Supplementary materials for: Efficient Coalition Structure Generation via Approximately Equivalent Induced Subgraph Games

Filippo Bistaffa*, Georgios Chalkiadakis†, and Alessandro Farinelli‡
*IIIA-CSIC, 08193 Cerdanyola, Spain
†School of Electrical and Computer Engineering, Technical University of Crete, 73100 Chania, Greece
‡Department of Computer Science, University of Verona, 37134 Verona, Italy
Emails: *filippo.bistaffa@iiia.csic.es, †gehalk@intelligence.tuc.gr, ‡alessandro.farinelli@univr.it

I. Theoretical Proofs

Here we provide the proofs not included in our manuscript due to space constraints. All proposition and equation numbers correspond to the ones in the manuscript, except for the ones introduced here, denoted as (1A), (2A), etc.

Proposition 2. *The worst-case complexity of solving the AE-ISG model when using ℓ_1 (resp. ℓ_2) is $\mathcal{O}\Big(\big(e + |\mathcal{FC}(\mathcal{G})|\big)^2 \cdot |\mathcal{FC}(\mathcal{G})|\Big)$ (resp. $\mathcal{O}\big(e^2 \cdot |\mathcal{FC}(\mathcal{G})|\big)$).*

Proof. Solving a LP problem requires $\mathcal{O}\big(x^2 \cdot c\big)$ operations, where x is the number of variables and c is the number of constraints [1]. With ℓ_1, the LP problem in

$$\begin{aligned} \text{minimise} \quad & \sum_{S \in \mathcal{FC}(\mathcal{G})} t_S, \\ \text{subject to} \quad & \left.\begin{aligned} \sum_{\substack{(i,j)\in E \\ \{i,j\}\subseteq S}} \{w_{i,j}\} - t_S \leq v(S), \\ \sum_{\substack{(i,j)\in E \\ \{i,j\}\subseteq S}} \{w_{i,j}\} + t_S \geq v(S), \end{aligned}\right\} \quad \forall S \in \mathcal{FC}(\mathcal{G}). \end{aligned} \tag{5}$$

has $e + |\mathcal{FC}(\mathcal{G})|$ variables and $2 \cdot |\mathcal{FC}(\mathcal{G})|$ constraints. Henceforth, solving such LP problem requires $\mathcal{O}\Big(\big(e + |\mathcal{FC}(\mathcal{G})|\big)^2 \cdot |\mathcal{FC}(\mathcal{G})|\Big)$. Solving a least-squares approximation problem requires $\mathcal{O}\big(b^2 \cdot a\big)$ operations, where a is the number of rows of A and b is the number of columns [1]. Hence, in the case of ℓ_2, solving the problem in

$$\text{minimise} \sum_{S \in \mathcal{FC}(\mathcal{G})} \left(\sum_{\substack{(i,j)\in E \\ \{i,j\}\subseteq S}} \{w_{i,j}\} - v(S) \right)^2. \tag{6}$$

requires $\mathcal{O}\big(e^2 \cdot |\mathcal{FC}(\mathcal{G})|\big)$ operations. □

Lemma 1. *Given $G = (\mathcal{G}, v)$, its approximate ISG AE-ISG(G), and the corresponding residual vector r, then*

$$V(CS^*_G) - V(CS^*_{AE}) \leq r(CS^*_{AE}) - r(CS^*_G), \quad \text{where } r(CS) = \sum_{S \in CS} r_S. \tag{15}$$

Proof. Given any CS, $r(CS) = W(CS) - V(CS)$ (Equation 1). Hence,

$$V(CS^*_G) - V(CS^*_{AE}) = W(CS^*_G) - r(CS^*_G) - V(CS^*_{AE}). \tag{1A}$$

Since CS^*_{AE} is the optimal CSG solution for AE-ISG(G), and we consider CSG as a maximisation problem, $W(CS^*_G) \leq W(CS^*_{AE})$. Thus, from (1A) we can infer

$$V(CS^*_G) - V(CS^*_{AE}) \leq W(CS^*_{AE}) - r(CS^*_G) - V(CS^*_{AE}). \tag{2A}$$

Since $W(CS^*_{AE}) - V(CS^*_{AE}) = r(CS^*_{AE})$, (2A) is equivalent to (15). □

Proposition 5. *Given a CFG G, its approximate ISG AE-ISG(G), and a coalition structure CS solution to the CSG problem of AE-ISG(G) with an optimality gap O_{CS}, then*

$$V(CS^*_G) - V(CS) \leq O_{CS} + r(CS) - min_r \tag{18}$$

Proof. Given any CS, $W(CS) = V(CS) + r(CS)$ (Equation 1). We rewrite

$$W(CS^*_{AE}) - W(CS) \leq O_{CS}, \tag{13}$$

as

$$V(CS^*_{AE}) + r(CS^*_{AE}) - V(CS) - r(CS) \leq O_{CS}. \tag{3A}$$

Since $V(CS^*_{AE}) \geq V(CS^*_G) - r(CS^*_{AE}) + min_r$ (Proposition 4), from (3A) we get

$$V(CS^*_G) - \cancel{r(CS^*_{AE})} + min_r + \cancel{r(CS^*_{AE})} - V(CS) - r(CS) \leq O_{CS}, \tag{4A}$$

which is equivalent to (18). □

II. KGC *vs* Heuristic GC Approach

We now compare one of the most notable GC approaches, i.e., one based on the concept of *dominant set* (DS) [2], with KGC. We compare the two approaches in terms of runtime and solution quality, following the same methodology discussed in the manuscript. As expected, results in Table I show that the heuristic approach is much faster in terms of runtime since it does not seek optimality. Of course, its performance is sub-optimal: the heuristic approach achieves a solution quality of 84% *wrt* the optimal solution computed by KGC. Moreover, we remark that by replacing KGC with a heuristic approach one loses all the theoretical quality guarantees we provide in Section V-A of our manuscript, which rely on the optimality of the solution of the GC problem.

n	KGC runtime (s)	DS runtime (s)	DS quality (%)
50	9.96	0.00	0.84
100	517.08	0.02	0.85
200	7330.00	0.04	0.84

TABLE I
RESULTS OF THE COMPARISON.

REFERENCES

[1] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2004.
[2] M. Pavan and M. Pelillo, "Dominant sets and pairwise clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 1, pp. 167–172, 2006.