Pyramid-based Scatterplots Sampling for Progressive and Streaming Data Visualization

– Supplementary Material –

Xin Chen, Jian Zhang, Chi-Wing Fu, Jean-Daniel Fekete, Yunhai Wang

This supplemental material file provides three additional experimental results for our paper: (i) parameter selection, (ii) the case study of the Stock Price data, and (iii) the complete details about the comparative evaluation.

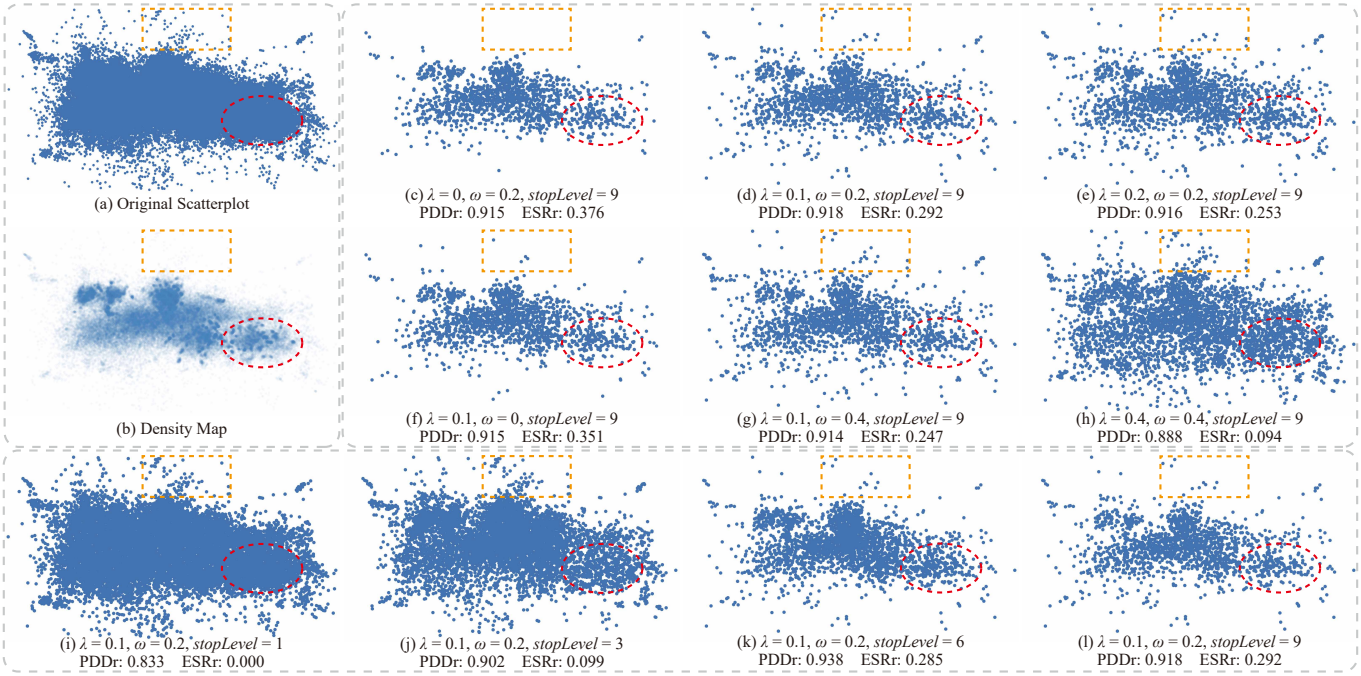


Fig. 1: Parameter analysis on the *Person Activity* dataset [4] with the PDDr and ESRr scores. (a) The opaque scatterplot of the whole 98K data points and (b) rendered as a transparent density map, where some major features are highlighted. The orange dashed boxes and red dashed circles reveal typical low- and medium-density regions, respectively. (c-h) Different sampled results produced by different λ and ω parameters with same *stopLevel*, where each result has around 1.6K data points; the associated specific parameter values and PDDr and ESRr scores are provided under each result. (i-l) Different sampled results produced by increasing *stopLevel* parameter, where the number of data points is decreasing and the associated PDDr and ESRr scores are provided.

1 Parameter Selection

Fig. 1 shows the influences of parameter settings in terms of the PDDr and ESRr scores on the *Person Activity* dataset [4], where the original scatterplot and density map are shown in Figs. 1(a,b).

Robustness: From Figs. 1(c-l), we can see that the PDDr scores of all results are around 0.9 except Fig. 1(i), indicating that our method is robust in the preservation of relative densities over the parameter choices. This observation is also consistent with the patterns shown in Fig. 1(b), where the densities of the central high-density regions are much larger than those in the surrounding low-density regions.

- *X. Chen, Y. Wang are with Shandong University. E-mail: chenxin199634, cloudseawang@gmail.com.*
- *J. Zhang is with CNIC, CAS, China. E-mail: zhangjian@sccas.cn.*
- *C.-W. Fu is with Dept. of Computer Science and Engineering, The Chinese University of Hong Kong, Hong Kong, China. E-mail: cwfu@cse.cuhk.edu.hk*
- *J.-D. Fekete is with University Paris-Saclay, CNRS, Inria, LISN, France. E-mail: Jean-Daniel.Fekete@inria.fr*
- *Y. Wang is the corresponding author.*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

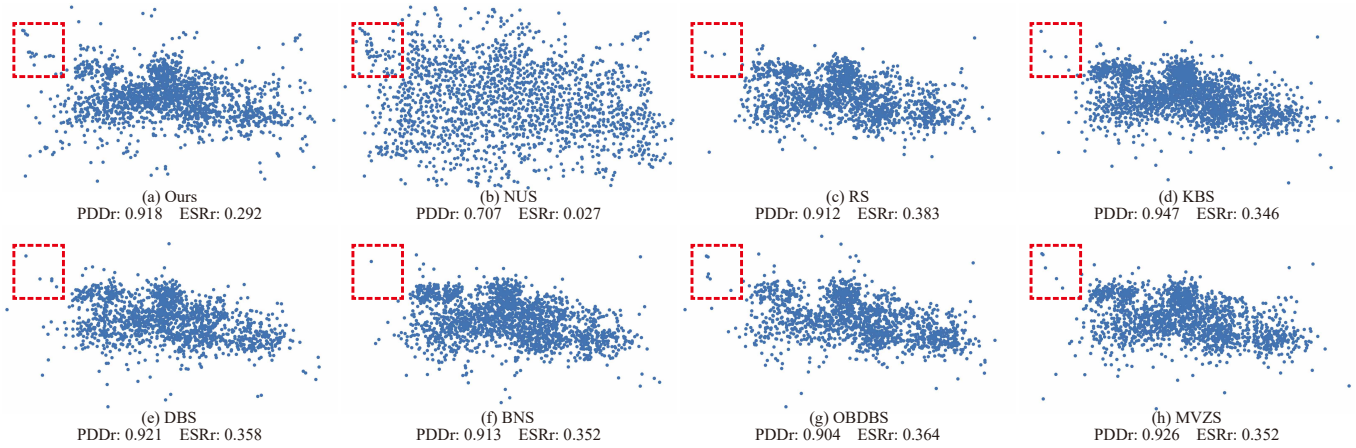


Fig. 2: Comparison of our method and the competing methods on the *Person Activity* dataset [4] with the PDDr and ESRr scores. While our method is competitive with these state-of-the-art methods in preserving relative densities, the red dashed boxes shows that our method can visually maintain the structure in the low-density regions more effectively.

λ **&** ω: In contrast, the ESRr score decreases as λ increases (see Figs. 1(c,d,e)) or as ω increases (see Figs. 1(f,d,g)), where more low-density regions are preserved in the corresponding results. In our experiment, we empirically choose $\lambda = 0.1$ and $\omega = 0.2$, because they lead to the largest PDDr score and a reasonable ESRr value.

stopLevel: Figs. 1(i,j,k,l) show the influence of the *stopLevel* parameter, where a small *stopLevel* value yields small ESRr and PDDr scores except the PDDr score in Fig. 1(l). This is reasonable since a small *stopLevel* value leads to more low-density regions to be preserved and thus the ESRr and PDDr become small. However, computing the PDDr score further depends on the size of the region to be compared. In this example, the region size 40×40 pixels is close to the sub-region size of the *stopLevel* 6, resulting that the PDDr value at *stopLevel* 6 becomes the largest.

Competing methods: Fig. 2 compares the results generated by our method and the competing methods on this dataset. We can see that the PDDr score of our method is very close to those of random sampling (RS), KD-tree-based sampling (KBS) [2], density-based sampling (DBS) [7], blue noise sampling (BNS) [3], and multi-view z-order sampling (MVZS) [5]. Outlier biased density-based sampling (OBDBS) [8] is slightly worse than our method, while non-uniform sampling (NUS) [1] is the worst. In contrast, NUS performs the best in ESRr, while our method is ranked the second and its score is smaller than all the rest methods. This observation is consistent with the finding revealed by our comparative evaluation presented in the paper.

The results indicate that our pyramid-based sampling is robust for preserving relative densities. Thus, we suggest users to preserve outliers by adjusting λ, ω and *stopLevel* for balancing the data and visibility densities. Our experiments show that setting λ and ω to be around 0.1 and 0.2 is good for most data. A small *stopLevel* may help to show more outliers by adding more samples to low density regions, whereas the proper level can be adjusted for different data sets and user preferences.

2 Case Study: Stock Price

Next, we present two additional density maps of the Stock Price case study in Fig. 3, which is about how the Sep. 11 attacks affected the stock market on the relationship between the stock volume and stock percentage change. We compared two scatterplots (see Figs. 3(a,b)) generated from the dataset for two different time-ranges: (i) from Aug. 11, 2001 to Sep. 10, 2001 (see Fig. 3(a)) and (ii) the whole Sep. 2001 (see Fig. 3(b)). As we can see, the changes in stock percentage have much larger variations in Fig. 3(b) than those in Fig. 3(a). Also, more points with negative stock changes can be observed than those with positive changes. Based on these observations, we see that the Sep. 11 attacks resulted in a negative effect and many stocks show large fluctuations. However, the streaming sampling results in Figs. 3(c,d) show that the horizontal major trend keeps stable around 0% in both ranges, as shown in Figs. 3(e,f). Further comparing the samples in the red and orange boxes reveals more stocks with negative stock changes and large fluctuations in Sep. 2001. Therefore, we conclude that our streaming sampling method produces faithful visualizations.

3 Comparative Evaluation

Finally, we elaborate on the comparative evaluation of our static sampling with 12 synthetic datasets created by mixing Gaussian distributions and a uniform distribution, and 28 real datasets collected from Kaggle [6] and the UCI data repository [4]. Specifically, we provide screenshots of the original scatterplot and sampled results generated by our approach (PBS) and seven competitive sampling methods: random sampling (RS), blue noise sampling (BNS) [3], density-based sampling (DBS) [7], non-uniform sampling (NUS) [1], outlier biased density-based sampling (OBDBS) [8], multi-view z-order sampling (MVZS) [5], and KD-tree-based sampling (KBS) [2] with their corresponding scores.

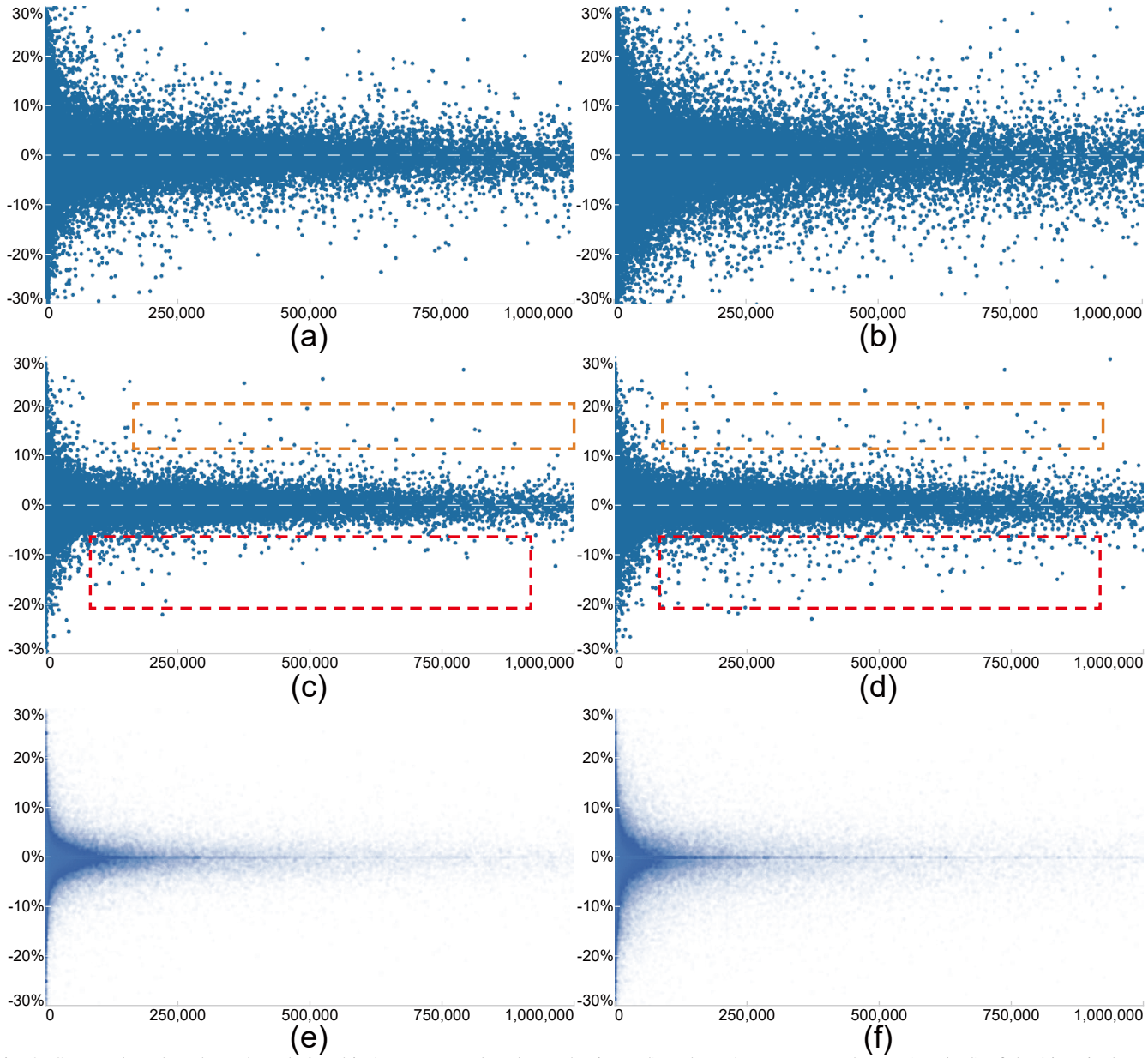
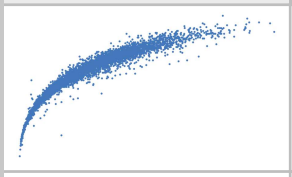
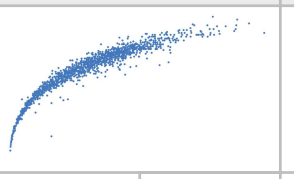
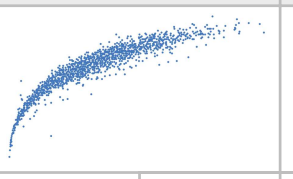
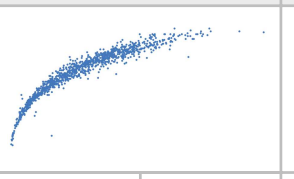
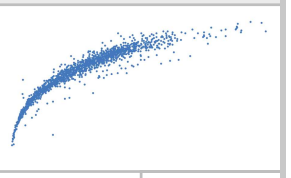
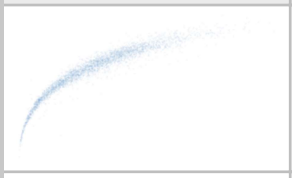
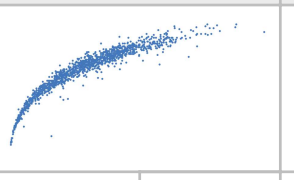
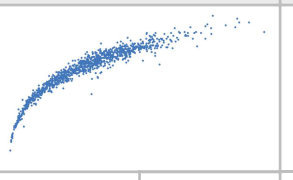
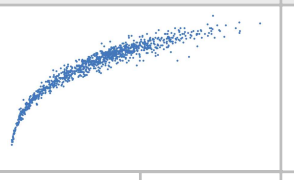
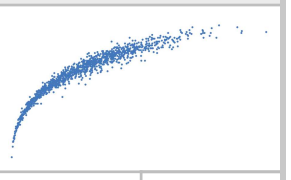
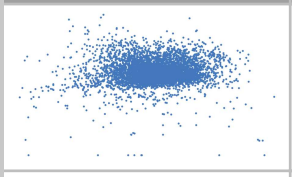
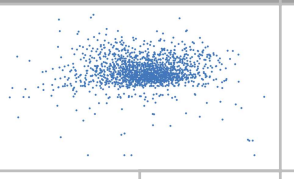
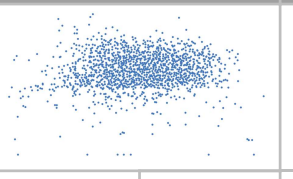
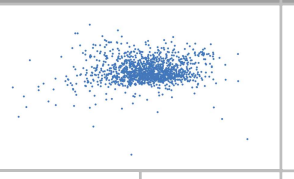
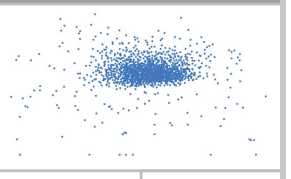
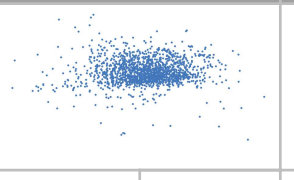
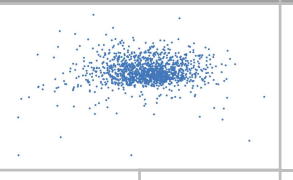
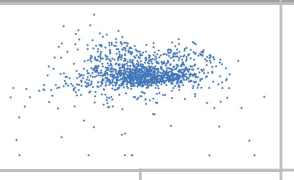
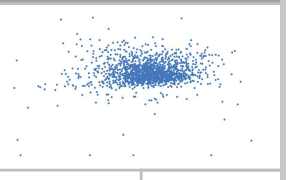
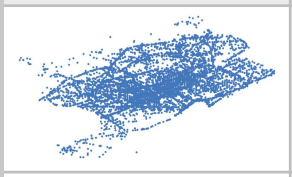
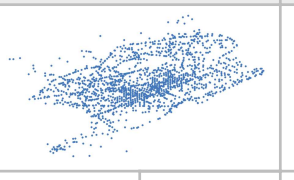
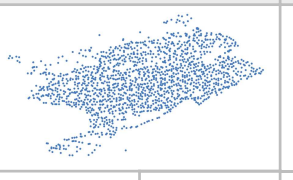
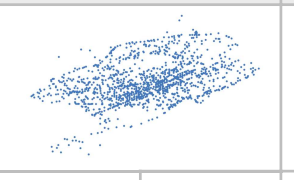
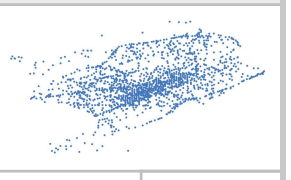
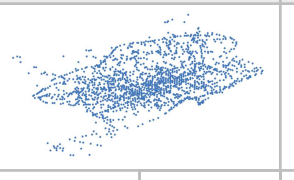
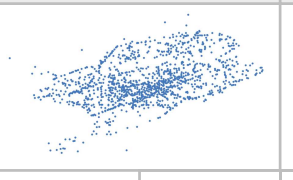
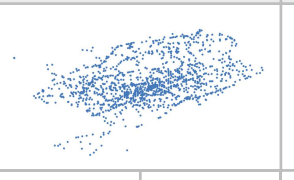
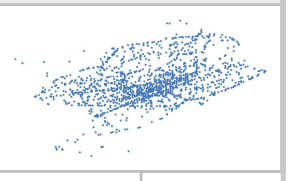
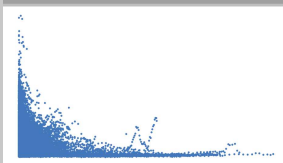
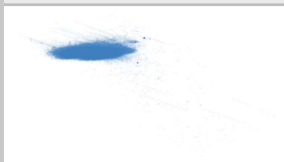
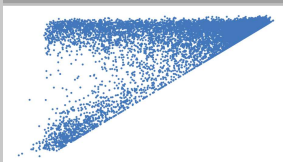
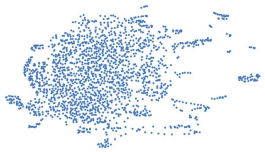
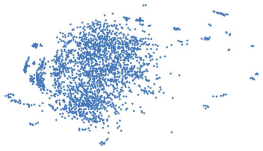
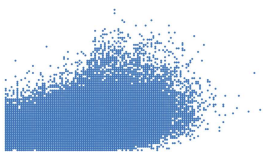
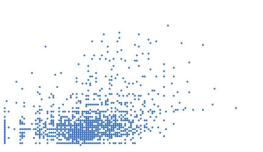
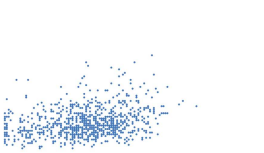
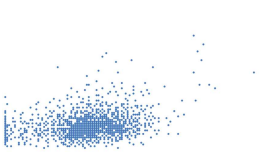
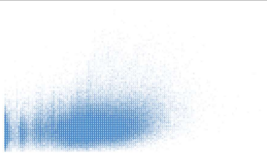
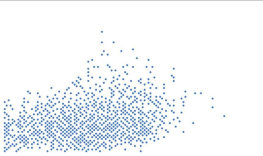
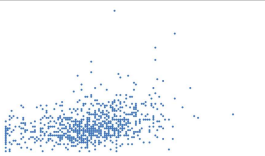
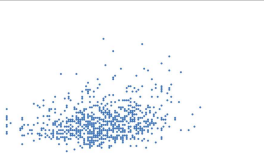
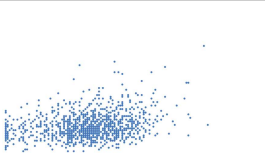
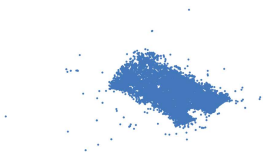
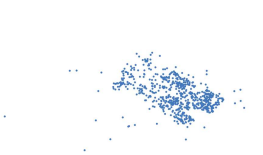
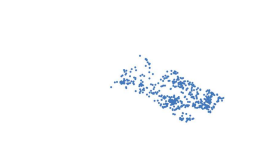
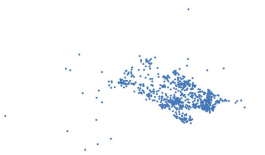
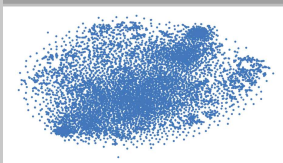
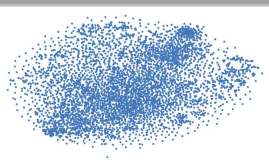
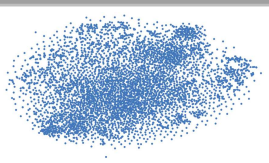
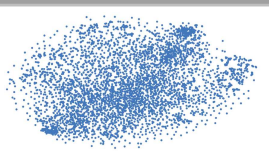
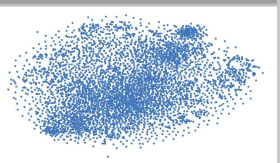
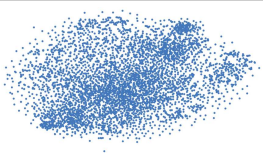
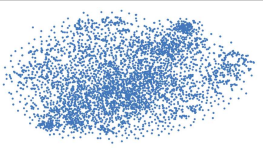
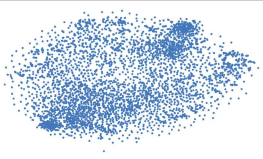
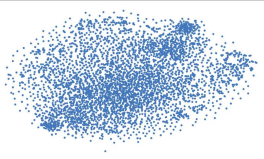
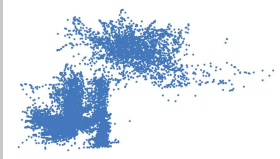
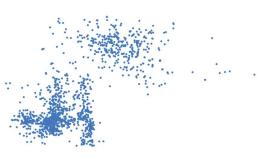
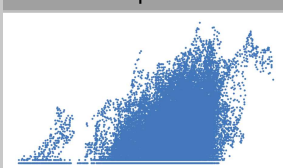
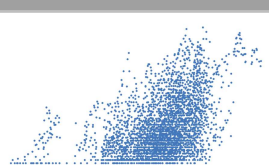
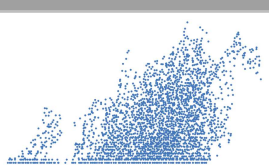
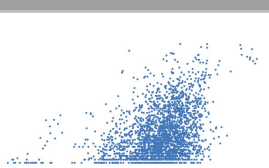
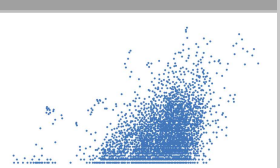
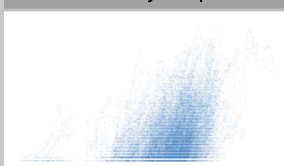
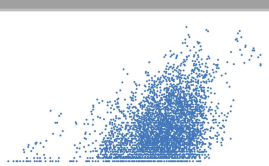
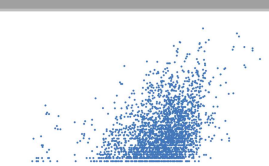
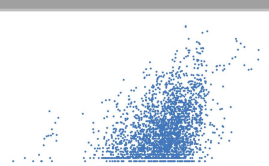
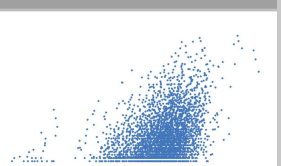


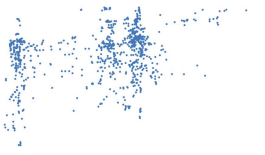
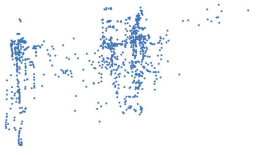
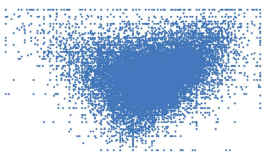
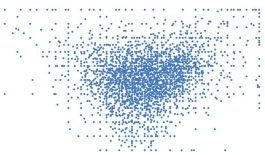
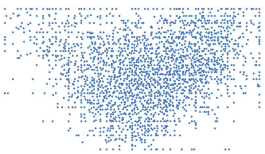
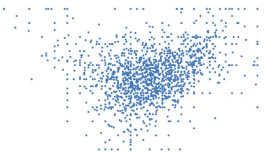
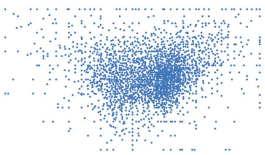
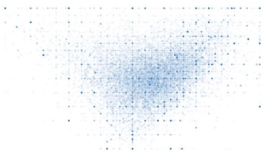
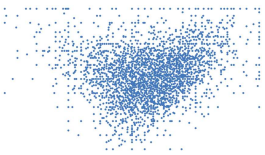
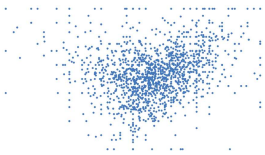
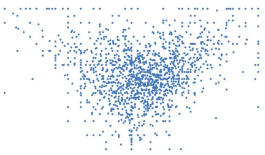
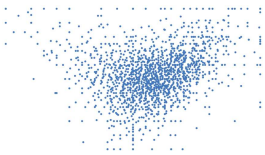
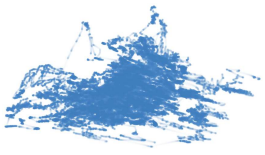
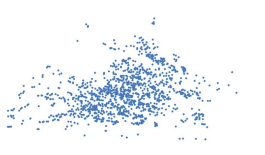
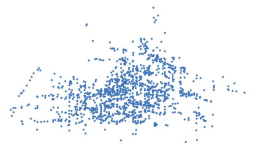
Fig. 3: Scatterplots that show the relationship between stock volume (horizontal) and stock percentage change (vertical) of the historical stock market dataset for two different time ranges: before the September 11 attacks (left column) and the whole September 2001 (right column). (a,b) the overplotted scatterplots of the original data; (c,d) streaming visualization results of our method from (a,b); and (e,f) the corresponding density maps of (a,b).

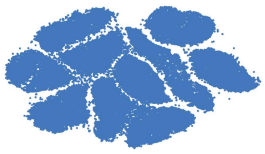
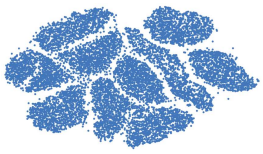
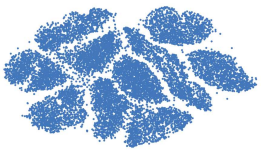
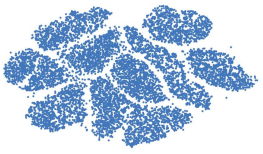
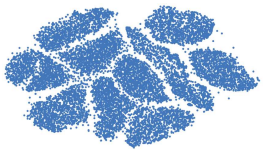
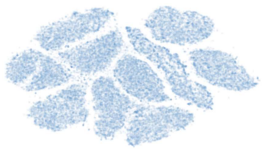
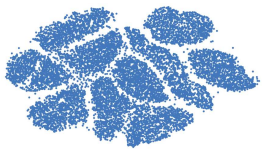
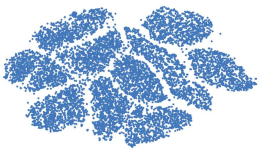
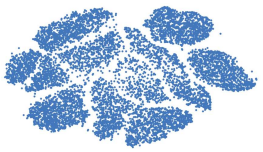
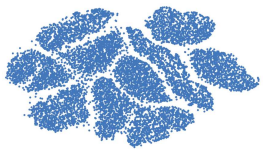
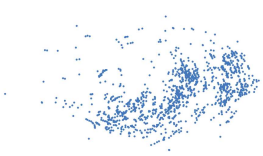
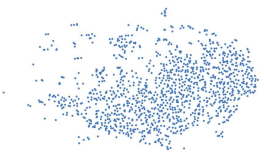
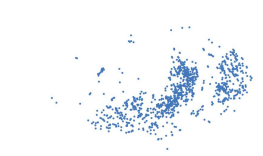
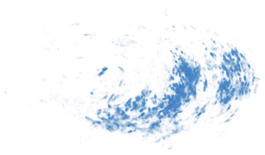
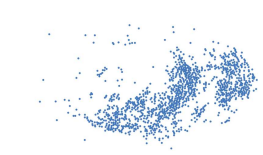
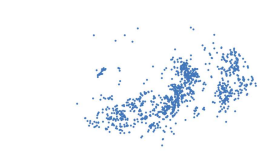
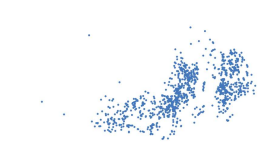
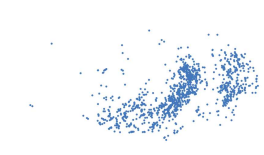
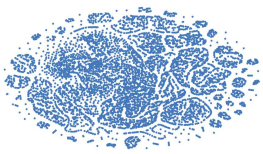
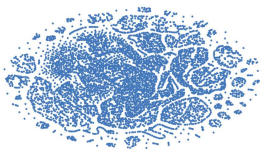
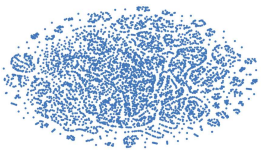
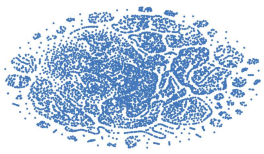
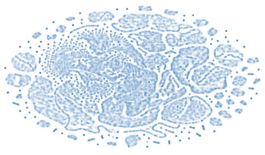
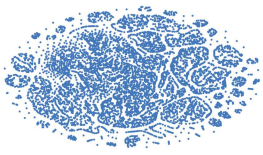
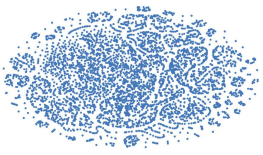
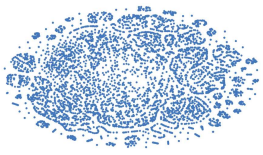
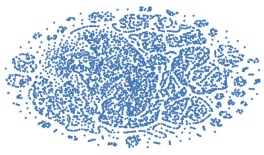
abalone										
Input	PBS		NUS		RS		KBS			
										
# points: 4177	PDDr: 0.93	ESRr: 0.04	PDDr: 0.91	ESRr: 0.00	PDDr: 0.93	ESRr: 0.06	PDDr: 0.94	ESRr: 0.03		
Density map	BNS		DBS		OBDBS		MVZS			
										
# samples: ~1091	PDDr: 0.94	ESRr: 0.05	PDDr: 0.93	ESRr: 0.07	PDDr: 0.91	ESRr: 0.05	PDDr: 0.95	ESRr: 0.06		
avila										
Input	PBS		NUS		RS		KBS			
										
# points: 5672	PDDr: 0.94	ESRr: 0.08	PDDr: 0.88	ESRr: 0.00	PDDr: 0.94	ESRr: 0.14	PDDr: 0.93	ESRr: 0.10		
Density map	BNS		DBS		OBDBS		MVZS			
										
# samples: ~1366	PDDr: 0.94	ESRr: 0.11	PDDr: 0.93	ESRr: 0.13	PDDr: 0.91	ESRr: 0.09	PDDr: 0.94	ESRr: 0.13		
Barcelona										
Input	PBS		NUS		RS		KBS			
										
# points: 10284	PDDr: 0.84	ESRr: 0.06	PDDr: 0.67	ESRr: 0.01	PDDr: 0.83	ESRr: 0.11	PDDr: 0.86	ESRr: 0.06		
Density map	BNS		DBS		OBDBS		MVZS			
										
# samples: ~1566	PDDr: 0.83	ESRr: 0.07	PDDr: 0.82	ESRr: 0.10	PDDr: 0.83	ESRr: 0.10	PDDr: 0.86	ESRr: 0.10		

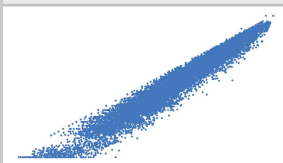
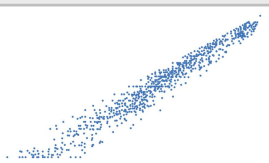
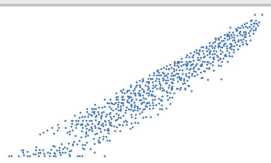
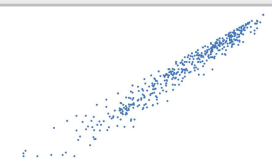
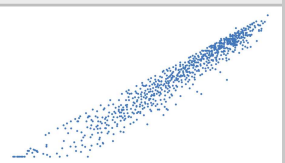
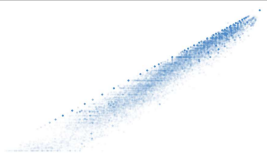
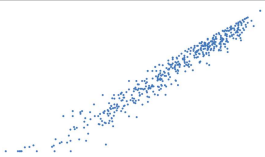
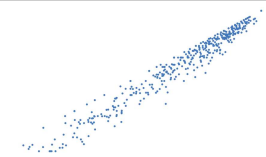
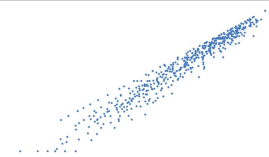
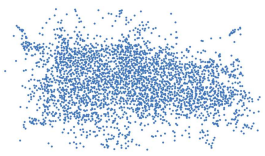
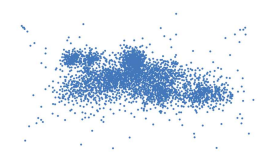
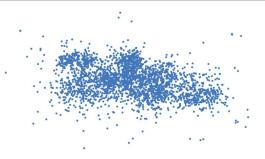
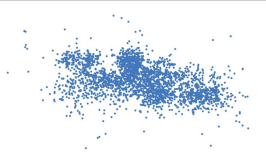
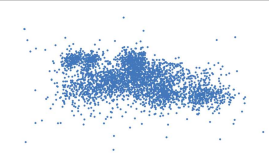
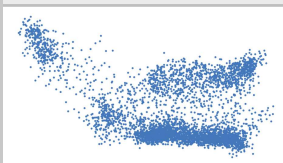
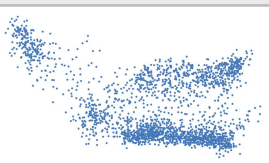
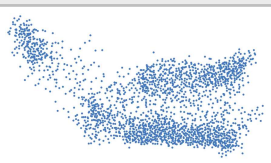
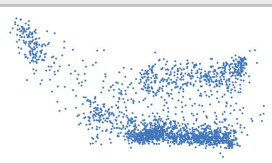
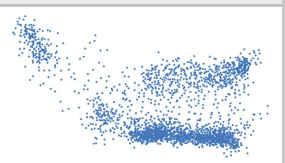
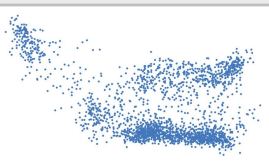
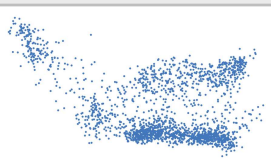
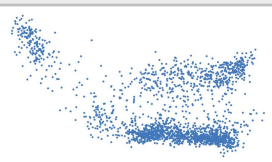
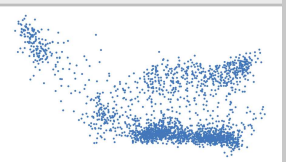
Beijing									
Input	PBS		NUS		RS		KBS		
									
# points: 41757	PDDr: 0.95	ESRr: 0.06	PDDr: 0.52	ESRr: 0.01	PDDr: 0.93	ESRr: 0.12	PDDr: 0.93	ESRr: 0.08	
Density map	BNS		DBS		OBDBS		MVZS		
									
# samples: ~366	PDDr: 0.96	ESRr: 0.09	PDDr: 0.93	ESRr: 0.13	PDDr: 0.93	ESRr: 0.11	PDDr: 0.96	ESRr: 0.12	
CreditCardFraud									
Input	PBS		NUS		RS		KBS		
									
# points: 284807	PDDr: 0.97	ESRr: 0.17	PDDr: 0.65	ESRr: 0.02	PDDr: 0.98	ESRr: 0.26	PDDr: 0.98	ESRr: 0.26	
Density map	BNS		DBS		OBDBS		MVZS		
									
# samples: ~452	PDDr: 0.98	ESRr: 0.24	PDDr: 0.97	ESRr: 0.26	PDDr: 0.96	ESRr: 0.25	PDDr: 0.98	ESRr: 0.26	
crowdsourced									
Input	PBS		NUS		RS		KBS		
									
# points: 10545	PDDr: 0.89	ESRr: 0.09	PDDr: 0.79	ESRr: 0.03	PDDr: 0.88	ESRr: 0.16	PDDr: 0.90	ESRr: 0.12	
Density map	BNS		DBS		OBDBS		MVZS		
									
# samples: ~1261	PDDr: 0.90	ESRr: 0.09	PDDr: 0.87	ESRr: 0.14	PDDr: 0.86	ESRr: 0.20	PDDr: 0.90	ESRr: 0.16	

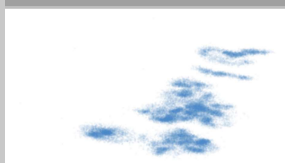
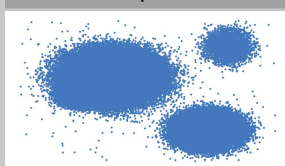
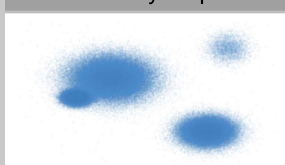
DailySports								
Input	PBS		NUS		RS		KBS	
								
# points: 540000	PDDr: 0.81	ESRr: 0.13	PDDr: 0.70	ESRr: 0.05	PDDr: 0.91	ESRr: 0.22	PDDr: 0.87	ESRr: 0.17
Density map	BNS		DBS		OBDBS		MVZS	
								
# samples: ~1890	PDDr: 0.75	ESRr: 0.15	PDDr: 0.90	ESRr: 0.21	PDDr: 0.89	ESRr: 0.22	PDDr: 0.91	ESRr: 0.21
diabetes								
Input	PBS		NUS		RS		KBS	
								
# points: 99493	PDDr: 0.87	ESRr: 0.26	PDDr: 0.46	ESRr: 0.05	PDDr: 0.87	ESRr: 0.33	PDDr: 0.90	ESRr: 0.31
Density map	BNS		DBS		OBDBS		MVZS	
								
# samples: ~844	PDDr: 0.84	ESRr: 0.23	PDDr: 0.86	ESRr: 0.34	PDDr: 0.83	ESRr: 0.33	PDDr: 0.89	ESRr: 0.33
Emergency_911Calls								
Input	PBS		NUS		RS		KBS	
								
# points: 332591	PDDr: 0.92	ESRr: 0.07	PDDr: 0.74	ESRr: 0.00	PDDr: 0.90	ESRr: 0.11	PDDr: 0.93	ESRr: 0.09
Density map	BNS		DBS		OBDBS		MVZS	
								
# samples: ~657	PDDr: 0.93	ESRr: 0.10	PDDr: 0.92	ESRr: 0.11	PDDr: 0.91	ESRr: 0.11	PDDr: 0.93	ESRr: 0.11

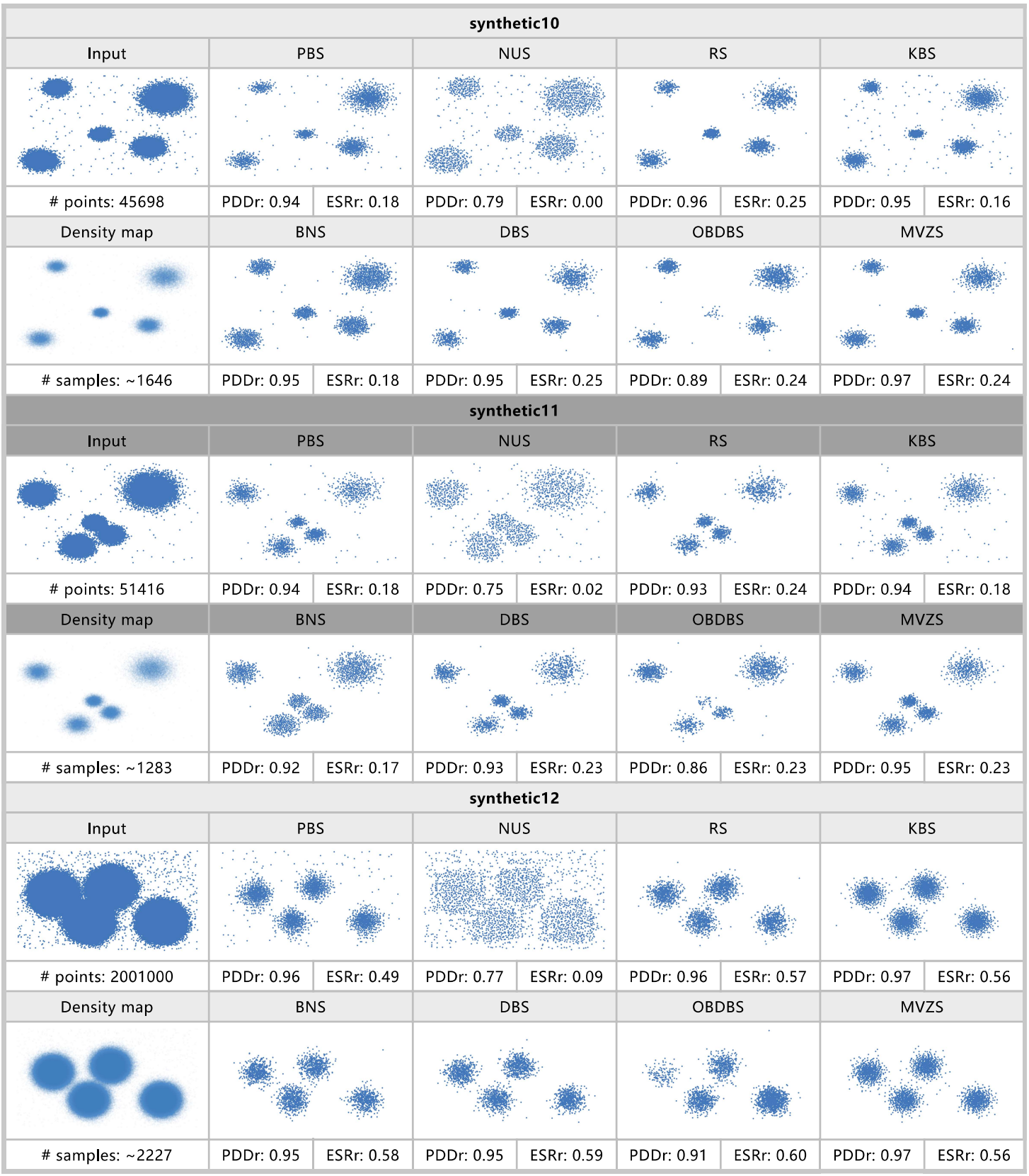
Epileptic									
Input	PBS		NUS		RS		KBS		
									
# points: 11500	PDDr: 0.88	ESRr: 0.01	PDDr: 0.88	ESRr: 0.01	PDDr: 0.86	ESRr: 0.03	PDDr: 0.90	ESRr: 0.01
Density map	BNS		DBS		OBDBS		MVZS		
									
# samples: ~4824	PDDr: 0.87	ESRr: 0.01	PDDr: 0.85	ESRr: 0.02	PDDr: 0.75	ESRr: 0.02	PDDr: 0.90	ESRr: 0.01
facial									
Input	PBS		NUS		RS		KBS		
									
# points: 12903	PDDr: 0.91	ESRr: 0.08	PDDr: 0.82	ESRr: 0.01	PDDr: 0.91	ESRr: 0.16	PDDr: 0.91	ESRr: 0.12
Density map	BNS		DBS		OBDBS		MVZS		
									
# samples: ~1448	PDDr: 0.92	ESRr: 0.10	PDDr: 0.91	ESRr: 0.14	PDDr: 0.91	ESRr: 0.15	PDDr: 0.91	ESRr: 0.14
forest_covertype									
Input	PBS		NUS		RS		KBS		
									
# points: 41802	PDDr: 0.92	ESRr: 0.08	PDDr: 0.84	ESRr: 0.01	PDDr: 0.93	ESRr: 0.18	PDDr: 0.93	ESRr: 0.14
Density map	BNS		DBS		OBDBS		MVZS		
									
# samples: ~2534	PDDr: 0.91	ESRr: 0.10	PDDr: 0.93	ESRr: 0.18	PDDr: 0.93	ESRr: 0.16	PDDr: 0.95	ESRr: 0.17

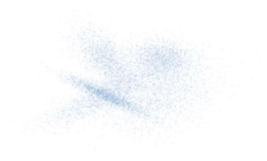
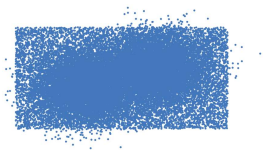
gesture									
Input	PBS		NUS		RS		KBS		
									
# points: 4845	PDDr: 0.87	ESRr: 0.07	PDDr: 0.78	ESRr: 0.01	PDDr: 0.88	ESRr: 0.13	PDDr: 0.87	ESRr: 0.08	
Density map	BNS		DBS		OBDBS		MVZS		
									
# samples: ~1038	PDDr: 0.87	ESRr: 0.08	PDDr: 0.88	ESRr: 0.12	PDDr: 0.85	ESRr: 0.13	PDDr: 0.90	ESRr: 0.13	
google_sentiment									
Input	PBS		NUS		RS		KBS		
									
# points: 37427	PDDr: 0.82	ESRr: 0.16	PDDr: 0.72	ESRr: 0.03	PDDr: 0.82	ESRr: 0.28	PDDr: 0.83	ESRr: 0.15	
Density map	BNS		DBS		OBDBS		MVZS		
									
# samples: ~2165	PDDr: 0.82	ESRr: 0.15	PDDr: 0.83	ESRr: 0.27	PDDr: 0.79	ESRr: 0.24	PDDr: 0.84	ESRr: 0.23	
HT_sensor									
Input	PBS		NUS		RS		KBS		
									
# points: 1565007	PDDr: 0.90	ESRr: 0.11	PDDr: 0.81	ESRr: 0.04	PDDr: 0.91	ESRr: 0.18	PDDr: 0.93	ESRr: 0.15	
Density map	BNS		DBS		OBDBS		MVZS		
									
# samples: ~2320	PDDr: 0.86	ESRr: 0.15	PDDr: 0.88	ESRr: 0.23	PDDr: 0.87	ESRr: 0.21	PDDr: 0.88	ESRr: 0.23	

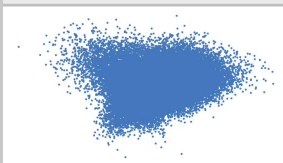
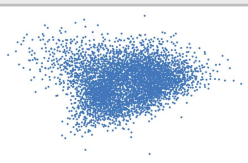
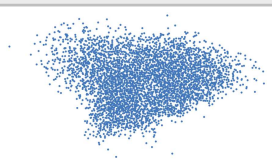
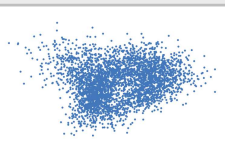
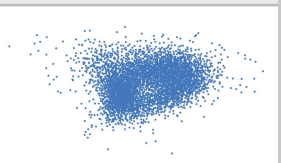
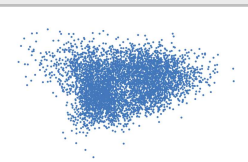
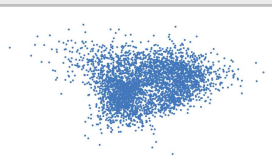
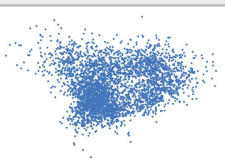
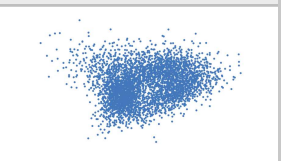
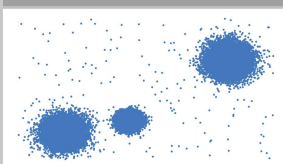
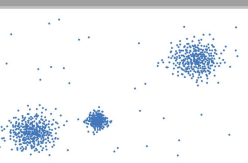
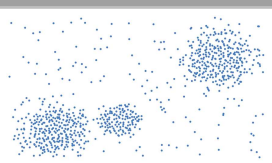
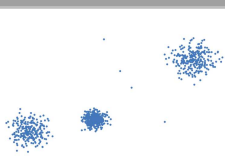
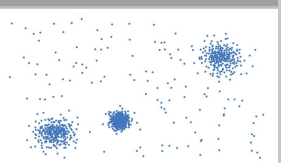
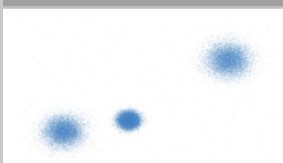
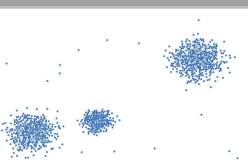
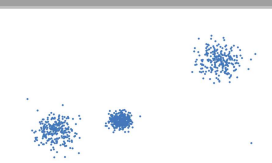
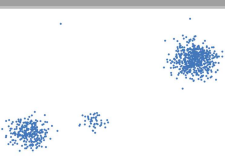
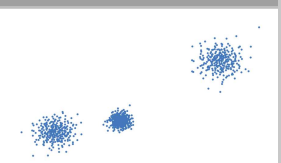
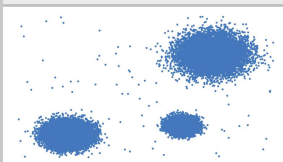
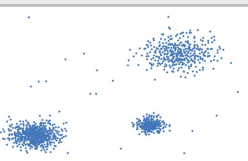
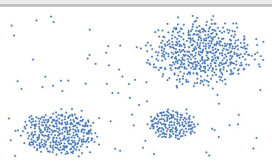
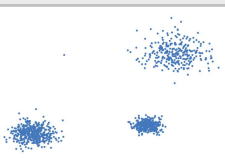
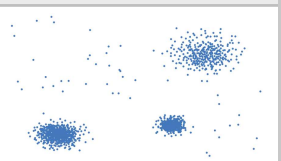
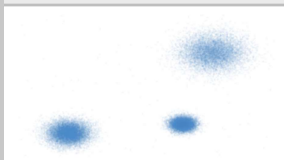
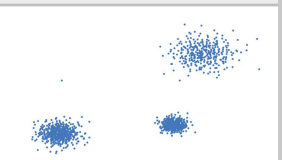
mnist									
Input	PBS		NUS		RS		KBS		
									
# points: 70000	PDDr: 0.84	ESRr: 0.02	PDDr: 0.87	ESRr: 0.02	PDDr: 0.83	ESRr: 0.03	PDDr: 0.87	ESRr: 0.02	
Density map	BNS		DBS		OBDBS		MVZS		
									
# samples: ~9534	PDDr: 0.83	ESRr: 0.02	PDDr: 0.81	ESRr: 0.04	PDDr: 0.70	ESRr: 0.03	PDDr: 0.90	ESRr: 0.03	
MoCap									
Input	PBS		NUS		RS		KBS		
									
# points: 74975	PDDr: 0.91	ESRr: 0.17	PDDr: 0.70	ESRr: 0.03	PDDr: 0.92	ESRr: 0.26	PDDr: 0.93	ESRr: 0.23	
Density map	BNS		DBS		OBDBS		MVZS		
									
# samples: ~1223	PDDr: 0.91	ESRr: 0.18	PDDr: 0.91	ESRr: 0.25	PDDr: 0.89	ESRr: 0.27	PDDr: 0.91	ESRr: 0.24	
NewsPopularity									
Input	PBS		NUS		RS		KBS		
									
# points: 93239	PDDr: 0.76	ESRr: 0.01	PDDr: 0.80	ESRr: 0.02	PDDr: 0.74	ESRr: 0.02	PDDr: 0.80	ESRr: 0.01	
Density map	BNS		DBS		OBDBS		MVZS		
									
# samples: ~6279	PDDr: 0.76	ESRr: 0.02	PDDr: 0.73	ESRr: 0.02	PDDr: 0.56	ESRr: 0.02	PDDr: 0.78	ESRr: 0.01	

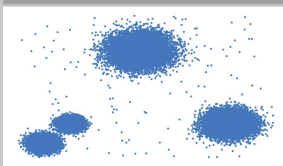
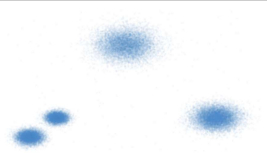
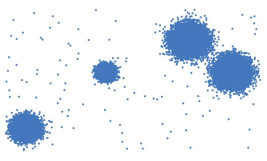
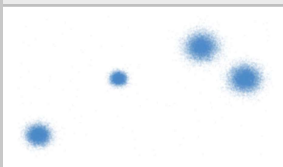
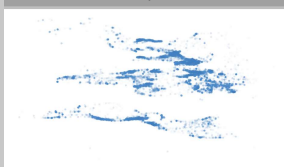
Nomao									
Input		PBS		NUS		RS		KBS	
									
# points: 34410		PDDr: 0.77	ESRr: 0.06	PDDr: 0.63	ESRr: 0.03	PDDr: 0.83	ESRr: 0.11	PDDr: 0.83	ESRr: 0.04
Density map		BNS		DBS		OBDBS		MVZS	
									
# samples: ~722		PDDr: 0.80	ESRr: 0.06	PDDr: 0.80	ESRr: 0.11	PDDr: 0.80	ESRr: 0.11	PDDr: 0.82	ESRr: 0.10
PersonActivity_4classes									
Input		PBS		NUS		RS		KBS	
									
# points: 98568		PDDr: 0.94	ESRr: 0.21	PDDr: 0.83	ESRr: 0.02	PDDr: 0.94	ESRr: 0.33	PDDr: 0.96	ESRr: 0.29
Density map		BNS		DBS		OBDBS		MVZS	
									
# samples: ~2878		PDDr: 0.93	ESRr: 0.24	PDDr: 0.95	ESRr: 0.31	PDDr: 0.93	ESRr: 0.32	PDDr: 0.95	ESRr: 0.32
satimage									
Input		PBS		NUS		RS		KBS	
									
# points: 4435		PDDr: 0.92	ESRr: 0.04	PDDr: 0.89	ESRr: 0.00	PDDr: 0.90	ESRr: 0.08	PDDr: 0.92	ESRr: 0.06
Density map		BNS		DBS		OBDBS		MVZS	
									
# samples: ~2033		PDDr: 0.91	ESRr: 0.07	PDDr: 0.91	ESRr: 0.08	PDDr: 0.90	ESRr: 0.09	PDDr: 0.92	ESRr: 0.08

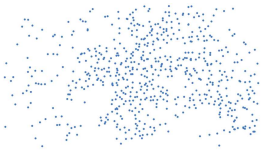
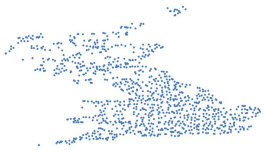
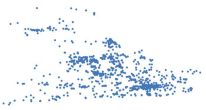
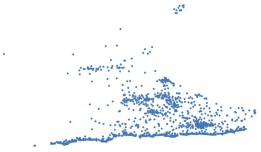
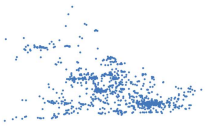
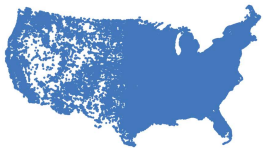
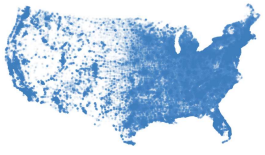
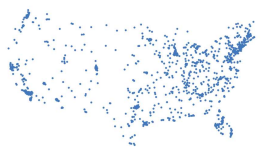
sensorless									
Input	PBS		NUS		RS		KBS		
									
# points: 42552	PDDr: 0.92	ESRr: 0.07	PDDr: 0.81	ESRr: 0.01	PDDr: 0.92	ESRr: 0.08	PDDr: 0.94	ESRr: 0.06
Density map	BNS		DBS		OBDBS		MVZS		
									
# samples: ~1986	PDDr: 0.92	ESRr: 0.06	PDDr: 0.93	ESRr: 0.08	PDDr: 0.92	ESRr: 0.08	PDDr: 0.95	ESRr: 0.09
Spotify									
Input	PBS		NUS		RS		KBS		
									
# points: 174389	PDDr: 0.93	ESRr: 0.15	PDDr: 0.84	ESRr: 0.02	PDDr: 0.92	ESRr: 0.20	PDDr: 0.95	ESRr: 0.16
Density map	BNS		DBS		OBDBS		MVZS		
									
# samples: ~3481	PDDr: 0.90	ESRr: 0.16	PDDr: 0.91	ESRr: 0.19	PDDr: 0.90	ESRr: 0.18	PDDr: 0.95	ESRr: 0.18
synthetic1									
Input	PBS		NUS		RS		KBS		
									
# points: 255147	PDDr: 0.94	ESRr: 0.28	PDDr: 0.78	ESRr: 0.03	PDDr: 0.95	ESRr: 0.38	PDDr: 0.96	ESRr: 0.34
Density map	BNS		DBS		OBDBS		MVZS		
									
# samples: ~2170	PDDr: 0.94	ESRr: 0.29	PDDr: 0.95	ESRr: 0.36	PDDr: 0.86	ESRr: 0.37	PDDr: 0.96	ESRr: 0.37

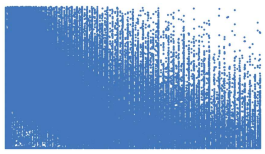
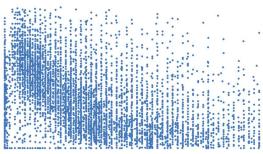
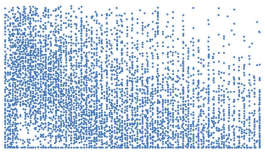
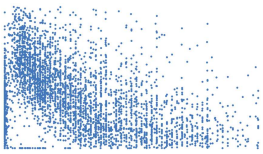
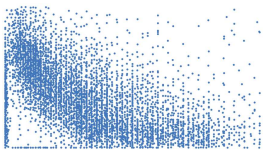
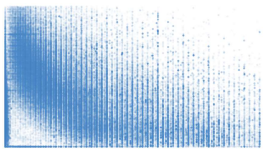
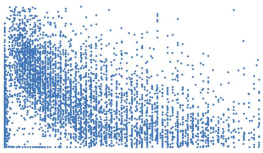
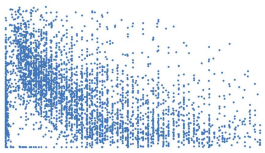
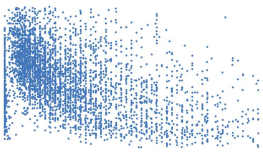
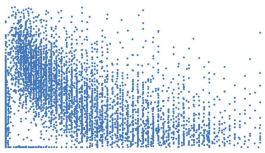

synthetic10
Input
PBS
NUS
RS
KBS
points: 45698
PDDr: 0.94
ESRr: 0.18
PDDr: 0.79
ESRr: 0.00
PDDr: 0.96
ESRr: 0.25
PDDr: 0.95
ESRr: 0.16
Density map
BNS
DBS
OBDBS
MVZS
samples: ~1646
PDDr: 0.95
ESRr: 0.18
PDDr: 0.95
ESRr: 0.25
PDDr: 0.89
ESRr: 0.24
PDDr: 0.97
ESRr: 0.24
synthetic11
Input
PBS
NUS
RS
KBS
points: 51416
PDDr: 0.94
ESRr: 0.18
PDDr: 0.75
ESRr: 0.02
PDDr: 0.93
ESRr: 0.24
PDDr: 0.94
ESRr: 0.18
Density map
BNS
DBS
OBDBS
MVZS
samples: ~1283
PDDr: 0.92
ESRr: 0.17
PDDr: 0.93
ESRr: 0.23
PDDr: 0.86
ESRr: 0.23
PDDr: 0.95
ESRr: 0.23
synthetic12
Input
PBS
NUS
RS
KBS
points: 2001000
PDDr: 0.96
ESRr: 0.49
PDDr: 0.77
ESRr: 0.09
PDDr: 0.96
ESRr: 0.57
PDDr: 0.97
ESRr: 0.56
Density map
BNS
DBS
OBDBS
MVZS
samples: ~2227
PDDr: 0.95
ESRr: 0.58
PDDr: 0.95
ESRr: 0.59
PDDr: 0.91
ESRr: 0.60
PDDr: 0.97
ESRr: 0.56

synthetic2									
Input	PBS		NUS		RS		KBS		
									
# points: 10000	PDDr: 0.94	ESRr: 0.05	PDDr: 0.89	ESRr: 0.00	PDDr: 0.93	ESRr: 0.11	PDDr: 0.94	ESRr: 0.05	
Density map	BNS		DBS		OBDBS		MVZS		
									
# samples: ~2806	PDDr: 0.94	ESRr: 0.08	PDDr: 0.93	ESRr: 0.08	PDDr: 0.89	ESRr: 0.09	PDDr: 0.95	ESRr: 0.10	
synthetic3									
Input	PBS		NUS		RS		KBS		
									
# points: 10000	PDDr: 0.91	ESRr: 0.05	PDDr: 0.88	ESRr: 0.00	PDDr: 0.91	ESRr: 0.10	PDDr: 0.94	ESRr: 0.06	
Density map	BNS		DBS		OBDBS		MVZS		
									
# samples: ~3138	PDDr: 0.92	ESRr: 0.07	PDDr: 0.92	ESRr: 0.08	PDDr: 0.87	ESRr: 0.06	PDDr: 0.94	ESRr: 0.09	
synthetic4									
Input	PBS		NUS		RS		KBS		
									
# points: 30000	PDDr: 0.91	ESRr: 0.03	PDDr: 0.88	ESRr: 0.01	PDDr: 0.90	ESRr: 0.06	PDDr: 0.93	ESRr: 0.03	
Density map	BNS		DBS		OBDBS		MVZS		
									
# samples: ~6035	PDDr: 0.90	ESRr: 0.04	PDDr: 0.90	ESRr: 0.05	PDDr: 0.82	ESRr: 0.04	PDDr: 0.93	ESRr: 0.04	

synthetic5									
Input	PBS		NUS		RS		KBS		
									
# points: 30000	PDDr: 0.93	ESRr: 0.08	PDDr: 0.86	ESRr: 0.00	PDDr: 0.94	ESRr: 0.14	PDDr: 0.95	ESRr: 0.12	
Density map	BNS		DBS		OBDBS		MVZS		
									
# samples: ~3545	PDDr: 0.93	ESRr: 0.11	PDDr: 0.93	ESRr: 0.13	PDDr: 0.88	ESRr: 0.11	PDDr: 0.95	ESRr: 0.15	
synthetic6									
Input	PBS		NUS		RS		KBS		
									
# points: 37606	PDDr: 0.96	ESRr: 0.21	PDDr: 0.80	ESRr: 0.02	PDDr: 0.95	ESRr: 0.27	PDDr: 0.93	ESRr: 0.14	
Density map	BNS		DBS		OBDBS		MVZS		
									
# samples: ~879	PDDr: 0.94	ESRr: 0.20	PDDr: 0.95	ESRr: 0.27	PDDr: 0.87	ESRr: 0.26	PDDr: 0.95	ESRr: 0.26	
synthetic7									
Input	PBS		NUS		RS		KBS		
									
# points: 37177	PDDr: 0.95	ESRr: 0.17	PDDr: 0.78	ESRr: 0.02	PDDr: 0.95	ESRr: 0.21	PDDr: 0.94	ESRr: 0.17	
Density map	BNS		DBS		OBDBS		MVZS		
									
# samples: ~1078	PDDr: 0.94	ESRr: 0.16	PDDr: 0.94	ESRr: 0.22	PDDr: 0.86	ESRr: 0.21	PDDr: 0.96	ESRr: 0.20	

synthetic8									
Input	PBS		NUS		RS		KBS		
									
# points: 40517	PDDr: 0.95	ESRr: 0.18	PDDr: 0.82	ESRr: 0.02	PDDr: 0.94	ESRr: 0.24	PDDr: 0.94	ESRr: 0.17	
Density map	BNS		DBS		OBDBS		MVZS		
									
# samples: ~1194	PDDr: 0.94	ESRr: 0.18	PDDr: 0.94	ESRr: 0.23	PDDr: 0.86	ESRr: 0.24	PDDr: 0.96	ESRr: 0.23	
synthetic9									
Input	PBS		NUS		RS		KBS		
									
# points: 37423	PDDr: 0.96	ESRr: 0.13	PDDr: 0.81	ESRr: 0.00	PDDr: 0.95	ESRr: 0.19	PDDr: 0.95	ESRr: 0.10	
Density map	BNS		DBS		OBDBS		MVZS		
									
# samples: ~1206	PDDr: 0.95	ESRr: 0.15	PDDr: 0.95	ESRr: 0.19	PDDr: 0.89	ESRr: 0.19	PDDr: 0.97	ESRr: 0.18	
terrorism									
Input	PBS		NUS		RS		KBS		
									
# points: 177133	PDDr: 0.90	ESRr: 0.18	PDDr: 0.62	ESRr: 0.05	PDDr: 0.89	ESRr: 0.26	PDDr: 0.91	ESRr: 0.20	
Density map	BNS		DBS		OBDBS		MVZS		
									
# samples: ~465	PDDr: 0.89	ESRr: 0.18	PDDr: 0.89	ESRr: 0.25	PDDr: 0.87	ESRr: 0.27	PDDr: 0.90	ESRr: 0.24	

TLC trip									
Input	PBS		NUS		RS		KBS		
									
# points: 1999697	PDDr: 0.98	ESRr: 0.32	PDDr: 0.61	ESRr: 0.10	PDDr: 0.98	ESRr: 0.49	PDDr: 0.98	ESRr: 0.47	
Density map	BNS		DBS		OBDBS		MVZS		
									
# samples: ~628	PDDr: 0.98	ESRr: 0.49	PDDr: 0.95	ESRr: 0.50	PDDr: 0.97	ESRr: 0.49	PDDr: 0.98	ESRr: 0.49	
uk_road_safety									
Input	PBS		NUS		RS		KBS		
									
# points: 2047256	PDDr: 0.94	ESRr: 0.02	PDDr: 0.88	ESRr: 0.00	PDDr: 0.96	ESRr: 0.04	PDDr: 0.90	ESRr: 0.04	
Density map	BNS		DBS		OBDBS		MVZS		
									
# samples: ~1012	PDDr: 0.95	ESRr: 0.05	PDDr: 0.95	ESRr: 0.04	PDDr: 0.95	ESRr: 0.04	PDDr: 0.95	ESRr: 0.05	
US_census									
Input	PBS		NUS		RS		KBS		
									
# points: 2000000	PDDr: 0.84	ESRr: 0.25	PDDr: 0.74	ESRr: 0.13	PDDr: 0.87	ESRr: 0.34	PDDr: 0.91	ESRr: 0.28	
Density map	BNS		DBS		OBDBS		MVZS		
									
# samples: ~1200	PDDr: 0.87	ESRr: 0.35	PDDr: 0.88	ESRr: 0.34	PDDr: 0.84	ESRr: 0.39	PDDr: 0.89	ESRr: 0.33	

vehicles									
Input	PBS		NUS		RS		KBS		
									
# points: 383430	PDDr: 0.90	ESRr: 0.17	PDDr: 0.79	ESRr: 0.05	PDDr: 0.89	ESRr: 0.29	PDDr: 0.91	ESRr: 0.22	
Density map	BNS		DBS		OBDBS		MVZS		
									
# samples: ~4027	PDDr: 0.90	ESRr: 0.24	PDDr: 0.89	ESRr: 0.27	PDDr: 0.82	ESRr: 0.26	PDDr: 0.92	ESRr: 0.27	

REFERENCES

[1] E. Bertini and G. Santucci. By Chance is Not Enough: Preserving Relative Density through Nonuniform Sampling. In *Proceedings of the International Conference on Information Visualisation*, pp. 622–629, 2004. doi: 10.1109/IV.2004.1320207

[2] X. Chen, T. Ge, J. Zhang, B. Chen, C.-W. Fu, O. Deussen, and Y. Wang. A Recursive Subdivision Technique for Sampling Multi-class Scatterplots. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):729–738, Jan 2020. doi: 10.1109/TVCG.2019.2934541

[3] R. L. Cook. Stochastic Sampling in Computer Graphics. *ACM Transactions on Graphics*, 5(1):51–72, 1986. doi: 10.1145/7529.8927

[4] D. Dua and C. Graff. UCI Machine Learning Repository. `https://archive.ics.uci.edu/ml`, 2017.

[5] R. Hu, T. Sha, O. van Kaick, O. Deussen, and H. Huang. Data Sampling in Multi-view and Multi-class Scatterplots via Set Cover Optimization. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):739–748, 2019. doi: 10.1109/TVCG.2019.2934799

[6] Kaggle Inc. Kaggle. `https://www.kaggle.com/`.

[7] C. R. Palmer and C. Faloutsos. Density Biased Sampling: An Improved Method for Data Mining and Clustering. In *Proceedings of the ACM SIGMOD International Conference on Management of Data*, pp. 82–92, 2000. doi: 10.1145/335191.335384

[8] S. Xiang, X. Ye, J. Xia, J. Wu, Y. Chen, and S. Liu. Interactive Correction of Mislabeled Training Data. In *Proceedings of the IEEE Conference on Visual Analytics Science and Technology*, pp. 57–68, 2019. doi: 10.1109/VAST47406.2019.8986943