IJCNN 2001 Neural Network Competition

Ford Research Laboratory

Danil Prokhorov
(dprokhor@ford.com)

gratefully acknowledging assistance of

Timothy Feldkamp
Lee Feldkamp
Kenneth Marko

Outline of presentation

- **Challenge 1 and its origin**
- **Entries and results**
- **Challenge 2 and its origin**
- **Entries, results and another data set**
- **No entries for Challenge 3 in spite of extended deadline June 15 and e-mail announcements!**

Generalization Ability Challenge (GAC)

Time series of 50000 samples is produced by physical system (10-cylinder internal combustion engine). Each sample k of time series consists of four inputs and one output (k = 1, 2, …, 50000).

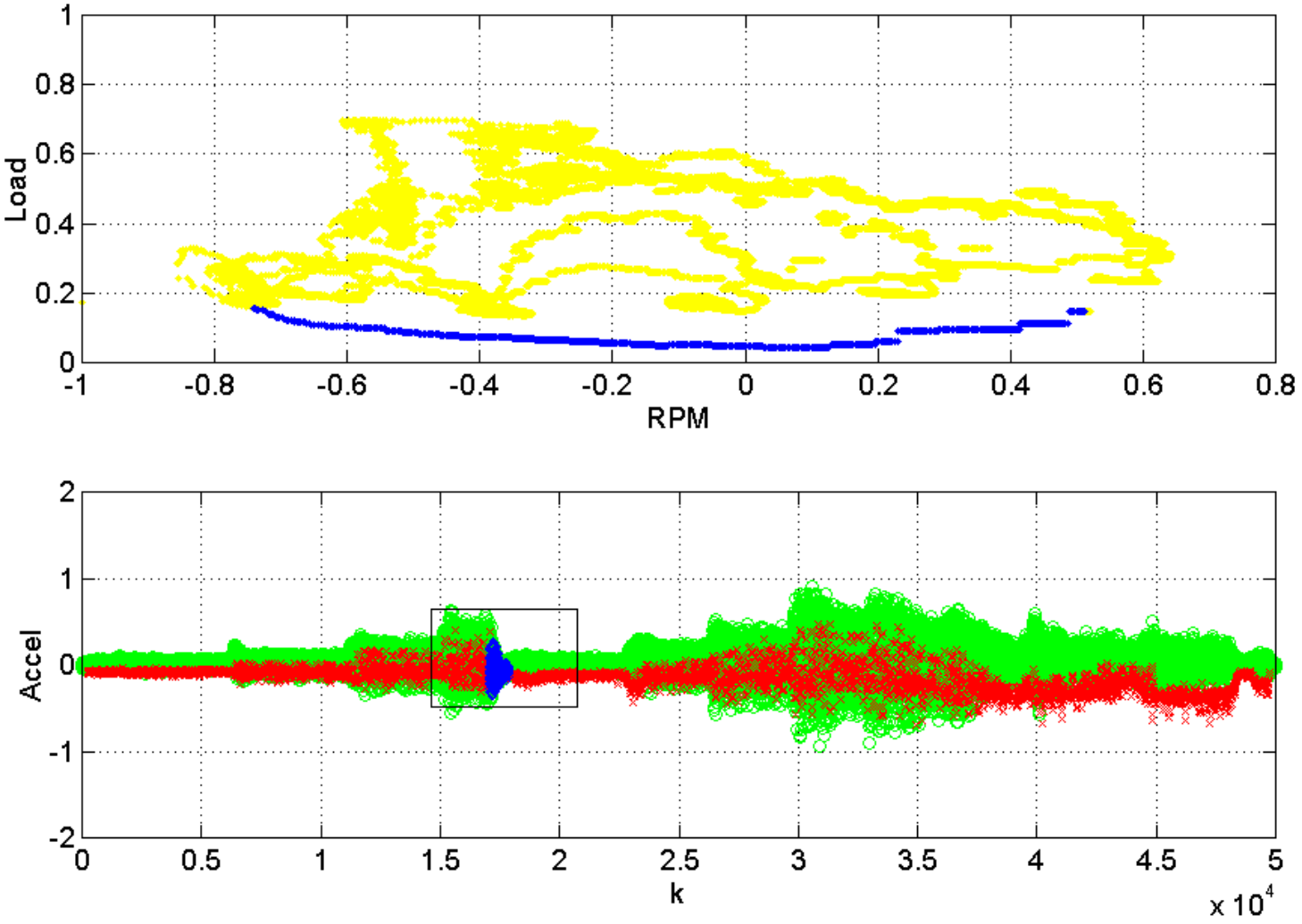
The first input $X_1(k)$ represents a binary synchronization pulse related to a natural periodicity in the system (cylinder identifier). The second and third inputs, $X_2(k)$ and $X_3(k)$, represent context (engine crankshaft speed in RPM and load). The fourth input $X_4(k)$ (crankshaft acceleration) has a more direct relationship with the bipolar output Y(k) (normal firing –1 or misfire +1, with normals dominating), but the relationship is complicated due to dynamics of the system (torsional oscillations of crankshaft depend on speed and load as well as presence of misfires prior to time k).

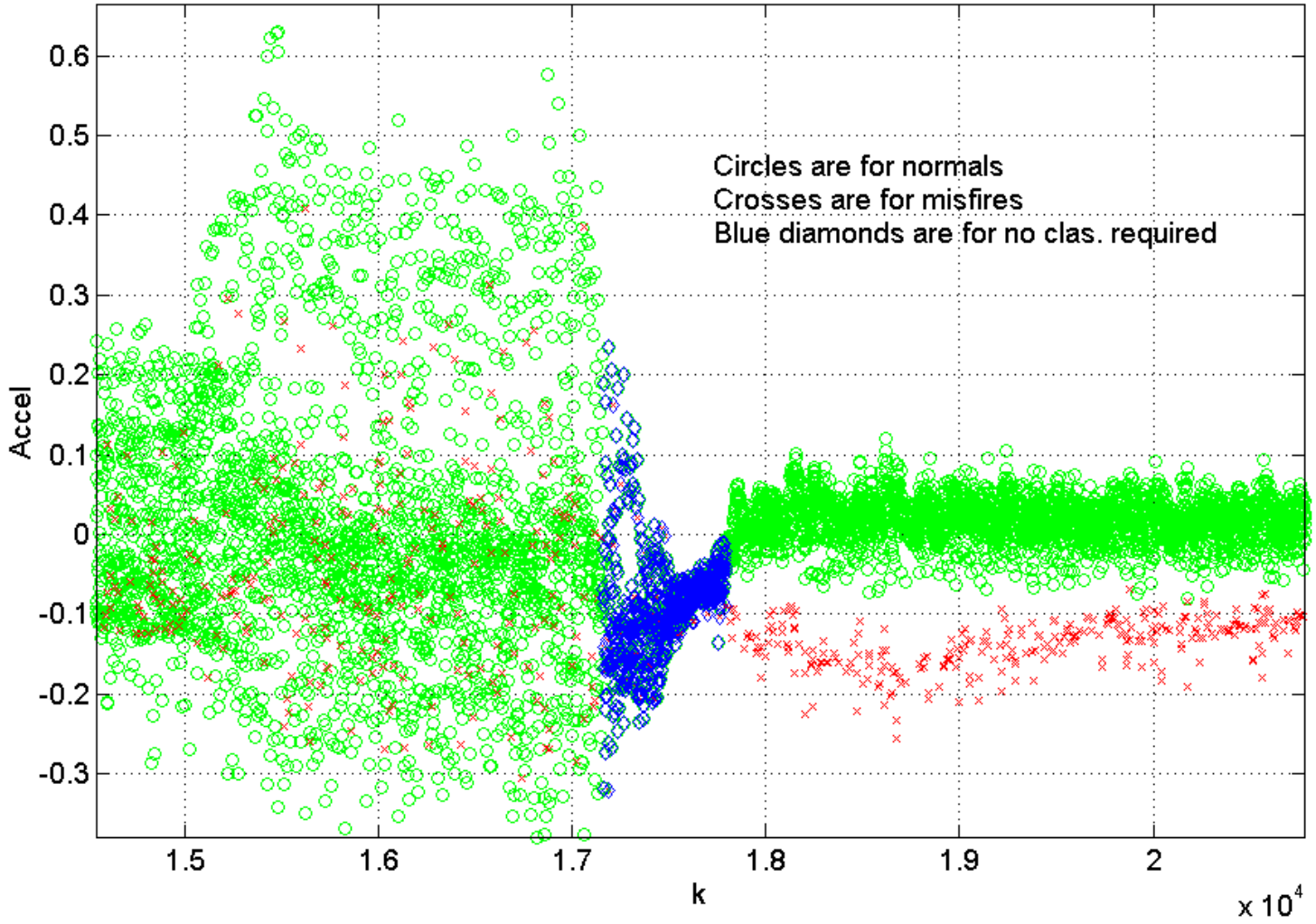
Additional (fifth) column in the data files indicates samples for which the predictions are to be counted. We evaluate accuracy of the prediction of the output made for the k-th sample only if the k-th value in the fifth column is equal to one. It is not mandated by the regulatory authorities to detect misfires in certain regions on the load-speed map. Such regions are denoted as zeros in the fifth column.

Given input vectors X, predict the bipolar output y(k) for all k with as few sign errors as possible (error counted only in those regions where the fifth column's values are equal to one). Any past or future information may be used when making predictions of Y(k). To evaluate the accuracy of the model, another time series (two files 50000 samples each) is provided with the output column missing.

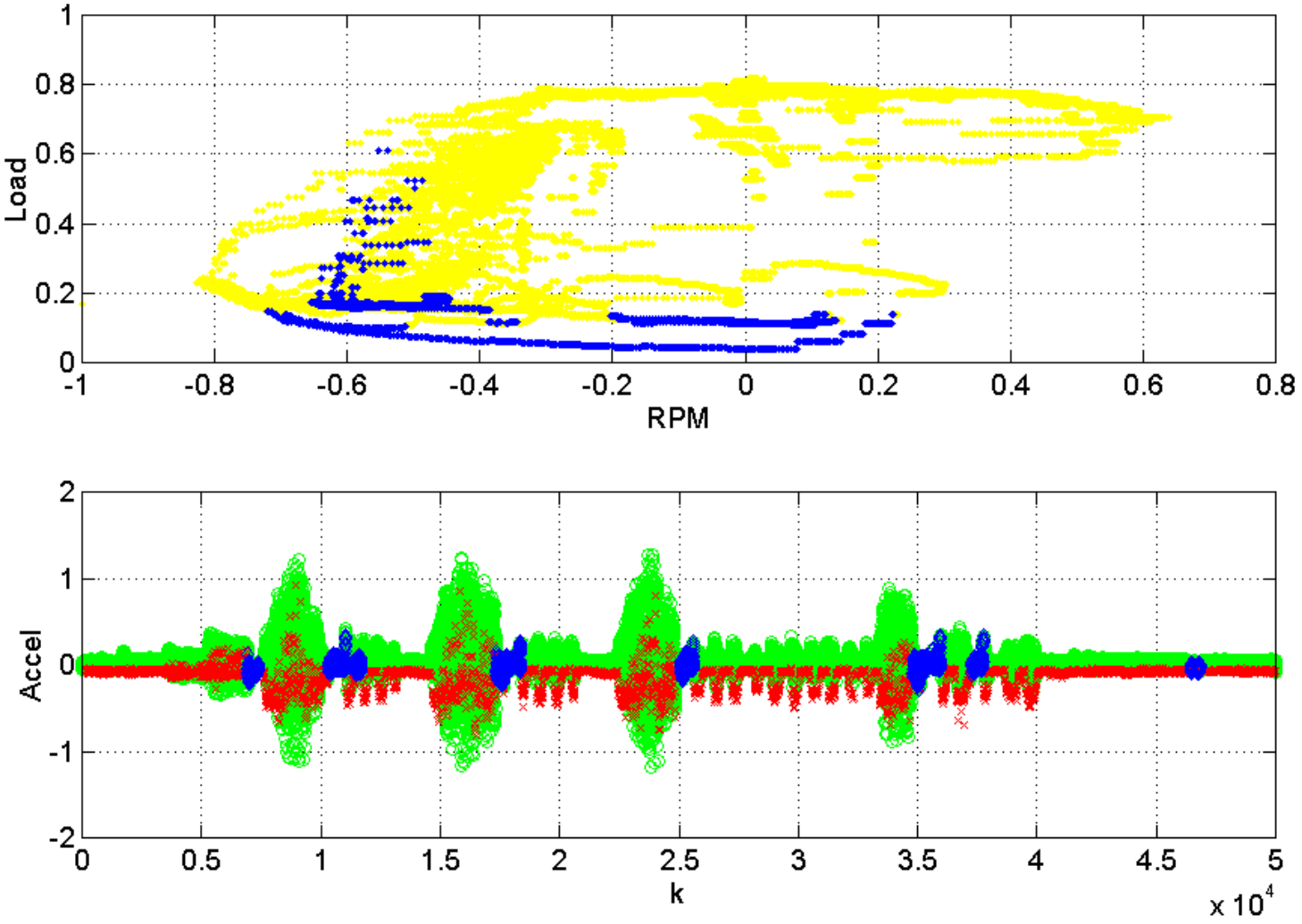
Our approach is based on RMLP trained via multi-stream EKF. For details see, e.g., Feldkamp and Puskorius, *Proc. of the IEEE*, November 1998.

Training set for GAC

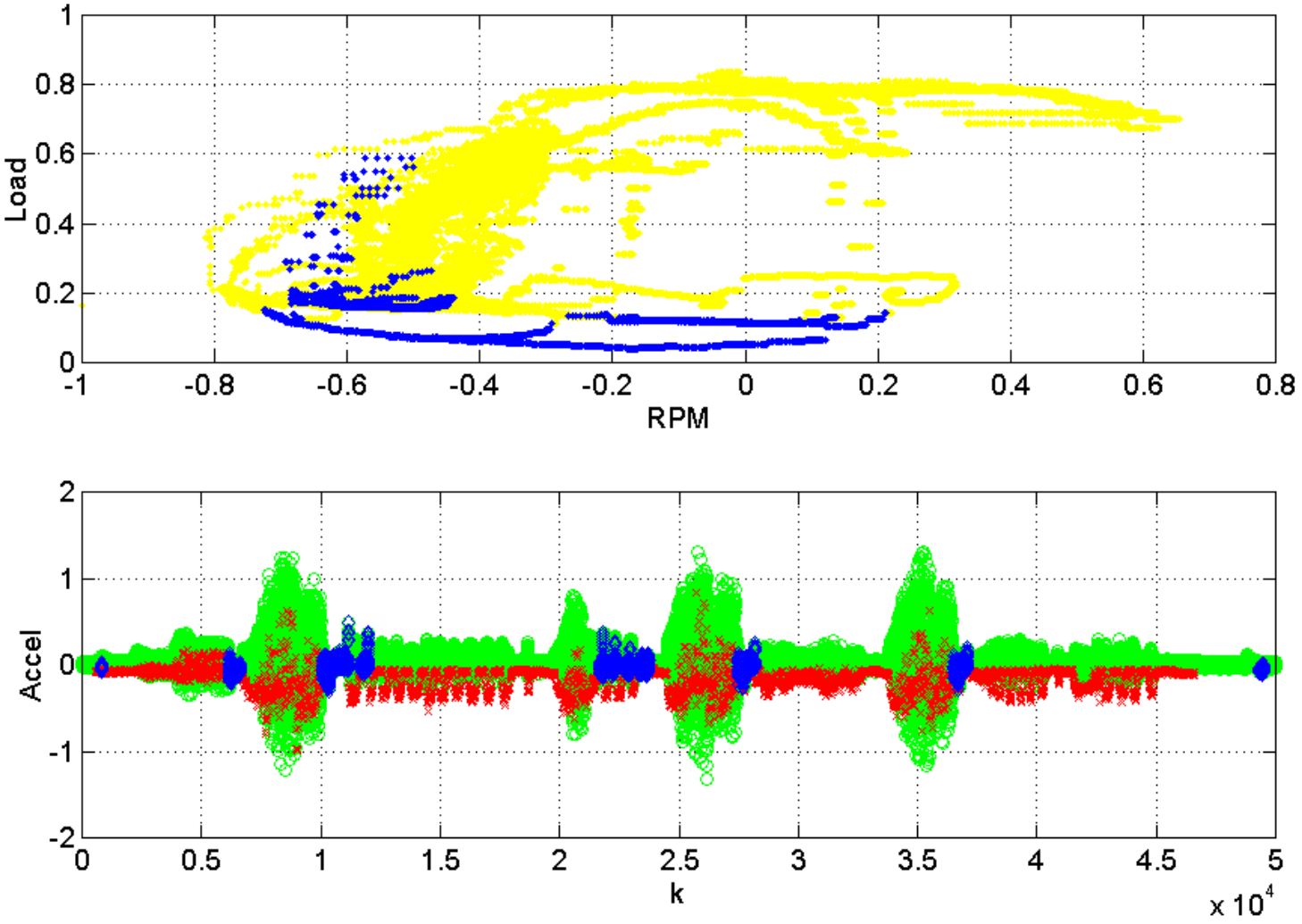



Circles are for normals
Crosses are for misfires
Blue diamonds are for no clas. required
Accel
k
x 10^4
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
1.5
1.6
1.7
1.8
1.9
2

Test set #1



Test set #2



ENTRIES AND RESULTS (GAC)

Name	Affiliation	Number of errors (gacte1)	Number of errors (gacte2)	Total
Chih-Chung Chang	National Taiwan University	656 ($\alpha<\beta$)	637 ($\alpha<\beta$)	1293
ANS and AD groups	Ford Research	~800 ($\alpha<\beta$)	~850 ($\alpha<\beta$)	~1650
Felix Heimes	BAE SYSTEMS Controls	1022 ($a>b$)	1244 ($a>b$)	2266
Zhihang Chen	University of Michigan-Dearborn	1344 ($\alpha<\beta$)	1070 ($\alpha<\beta$)	2414
Xiao Hu	University of Missouri-Rolla	1699 ($\alpha<\beta$)	1643 ($\alpha<\beta$)	3342
Huseyin Ince	University of Oklahoma-Norman	2658	2521	5179
Li Min Fu	University of Florida-Gainesville	3293	2911	6204
Thomas Hanselmann	University of Western Australia	5095	4820	9915
B. Yegnanarayana	Indian Inst. of Technology, Madras	7149	6409	13558
R. Ghosh	Griffith University, Australia	16258	10238	26496
Song Yang	Nanyang Tech Univ., Singapore	15368	15498	30866

- Out of 100K points, 8300 are no clas. required
- 1000 errors is 99%, 2000 errors is 98%
- Clas. as all normals is 91% accurate
- False alarm rate should be lower than misdetection rate ($\alpha<\beta$); not specified to the participants
- Multiple entries were not encouraged… but not prohibited either. Quantitative feedback for each entry was not provided until after the deadline. Should have either prohibited multiple entries or asked for no more than so many entries per team.

Text Decoding Challenge (TDC)

Three-dimensional time series encodes an unknown text message in English (8-bit ASCII coding; SPACE is replaced by NULL; all letters are capitals).

Given the binary sequence, we slide a window of length two across the entire sequence shifting by one from left to right to generate a pattern of two-bit numbers (right bit is the least significant bit). This pattern governs generation of four smooth nonlinear mappings

$$Y(k)=F_i(U_1(k),U_2(k)), \qquad i=00,01,10,11,$$

where $k=1,2,\ldots,T$,
$Y(k)$ **is element of output sequence,**
$U_1(k)$, $U_2(k)$ **are i.i.d. uniform random numbers in (-1,+1).**

For example (see Figure 1),

sequence (0)0110…

$F\_00(U_1(1),U_2(1))$ **to produce** $Y(1)$,
$F\_01(U_1(2),U_2(2))$ **to produce** $Y(2)$,
$F\_11(U_1(3),U_2(3))$ **to produce** $Y(3)$,
$F\_10(U_1(4),U_2(4))$ **to produce** $Y(4)$, **etc.**

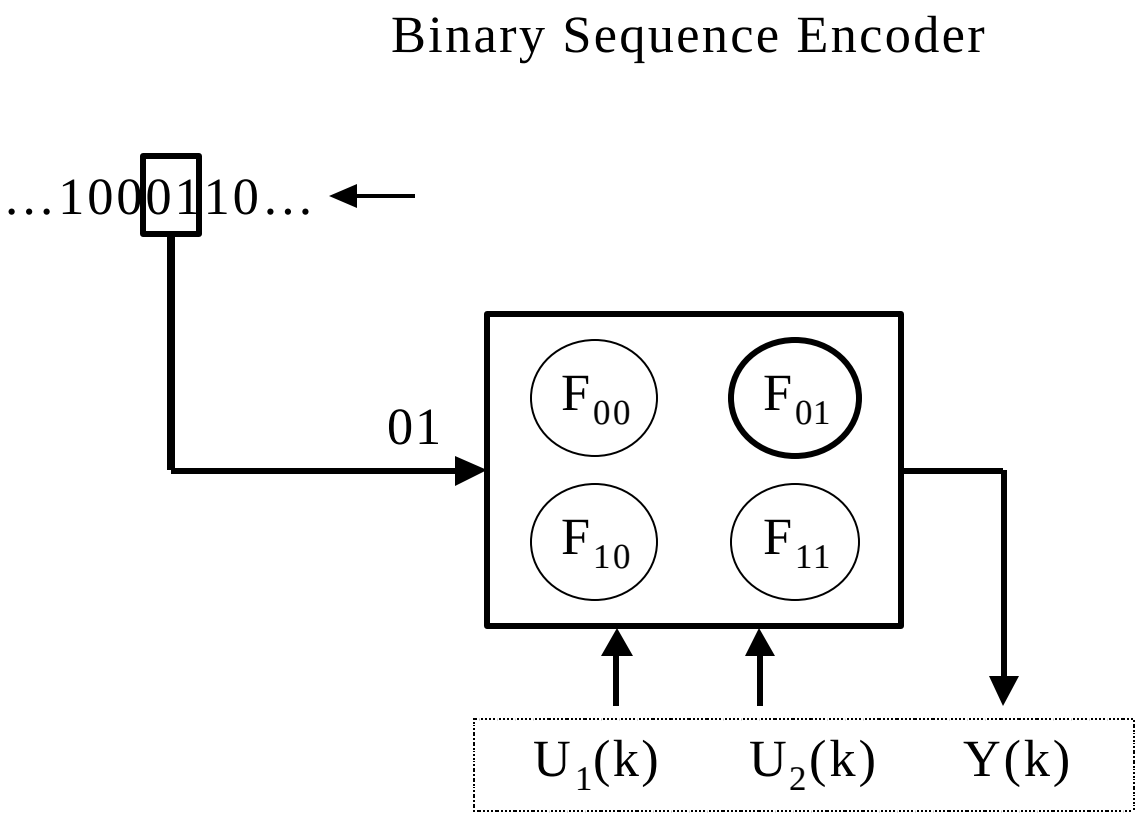

Binary Sequence Encoder
…10001110…
01
F00
F01
F10
F11
U1(k)
U2(k)
Y(k)

Figure 1

Text Decoding Challenge (TDC)

Set A has 16,280 data points (U_1 (k), U_2 (k), Y(k), bit(k)), and it is provided to help in developing a decoding procedure (encodes an excerpt from the U.S. Constitution).

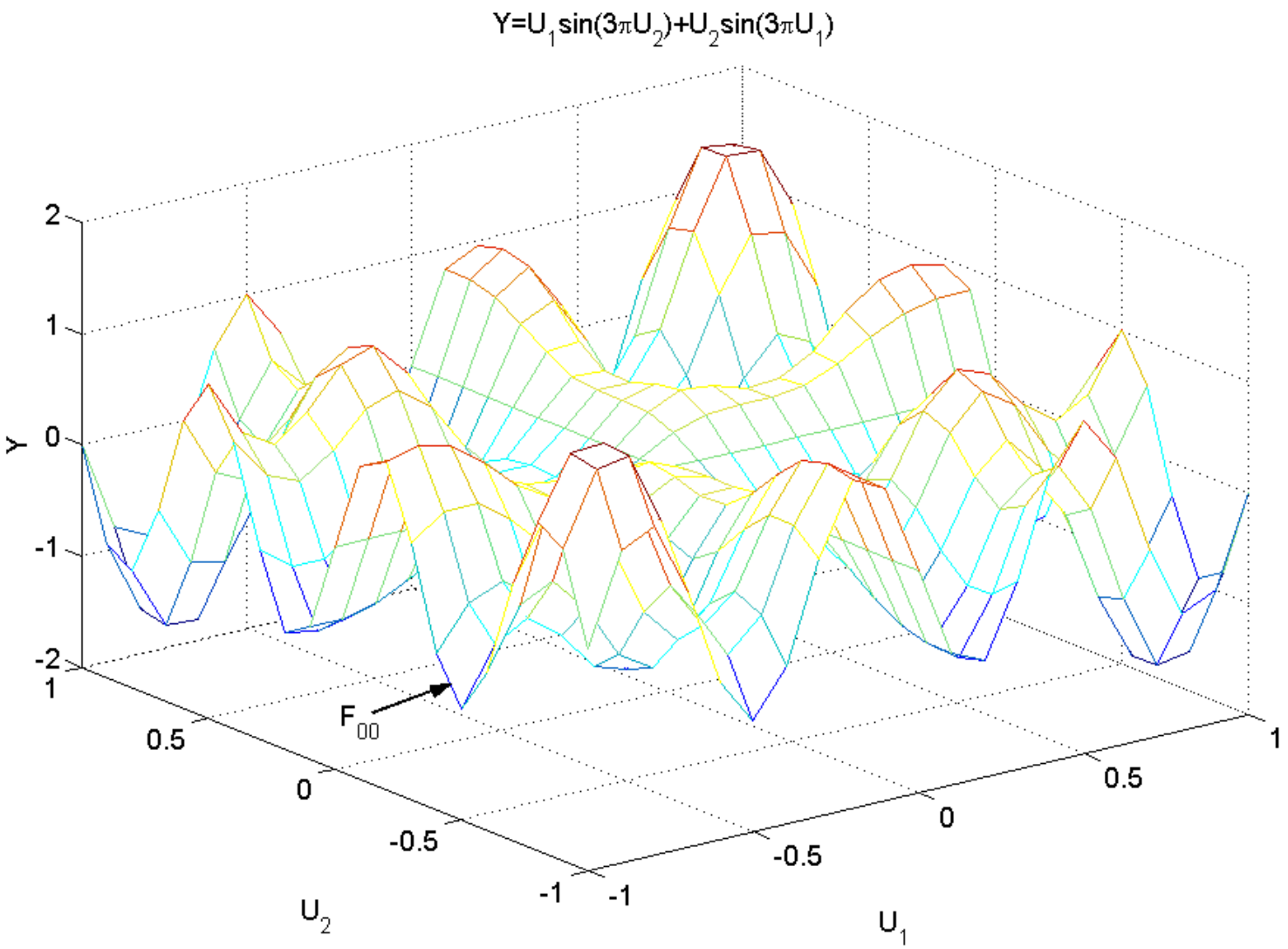
Set B has 30,720 data points (U_1 (k), U_2 (k), Y(k),?). Set B is missing the fourth column, a binary sequence for a text encoded by (U_1 (k), U_2 (k), Y(k)). The fraction of 1 bits in either of the binary sequences is approximately 0.135. You are required to recreate a binary sequence for the set B. One approach is to determine, for each k, which of the four functions was used to transform (U_1 (k), U_2 (k)) into Y(k). The winning entry is the one with the smallest number of errors in the sequence.

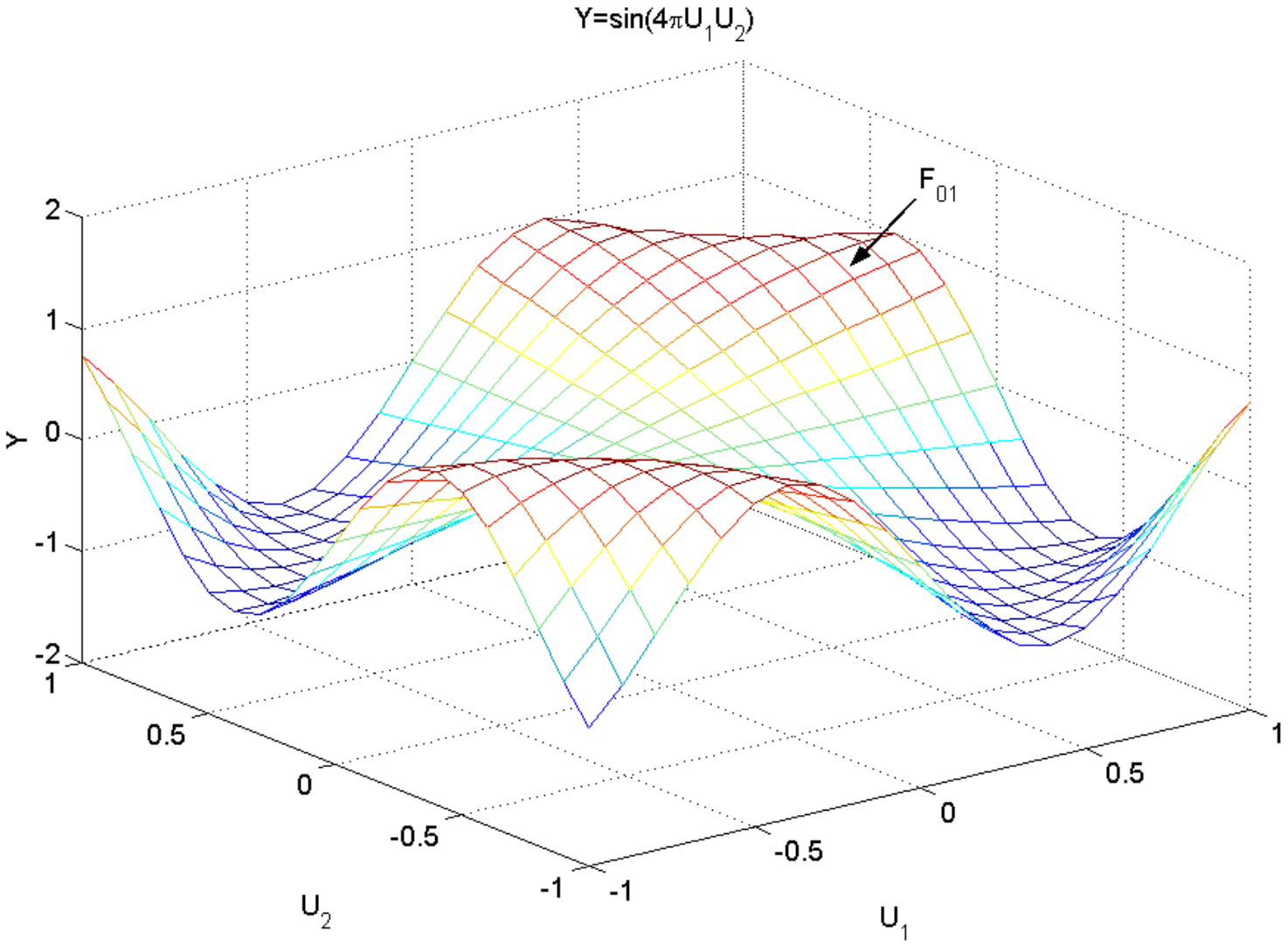
Motivation: misfire detection problem.

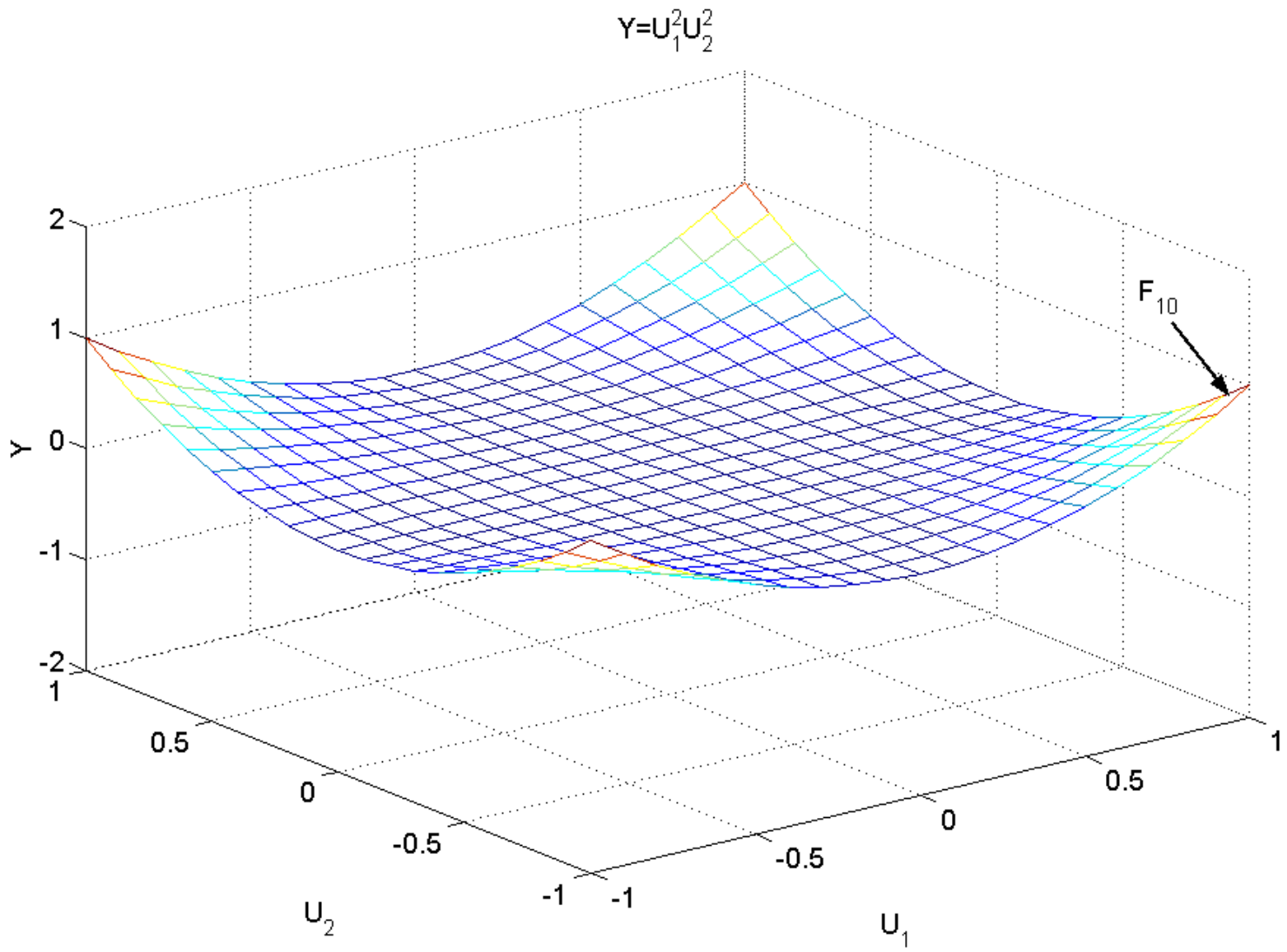
- **Assume normals as 0 and misfires as 1 in the sequence.**

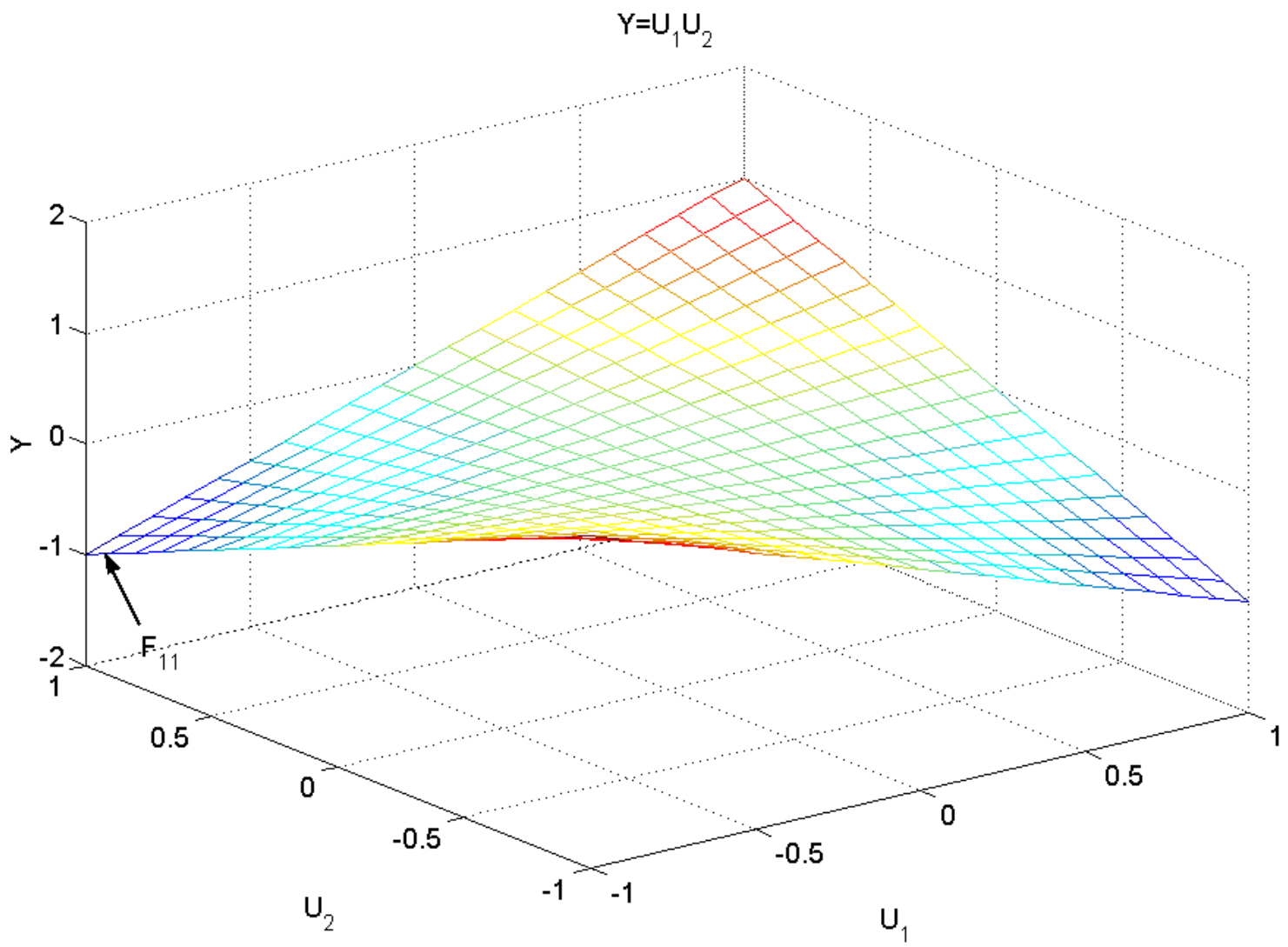
- **Given input variables X (load and speed), cylinder identity, misfire pattern and output acceleration Y, infer if misfire occurs at step k by learning models of accelerations for each cylinder and each relevant pattern of misfires.**

- **Though knowing whether sequence forms text helps in postprocessing, our approach does not need such knowledge. Likewise, ability to visualize data is not a prerequisite for its success. For details see Feldkamp et al., Approach to Adaptive Classification, In Haykin and Kosko (Eds.), *Intelligent Signal Processing*, IEEE Press, 2000.**

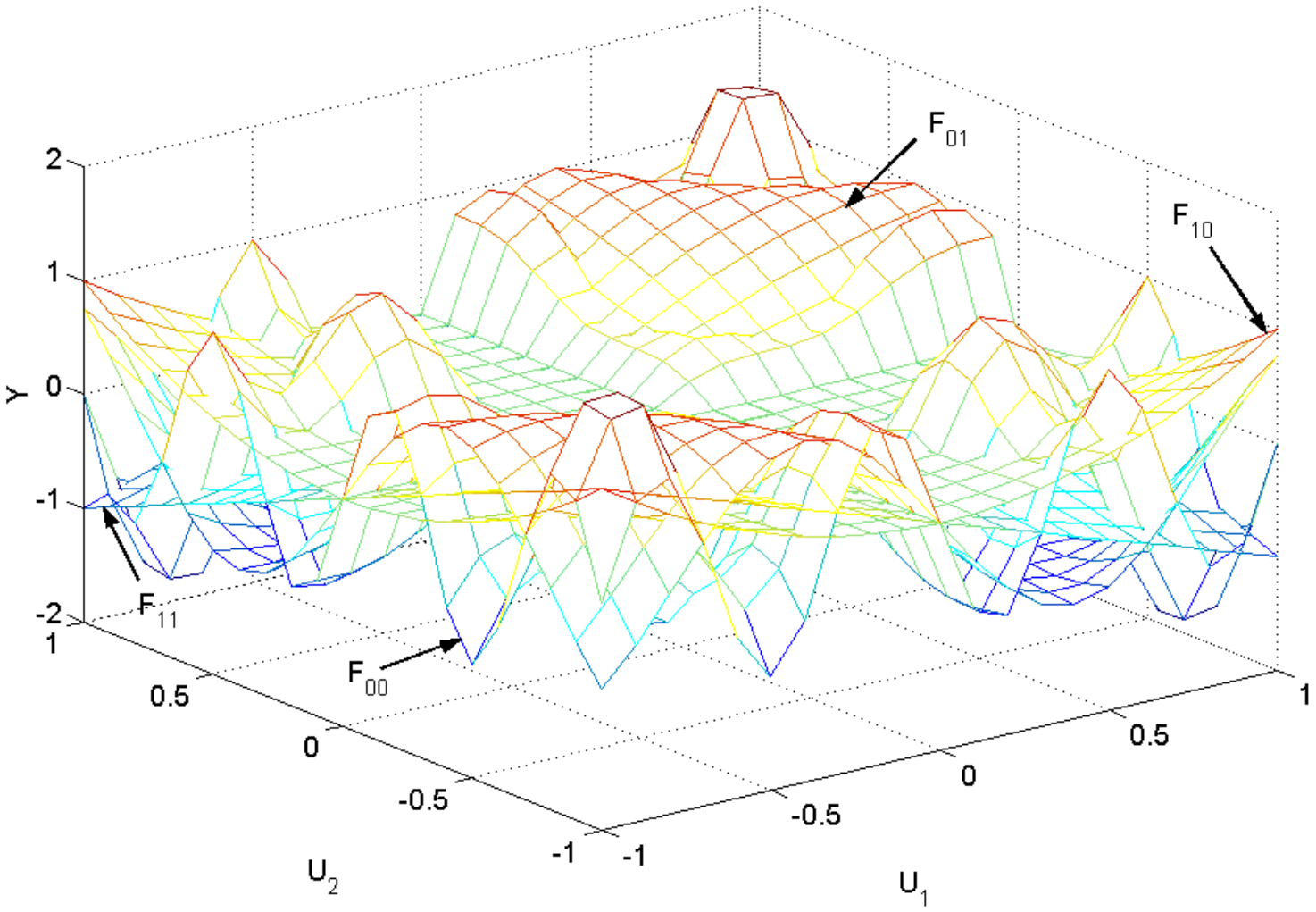

Y=U1sin(3πU2)+U2sin(3πU1)
Y
2
1
0
-1
-2
F00
1
0.5
0
-0.5
-1
U2
-1
-0.5
0
0.5
1
U1


Y=sin(4πU1U2)
F01
2
1
0
-1
-2
Y
1
0.5
0
-0.5
-1
U2
-1
-0.5
0
0.5
1
U1

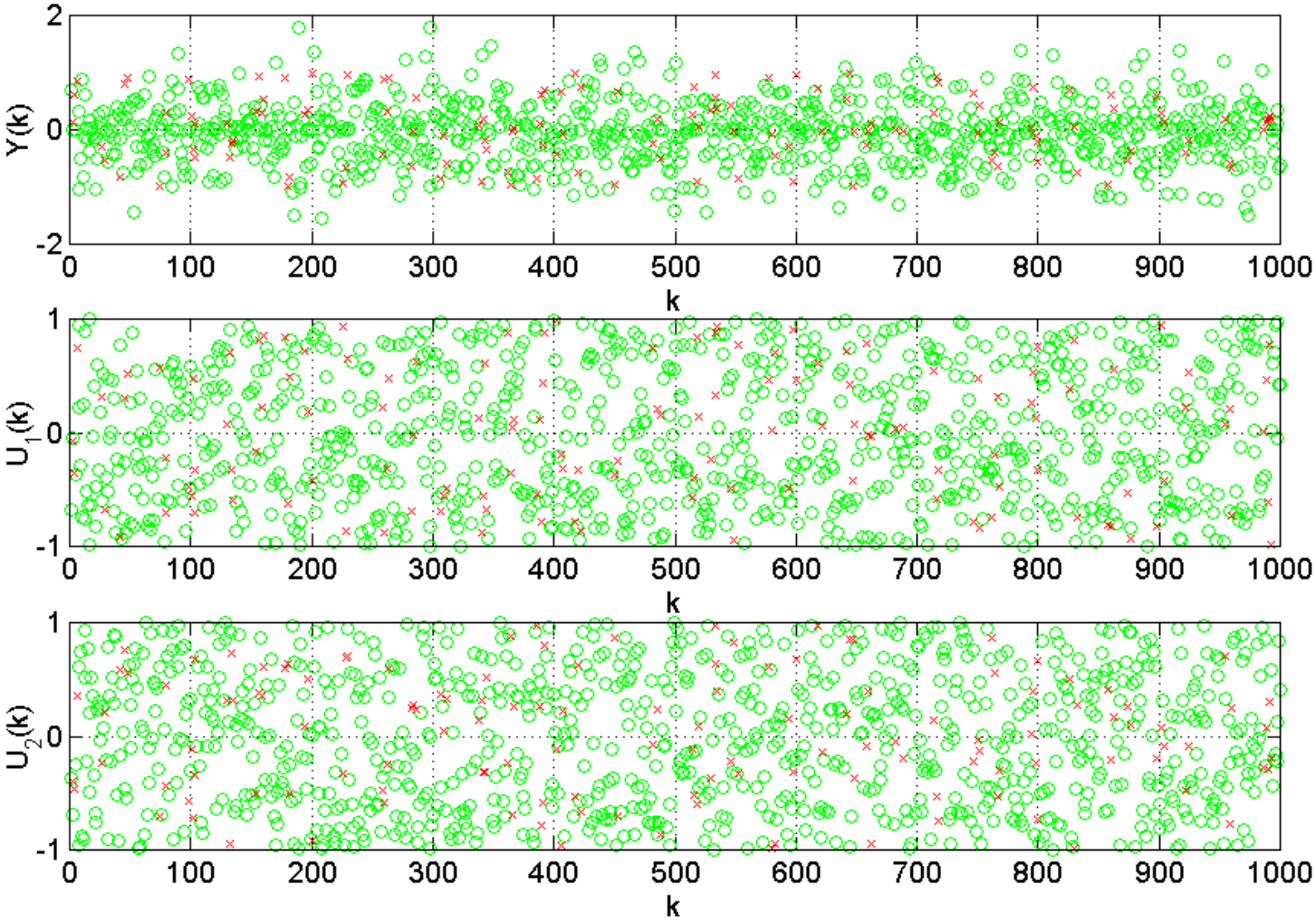

Y=U1^2U2^2
F10
Y
2
1
0
-1
-2
1
0.5
0
-0.5
-1
U2
-1
-0.5
0
0.5
1
U1


Y=U1U2
2
1
0
-1
-2
Y
F11
1
0.5
0
-0.5
-1
U2
-1
-0.5
0
0.5
1
U1

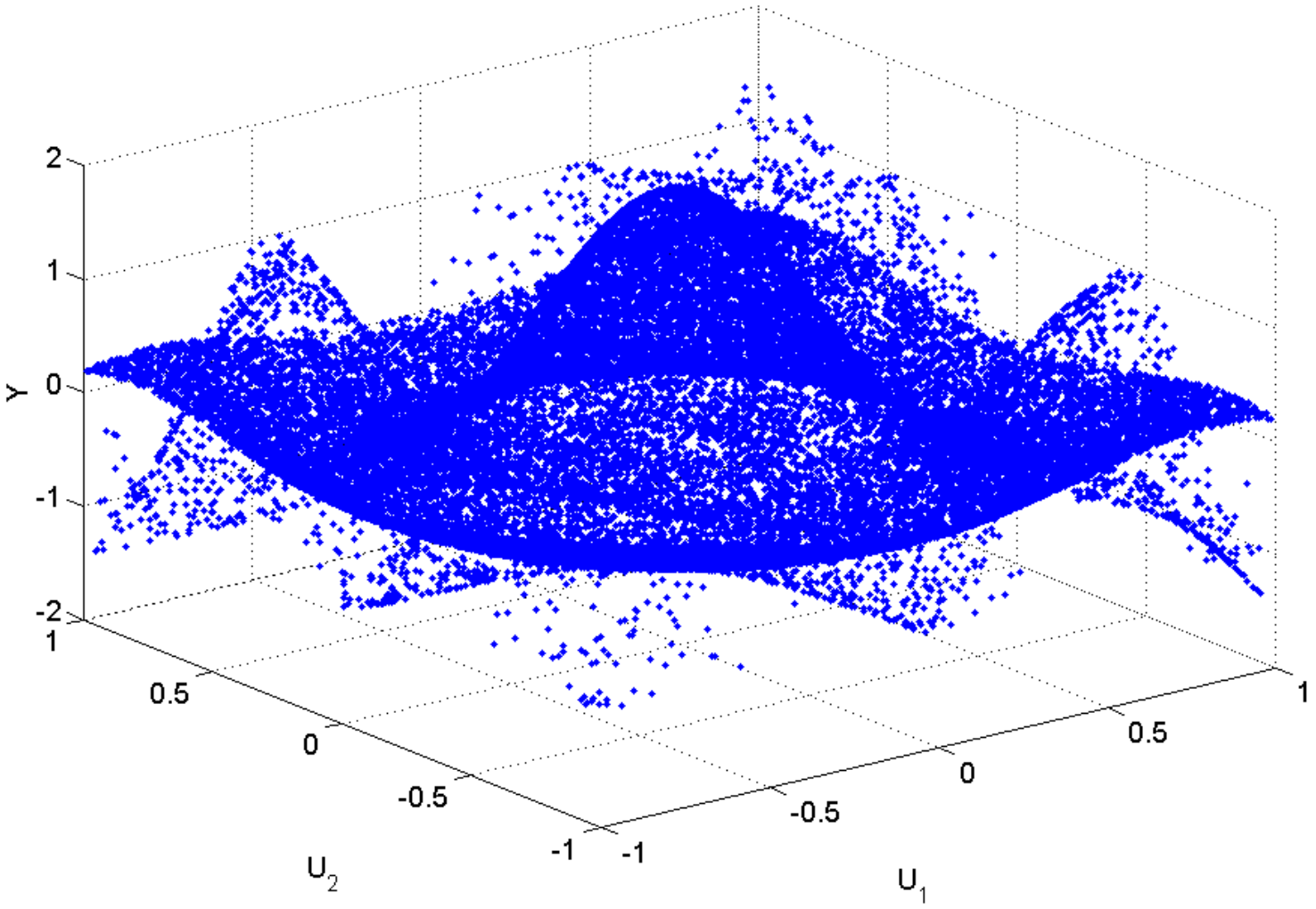
Training set of functions (set A)



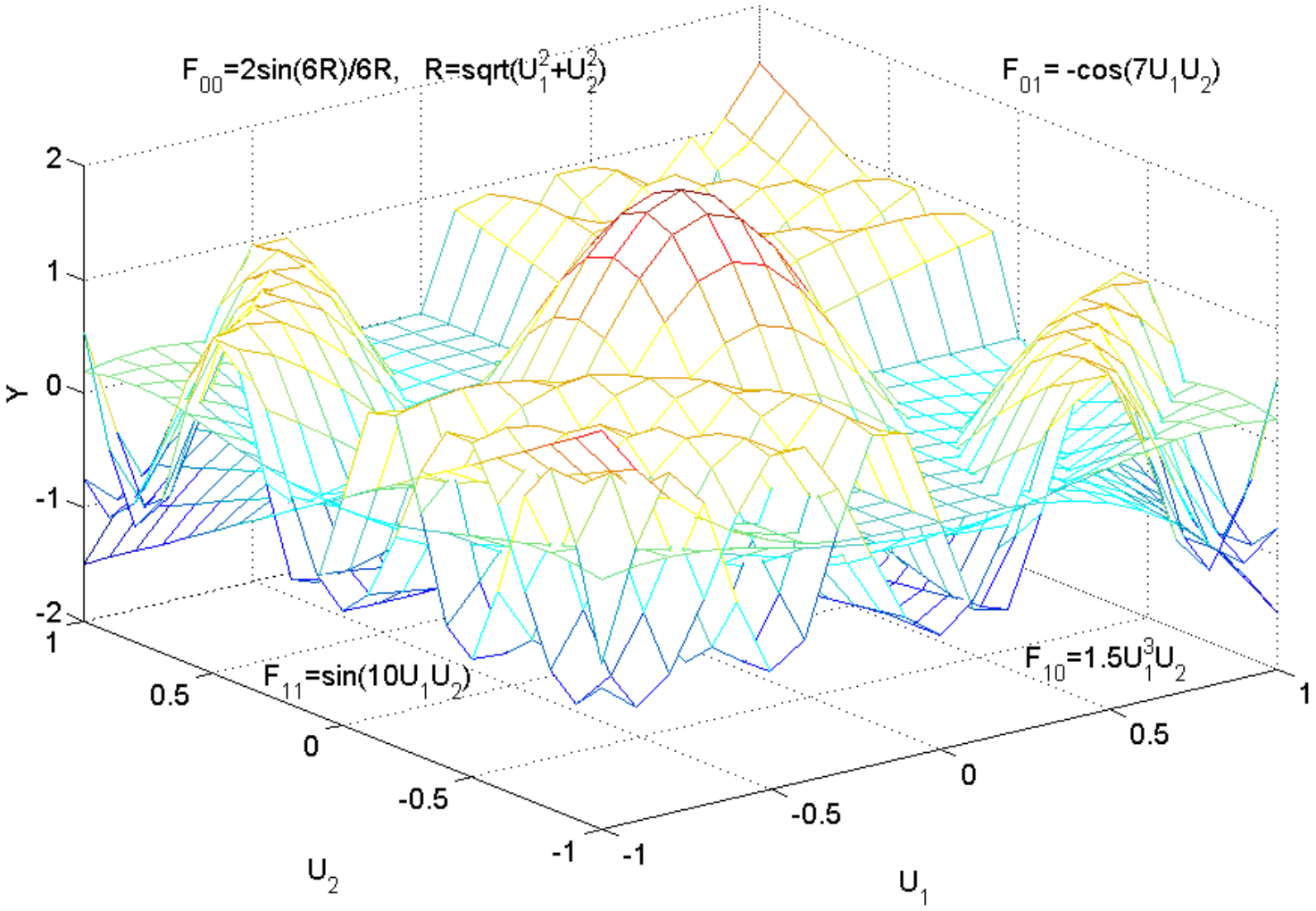
Fragment of sequence for set A (crosses are for ones, circles are for zeros)



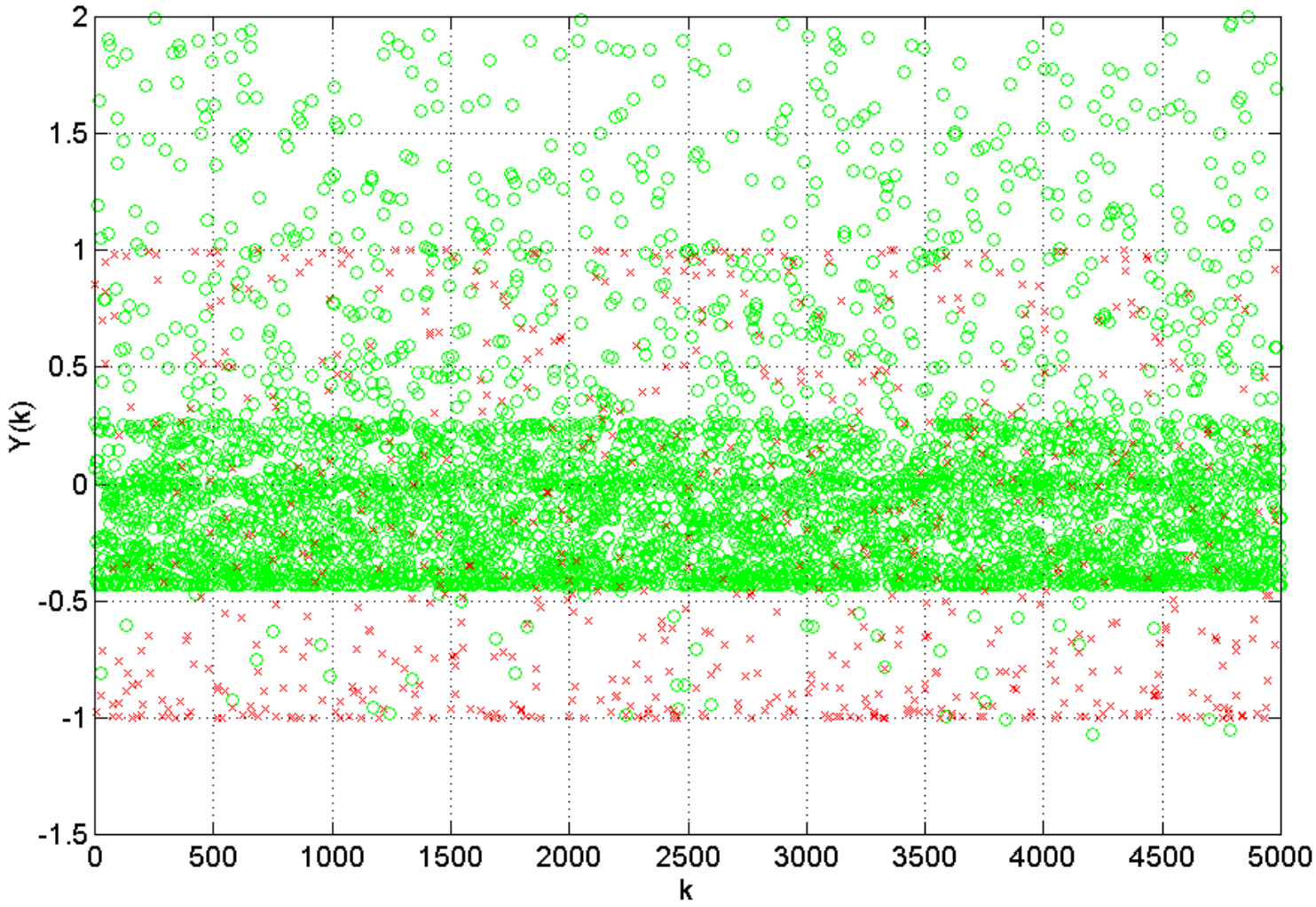
Test set of points (set B)



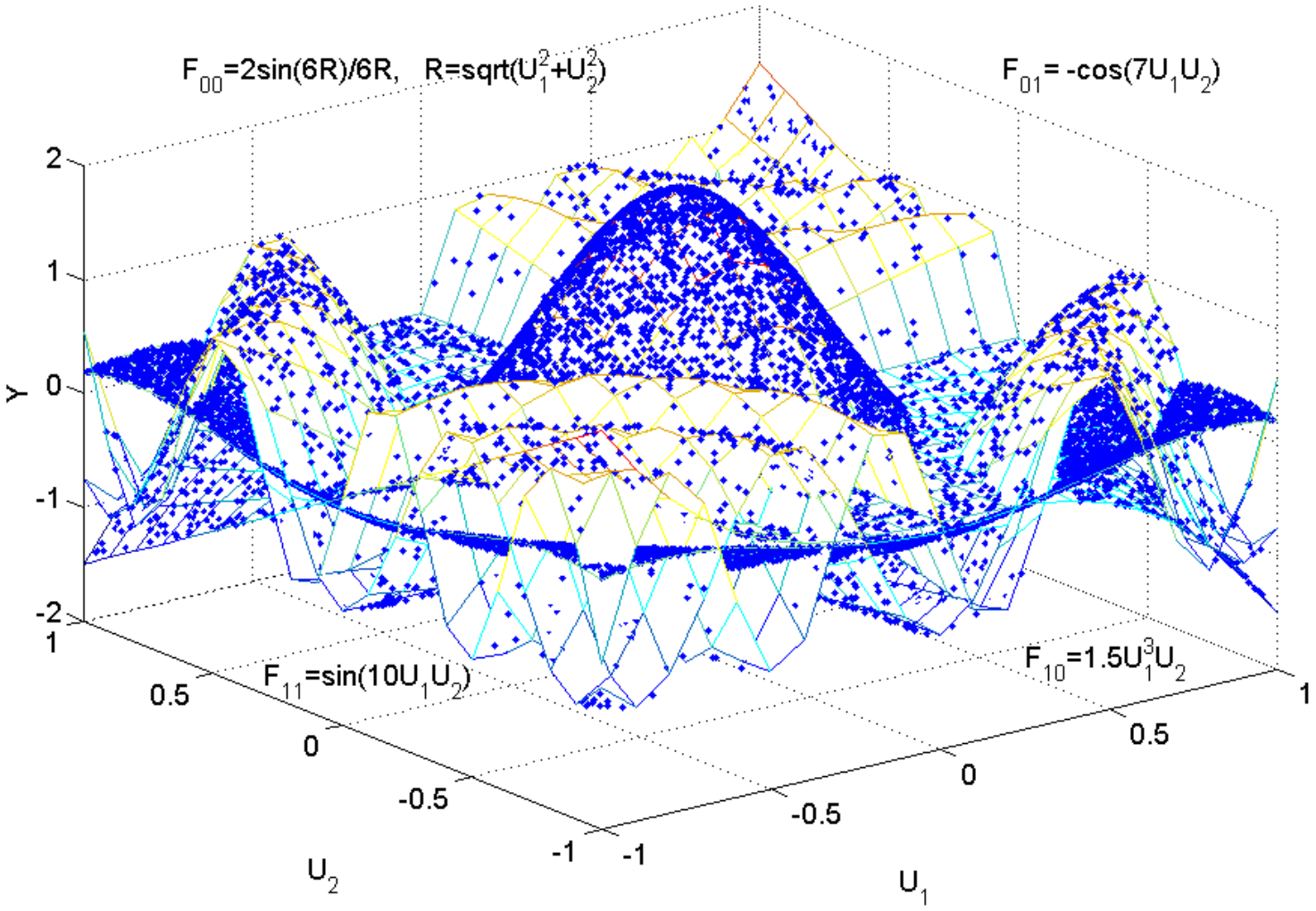
Test set of functions (set B)



Fragment of output sequence for set B



Test set of functions with points (set B)



ENTRIES AND RESULTS (TDC)*

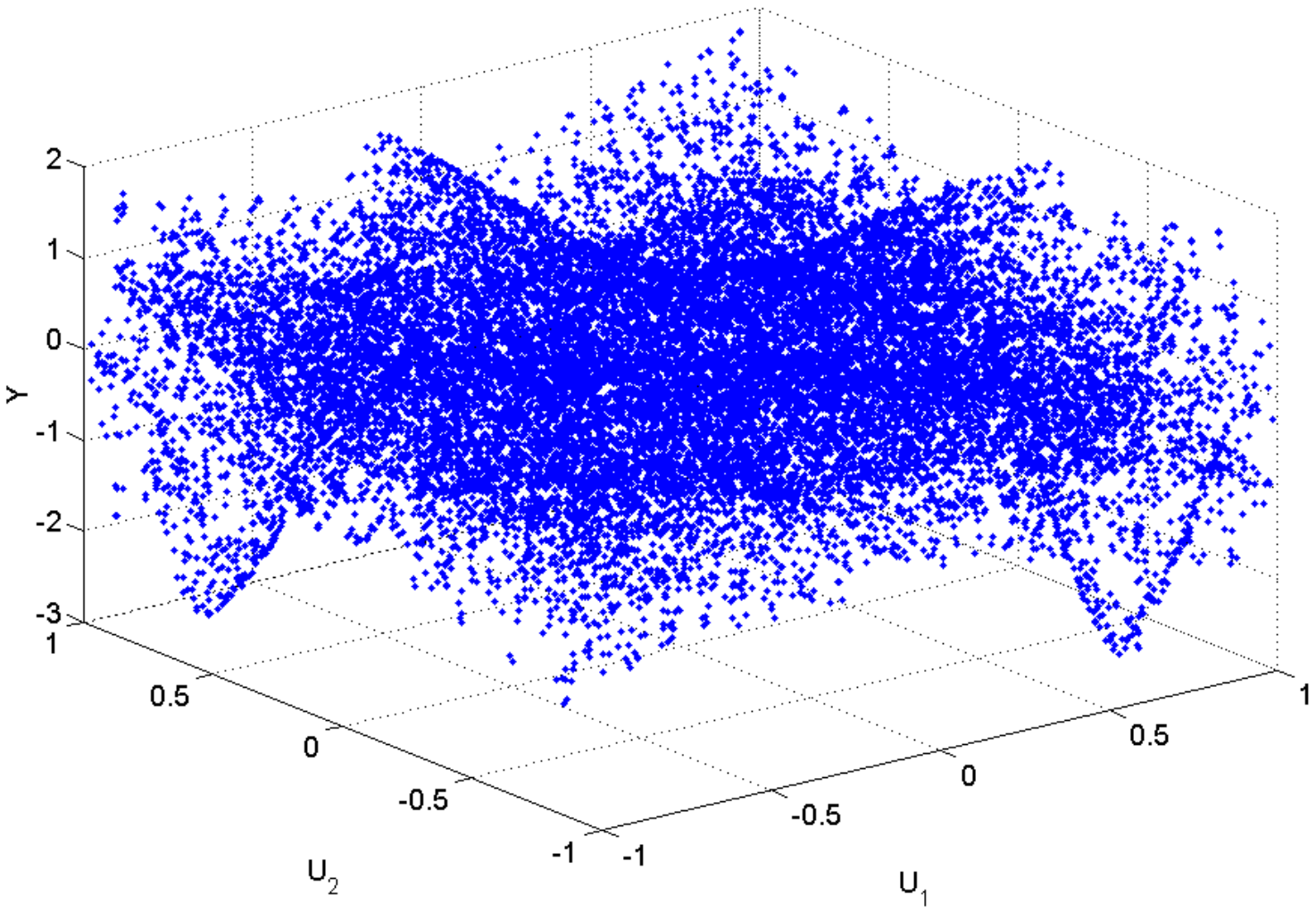
Name	Affiliation	Results
Chih-Chung Chang	National Taiwan University	0 errors (02/13)
Xindi Cai	University of Missouri-Rolla	0 errors (02/28)
Macarie Breazu	University "Lucian Blaga" of Sibiu, Romania	0 errors (04/12)

*Encoded text is an excerpt from "The Heart of a Dog" by Mikhail Bulgakov

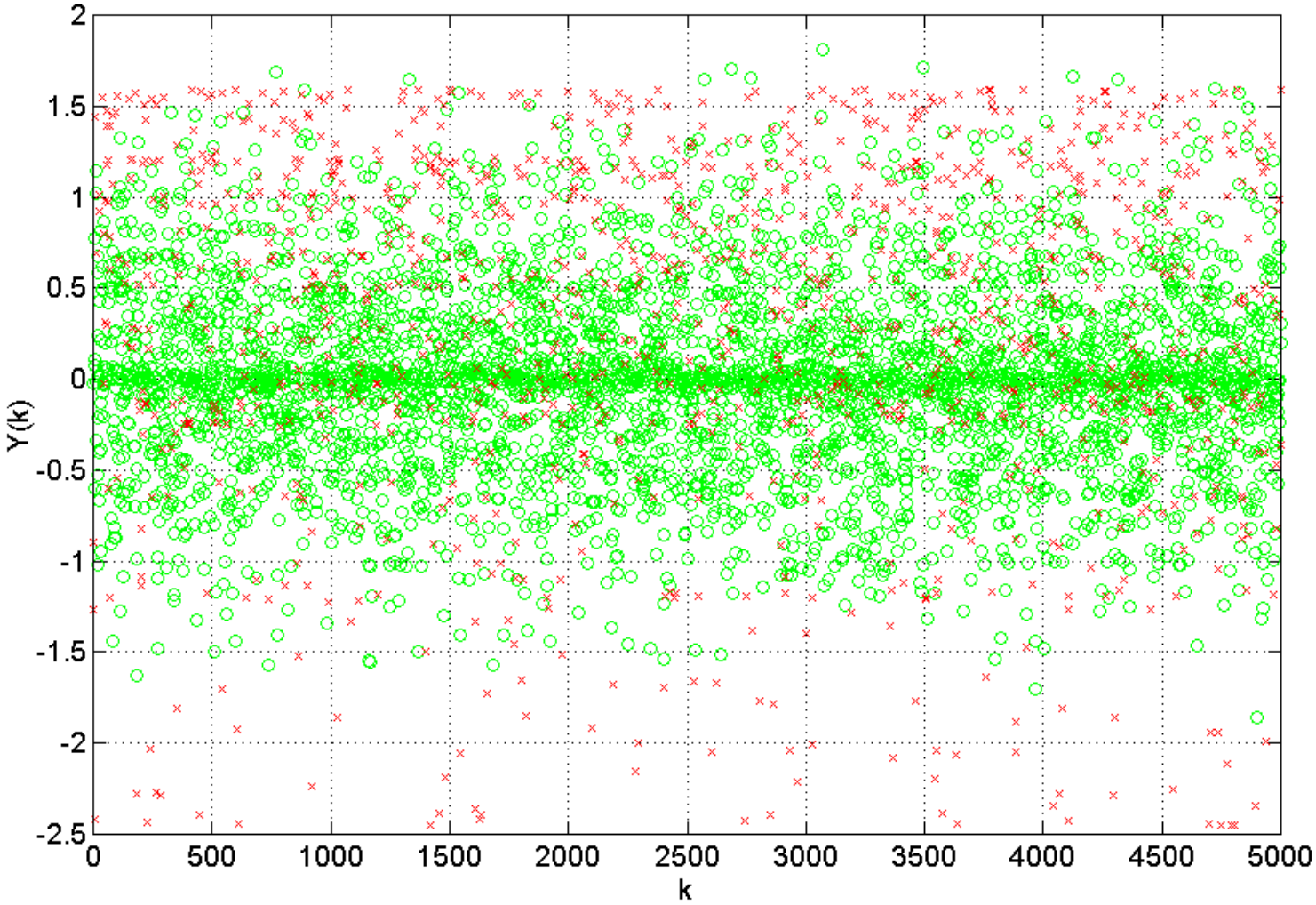
Set C was also provided, but entries were not made mandatory.

Set C does not encode any characters. Probability of ones in set C is 0.3 (not reported to the participants).

Test set of points (set C)



Fragment of output sequence for set C



Test set of functions with points (set C)