	5% CS						
Dataset	PCCC	PCCC-R	COPKM	CSC	DILS	LCC	KMEANS
Appendicitis	**0**	**0**	**0**	1	**0**	**0**	1
Breast Cancer	**0**	**0**	–	189	3	23	39
Bupa	**0**	**0**	–	70	1	6	58
Circles	**0**	**0**	–	49	**0**	6	31
Ecoli	**0**	**0**	**0**	70	**0**	**0**	21
Glass	**0**	**0**	**0**	36	**0**	**0**	11
Haberman	**0**	**0**	–	42	**0**	4	28
Hayesroth	**0**	**0**	**0**	14	**0**	1	6
Heart	**0**	**0**	–	49	**0**	1	18
Ionosphere	**0**	**0**	–	61	1	8	23
Iris	**0**	**0**	**0**	15	**0**	**0**	2
Led7Digit	**0**	**0**	**0**	210	**0**	–	18
Monk2	**0**	**0**	–	123	1	15	64
Moons	**0**	**0**	**0**	44	**0**	1	15
Movement Libras	**0**	**0**	**0**	100	**0**	**0**	6
Newthyroid	**0**	**0**	**0**	29	**0**	**0**	11
Saheart	**0**	**0**	–	122	6	25	55
Sonar	**0**	**0**	–	25	**0**	1	14
Soybean	**0**	**0**	**0**	1	**0**	**0**	0
Spectfheart	**0**	**0**	–	32	**0**	2	29
Spiral	**0**	**0**	–	51	**0**	5	29
Tae	**0**	**0**	**0**	15	**0**	**0**	7
Vehicle	**0**	**0**	–	603	7	87	204
Wine	**0**	**0**	**0**	25	**0**	**0**	2
Zoo	**0**	**0**	**0**	2	**0**	**0**	0
Mean	**0**	**0**	**0***	79	1	8*	28

*Nan values (–) are ignored when computing the sum.

Table W17: Average number of cannot-link constraint violations of the PCCC and the PCCC-R algorithms and the four state-of-the-art algorithms (COPKM, CSC, DILS, LCC) for the constraint sets of size 5% CS. The lowest values are stated in bold. The column KMEANS reports the average number of cannot-link constraint violations obtained with the k-means algorithm. The hyphen indicates that the respective algorithm returned no solution within the time limit of 1,800 seconds.