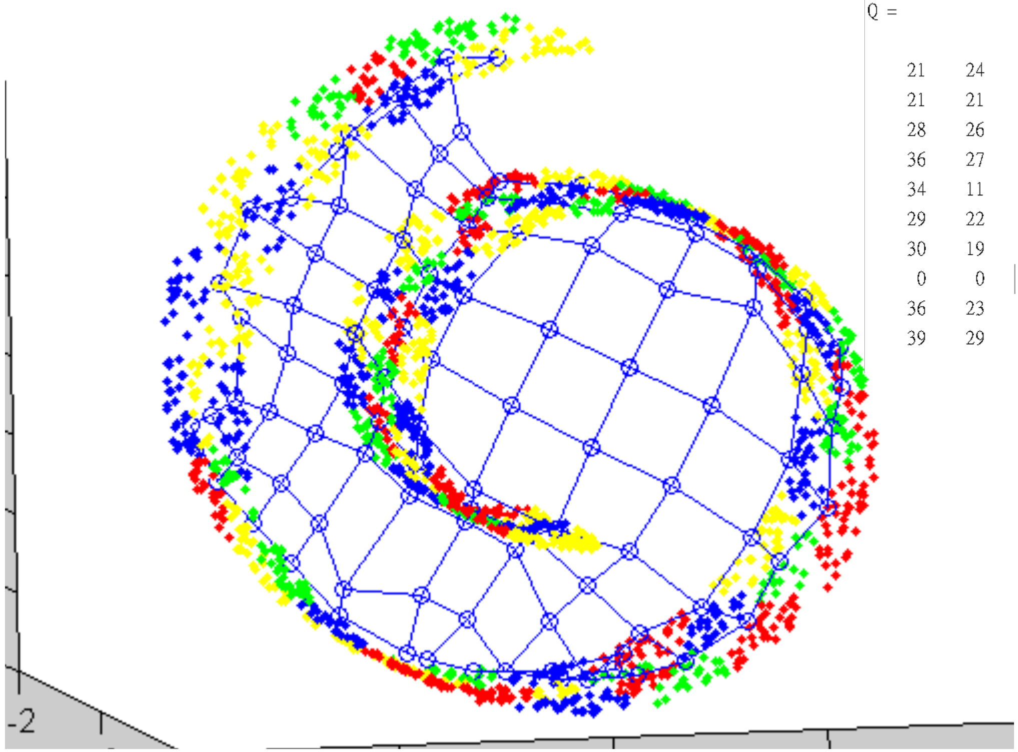
Dimensionality Reduction
Visualization and approximation



Q =

21	24	26	36	17	30	22	20	8	29
21	21	31	28	34	28	0	30	12	23
28	26	30	0	0	0	0	0	26	24
36	27	0	0	0	62	26	22	0	25
34	11	13	0	0	39	29	32	0	33
29	22	36	28	32	33	20	29	28	25
30	19	16	16	19	27	16	27	0	34
0	0	0	31	15	22	26	0	0	22
36	23	24	0	0	0	0	0	16	25
39	29	29	17	38	25	36	27	15	35

Dimensionality Reduction and Visualization

Self-organization and Lattice-connected Gaussian mixtures

Self-organization and associative memory

Teuvo Kohonen

Self-Organization and Associative Memory

Third Edition

With 100 Figures

Springer-Verlag Berlin Heidelberg New York London Paris Tokyo Hong Kong

Teuvo Kohonen

0 書評

Springer-Verlag, 1988 - 312 頁

An analogue approach to the travelling salesman problem using an elastic net method

Richard Durbin* & David Willshaw†

* MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH and King's College Research Centre, Cambridge CB2 1ST, UK
† Department of Zoology, University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, UK

The travelling salesman problem[1] is a classical problem in the field of combinatorial optimization, concerned with efficient methods for maximizing or minimizing a function of many independent variables. Given the positions of *N* cities, which in the simplest case lie in the plane, what is the shortest closed tour in which each city can be visited once? We describe how a parallel analogue algorithm, derived from a formal model[2–3] for the establishment of topographically ordered projections in the brain[4–10], can be applied to the travelling salesman problem[1,11,12]. Using an iterative procedure, a circular closed path is gradually elongated non-uniformly until it eventually passes sufficiently near to all the cities to define a tour. This produces shorter tour lengths than another recent parallel analogue algorithm[13], scales well with the size of the problem, and is naturally extendable to a large class of optimization problems involving topographic mappings between geometrical structures[14].

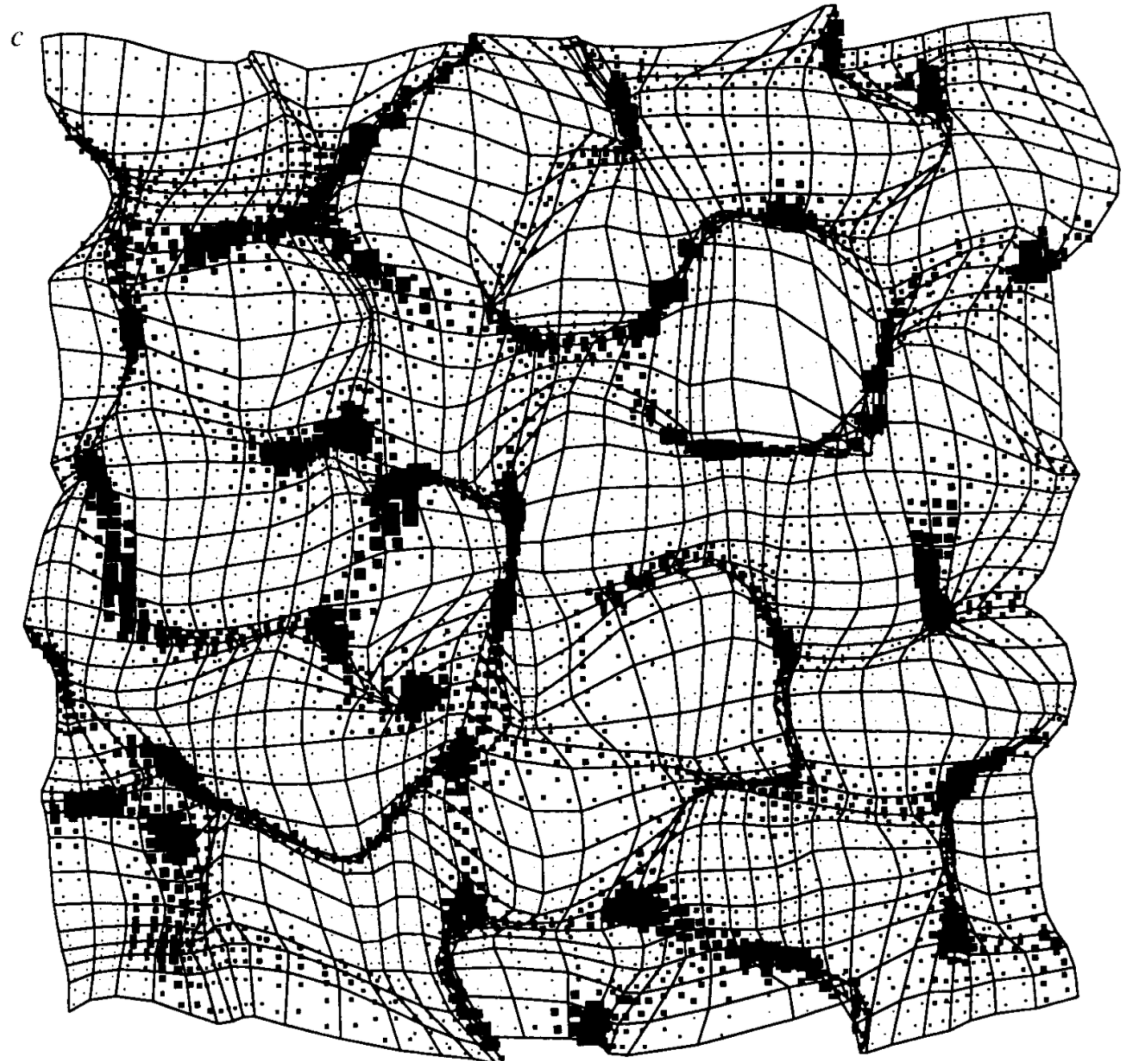
Although easy to state, the travelling salesman problem (TSP)

A dimension reduction framework for understanding cortical maps

Richard Durbin* & Graeme Mitchison†

* Department of Psychology, Stanford University, California 94305, USA
† Physiological Laboratory, Downing Street, Cambridge CB2 3EG, UK,
and The Research Centre, King's College, Cambridge CB2 1ST, UK

WE argue that cortical maps, such as those for ocular dominance, orientation and retinotopic position in primary visual cortex[1], can be understood in terms of dimension-reducing mappings from many-dimensional parameter spaces to the surface of the cortex. The goal of these mappings is to preserve as far as possible neighbourhood relations in parameter space so that local computations in parameter space can be performed locally in the cortex. We have found that, in a simple case[2], certain self-organizing models[3,4] generate maps that are near-optimally local, in the sense that they come close to minimizing the neuronal wiring required for local operations. When these self-organizing models are applied to the task of simultaneously mapping retinotopic position and orientation, they produce maps with orientation vortices resem-



Durbin and Mitchison 1990

Self-Organization Using Potts Models

Cheng-Yuan, Liou; Jiann-Ming, Wu

Self-organization map; Neural network; Potts model; Elastic ring; Mean field annealing; Hairy model

In this work, we use Potts neurons for the competitive mechanism in a self-organization model. We obtain new algorithms on the basis of a Potts neural network for coherent mapping, and we remodel the Durbin algorithm and the Kohonen algorithm with mean field annealing. The resulting dimension-reducing mappings possess a highly reliable topology preservation such that the nearby elements in the parameter space are ordered as similarly as possible on the cortex-like map, and the objective function costs between neighboring cortical points are as smooth as possible. The proposed Potts neural network contains two sets of interactive dynamics for two kinds of mappings, one from the parameter space to the cortical space and the other in the reverse way. We present a theoretical approach to developing self-organizing algorithms with a novel decision principle for competitive learning. We find that one Potts neuron is able to implement the Kohonen algorithm. Both implementation and simulation results are encouraging.

PERGAMON

Neural Networks 15 (2002) 337–347

www.elsevier.com/locate/neunet

Contributed article

Learning generative models of natural images

Jiann-Ming Wu*, Zheng-Han Lin

Department of Applied Mathematics, National Donghwa University, Hualien, Taiwan, ROC

Received 7 May 2001; revised 17 December 2001; accepted 17 December 2001

Abstract

This work proposes an unsupervised learning process for analysis of natural images. The derivation is based on a generative model, a stochastic coin-flip process directly operating on many disjoint multivariate Gaussian distributions. Following the maximal likelihood principle and using the Potts encoding, the goodness-of-fit of the generative model to tremendous patches randomly sampled from natural images is quantitatively expressed by an objective function subject to a set of constraints. By further combination of the objective function and the minimal wiring criterion, we achieve a mixed integer and linear programming. A hybrid of the mean field annealing and the gradient descent method is applied to the mathematical framework and produces three sets of interactive dynamics for the learning process. Numerical simulations show that the learning process is effective for extraction of orientation, localization and bandpass features and the generative model can make an ensemble of a sparse code for natural images.

Keywords: Neural networks; Cortical maps; Elastic net; Potts model; Self-organization; Unsupervised learning; Natural images

標題: **Global and local feature extraction by natural elastic nets**
作者: Wu, JM; Lin, ZH
來源: IEICE TRANSACTIONS ON INFORMATION AND SYSTEMS 卷: **E87D** 期: **9** 頁碼: **2267-227**
版日期: **SEP 2004**
被引用次數: **0** (來自 Web of Science)
[檢視摘要]

標題: **Learning generative models of natural images**
作者: Wu, JM; Lin, ZH
來源: NEURAL NETWORKS 卷: **15** 期: **3** 頁碼: **337-347** 文獻號碼: **PII S0893-6080(02)00018-7**
10.1016/S0893-6080(02)00018-7 出版日期: **APR 2002**
被引用次數: **5** (來自 Web of Science)
全文 [檢視摘要]

標題: **Self-organization using Potts models**
作者: Liou, CY; Wu, JM
來源: NEURAL NETWORKS 卷: **9** 期: **4** 頁碼: **671-684** DOI: **10.1016/0893-6080(95)00111-5** 出
期: **JUN 1996**
被引用次數: **9** (來自 Web of Science)
全文 [檢視摘要]

Neurocomputing 74 (2011) 2228–2240

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Annealed Kullback–Leibler divergence minimization for generalized TSP, spot identification and gene sorting

Jiann-Ming Wu *, Pei-Hsun Hsu

Department of Applied Mathematics, National Dong Hwa University, Shoufeng, Hualien 974, Taiwan

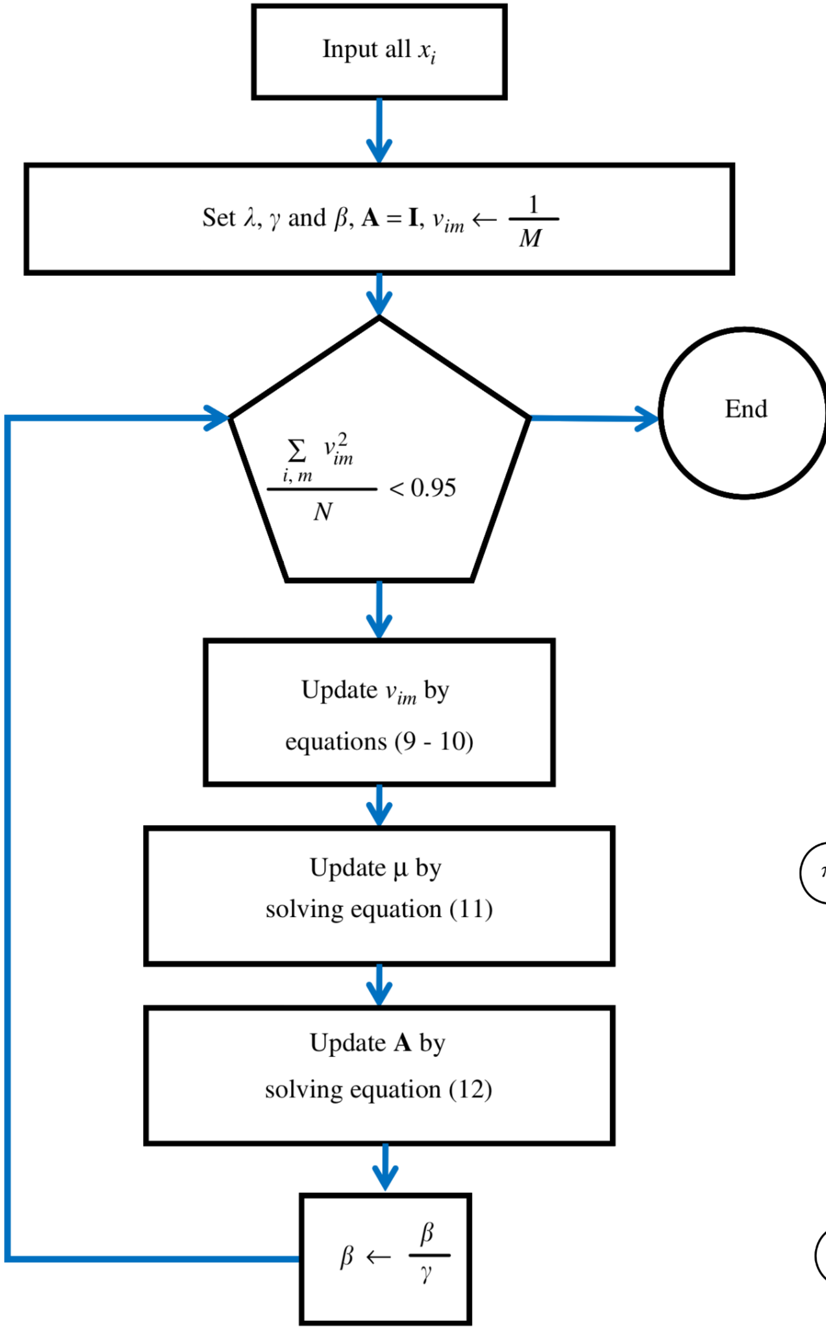
A R T I C L E   I N F O

Article history:
Received 11 August 2010
Received in revised form
27 December 2010
Accepted 2 March 2011
Communicated by L. Kurgan
Available online 5 April 2011

Keywords:
Maximal fitting and minimal wiring principle
Mean field annealing
Mixed integer programming
Gene sorting
Dimensionality reduction
Data visualization

A B S T R A C T

This work explores learning LCGM (lattice-connected Gaussian mixture) models by annealed Kullback–Leibler (KL) divergence minimization for a hybrid of topological and statistical pattern analysis. The KL divergence measures the general criteria of learning an LCGM model that is composed of a lattice of multivariate Gaussian units. A planar lattice emulates topological order of cortex-like neighboring relations and built-in parameters of connected Gaussian units represent statistical features of unsupervised data. Learning an LCGM model involves collateral optimization tasks of resolving mixture combinatorics and extracting geometric features from high-dimensional patterns. Under assumption that mixture combinatorics encoded by Potts variables obey the Boltzmann distribution, approximating their joint probability by the product of individual probabilities is qualified by the KL divergence whose minimization under physical-like deterministic annealing faithfully optimizes involved mixture combinatorics and geometric features. Numerical simulations show the proposed annealed KL divergence minimization is effective and reliable for solving generalized TSP, spot identification, self-organization and visualization and sorting of yeast gene expressions.



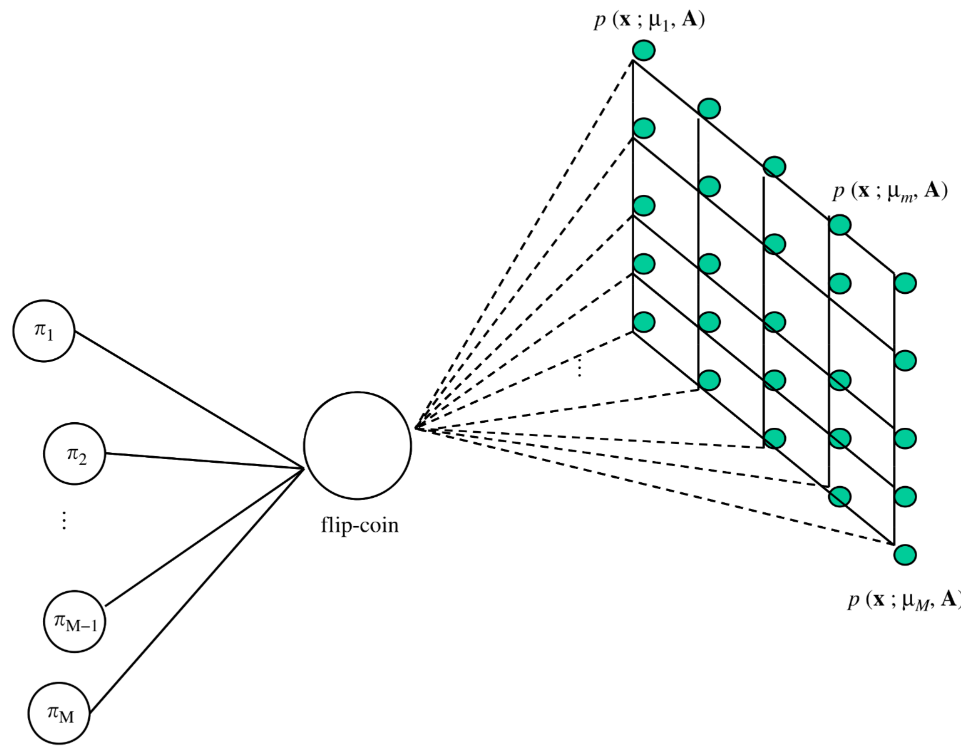
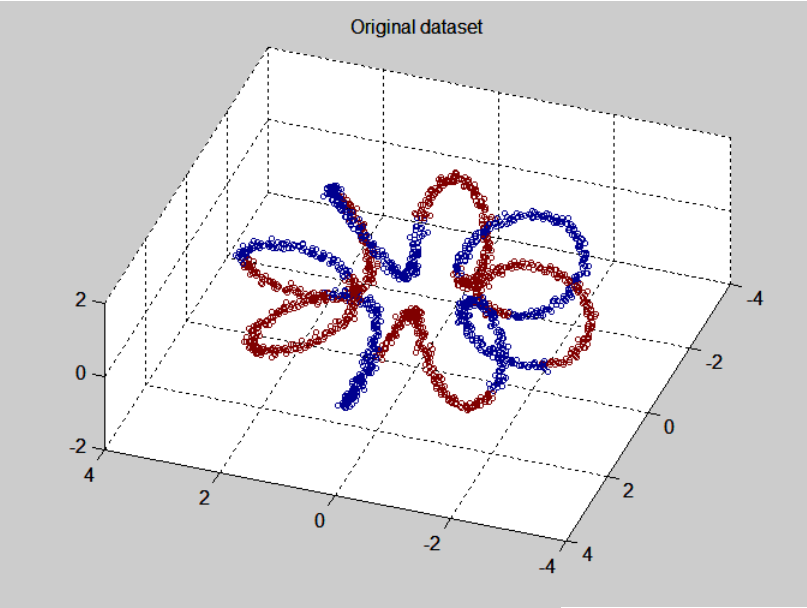
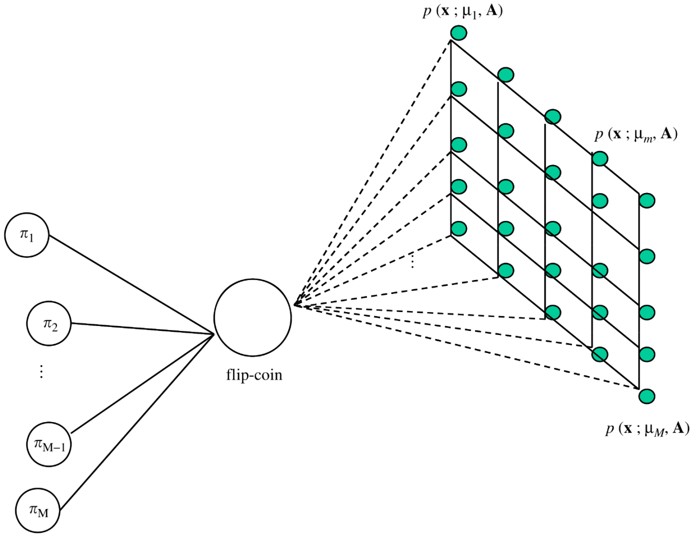
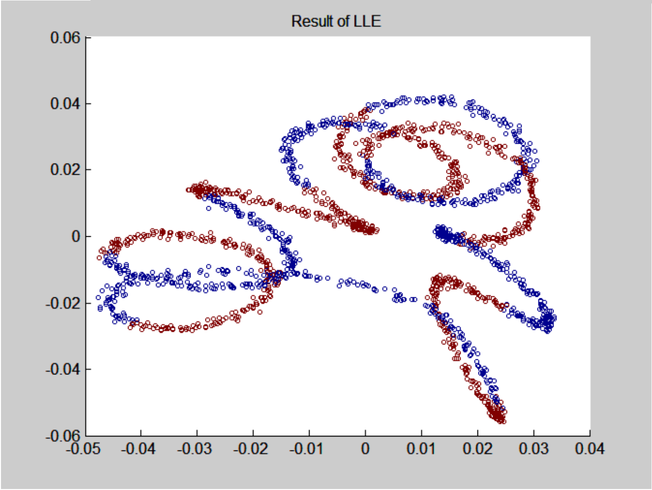
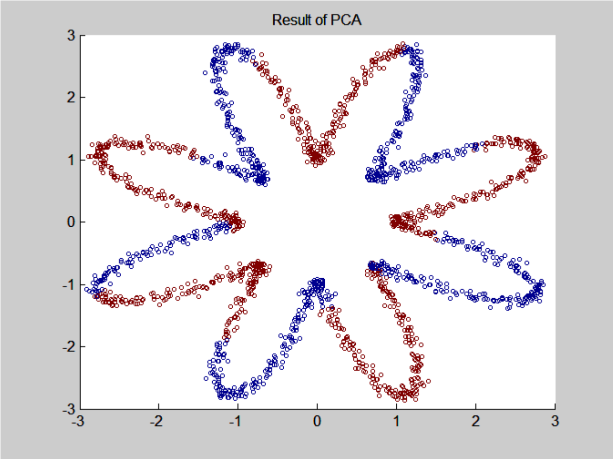


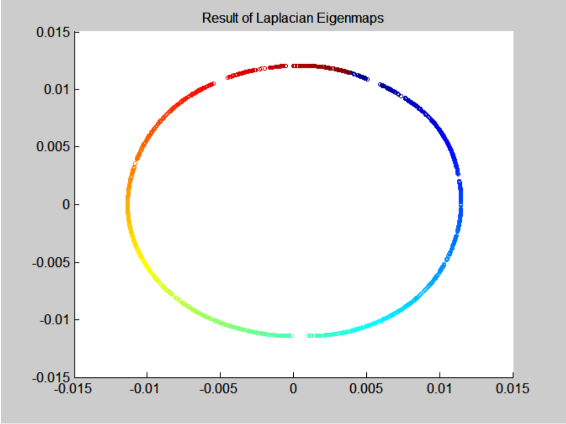
Fig. 2. The flow chart of learning an LCGM model by annealed KL minimization.

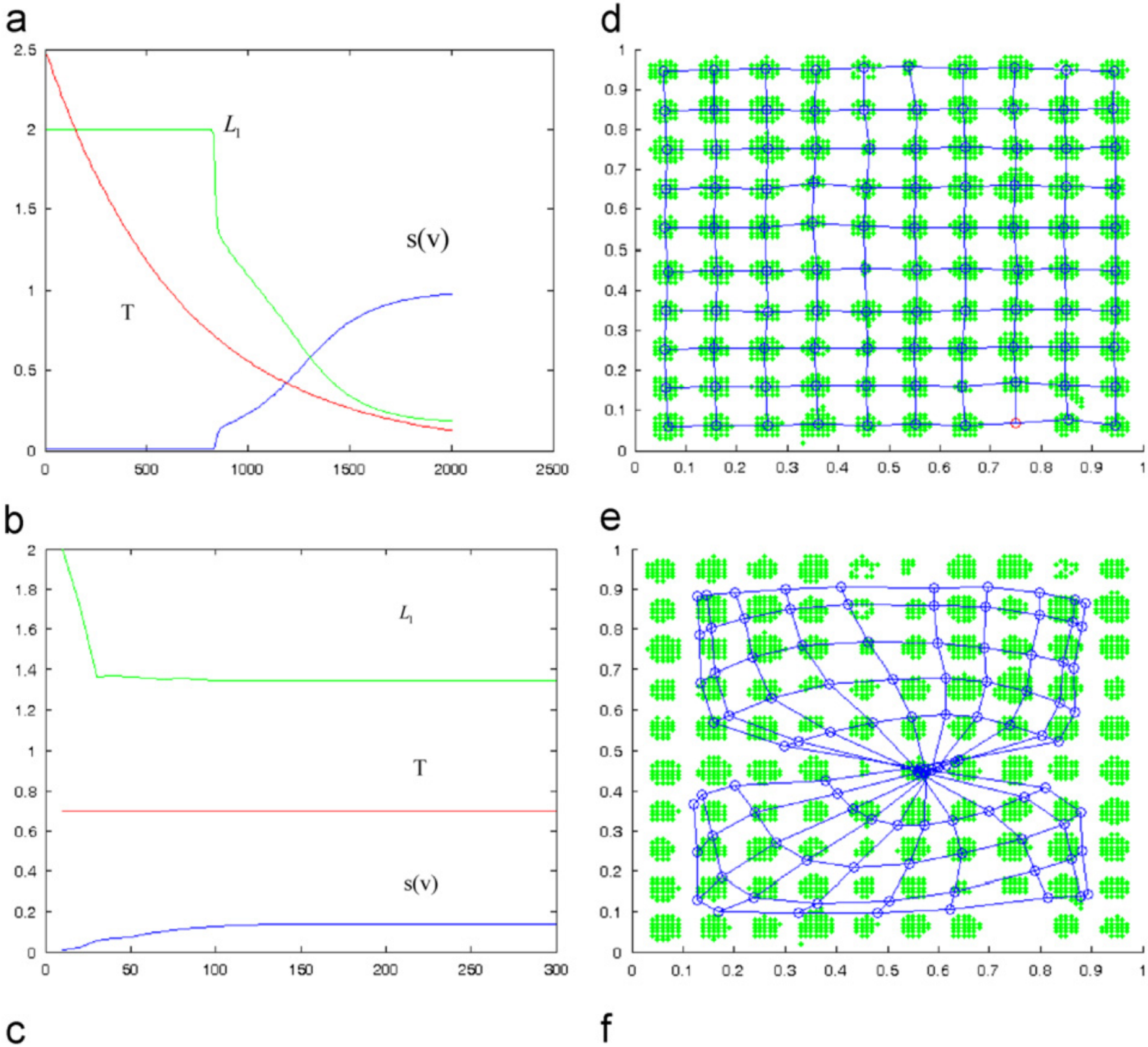

Original dataset

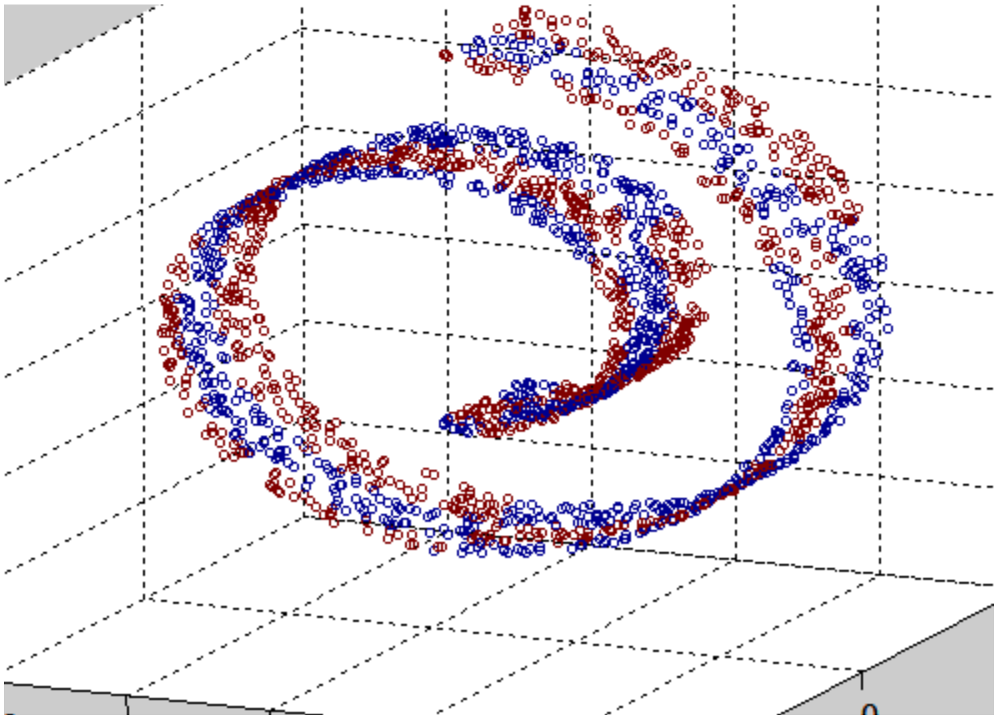

p (**x** ; μ_1, **A**)
p (**x** ; μ_m, **A**)
p (**x** ; μ_M, **A**)
π_1
π_2
π_{M-1}
π_M
flip-coin

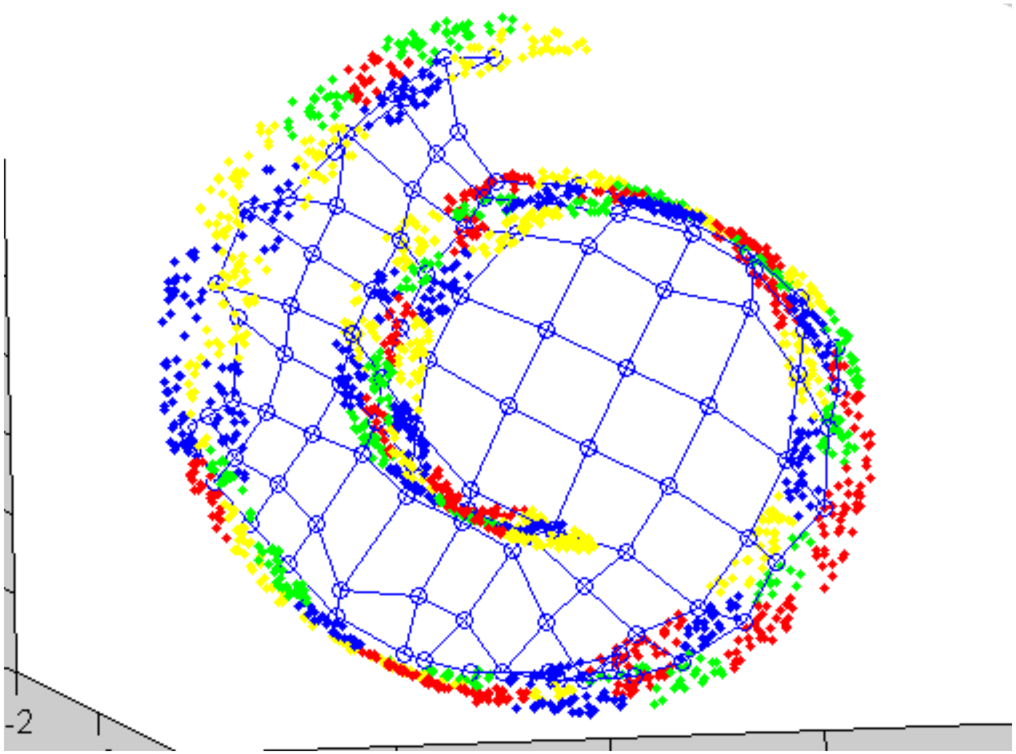

Result of LLE

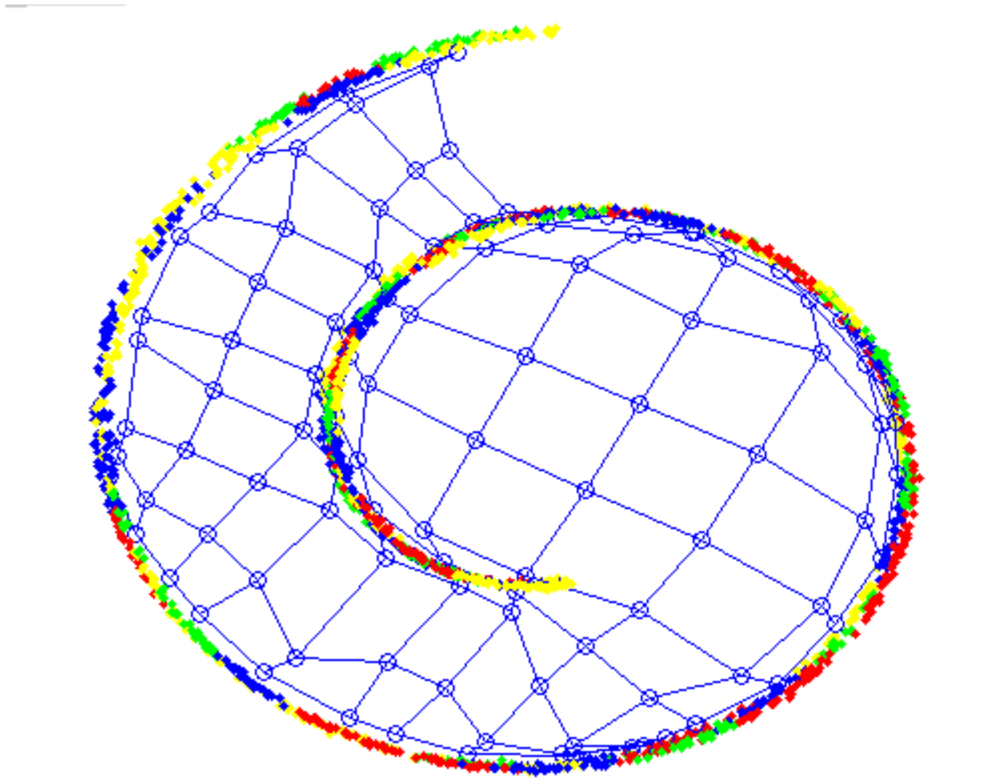

Result of PCA


Result of Laplacian Eigenmaps


a
L_1
s(v)
T
b
L_1
T
s(v)
c
d
e
f

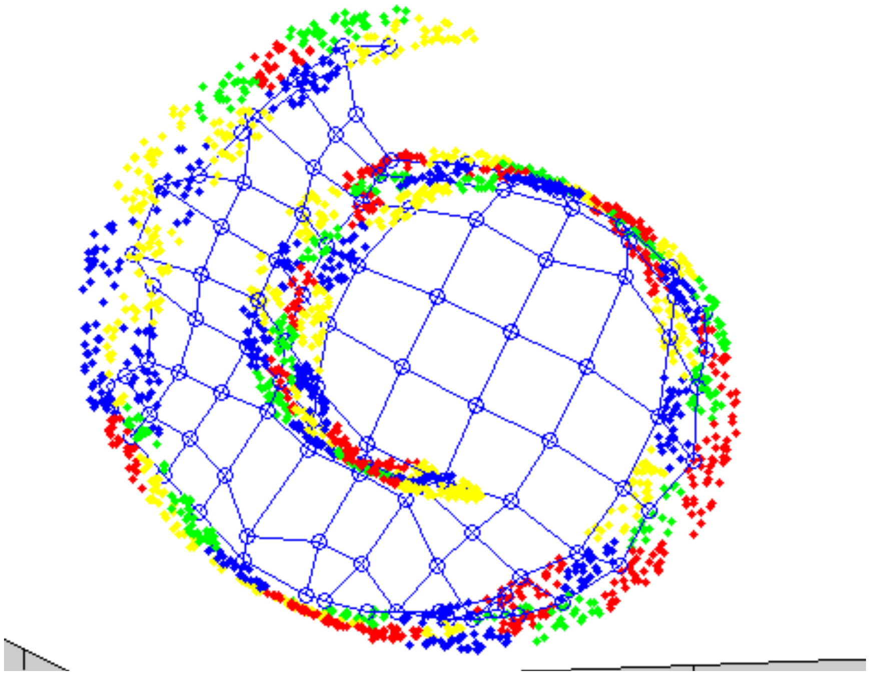



-2

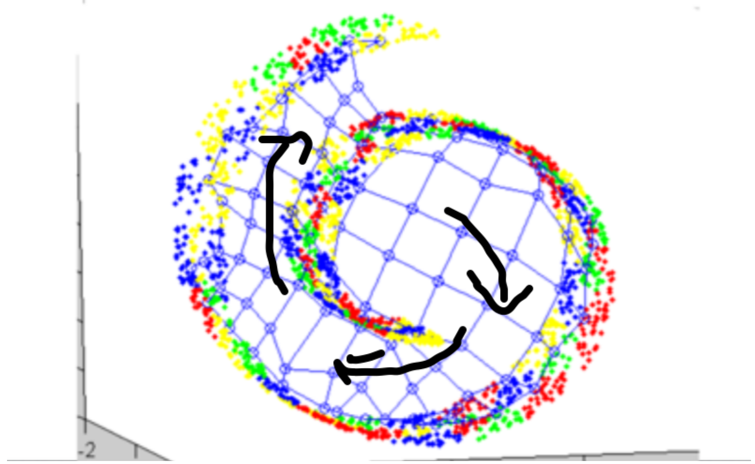


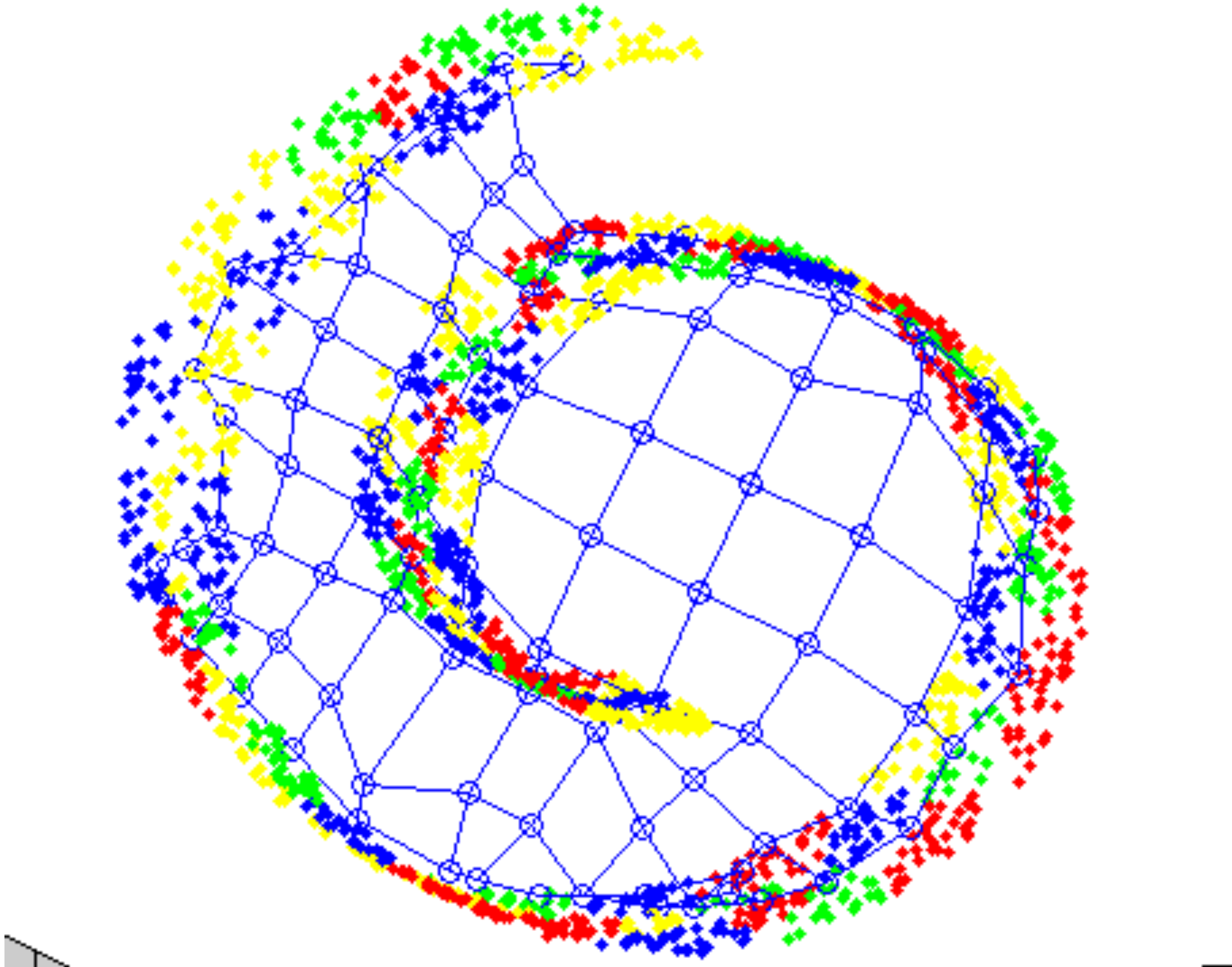
Q =

21	24	26	36	17	30	22	20	8	29
21	21	31	28	34	28	0	30	12	23
28	26	30	0	0	0	0	0	26	24
36	27	0	0	0	62	26	22	0	25
34	11	13	0	0	39	29	32	0	33
29	22	36	28	32	33	20	29	28	25
30	19	16	16	19	27	16	27	0	34
0	0	0	31	15	22	26	0	0	22
36	23	24	0	0	0	0	0	16	25
39	29	29	17	38	25	36	27	15	35



21	24	26	36	17	30	22	20	8	29
21	21	31	28	34	28	0	30	12	23
28	26	30	0	0	0	0	0	26	24
36	27	0	0	0	62	26	22	0	25
34	11	13	0	0	39	29	32	0	33
29	22	36	28	32	33	20	29	28	25
30	19	16	16	19	27	16	27	0	34
0	0	0	31	15	22	26	0	0	22
36	23	24	0	0	0	0	0	16	25
39	29	29	17	38	25	36	27	15	35





Strategies

Annealed Clustering

Self-organization map, Lattice-structured Gaussian mixtures

Set-valued mapping

Principle component analysis

Set-valued mapping

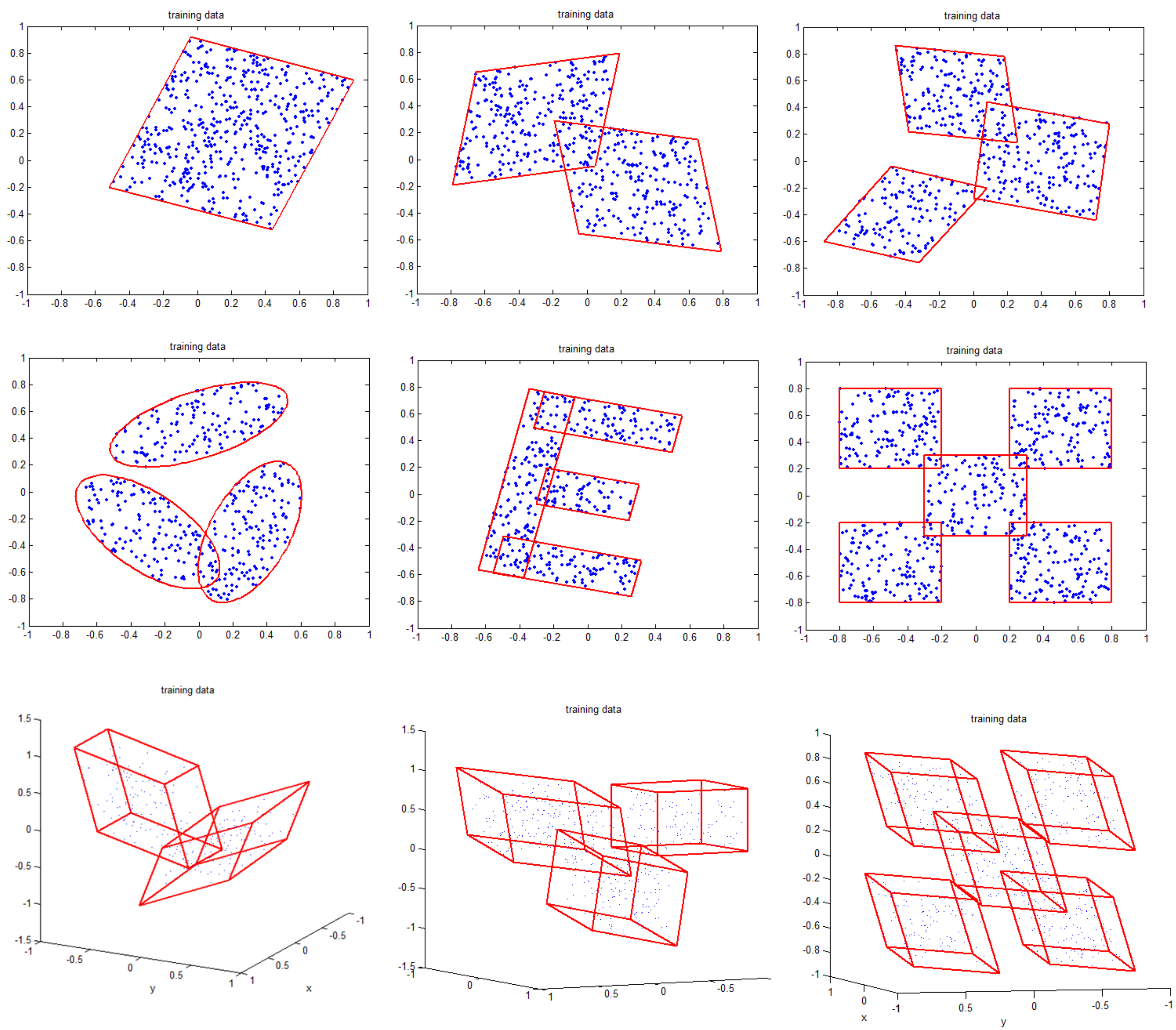
Mixtures of independent samples from many joined elementary single-valued mappings

Computations

Density support analysis

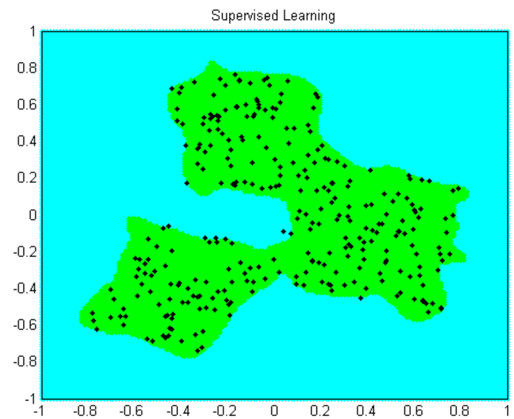
Decomposition to disjoint subsets that could be approximated by learning adaptive single-valued mappings

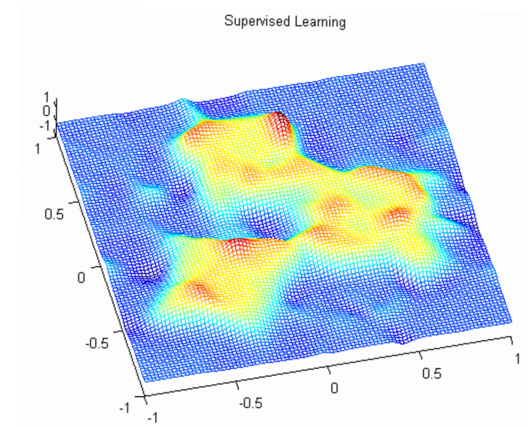
Swiss 2


training data
training data
training data
training data
training data
training data
training data
y
x
training data
training data
x
y

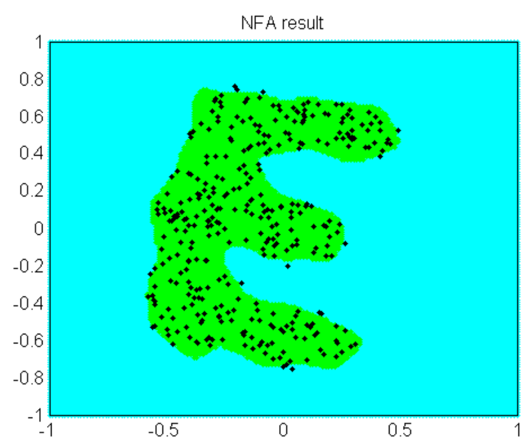
Left : The result of neural support approximation.
Right : The mesh of neural function via supervised learning.

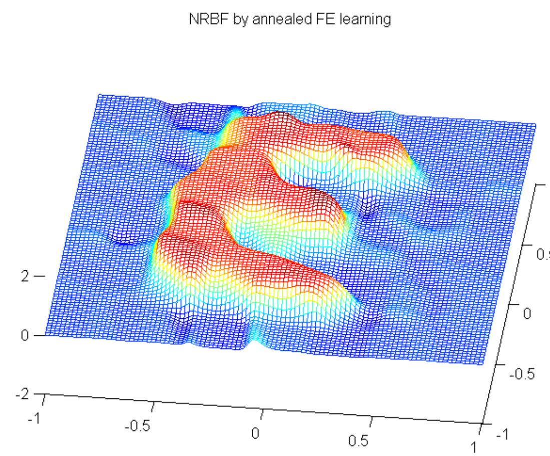
Example 3
M=41, K=3

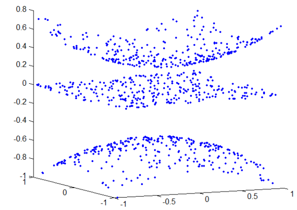




Example 5
M=41,K=2





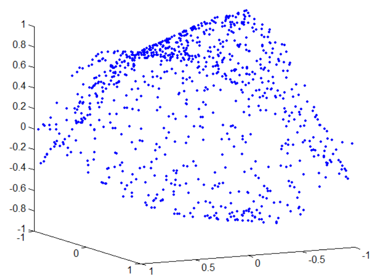


$F_1 = \{f_1, f_2, f_3\}$

$f_1(x_1, x_2) = x_1^2 + x_2^2 + 1$

$f_2(x_1, x_2) = -x_1^2 - x_2^2 - 2$

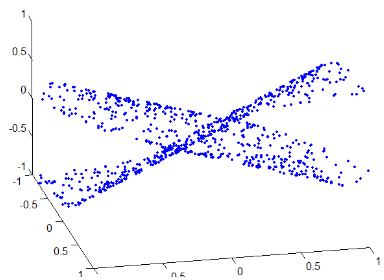
$f_3(x_1, x_2) = 0.1x_1 + 0.1x_2$



$F_2 = \{f_1, f_2\}$

$f_1(x_1, x_2) = \sin(x_1 + x_2)$

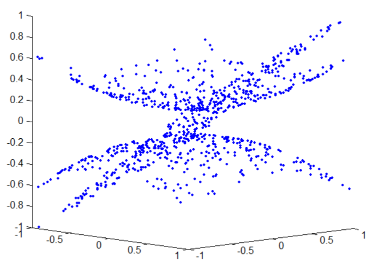
$f_2(x_1, x_2) = \cos(x_1 - x_2)$



$F_3 = \{f_1, f_2\}$

$f_1(x_1, x_2) = \tanh(x_1 + x_2)$

$f_2(x_1, x_2) = \tanh(x_1 - x_2)$



$F_4 = \{f_1, f_2, f_3\}$

$f_1(x_1, x_2) = x_1^2 + x_2^2 + 0.5$

$f_2(x_1, x_2) = -x_1^2 - x_2^2 - 0.5$

$f_3(x_1, x_2) = 2x_1 + 2x_2$

Figure 1

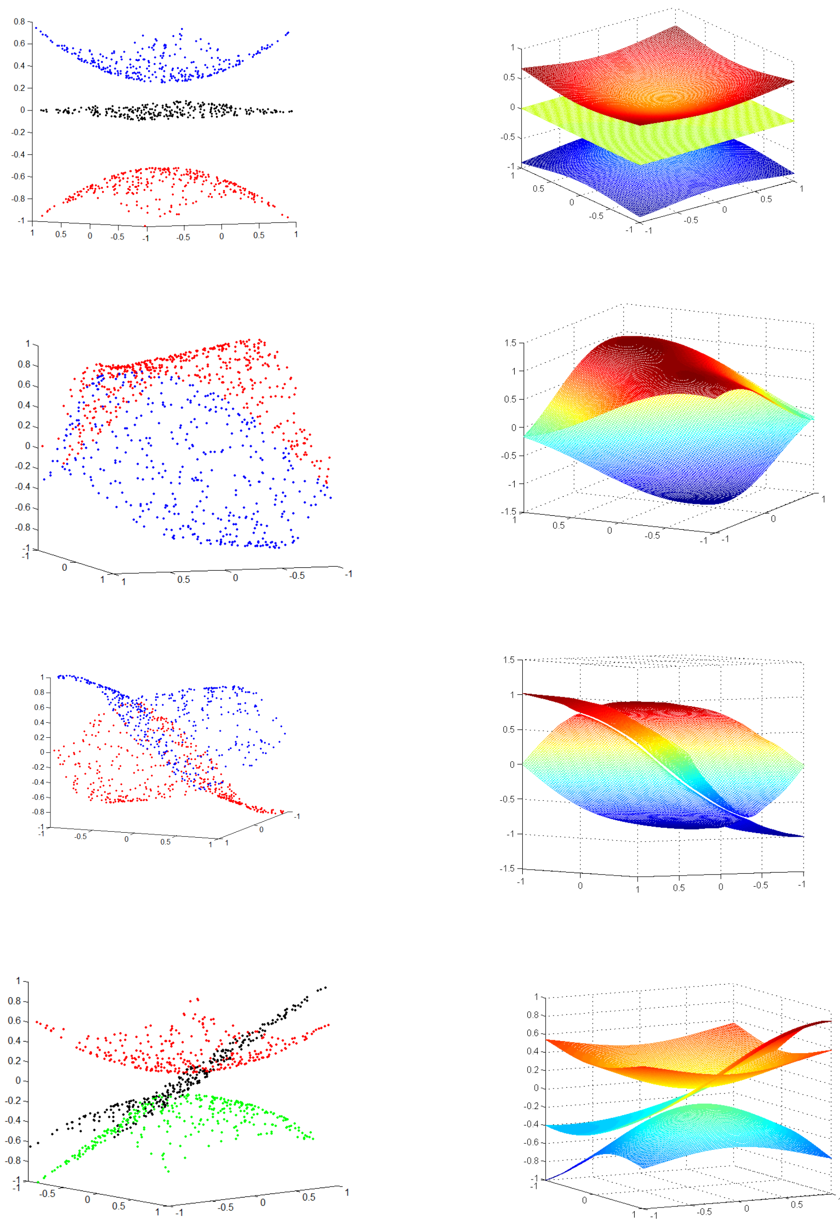
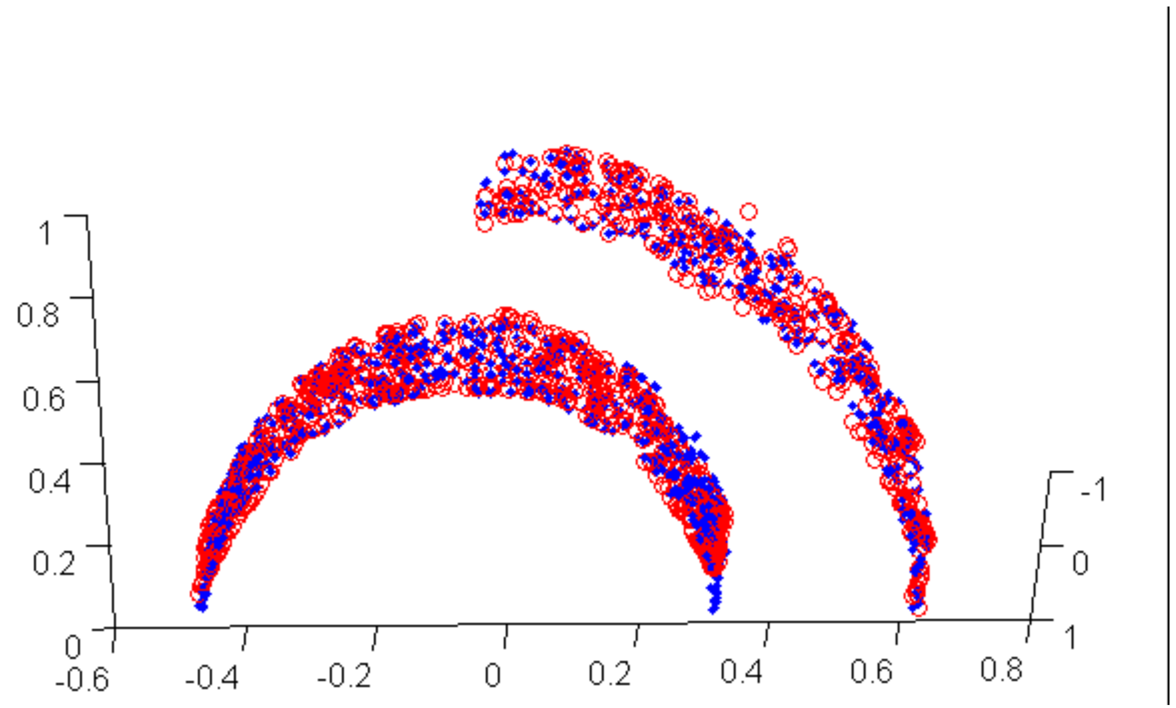
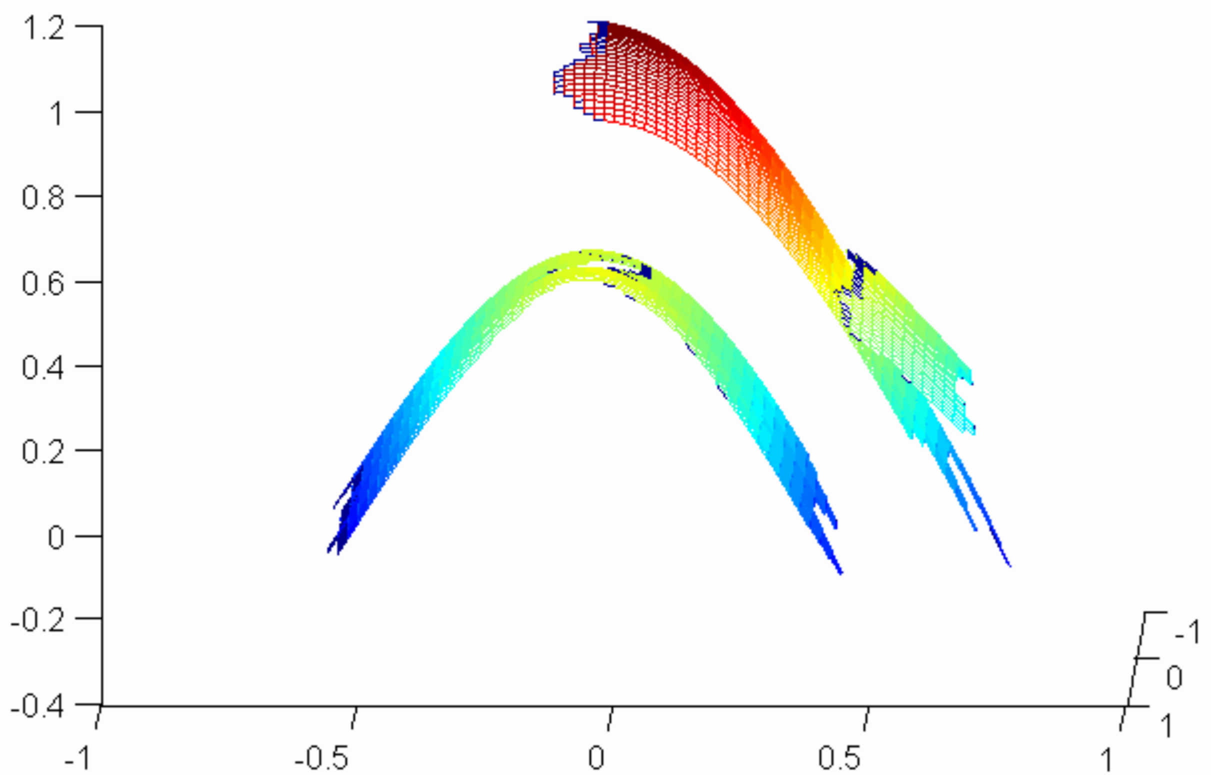


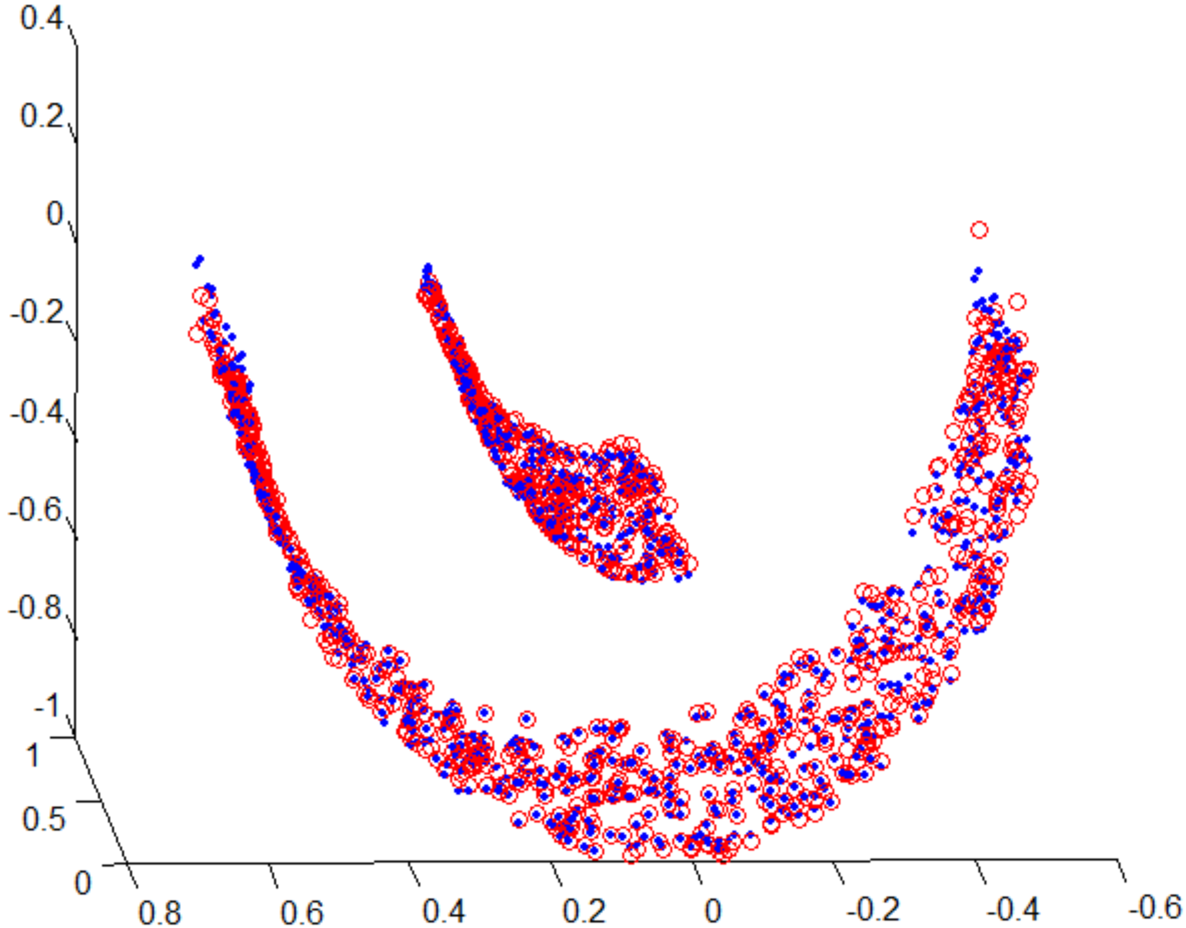
Figure 3



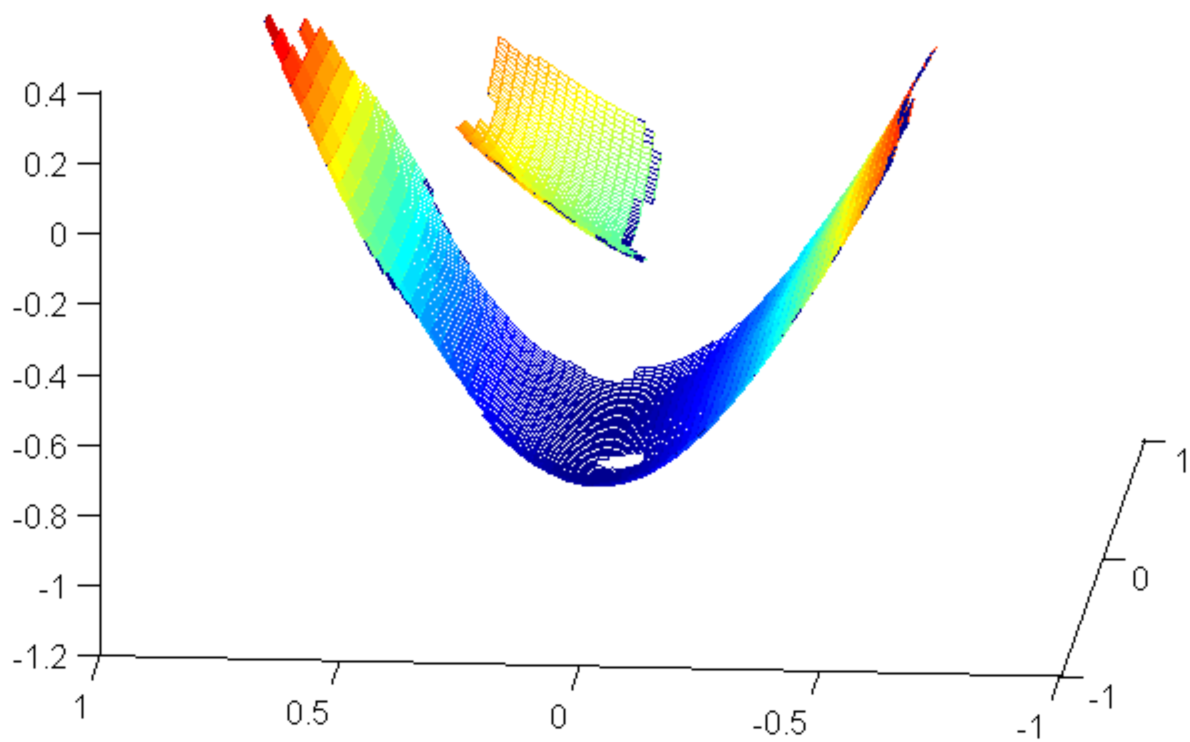
1
0.8
0.6
0.4
0.2
0
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
-1
0
1

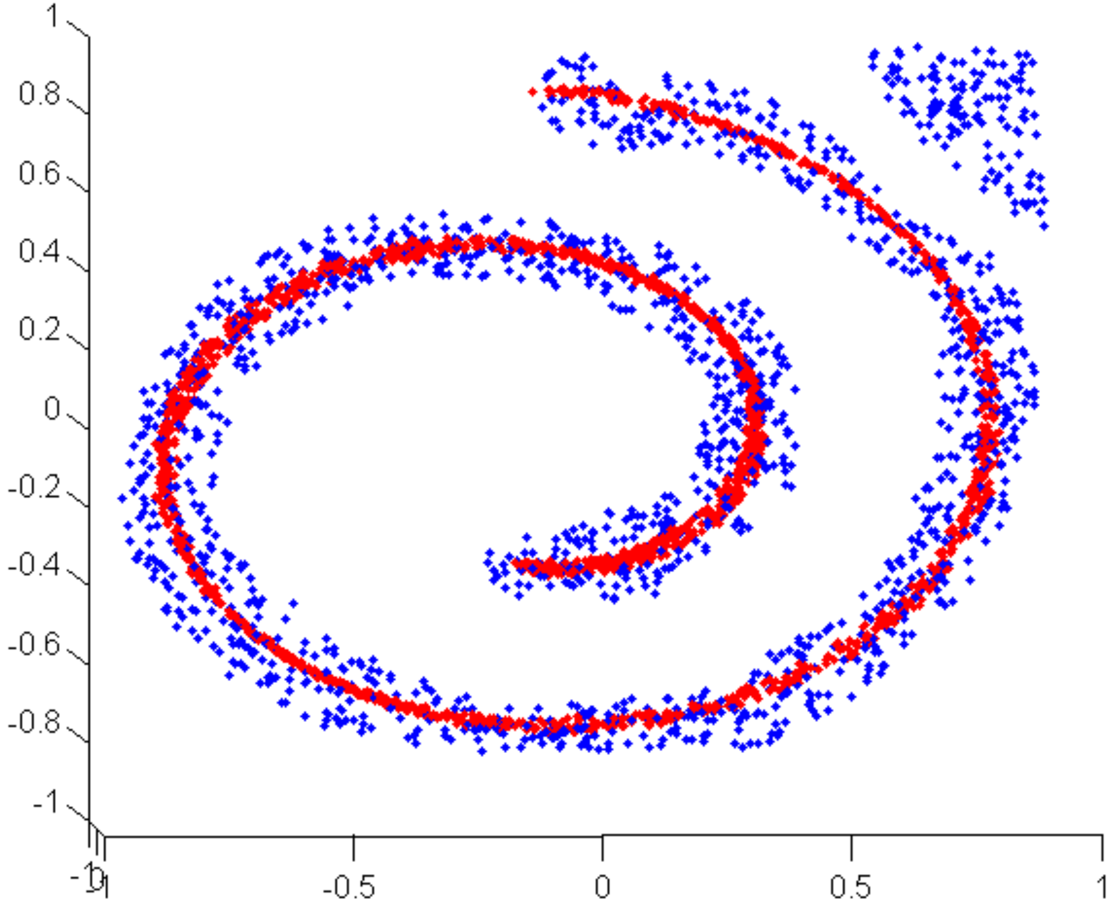


1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-1
-0.5
0
0.5
1
-1
0
1



0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
1
0.5
0
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6


0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
1
0.5
0
-0.5
-1
1
0
-1


1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1
-0.5
0
0.5
1