As suggested, we ran experiments with the same configurations and constraints on a collection of real-life event logs. Overall characteristics are shown in Table I.

TABLE I: Properties of the real-life log collection.

Property	Min.	Max.	Avg.	Med.
Nodes in DFG (|N|)	4	51	25	24
Edges in DFG (|E|)	10	772	236	125
Trace variants	1	22,632	4,344	846
Min. Trace length (|t|)	1	52	8	3
Max. Trace length (|t|)	15	990	171	90

A detailed overview of the characteristics of the individual logs is shown in Table II.

TABLE II: Properties of the real-life log collection.

Log	|E|	|N|	Trace var.	Max. |t|	Min. |t|
[1]	180	24	5,946	58	7
[2]	164	29	20	118	52
[3]	115	16	846	185	3
[4]	99	27	116	25	16
[5]	10	4	183	35	1
[6]	125	24	4,366	175	3
[7]	772	39	22,632	178	1
[8]	70	11	231	20	2
[9]	498	42	11,973	990	1
[10]	553	51	1,478	90	3
[11]	14	8	1	15	15

For the numerical attribute the `time:timestamp`-attribute was used to acquire durations for all logs. For the categorical attribute we used the `org:resource`-attribute was used. The healthcare process [3] is an exception, as it does not contain a `org:resource`-attribute, we used the `org:group`-attribute instead.

Table III presents the results obtained using the *Exh* configuration of GECCO for the same constraints as used for the synthetic data on the real-life event logs (cf. Table V in the paper for the synthetic data).

TABLE III: Results for *Exh*, averaged over solved cases.

Measure	*AC*	*AN*	*MC*	*MN*	*NN*	*Gr*	*Avg*
Solved	100%	20%	0%	0%	20%	100%	40%
Size red.	.67	.53	0	0	.53	.63	.63
Compl. red.	.57	.21	0	0	.21	.53	.49
Silhouette	.14	.23	0	0	.23	.10	.14
Runtime	53m	4s	-	-	4s	55m	46m

Table IV depicts the evaluation results for the different configurations. As was done for TABLE VI in the paper, we here average the results over all cases to avoid rewarding configurations that failed to solve an abstraction problem.

Table V shows the results averaged over all cases, as suggested by the reviewers for the synthetic data as well.

Note. The imposed constraints were initially chosen according to the characteristics of the synthetic data collection. Given the considerably different characteristics of the real-life logs, these same constraints yield many intractable abstraction problems (especially for the monotonic MC and MN constraints), resulting in a low number of solved tasks. For instance, certain logs have very few resource roles, which means it is often impossible to find groups of events performed by 2 or more different roles. While this intractability is reflected in the obtained results, the results still clearly indicate comparable trends as for the synthetic data collection. As an example we already ran additional experiments using a less restrictive monotonic constraint "[g.duration > 1 minute] in at least 80% of cases". This yielded 5 solved cases for *Exh*, 3 for DFG_∞, and 2 for DFG_{100}, as compared to no solved cases for *MN* from the paper for all configurations.

Also, it was not possible to acquire results for BL_{LPM} using the same configurations, due to timed outs for all cases.

TABLE IV: Results averaged over constraints using different configurations and the greedy baseline considering all problems when computing averages.

Measure	*Exh*	DFG_∞	DFG_{100}	BL_G
Solved	40%	40%	40%	30%
Size red.	.25	.25	.21	.13
Compl. red.	.20	.19	.16	.07
Silhouette	.06	.07	.04	.00
Runtime	43m	42m	20m	5m

TABLE V: Results averaged over constraints using different configurations and the greedy baseline considering only solved problems when computing averages.

Measure	*Exh*	DFG_∞	DFG_{100}	BL_G
Count	25	25	25	17
Size red.	.63	.63	.53	.45
Compl. red.	.49	.46	.40	.25
Silhouette	.14	.18	.10	-.02
Runtime	46m	44m	29m	11m

REFERENCES

[1] B. van Dongen, "Bpi challenge 2017," Feb 2017. [Online]. Available: https://data.4tu.nl/articles/dataset/BPI_Challenge_2017/12696884/1

[2] J. Munoz-Gama, R. R. de la Fuente, M. M. Sepúlveda, and R. R. Fuentes, "Conformance checking challenge 2019 (ccc19)," Feb 2019. [Online]. Available: https://data.4tu.nl/articles/dataset/Conformance_Checking_Challenge_2019_CCC19_/12714932/1

[3] F. Mannhardt, "Sepsis cases - event log," Dec 2016. [Online]. Available: https://data.4tu.nl/articles/dataset/Sepsis_Cases_-_Event_Log/12707639/1

[4] J. Buijs, "Environmental permit application process ('wabo'), coselog project," May 2014. [Online]. Available: https://data.4tu.nl/collections/Environmental_permit_application_process_WABO_CoSeLoG_project/5065529/1

[5] W. Steeman, "Bpi challenge 2013, closed problems," Apr 2013. [Online]. Available: https://data.4tu.nl/articles/dataset/BPI_Challenge_2013_closed_problems/12714476/1

[6] B. van Dongen, "Bpi challenge 2012," Apr 2012. [Online]. Available: https://data.4tu.nl/articles/dataset/BPI_Challenge_2012/12689204/1

[7] ——, "Bpi challenge 2014: Activity log for incidents," Apr 2014. [Online]. Available: https://data.4tu.nl/articles/dataset/BPI_Challenge_2014_Activity_log_for_incidents/12706424/1

[8] M. M. de Leoni and F. Mannhardt, "Road traffic fine management process," Feb 2015. [Online]. Available: https://data.4tu.nl/articles/dataset/Road_Traffic_Fine_Management_Process/12683249/1
[9] B. van Dongen, "Bpi challenge 2019," Jan 2019. [Online]. Available: https://data.4tu.nl/articles/dataset/BPI_Challenge_2019/12715853/1
[10] ——, "Bpi challenge 2020: Travel permit data," Mar 2020. [Online]. Available: https://data.4tu.nl/articles/dataset/BPI_Challenge_2020_Travel_Permit_Data/12718178/1
[11] A. Djedović, "Credit requirement event logs," Sep 2017. [Online]. Available: https://data.4tu.nl/articles/dataset/Credit_Requirement_Event_Logs/12693005/1