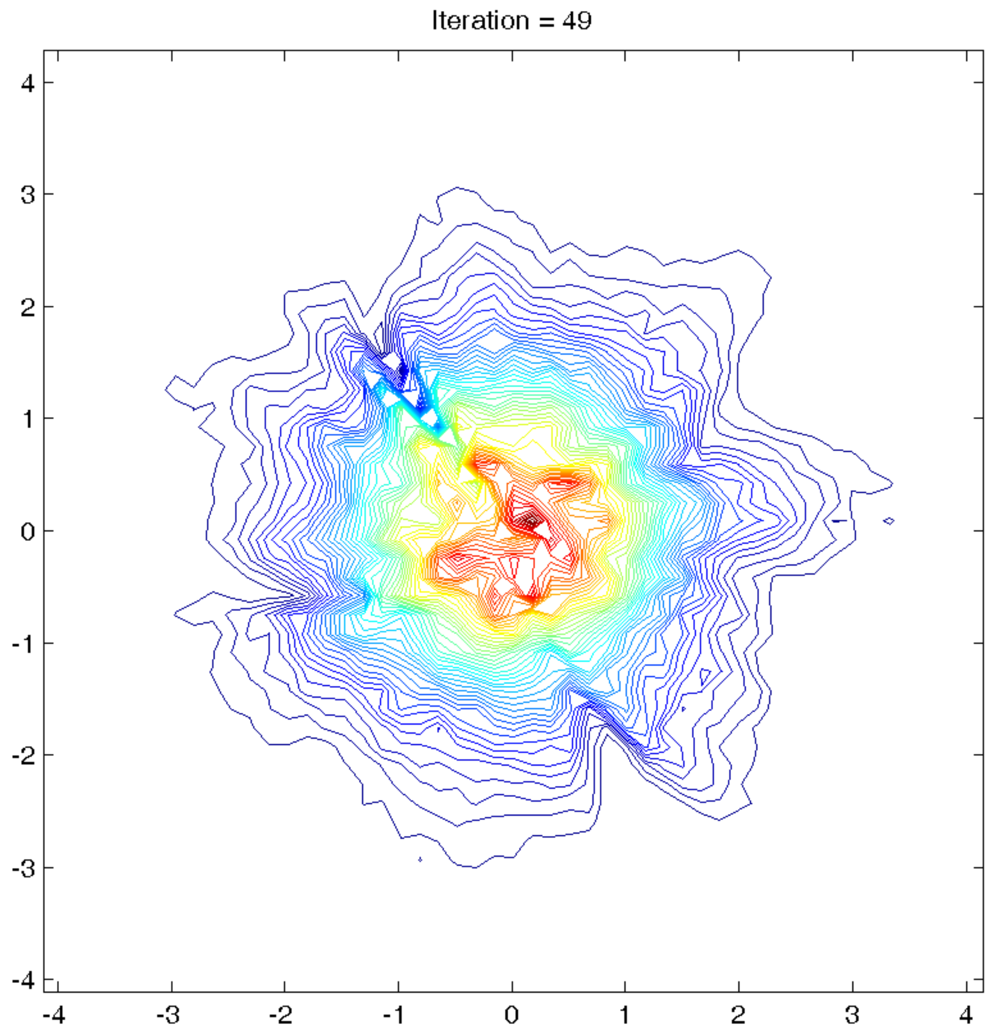
G-PCA



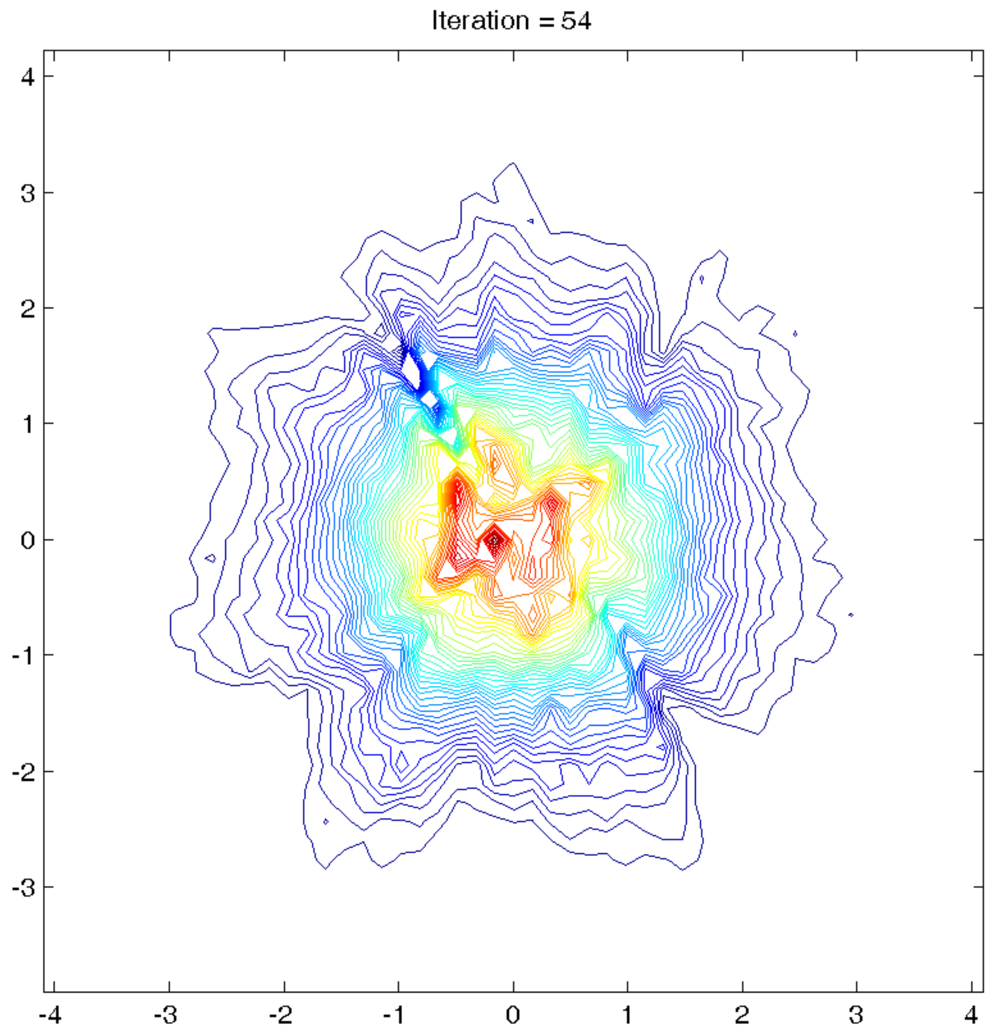
G-PCA



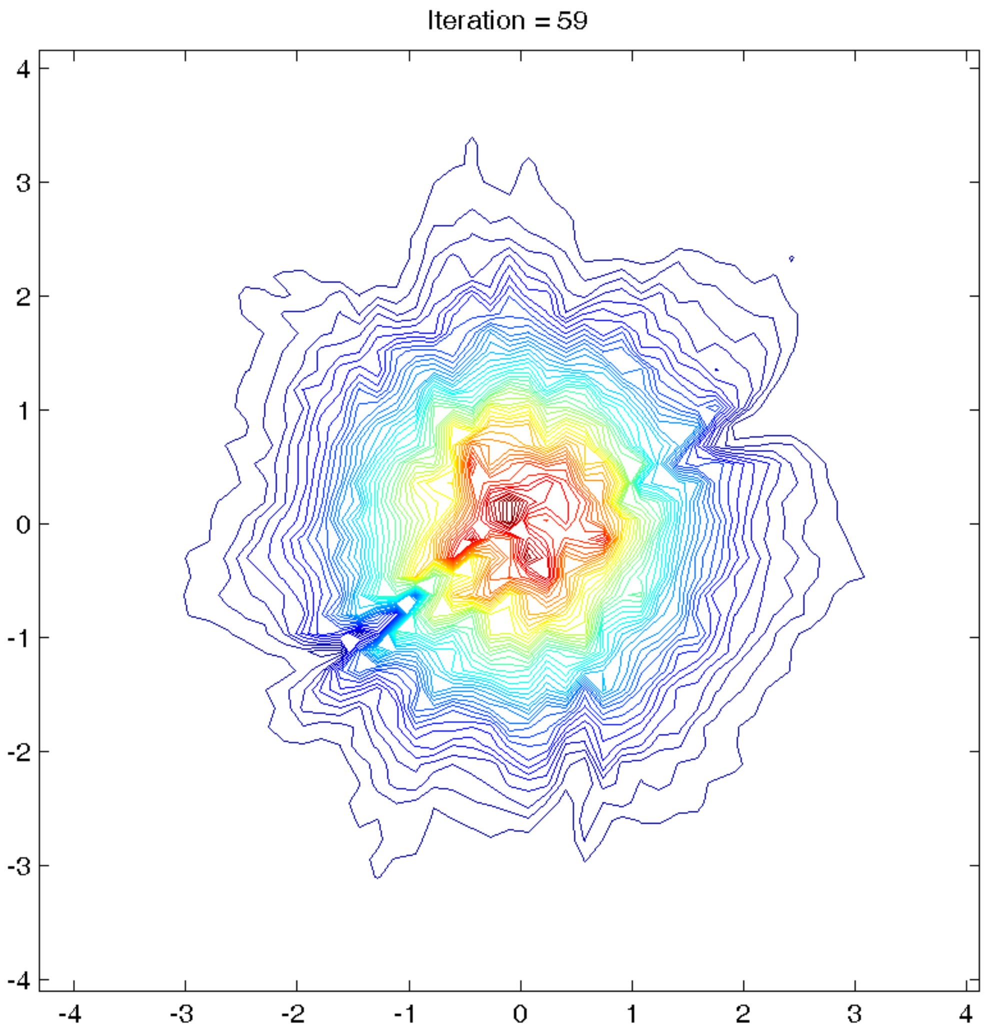
G-PCA



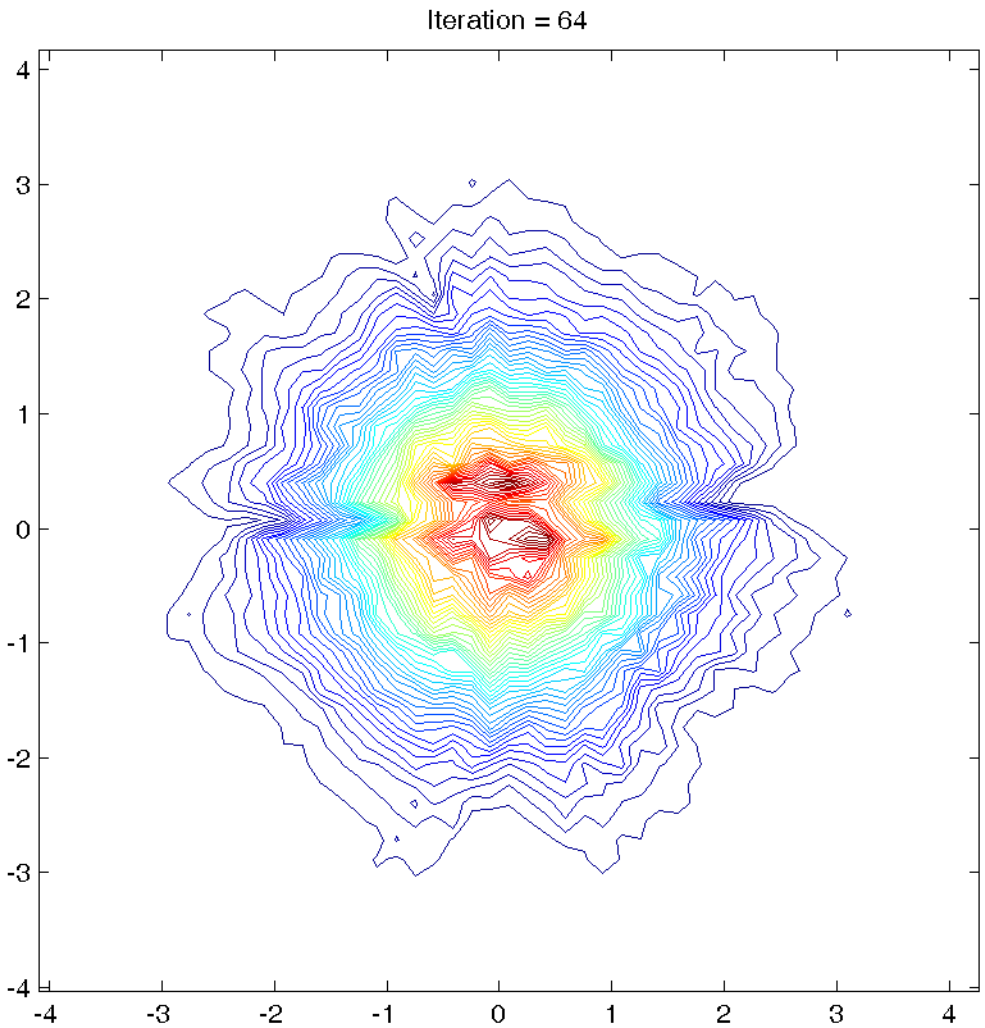
G-PCA



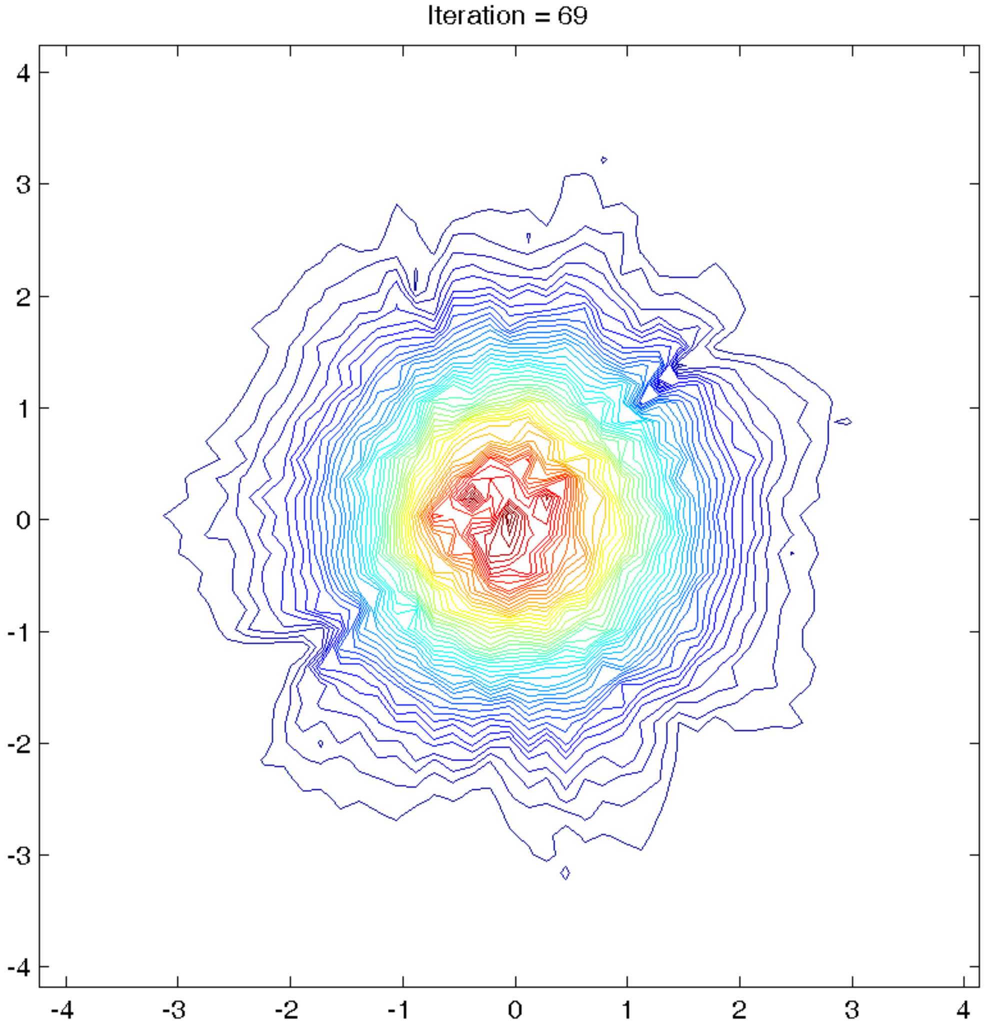
G-PCA



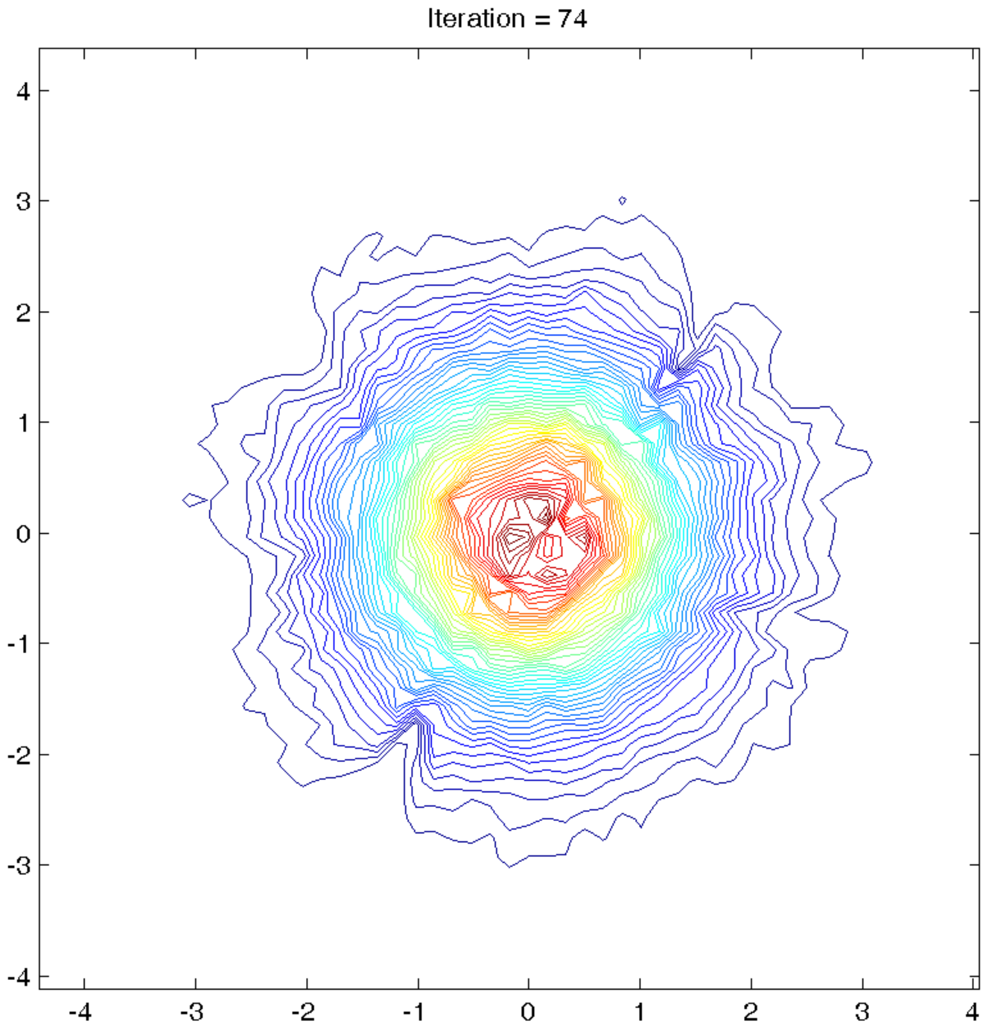
G-PCA



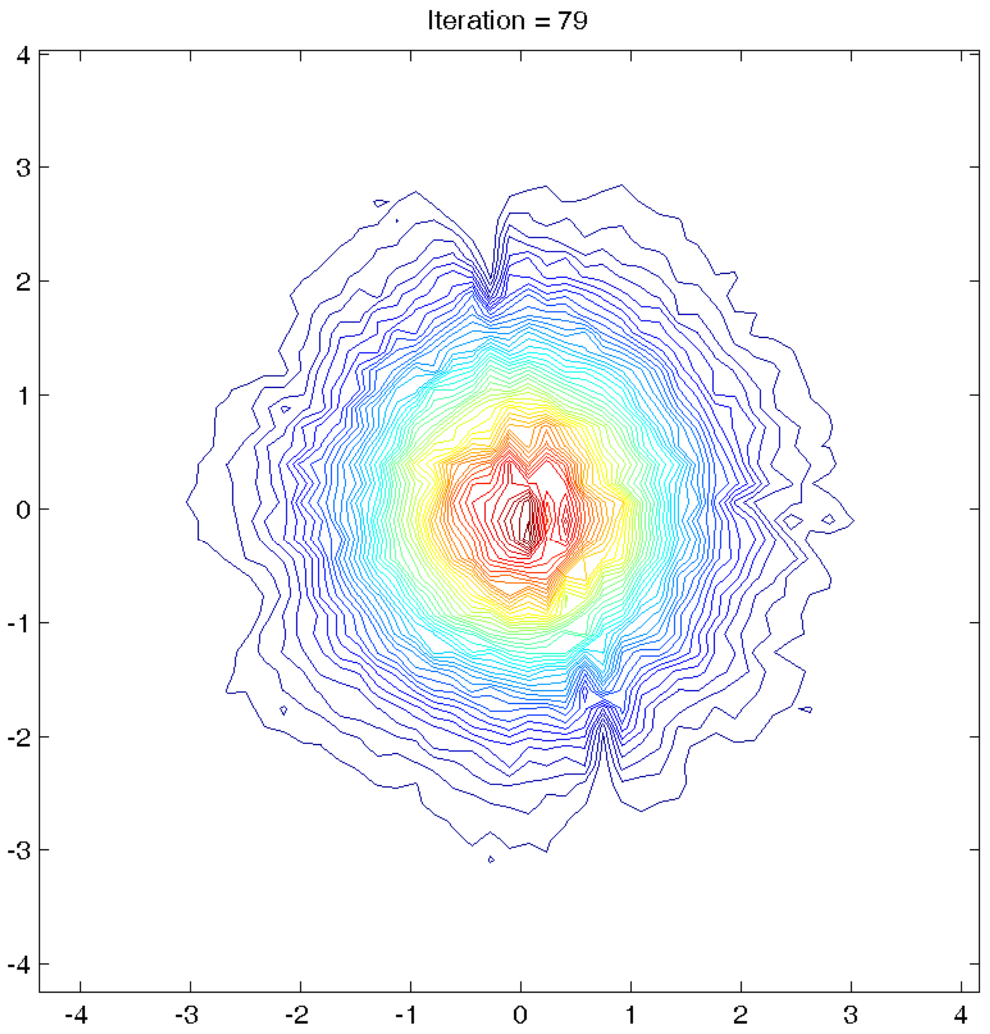
G-PCA



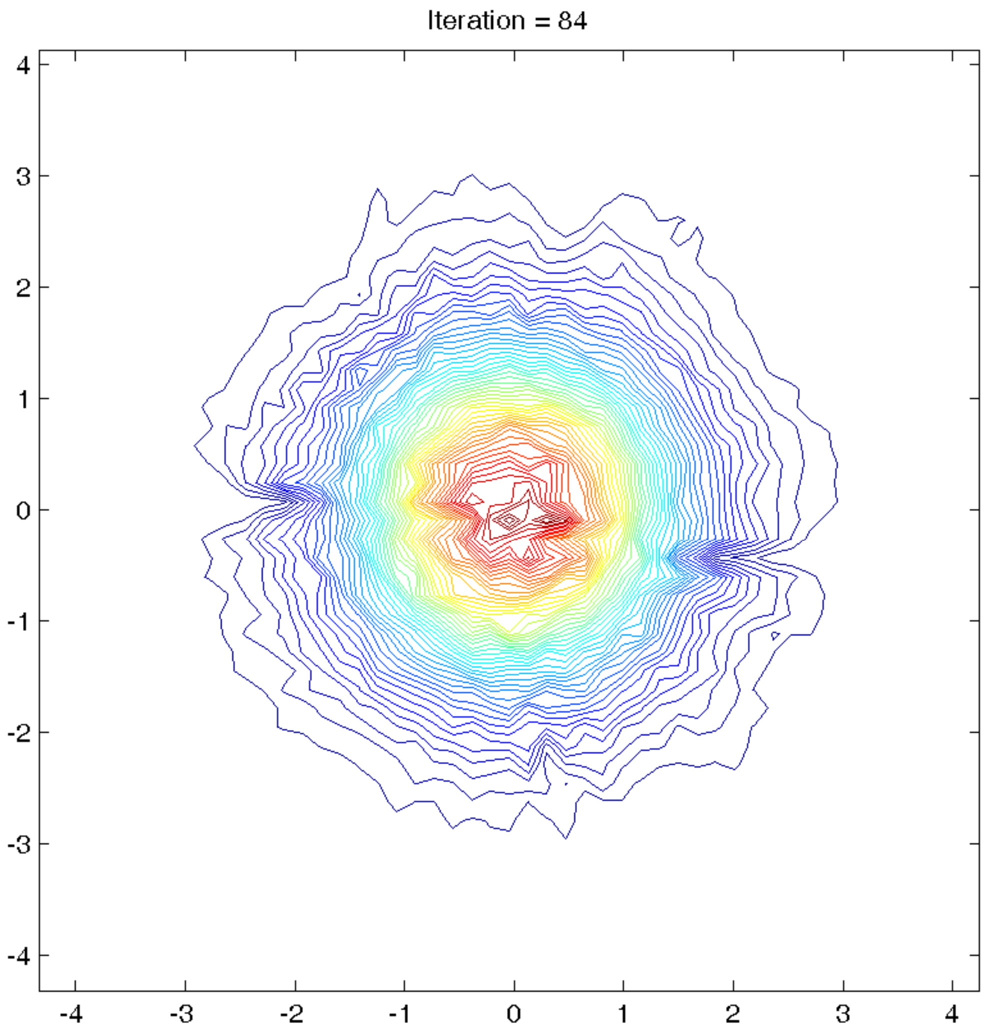
G-PCA



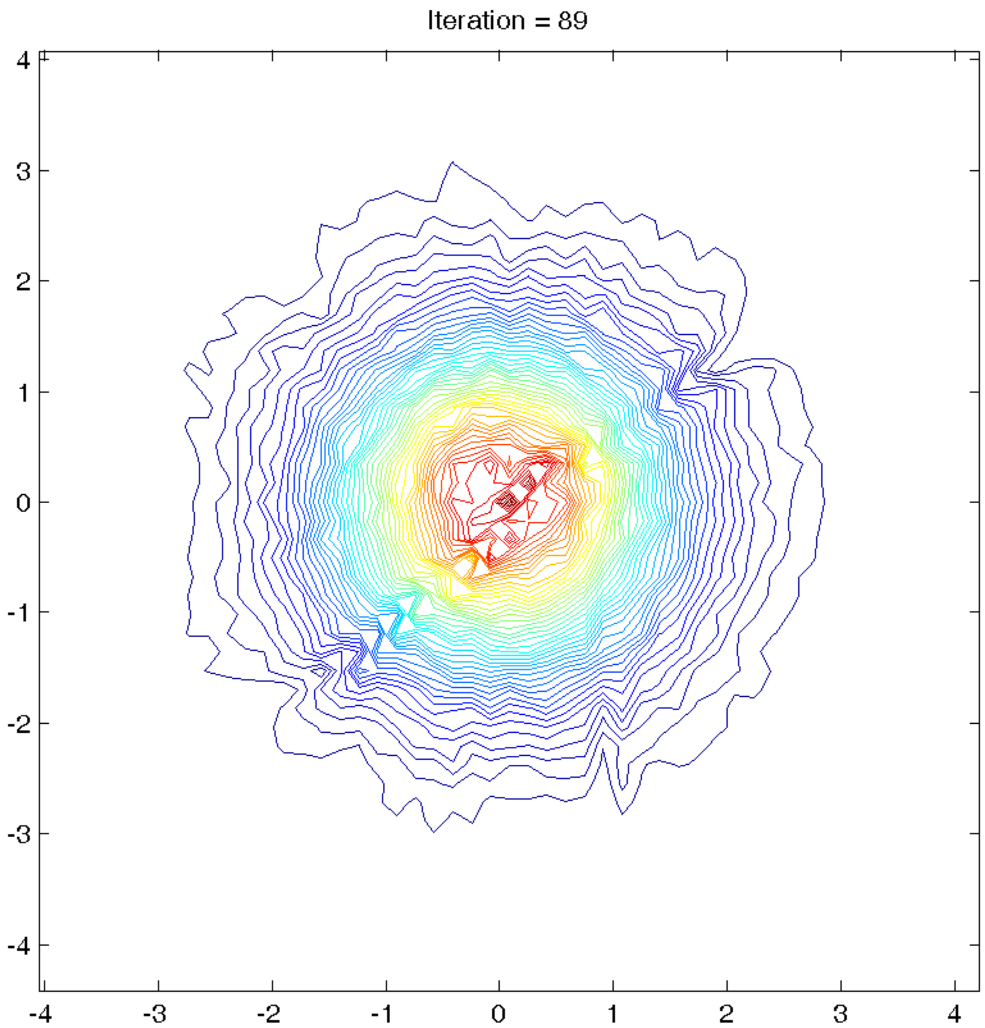
G-PCA



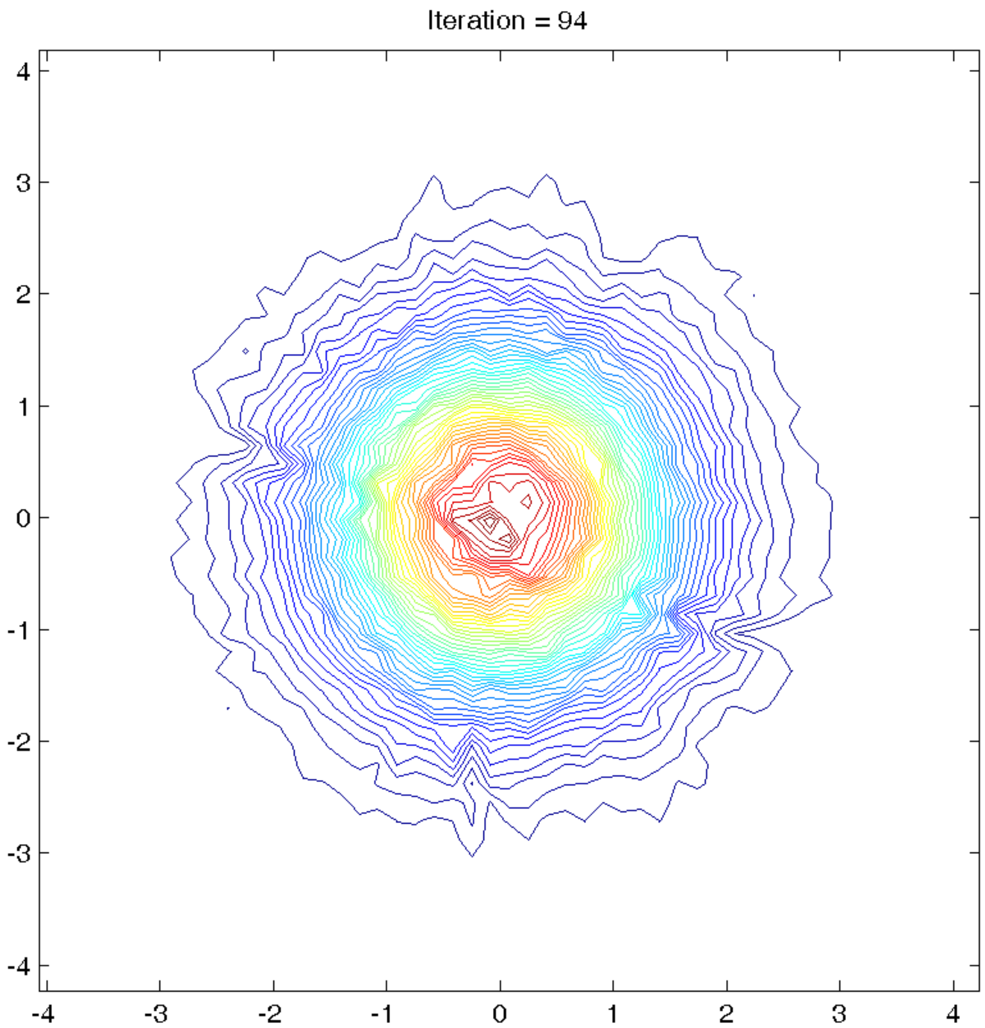
G-PCA



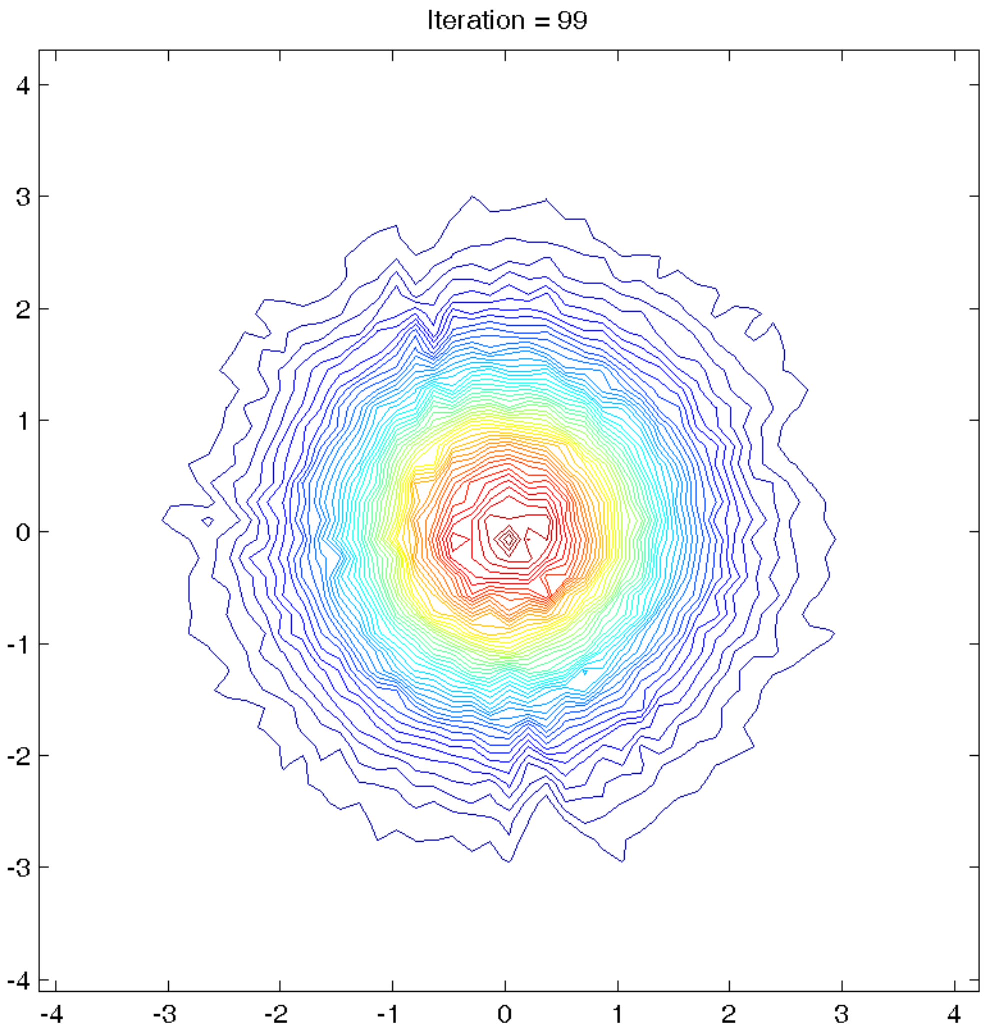
G-PCA



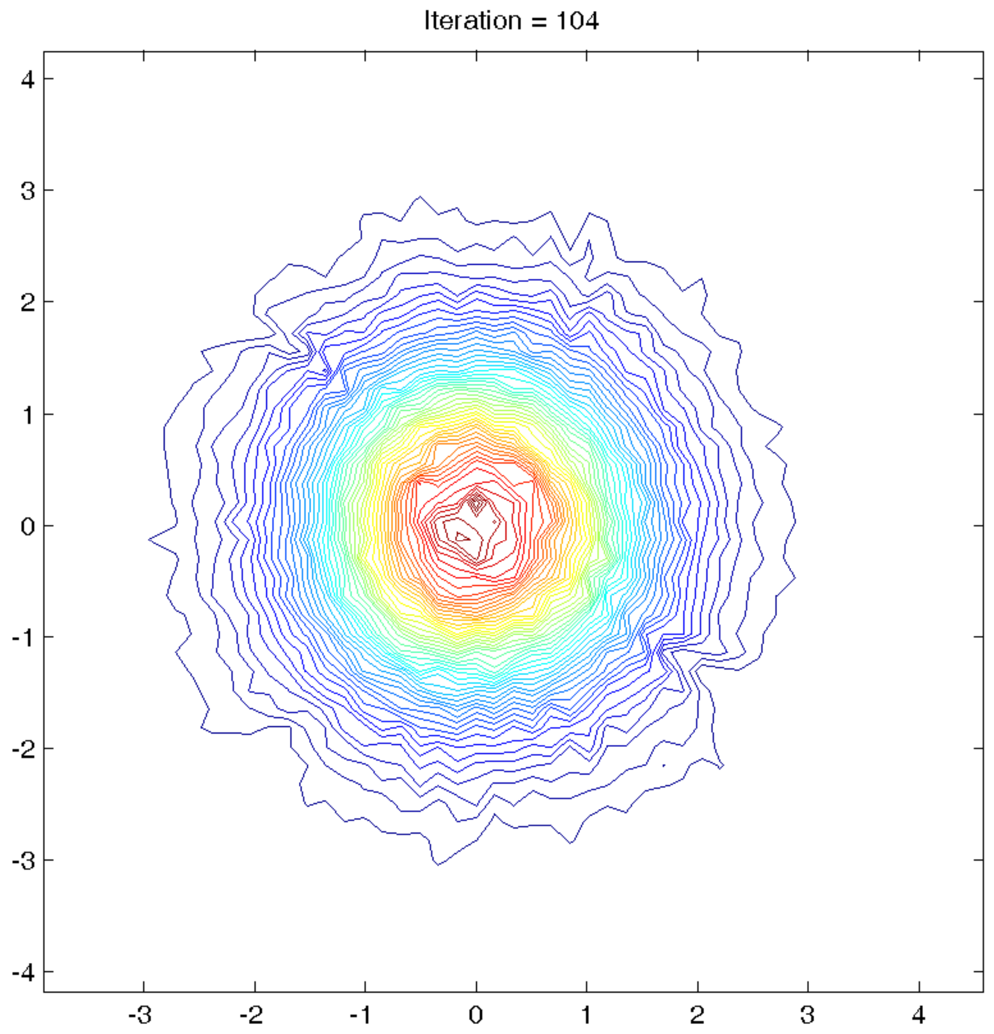
G-PCA



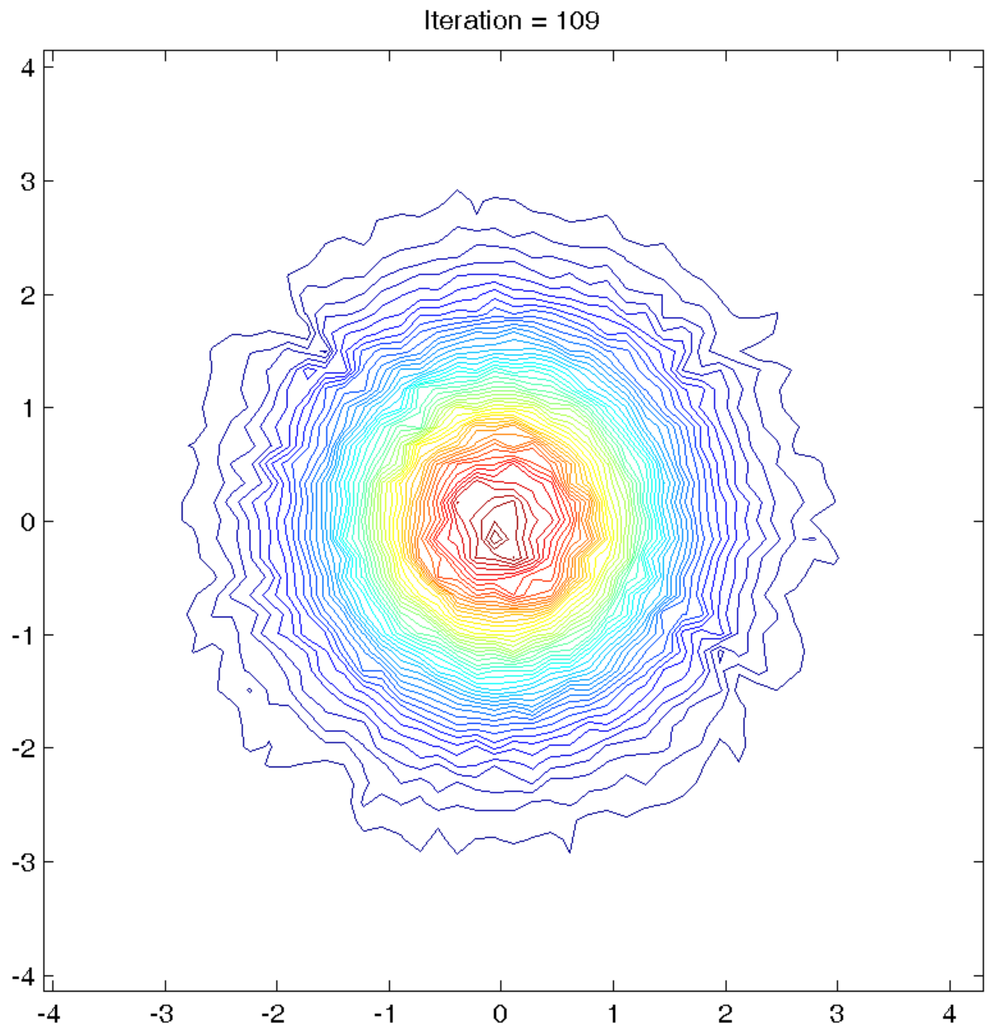
G-PCA



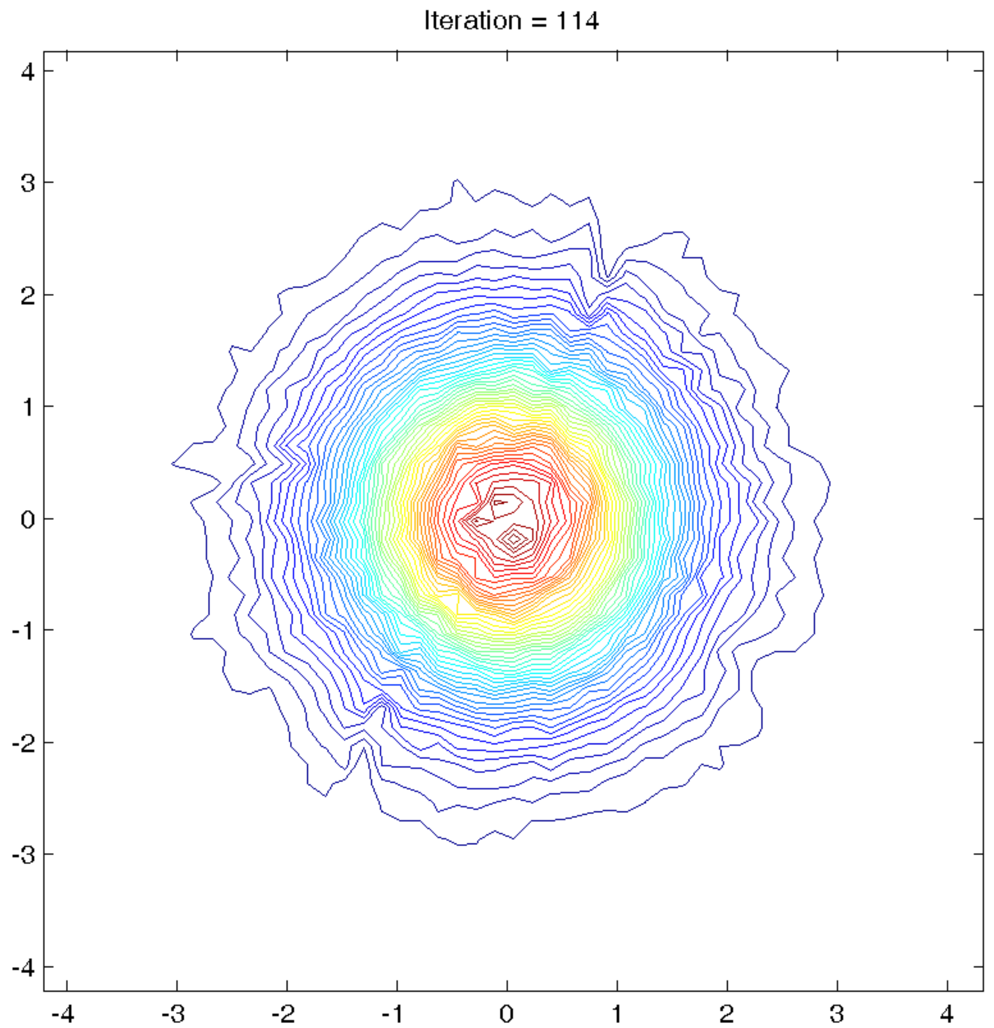
G-PCA



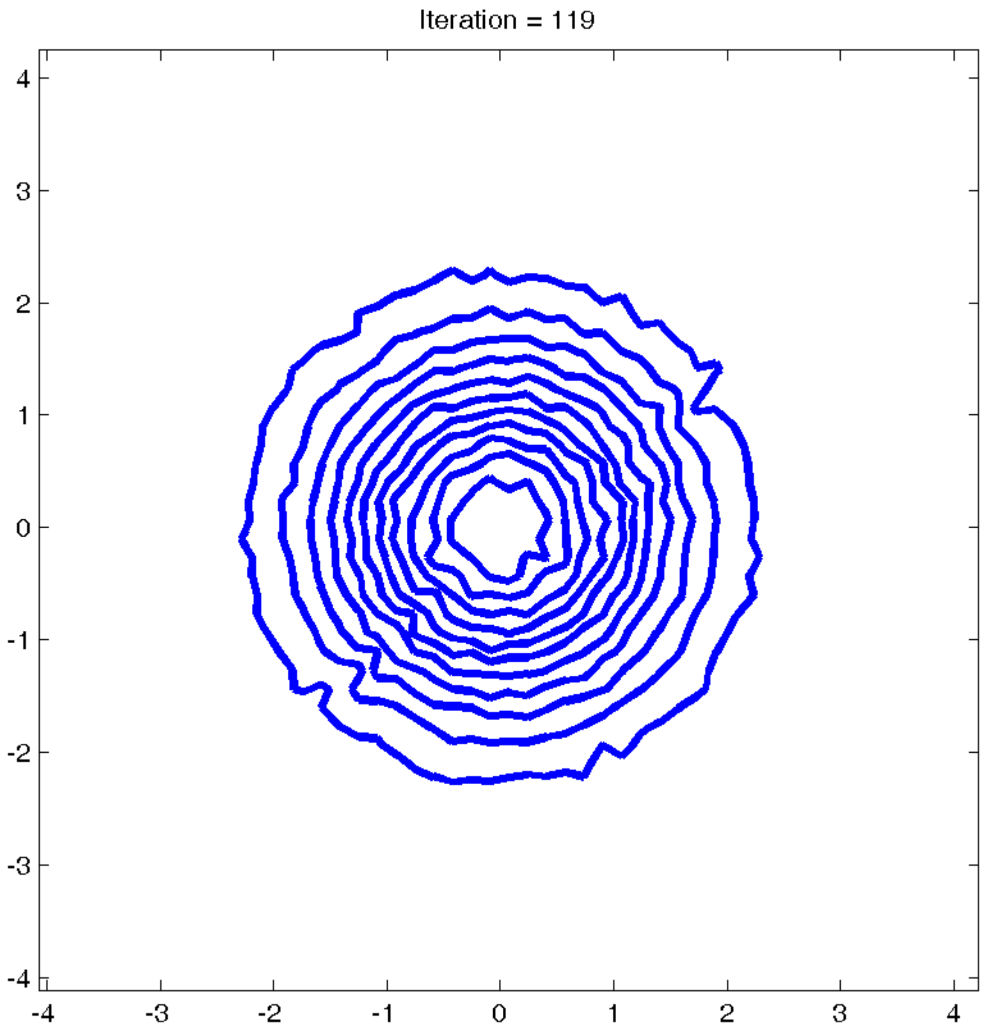
G-PCA



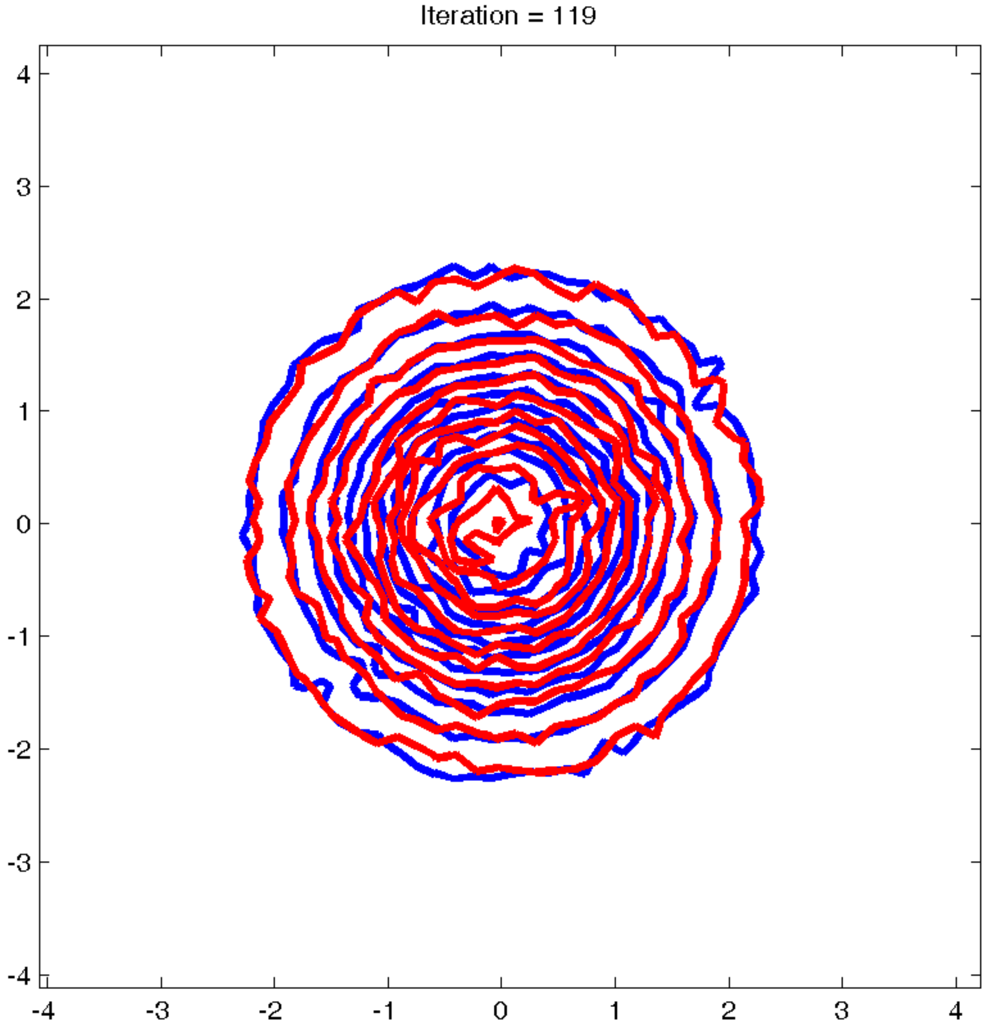
G-PCA



G-PCA



G-PCA

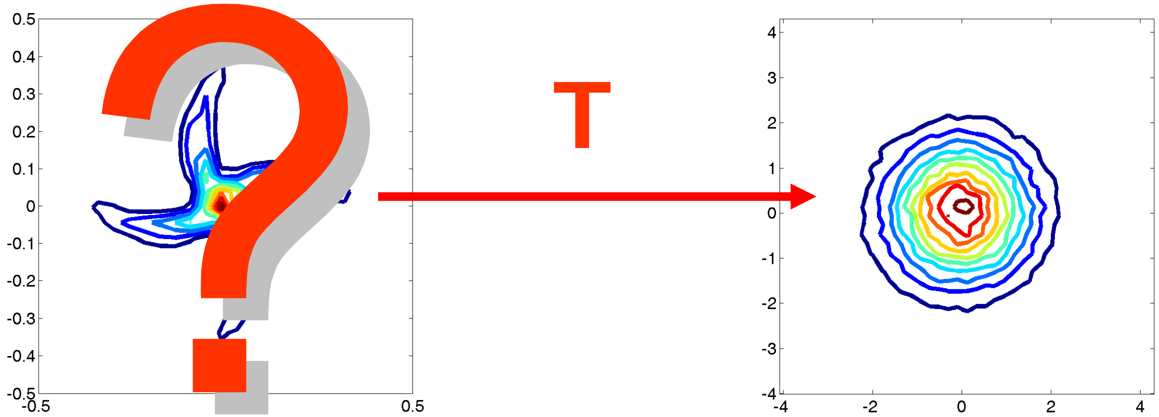


OUR APPROACH

$$Y = T(X)$$

$$P_X = P_Y * |J_T|$$

- From P(x) to P(y) (Gaussian)



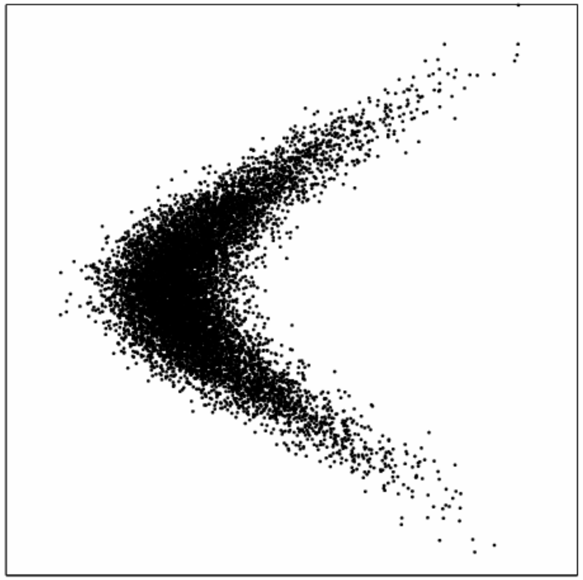
GPCA Jacobian

$$\nabla_{\mathbf{x}} \mathcal{G} = \prod_{k=1}^{N} \mathbf{B}_{(k)} \cdot \nabla_{\mathbf{x}^{(k)}} \Psi_{(k)}$$

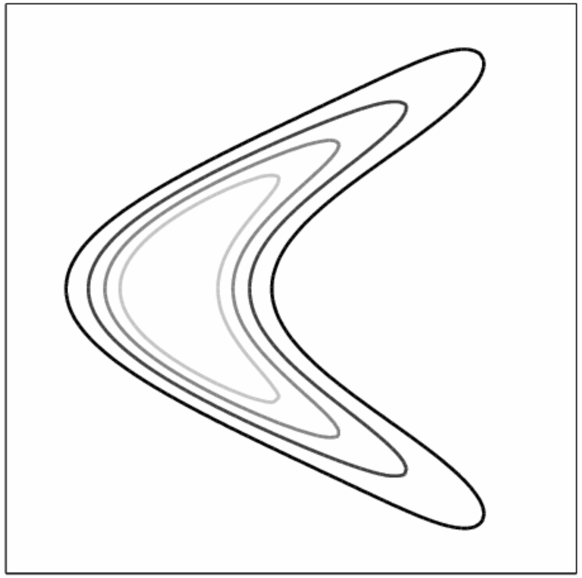
$$\nabla_{\mathbf{x}^{(k)}} \Psi_{(k)} = \begin{pmatrix} \frac{\partial \Psi^1_{(k)}}{\partial x^{(k)}_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{\partial \Psi^d_{(k)}}{\partial x^{(k)}_d} \end{pmatrix}$$

$$\frac{\partial \Psi^i_{(k)}}{\partial x^{(k)}_i} = \frac{\partial G}{\partial u} \cdot \frac{\partial u}{\partial x^{(k)}_i} = \left(\frac{\partial G^{-1}}{\partial x_i} \right)^{-1} \cdot p_i(x^{(k)}_i) = g(\Psi^i_{(k)}(x^{(k)}_i))^{-1} \cdot p_i(x^{(k)}_i)$$

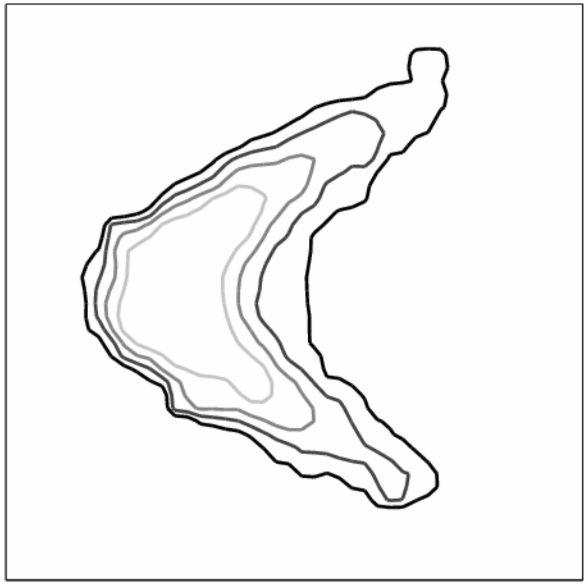
G-PCA – PDF estimation



DATA



ANALYTICAL PDF



G-PCA APPROACH

EXPERIMENTS

Gaussianization-PCA	**vs**	**SVDD**
(PDF)		**(Boundary)**

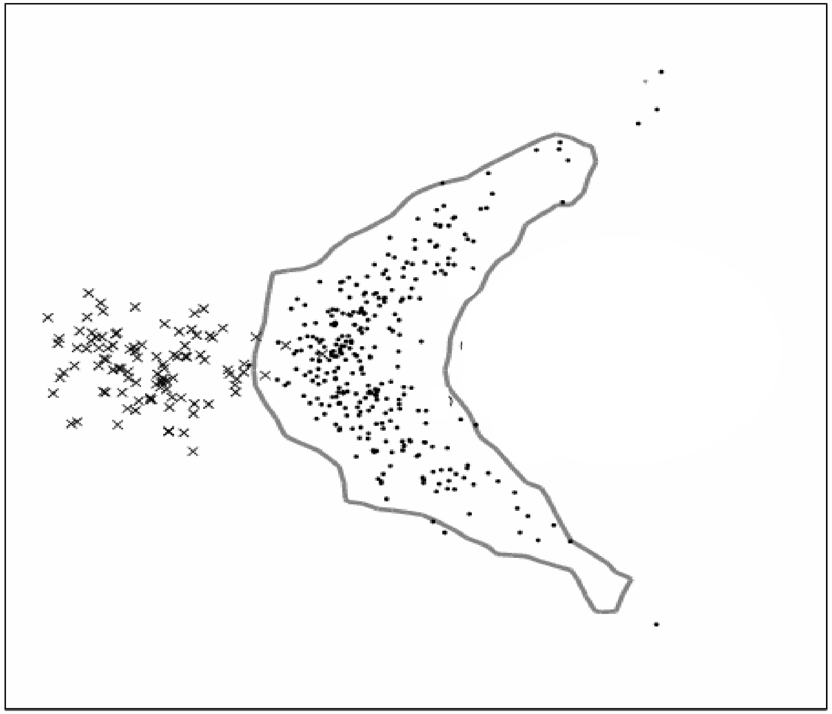
- 2D EXAMPLE (TOY EXAMPLE)

- 10D EXPERIMENT (REAL DATA EXAMPLE)

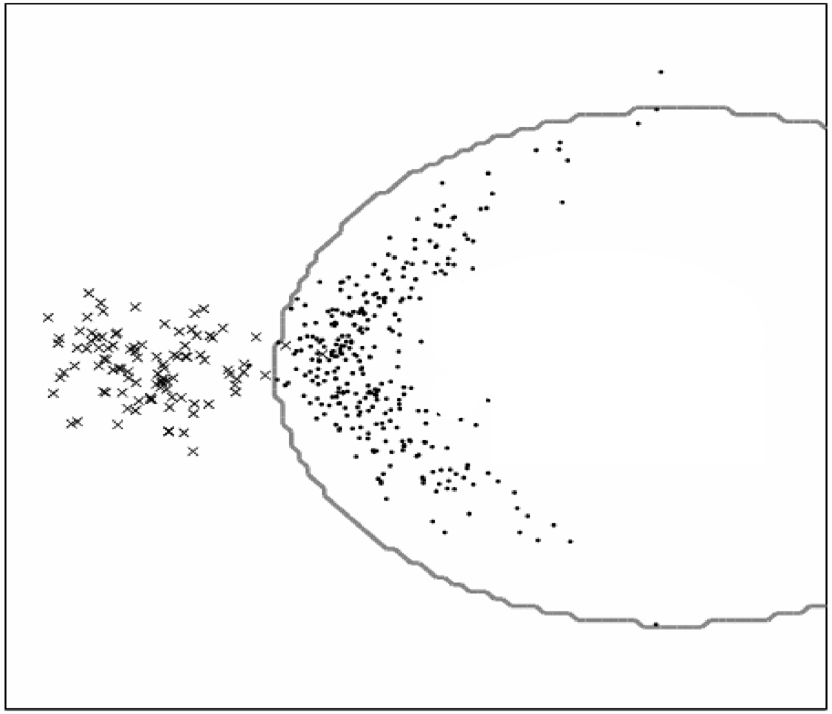
- TRAINING

Parameters: obtained to maximizing Kappa

TOY EXAMPLE GPCA vs SVDD



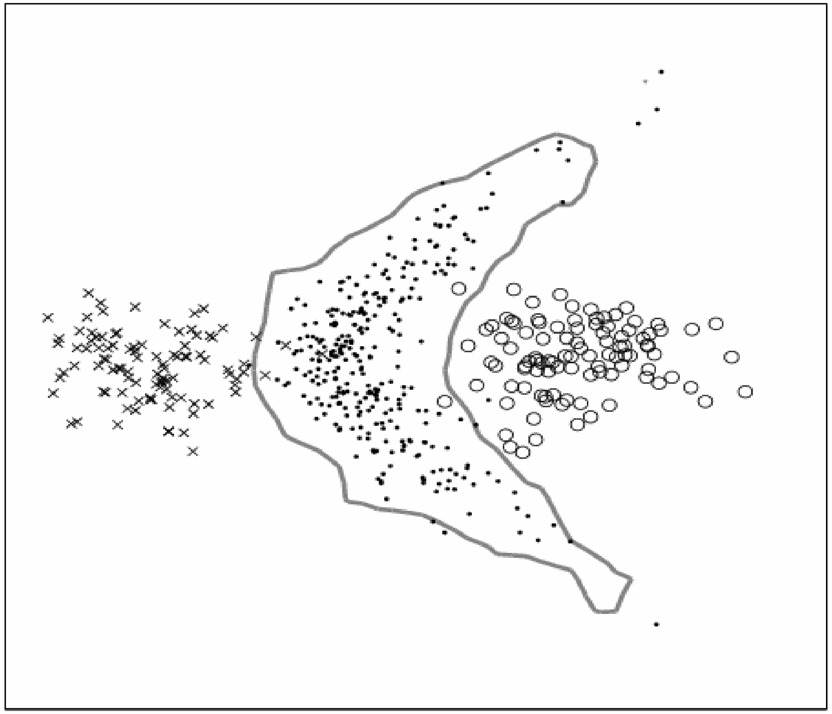
G-PCA



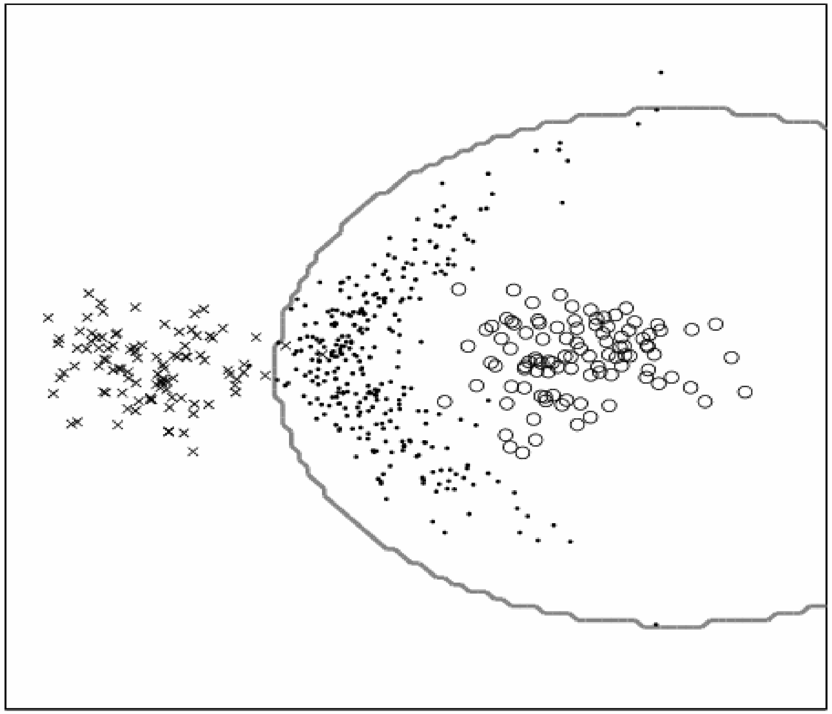
SVDD

- PROBLEM WITH NO TARGET SAMPLES
- PROBLEM WITH THE REPRESENTATIVITY OF NO TARGET SAMPLES

TOY EXAMPLE GPCA vs SVDD



G-PCA



SVDD

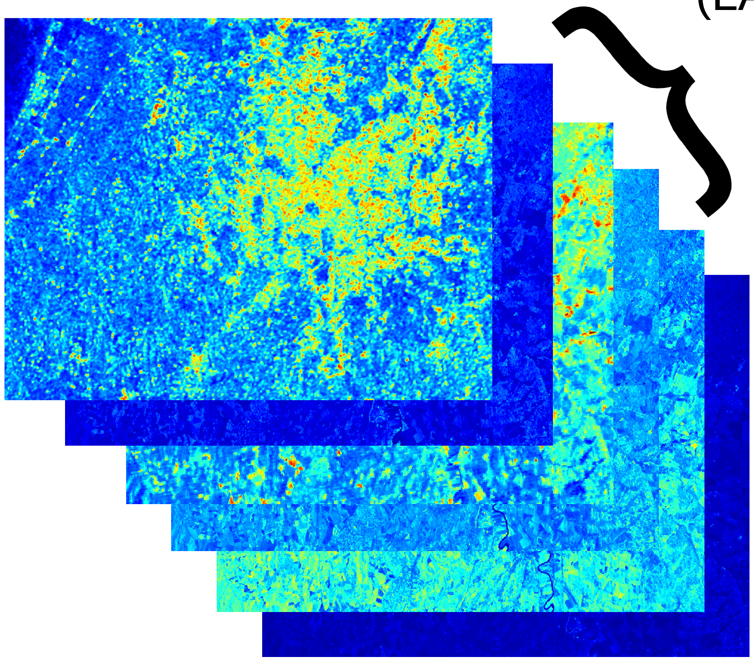
- PROBLEM WITH NO TARGET SAMPLES
- PROBLEM WITH THE REPRESENTATIVITY OF NO TARGET SAMPLES

10D EXPERIMENT RESULTS

- REAL DATA: 3 MULTISOURCE IMAGES

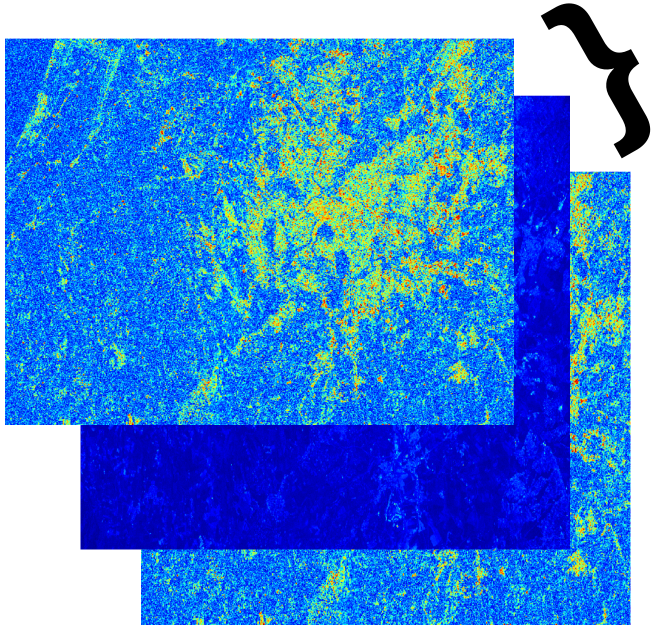
7 OPTIC BANDS

(LANDSAT)

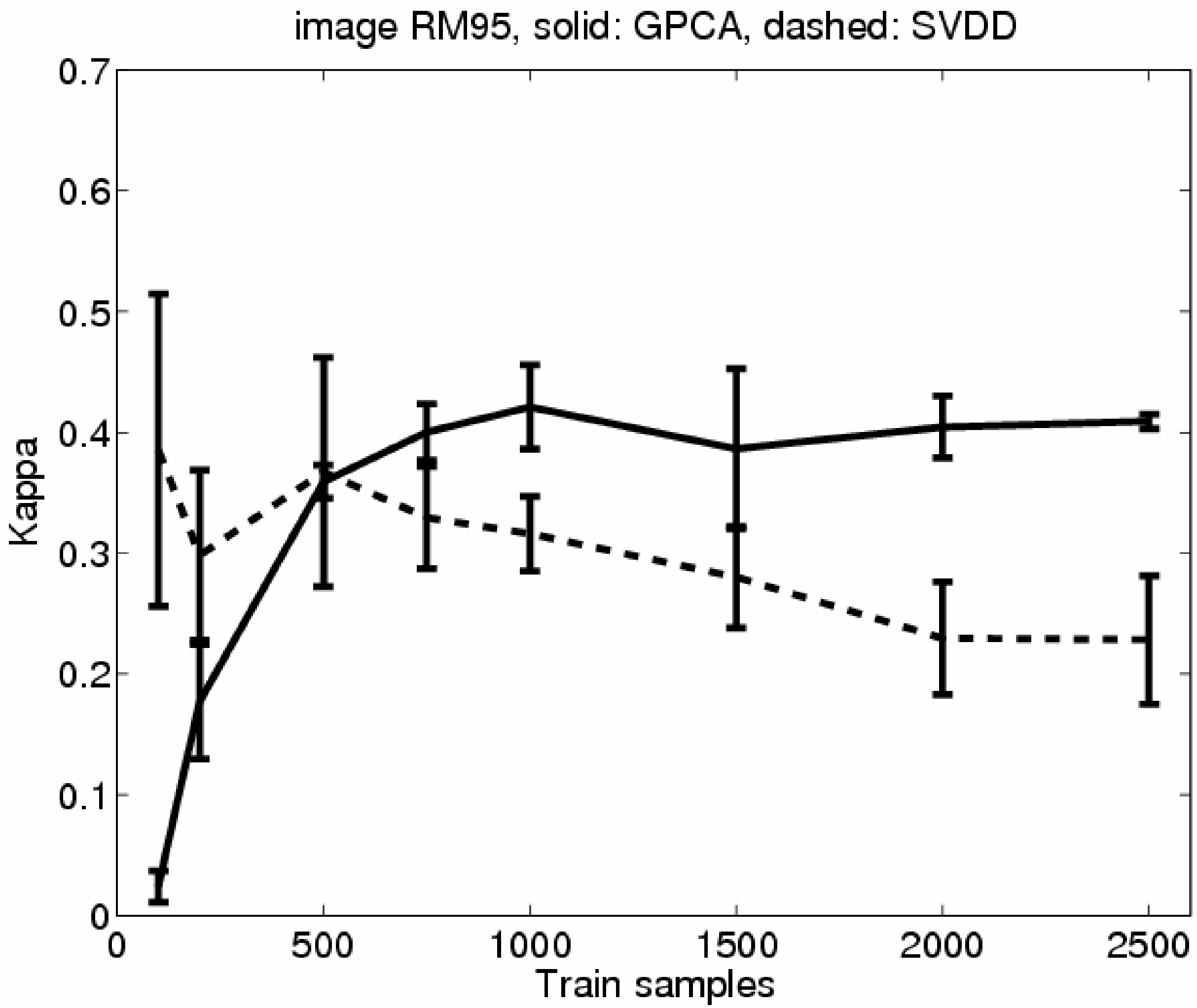


3 SAR BANDS

(ERS)



10D EXAMPLE RESULTS



Test region:

930 x 1440 pixels

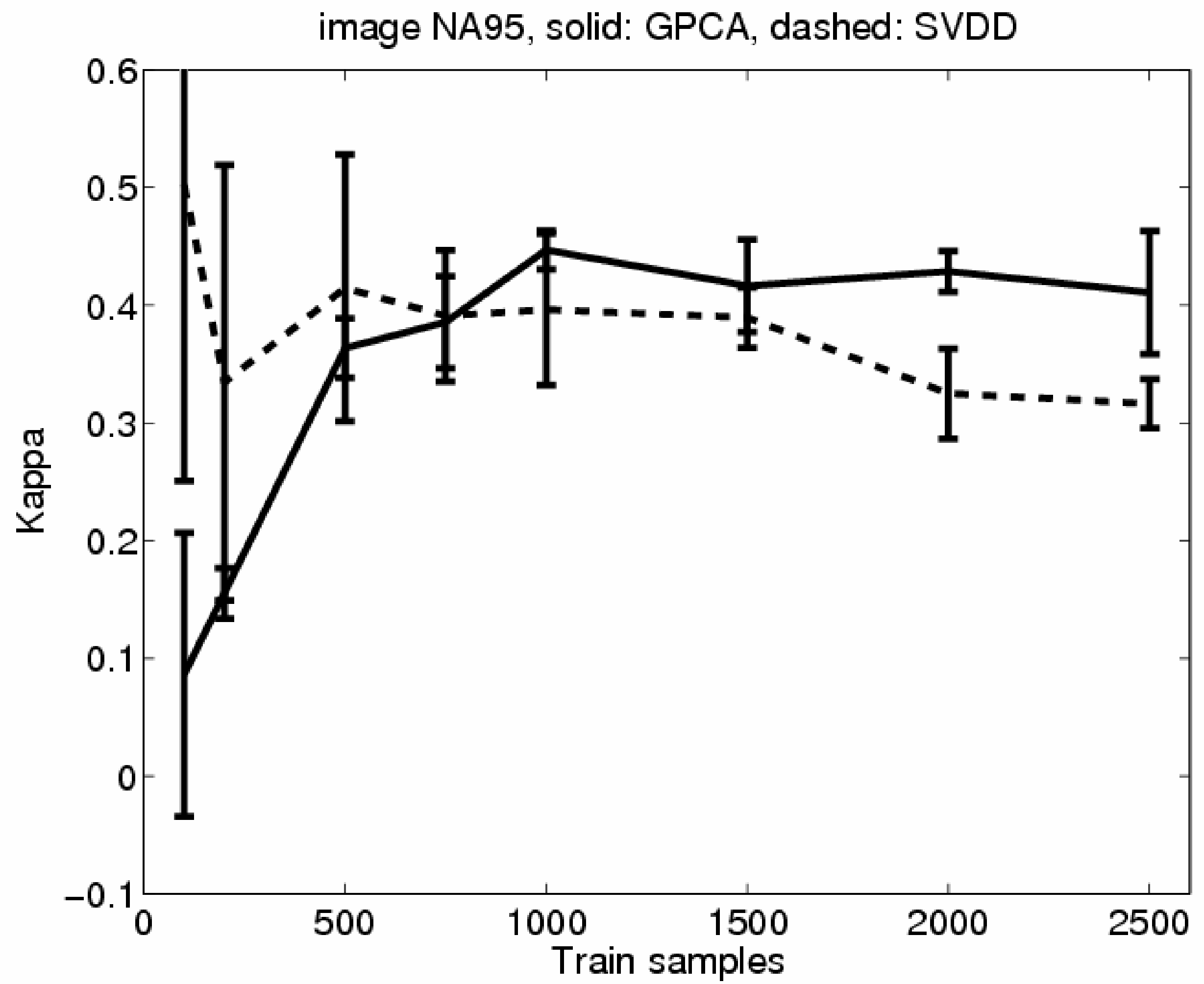
Test samples:

100.000

No-target samples:

10

10D EXAMPLE RESULTS



Test region:

930 x 1440 pixels

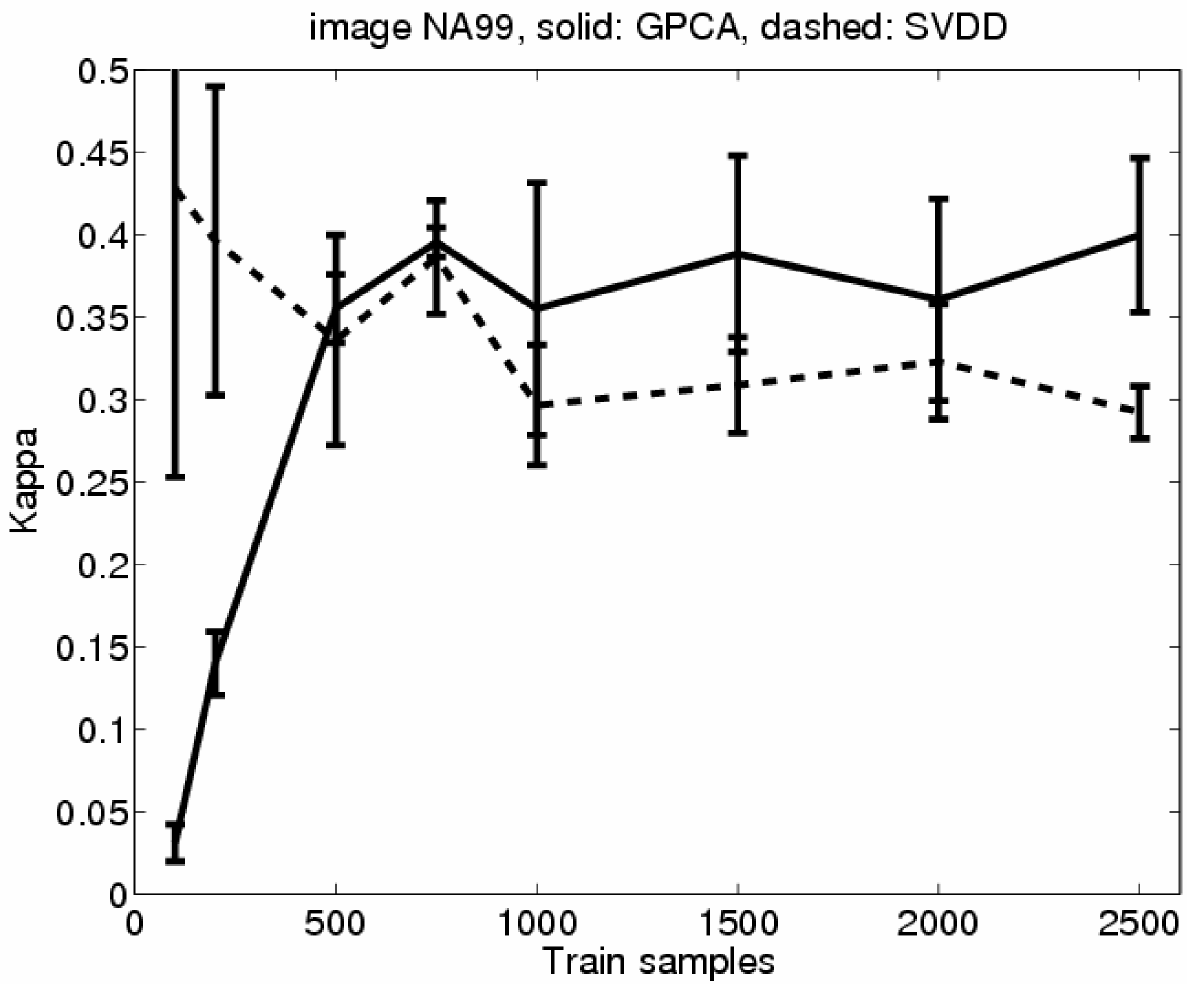
Test samples:

100.000

No-target samples:

10

10D EXAMPLE RESULTS



Test region:

930 x 1440 pixels

Test samples:

100.000

No-target samples:

10

10D EXAMPLE RESULTS

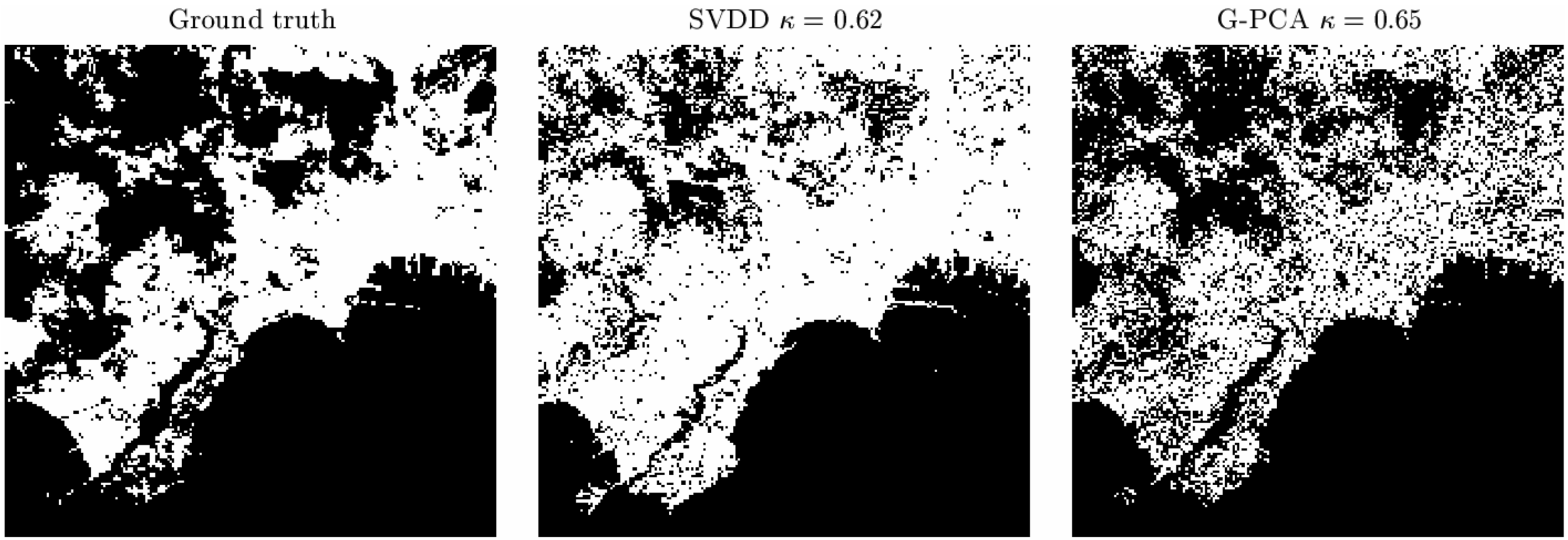
Test region: 200 x 200 pixels

Test samples: 40.000

No-target samples: 10

Target samples: 2.000

Image NA99



CONCLUSIONS

- Gaussianization-PCA can be used to characterize the PDF of a data without assuming a explicit model.
- Boundary estimation (SVDD) vs PDF estimation (GPCA):
  - Boundary estimation method works better with few samples, however it needs samples of non-target class.
  - PDF estimation characterizes better the target class without non-target class information.

FURTHER WORK

- Include spatial information.
- Test the performance of the algorithms with different number of non-target samples.
- Compare with other one-class methods (Boundary and PDF).