Hyper-parameter selection

Table 1: Hyper-parameters for ParDFROCC.

Dataset	m	ε	β	Sparsity (s)	Kernel	CPU cores
A1	10	0.1	2	1	Linear	8
A2	10	0.1	2	1	Polynomial	8
A3	100	0.1	2	1	Polynomial	8
A4	50	0.1	2	1	Linear	8
B1	500	0.1	2	1	Linear	8
B2	500	0.1	2	1	Polynomial	8
B3	1000	0.01	2	0.1	Linear	8
B4	5000	0.1	2	0.1	Linear	8
B5	5000	0.1	2	0.1	Linear	8
B6	5000	0.1	2	0.1	Linear	8
B7	10000	0.01	4	0.01	Linear	8
C1	10000	0.1	2	0.01	Linear	8
C2	10000	0.1	4	0.01	Linear	8

The hyper-parameters for FROCC were chosen by a grid search on validation set. Table 1 tabulates the optimal hyper-parameters used for ParDFROCC in our experiments. As it can be observed, for low dimensions sparsity offers no benefit as expected, and plays an increasingly important role as the dimensions of the data increase.

Table 2: Hyperparameters for baselines.

Dataset	OC-SVM		IsoForest	HBOS		LODA		βVAE				OCGAN	
	γ	ν	Estimators	Bins	α	Bins	Random Cuts	Encoder	Decoder	Hidden Activation	O/P Activation	Generator LR	Discriminant LR
A1	1.25e-01	0.5	50	10	0.01	10	100	[32, 16]	[16, 32]	relu	sigmoid	0.01	0.1
A2	1.00e-01	0.5	50	100	0.01	100	100	[32, 16]	[16, 32]	relu	sigmoid	0.01	0.1
A3	2.00e-02	0.5	50	100	0.01	100	100	[32, 16]	[16, 32]	relu	sigmoid	0.01	0.1
A4	2.86e-02	0.5	50	10	0.01	10	100	[32, 16]	[16, 32]	relu	sigmoid	0.01	0.1
B1	1.00e-02	0.5	50	100	0.01	100	100	[32, 16]	[16, 32]	relu	sigmoid	0.01	0.1
B2	8.70e-03	0.5	50	1000	0.01	1000	100	[32, 16]	[16, 32]	selu	sigmoid	0.001	0.01
B3	1.28e-03	0.4	100	100	0.1	100	1000	[128, 64, 32]	[32, 64, 128]	leaky relu	tanh	0.001	0.01
B4	3.26e-04	0.4	100	100	0.1	100	500	[128, 64, 32]	[32, 64, 128]	leaky relu	tanh	0.001	0.01
B5	3.26e-04	0.4	100	100	0.1	100	500	[128, 64, 32]	[32, 64, 128]	leaky relu	tanh	0.001	0.01
B6	3.26e-04	0.4	100	100	0.1	100	500	[128, 64, 32]	[32, 64, 128]	leaky relu	tanh	0.001	0.01
B7	3.02e-05	0.4	200	100	0.1	100	1000	[256, 128, 64, 32]	[32, 64, 128, 256]	leaky relu	tanh	0.001	0.01
C1	3.13e-07	0.3	500	1000	0.1	1000	5000	[512, 256, 128, 64, 32]	[32, 64, 128, 256, 512]	elu	tanh	0.0001	0.001
C2	6.02e-08	0.3	500	1000	0.1	1000	5000	[512, 256, 128, 64, 32]	[32, 64, 128, 256, 512]	elu	tanh	0.0001	0.001

Table 2 lists the hyper-parameters for baseline methods. The model archi-

tecture and hyper-parameters were selected as provided by authors when available. Otherwise a grid-search was performed for optimal parameter selection. For DROCC, we used the settings suggested by the authors.