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

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

Table W13: 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.