



Master-Thesis

High Quality Hypergraph Partitioning via Max-Flow-Min-Cut Computations

Tobias Heuer

Advisors: Prof. Dr. Peter Sanders
M. Sc. Sebastian Schlag

Institute of Theoretical Informatics, Algorithmics
Department of Informatics
Karlsruhe Institute of Technology

Date of submission: 05.01.2018

Contents

1 Experimental Results	3
1.1 Flow Algorithms and Networks	3
1.2 Configuring direct k -way Flow-based Refinement	7
1.3 Speed-Up Heuristics	9
1.4 Comparison with other Hypergraph Partitioners	12

1 Experimental Results

1.1 Flow Algorithms and Networks

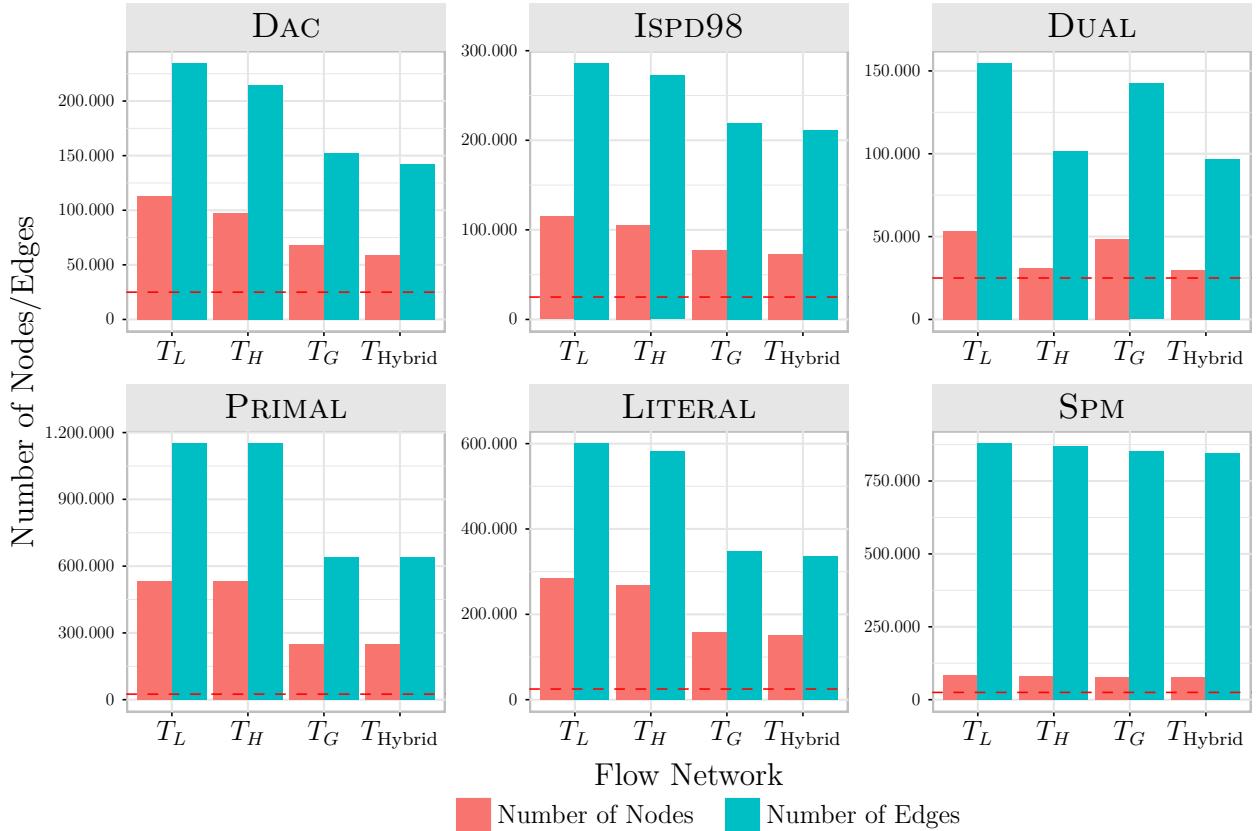


Figure 1: Comparison of the number of nodes and edges induced by flow problems of size $|V'| = 25000$ on our flow networks for different benchmark types. The red dashed lines indicates 25000 nodes.

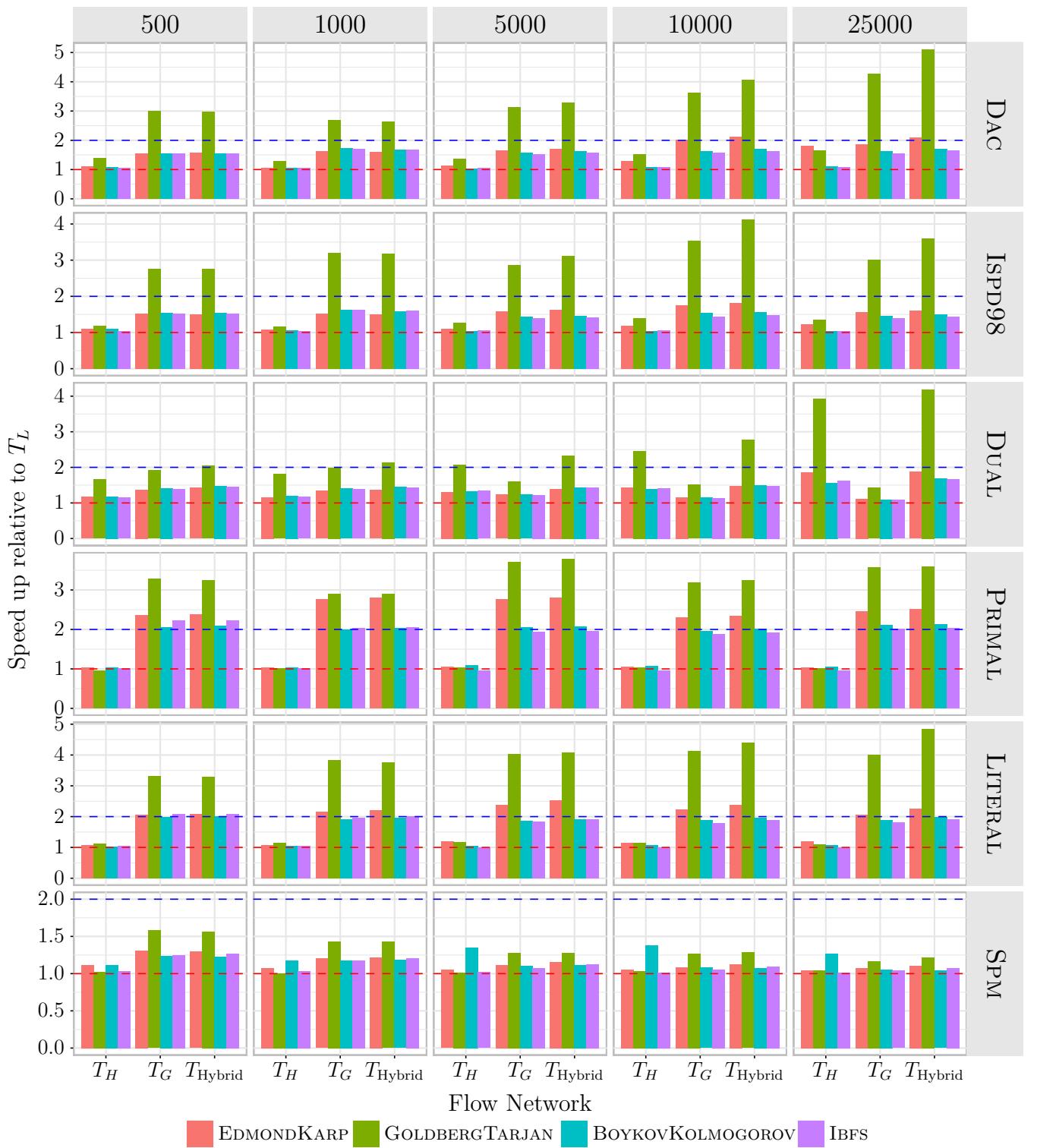


Figure 2: Speed-ups of the flow algorithms on different flow networks relative to execution on T_L for different problem sizes and types. The red resp. blue dashed line indicate a speed-up of 1 resp. 2.

$ V' $	IBFS $t[ms]$	BOYKOV $t[\%]$	GOLDBERG $t[\%]$	TARJAN $t[\%]$	EDMOND $t[\%]$	KARP $t[\%]$
500	0.81	1.79	24.56		-7.36	
1000	1.91	12.92	26.88		13.90	
5000	13.52	38.11	63.68		108.43	
10000	28.40	53.52	92.56		182.79	
25000	64.18	50.19	95.25		157.03	

Table 1: Average running times of our maximum flow algorithms on flow network T_{Hybrid} . Note, all values in the table are in percentage relative to the running time of the IBFS algorithm. In each line the fastest variant is marked bold.

Instance	$ V' $	IBFS $t[ms]$	BOYKOV-KOLMOGOROV $t[\%]$	GOLDBERG-TARJAN $t[\%]$	EDMOND-KARP $t[\%]$
ALL	500	0.81	1.79	24.56	-7.36
	1000	1.91	12.92	26.88	13.90
	5000	13.52	38.11	63.68	108.43
	10000	28.40	53.52	92.56	182.79
	25000	64.18	50.19	95.25	157.03
DAC	500	0.29	-10.46	11.79	-34.02
	1000	0.71	-9.51	18.18	-35.86
	5000	5.16	-5.94	12.21	-17.03
	10000	10.23	-1.30	33.46	32.26
	25000	26.05	2.14	112.29	67.77
ISPD98	500	0.42	-11.77	23.10	-13.55
	1000	0.98	-11.76	14.61	-12.01
	5000	7.33	-1.06	65.21	41.04
	10000	15.19	5.27	209.29	164.02
	25000	41.50	15.04	782.45	525.31
DUAL	500	0.24	-12.17	1.24	-42.11
	1000	0.51	-11.40	3.32	-43.34
	5000	3.41	-7.46	7.30	-43.50
	10000	6.65	-2.56	21.99	-37.53
	25000	15.63	3.86	65.81	-17.12
PRIMAL	500	1.45	24.48	58.97	143.23
	1000	3.51	22.79	66.63	200.62
	5000	26.34	23.07	122.09	589.46
	10000	48.87	14.37	110.18	484.42
	25000	102.32	17.38	237.90	852.30
LITERAL	500	0.66	0.55	61.36	38.80
	1000	1.63	8.06	51.86	84.71
	5000	11.55	13.41	94.09	227.94
	10000	25.28	14.81	130.87	375.89
	25000	55.67	7.27	125.34	367.70
SPM	500	1.60	3.50	8.07	-39.46
	1000	3.79	34.29	14.28	-14.95
	5000	26.30	121.86	65.34	105.19
	10000	60.57	196.95	103.28	256.18
	25000	140.35	169.78	14.70	62.44

Table 2: Average running times of our maximum flow algorithms on flow network T_{Hybrid} . Note, all values in the table are in percentage relative to the running time of the IBFS algorithm. In each line the fastest variant is marked bold.

1.2 Configuring direct k -way Flow-based Refinement

M2 - BOYKOV-KOLMOGOROV									
Config.	(+F,-M,-FM)		(+F,+M,-FM)		(+F,+M,+FM)		CONSTANT128		
α'	Avg[%]	$t[s]$	Avg[%]	$t[s]$	Avg[%]	$t[s]$	Avg[%]	$t[s]$	
1	-6.10	13.51	-5.62	14.22	0.23	15.19	0.53	55.75	
2	-3.20	16.89	-2.08	18.23	0.74	17.97	1.09	87.93	
4	-1.82	22.23	-0.20	24.29	1.21	22.50	1.61	144.42	
8	-0.85	31.49	0.98	34.43	1.71	30.58	2.16	257.41	
16	-0.20	48.66	1.75	53.23	2.21	45.04	2.69	498.29	
Ref.	(-F,-M,+FM)		6373.88	13.73					
M1 - GOLDBERG-TARJAN									
Config.	(+F,-M,-FM)		(+F,+M,-FM)		(+F,+M,+FM)		CONSTANT128		
α'	Avg[%]	$t[s]$	Avg[%]	$t[s]$	Avg[%]	$t[s]$	Avg[%]	$t[s]$	
1	-15.48	12.94	-15.26	13.29	0.14	14.99	0.32	67.38	
2	-10.50	16.07	-10.12	16.93	0.36	16.93	0.62	139.21	
4	-5.98	21.22	-5.08	23.01	0.67	20.76	1.03	274.60	
8	-3.22	30.73	-1.64	33.72	1.25	28.65	1.67	558.81	
16	-1.52	50.89	0.51	56.39	1.87	46.17	2.44	1220.92	
Ref.	(-F,-M,+FM)		6373.88	13.73					

Table 3: Table contains results for different configurations of our flow-based refinement framework for increasing α' . The quality in column *Avg.* is relative to our baseline configuration (-F,-M,+FM).

1 EXPERIMENTAL RESULTS

Config.	(+F,-M,-FM)		(+F,+M,-FM)		(+F,+M,+FM)	
α'	M1 - Avg [%]	M2 - Avg [%]	M1 - Avg [%]	M2 - Avg [%]	M1 - Avg [%]	M2 - Avg [%]
1	-15.48	-6.10	-15.26	-5.62	0.14	0.23
2	-10.50	-3.20	-10.12	-2.08	0.36	0.74
4	-5.98	-1.82	-5.08	-0.20	0.67	1.21
8	-3.22	-0.85	-1.64	0.98	1.25	1.71
16	-1.52	-0.20	0.51	1.75	1.87	2.21
Ref.	(-F,-M,+FM)		6373.88			

Table 4: Comparison on quality of our framework with different source and sink set modeling approaches. M1 represents the approach of Sanders and Schulz and M2 is our new variant.

M2 - BOYKOV-KOLMOGOROV			
Config.	(+F,-M,-FM)	(+F,+M,-FM)	(+F,+M,+FM)
α'	Avg[%]	Avg[%]	Avg[%]
1	-6.08	-5.57	0.23
2	-3.22	-2.10	0.72
4	-1.90	-0.26	1.18
8	-0.91	0.92	1.67
16	-0.29	1.65	2.15
Ref.	(-F,-M,+FM)	6376.03	
M1 - GOLDBERG-TARJAN			
Config.	(+F,-M,-FM)	(+F,+M,-FM)	(+F,+M,+FM)
α'	Avg[%]	Avg[%]	Avg[%]
1	-15.45	-15.19	0.17
2	-10.51	-10.08	0.38
4	-5.97	-5.10	0.68
8	-3.22	-1.66	1.26
16	-1.67	0.46	1.84
Ref.	(-F,-M,+FM)	6376.03	

Table 5: Table contains results of the effectiveness test for different configurations of our flow-based refinement framework for increasing α' . The quality in column *Avg.* is relative to our baseline configuration (-F,-M,+FM).

1.3 Speed-Up Heuristics

Variant	M2 - BOYKOV-KOLMOGOROV			
	Avg [%]	Min [%]	$t_{\text{flow}}[s]$	$t[s]$
KaHyPar-CA	7077.20	6820.17	-	29.26
KaHyPar-MF	-2.48	-2.13	51.76	81.02
KaHyPar-MF _(R1)	-2.41	-2.05	41.21	70.47
KaHyPar-MF _(R1,R2)	-2.40	-2.04	35.56	64.82
KaHyPar-MF _(R1,R2,R3)	-2.41	-2.05	26.64	55.90
Variant	M1 - GOLDBERG-TARJAN			
	Avg [%]	Min [%]	$t_{\text{flow}}[s]$	$t[s]$
KaHyPar-CA	7077.20	6820.17	-	29.26
KaHyPar-MF	-2.13	-1.80	52.28	81.54
KaHyPar-MF _(R1)	-2.05	-1.74	41.48	70.74
KaHyPar-MF _(R1,R2)	-2.05	-1.73	35.27	64.54
KaHyPar-MF _(R1,R2,R3)	-2.04	-1.75	27.62	56.88

Table 6: Results of our flow-based refinement framework with different speedup heuristics.

Partitioner	M2 - BOYKOV-KOLMOGOROV						
	Running Time $t[s]$						
	ALL	DAC	ISPD98	PRIMAL	LITERAL	DUAL	SPM
KaHyPar-CA	29.26	343.40	21.57	36.44	56.49	58.75	11.31
KaHyPar-MF	81.02	610.48	69.84	107.72	164.68	127.09	33.81
KaHyPar-MF _(R1)	70.47	526.26	55.56	91.10	136.17	113.87	30.62
KaHyPar-MF _(R1,R2)	64.82	503.04	47.31	84.65	123.63	101.41	28.95
KaHyPar-MF _(R1,R2,R3)	55.90	452.27	39.03	71.64	105.20	89.07	25.24
Partitioner	M1 - GOLDBERG-TARJAN						
	Running Time $t[s]$						
	ALL	DAC	ISPD98	PRIMAL	LITERAL	DUAL	SPM
KaHyPar-CA	29.26	343.40	21.57	36.44	56.49	58.75	11.31
KaHyPar-MF	81.54	699.18	75.97	114.67	185.56	143.93	28.74
KaHyPar-MF _(R1)	70.74	600.87	59.69	94.90	150.56	128.67	26.47
KaHyPar-MF _(R1,R2)	64.54	573.41	50.28	88.11	134.84	113.59	24.80
KaHyPar-MF _(R1,R2,R3)	56.88	526.86	43.32	74.76	116.79	101.76	22.31

Table 7: Comparing the average running time of KaHyPar-MF with KaHyPar-CA.

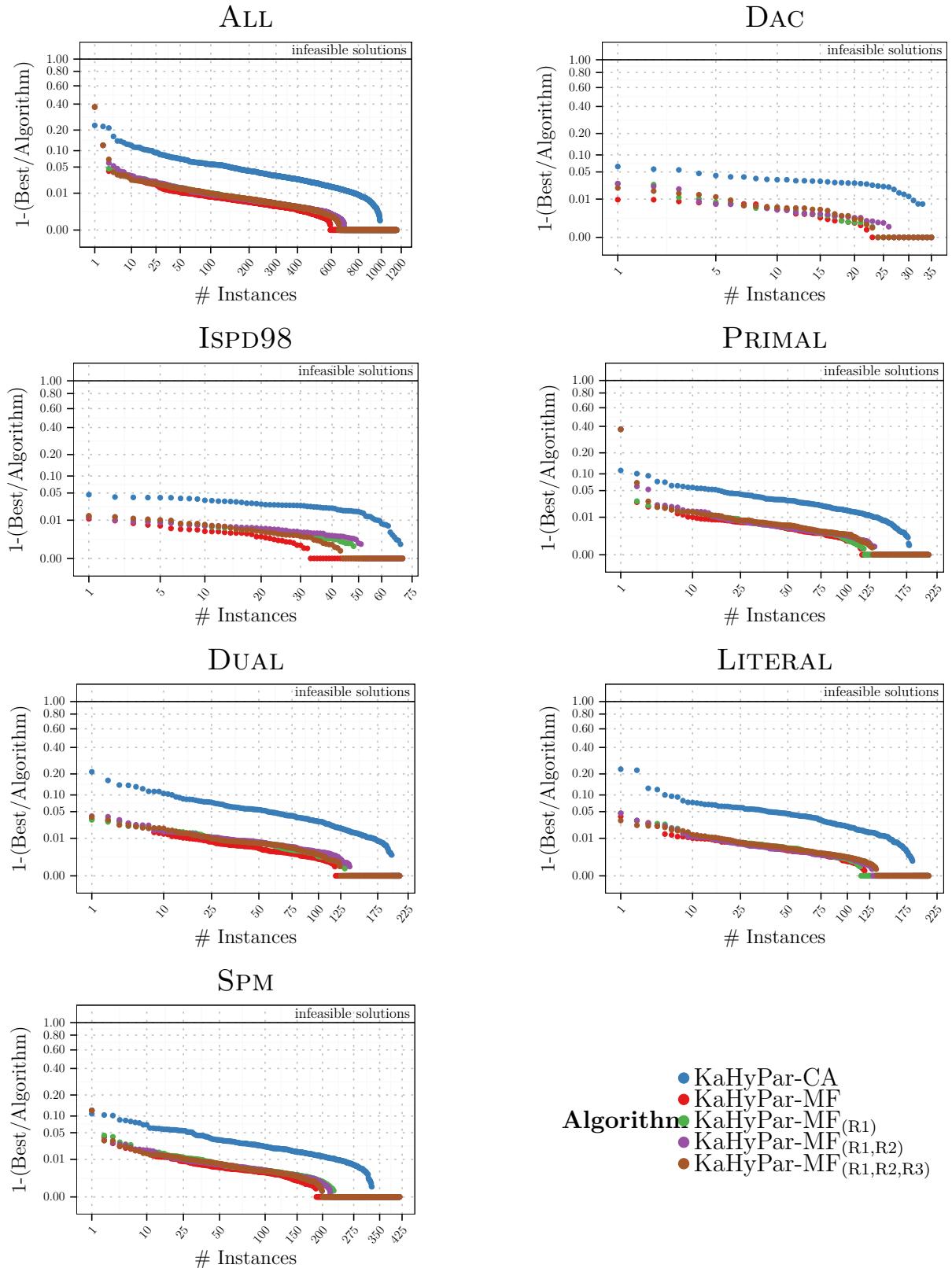


Figure 3: Min-Cut performance plots comparing KaHyPar-MF with KaHyPar-CA. M2 - BOYKOV-KOLMOGOROV

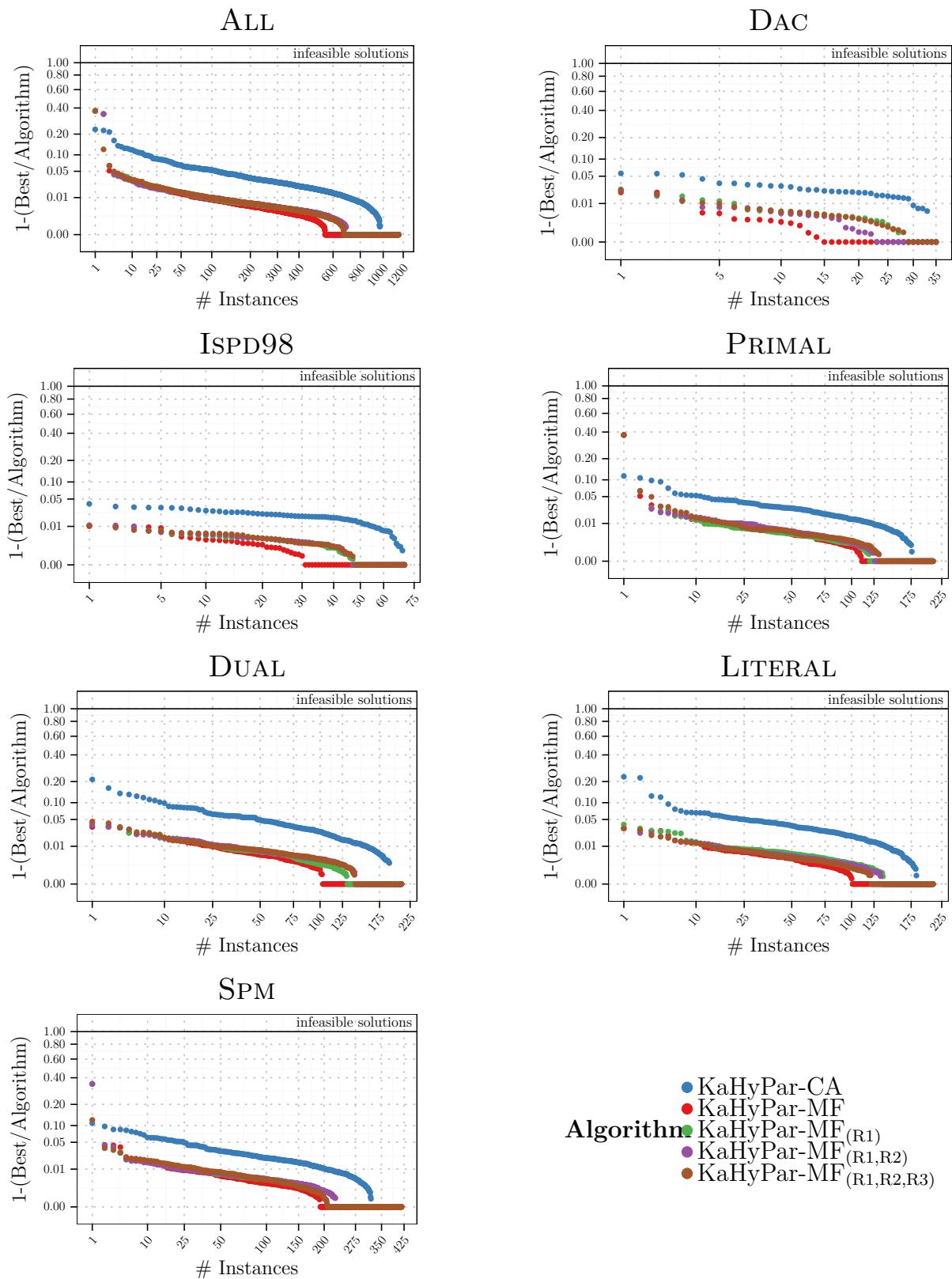


Figure 4: Min-Cut performance plots comparing KaHyPar-MF with KaHyPar-CA. M1 - GOLDBERG-TARJAN

1.4 Comparison with other Hypergraph Partitioners

Partitioner	M2 - BOYKOV-KOLMOGOROV						
	Running Time $t[s]$						
	ALL	DAC	ISPD98	PRIMAL	LITERAL	DUAL	SPM
KaHyPar-MF	63.33	505.75	21.10	71.81	140.53	98.06	32.64
KaHyPar-CA	30.85	368.97	12.35	32.91	64.65	68.27	13.63
hMetis-R	78.47	446.36	29.03	66.25	142.12	200.36	40.64
hMetis-K	57.36	240.92	23.18	44.23	94.89	125.55	35.08
PaToH-Q	5.84	28.34	1.89	6.90	9.24	10.57	3.35
PaToH-D	1.21	6.45	0.35	1.12	1.58	2.87	0.75
Partitioner	M1 - GOLDBERG-TARJAN						
	Running Time $t[s]$						
	ALL	DAC	ISPD98	PRIMAL	LITERAL	DUAL	SPM
KaHyPar-MF	62.24	637.58	22.29	71.63	140.84	106.24	29.61
KaHyPar-CA	31.05	368.97	12.35	32.91	64.65	68.27	13.91
hMetis-R	79.23	446.36	29.03	66.25	142.12	200.36	41.79
hMetis-K	57.86	240.92	23.18	44.23	94.89	125.55	35.95
PaToH-Q	5.89	28.34	1.89	6.90	9.24	10.57	3.42
PaToH-D	1.22	6.45	0.35	1.12	1.58	2.87	0.77

Table 8: Comparing the average running time of KaHyPar-MF with KaHyPar-CA and other hypergraph partitioners.

Partitioner	M2 - BOYKOV-KOLMOGOROV						
	Average $\lambda - 1$						
	ALL	DAC	ISPD98	PRIMAL	LITERAL	DUAL	SPM
KaHyPar-MF	7727.97	17480.70	5644.61	15863.61	15769.49	3038.31	6027.03
KaHyPar-CA	2.49	3.11	2.20	2.08	2.74	3.51	2.08
hMetis-R	15.79	3.64	1.62	2.07	2.80	43.44	19.77
hMetis-K	15.25	8.46	1.38	4.05	9.29	28.65	19.54
PaToH-Q	9.51	13.57	7.92	12.14	13.35	8.80	6.79
PaToH-D	16.77	23.75	15.08	18.27	21.54	18.31	12.92
Partitioner	M1 - GOLDBERG-TARJAN						
	Average $\lambda - 1$						
	ALL	DAC	ISPD98	PRIMAL	LITERAL	DUAL	SPM
KaHyPar-MF	7819.11	17590.10	5671.37	15923.74	15844.61	3061.94	6165.74
KaHyPar-CA	2.03	2.47	1.72	1.69	2.25	2.71	1.75
hMetis-R	15.21	2.99	1.14	1.69	2.31	42.33	19.22
hMetis-K	14.71	7.78	0.90	3.66	8.77	27.66	19.09
PaToH-Q	8.98	12.86	7.41	11.72	12.81	7.96	6.37
PaToH-D	16.21	22.98	14.54	17.83	20.97	17.40	12.50

Table 9: Comparison of average $(\lambda - 1)$ metric of KaHyPar-MF with KaHyPar-CA and other partitioners on different benchmark types. The results are in percentage relative to KaHyPar-MF.

Partitioner	M2 - BOYKOV-KOLMOGOROV						
	Running Time $t[s]$						
	$k = 2$	$k = 4$	$k = 8$	$k = 16$	$k = 32$	$k = 64$	$k = 128$
KaHyPar-MF	23.56	39.84	56.44	68.69	86.73	108.35	125.74
KaHyPar-CA	12.68	17.02	23.70	30.78	41.38	56.95	76.05
hMetis-R	27.87	51.03	73.94	90.09	107.94	127.26	147.72
hMetis-K	25.47	31.92	42.06	52.87	73.30	108.15	151.63
PaToH-Q	1.93	3.58	5.39	6.95	8.32	9.97	11.34
PaToH-D	0.43	0.76	1.11	1.40	1.69	2.00	2.27
Partitioner	M1 - GOLDBERG-TARJAN						
	Running Time $t[s]$						
	$k = 2$	$k = 4$	$k = 8$	$k = 16$	$k = 32$	$k = 64$	$k = 128$
KaHyPar-MF	22.13	38.51	55.04	67.83	85.75	108.97	128.04
KaHyPar-CA	12.68	17.16	23.88	31.01	41.69	57.35	76.61
hMetis-R	27.87	51.59	74.74	91.09	109.13	128.66	149.34
hMetis-K	25.47	32.27	42.50	53.41	74.00	109.12	152.92
PaToH-Q	1.93	3.61	5.44	7.01	8.40	10.06	11.44
PaToH-D	0.43	0.77	1.12	1.42	1.71	2.02	2.29

Table 10: Comparing the average running time of KaHyPar-MF with KaHyPar-CA and other partitioners for different values of k .

Partitioner	M2 - BOYKOV-KOLMOGOROV						
	Average $\lambda - 1$						
	$k = 2$	$k = 4$	$k = 8$	$k = 16$	$k = 32$	$k = 64$	$k = 128$
KaHyPar-MF	1057.94	3105.80	5988.27	9292.89	14582.34	21735.78	31477.02
KaHyPar-CA	2.32	2.62	2.86	2.73	2.55	2.28	2.08
hMetis-R	27.19	18.98	16.97	15.81	12.86	10.70	8.49
hMetis-K	27.59	17.84	15.75	15.58	11.78	10.25	8.48
PaToH-Q	11.74	9.14	9.14	10.01	9.38	9.05	8.10
PaToH-D	15.29	16.60	19.17	19.91	16.15	15.78	14.52

Partitioner	M1 - GOLDBERG-TARJAN						
	Average $\lambda - 1$						
	$k = 2$	$k = 4$	$k = 8$	$k = 16$	$k = 32$	$k = 64$	$k = 128$
KaHyPar-MF	1064.06	3147.96	6062.80	9406.00	14756.03	21978.89	31820.94
KaHyPar-CA	1.73	2.06	2.36	2.28	2.11	1.90	1.73
hMetis-R	26.46	18.26	16.34	15.25	12.33	10.23	8.08
hMetis-K	26.86	17.19	15.18	15.06	11.29	9.83	8.10
PaToH-Q	11.10	8.50	8.57	9.49	8.87	8.60	7.70
PaToH-D	14.62	15.94	18.55	19.34	15.62	15.31	14.09

Table 11: Comparison of average $(\lambda - 1)$ metric of KaHyPar-MF with KaHyPar-CA and other partitioners for different values of k . The results are in percentage relative to KaHyPar-MF.

1 EXPERIMENTAL RESULTS

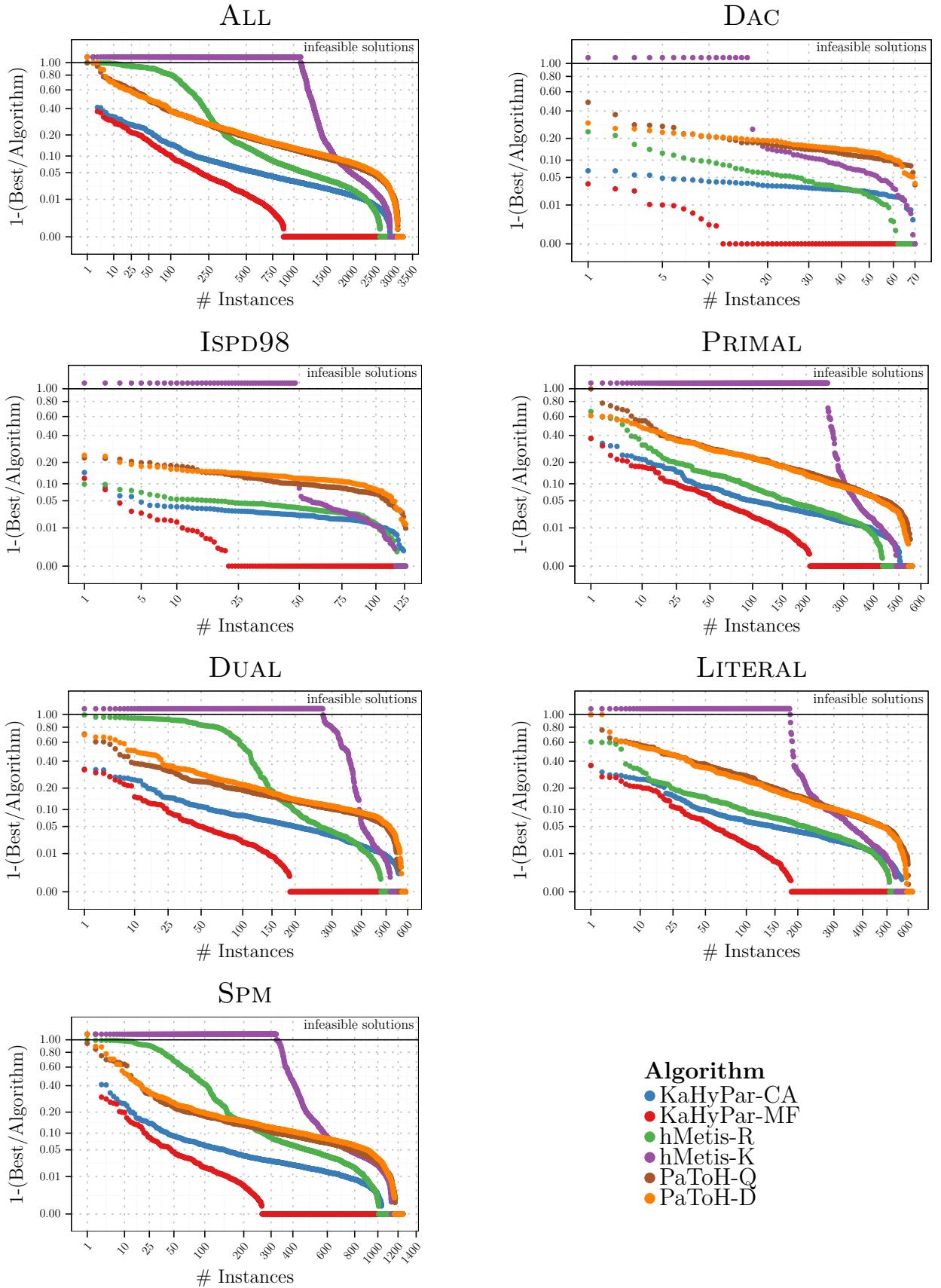


Figure 5: Min-Cut performance plots comparing KaHyPar-MF with KaHyPar-CA and other partitioners. M2 - BOYKOVKOLMOGOROV

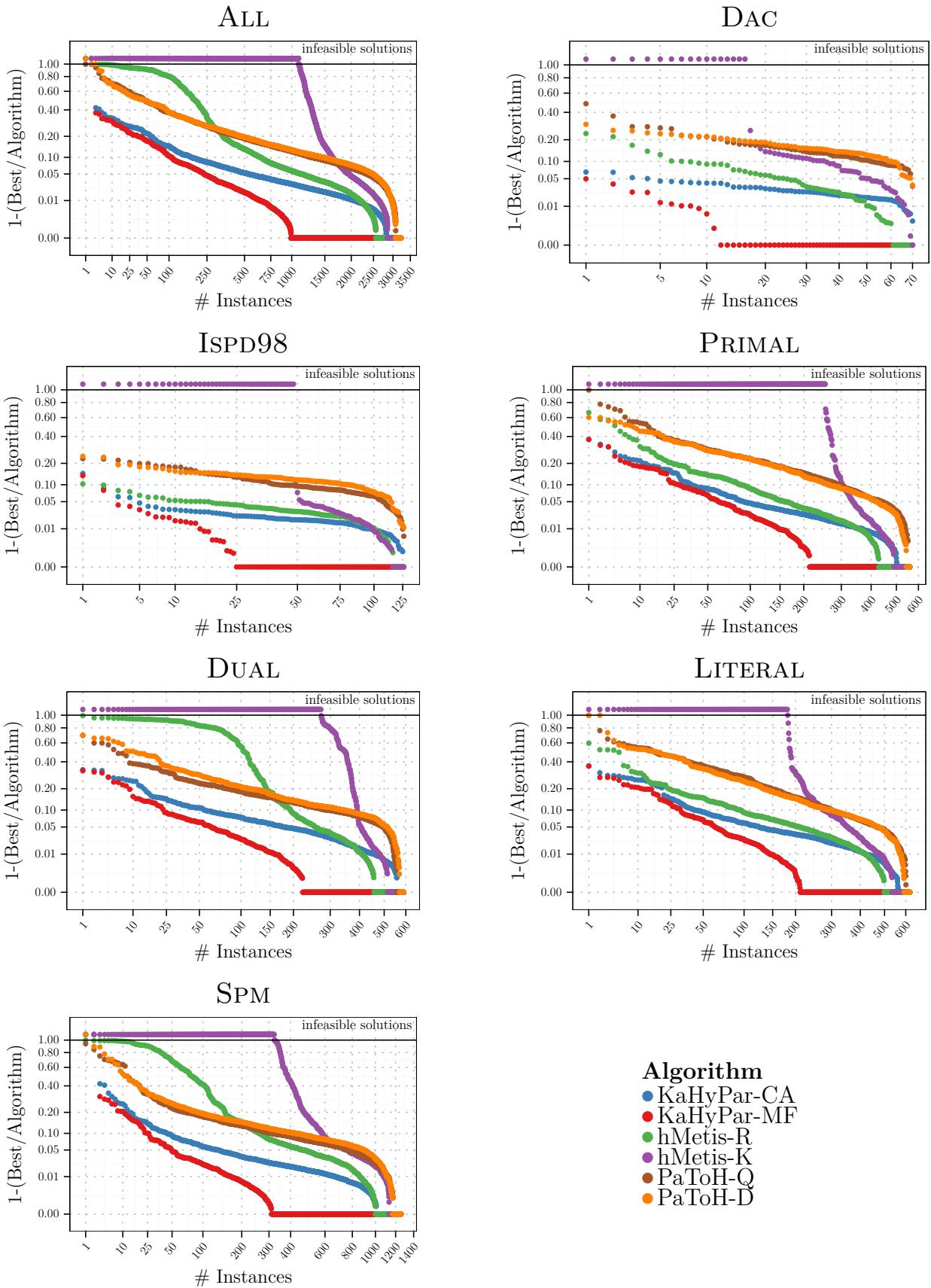


Figure 6: Min-Cut performance plots comparing KaHyPar-MF with KaHyPar-CA and other partitioners. M1 - GOLDBERG-TARJAN

1 EXPERIMENTAL RESULTS

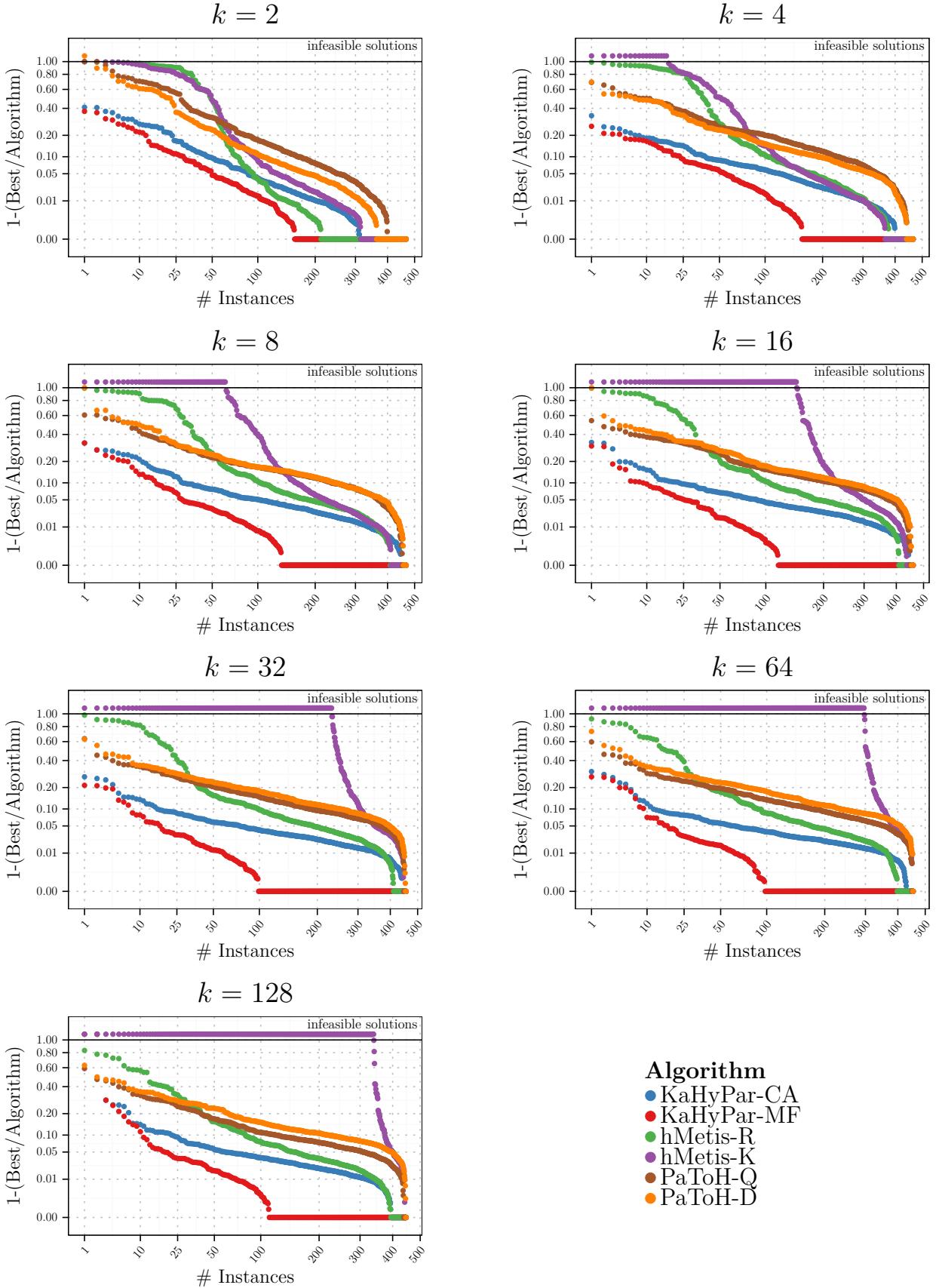


Figure 7: Min-Cut performance plots comparing KaHyPar-MF with KaHyPar-CA and other partitioners for different values of k . M2 - BOYKOVKOLMOGOROV

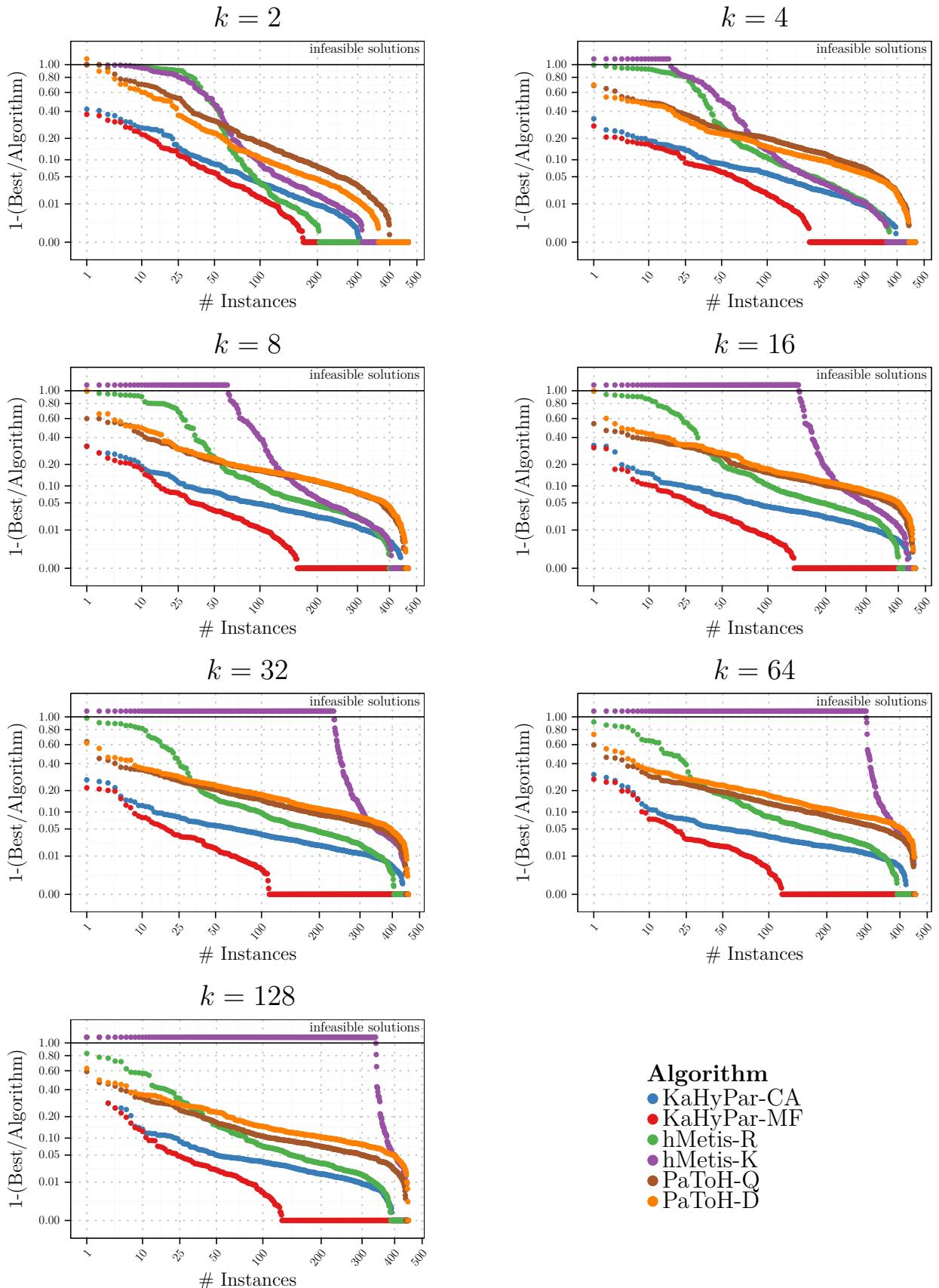


Figure 8: Min-Cut performance plots comparing KaHyPar-MF with KaHyPar-CA and other partitioners for different values of k . M1 - GOLDBERG-TARJAN