Combinatorial persistency criteria for multicut and max-cut
Supplementary material

Jan-Hendrik Lange
Max Planck Institute for Informatics
Saarland University

Bjoern Andres
Max Planck Institute for Informatics
Bosch Center for AI
University of Tübingen

Paul Swoboda
Max Planck Institute for Informatics

A. Proofs

Proof of Lemma 1.

Proof. Let x be an optimal solution of (P). Then $x^* = p(x)$ is also optimal and $x_i^* = \beta$. □

Proof of Lemma 2.

Proof. (i) Let $x \in \mathsf{MC}$ and assume that $z = p_{\delta(U)}(x) \notin \mathsf{MC}$. Then there exists a cycle C with exactly one cut edge in z, i.e. $z_f = 1$ for some $f \in E_C$ and $z_e = 0$ for all $e \in E_C \setminus \{f\}$. It holds that $x_f = 1$ and thus C crosses $\delta(U)$ exactly once, which is impossible.

(ii) Let $x \in \mathsf{MC}$ and assume that $z = p_U(x) \notin \mathsf{MC}$. Then there is a cycle C with $z_f = 1$ for some $f \in E_C$ and $z_e = 0$ for all $e \in E_C \setminus \{f\}$. Since $z \leq x$ there exists an edge $uv = g \in E_C$, $g \neq f$ with $x_g = 1$ and $z_g = 0$. Then, according to (1), there exists a uv-path P such that $x_e = 0$ for all $e \in E_P \setminus E(U)$. Replace the cycle C with the cycle induced by $E_C \triangle (E_P \cup \{g\})$. Repeating this argument for all such edges $g \in E_C$ yields a path P' connecting the endpoints of f such that $x_e = 0$ for all $e \in E_{P'} \setminus E(U)$, which is a contradiction to $z_f = 1$. □

Proof of Theorem 1.

Proof. First, we show that the mapping

$$p(x) = \begin{cases} p^{\triangle}_{\delta(U)}(x) & \text{if } x_f \neq \beta \\ x & \text{else} \end{cases}$$

is improving for the MAX-CUT problem. Let $x \in \mathsf{CUT}$, $z = p(x)$ and suppose $x_f \neq \beta$. It holds that

$$\begin{aligned} \langle \theta, z \rangle - \langle \theta, x \rangle &= \theta_f (z_f - x_f) + \sum_{e \in \delta(U) \setminus \{f\}} \theta_e (z_e - x_e) \\ &= -|\theta_f| + \sum_{e \in \delta(U) \setminus \{f\}} \theta_e (z_e - x_e) \\ &\leq -|\theta_f| + \sum_{e \in \delta(U) \setminus \{f\}} |\theta_e| \\ &\leq 0. \end{aligned}$$

Similarly, for $\beta = 0$, we show that the mapping

$$p(x) = \begin{cases} (p_f \circ p_{\delta(U)})(x) & \text{if } x_f \neq \beta \\ x & \text{else} \end{cases}$$

is improving for the MULTICUT problem. Let $x \in \mathsf{MC}$, $z = p(x)$ and suppose $x_f \neq \beta$. It holds that

$$\begin{aligned} \langle \theta, z \rangle - \langle \theta, x \rangle &= \theta_f (0 - 1) + \sum_{e \in \delta(U) \setminus \{f\}} \theta_e (z_e - x_e) \\ &\leq -|\theta_f| + \sum_{e \in \delta(U) \setminus \{f\}} |\theta_e| \\ &\leq 0. \end{aligned}$$

Finally, for $\beta = 1$, we show that the mapping

$$p(x) = \begin{cases} p_{\delta(U)}(x) & \text{if } x_f \neq \beta \\ x & \text{else} \end{cases}$$

is improving for multicut. Let $x \in \mathsf{MC}$, $z = p(x)$ and suppose $x_f \neq \beta$. It holds that

$$\begin{aligned} \langle \theta, z \rangle - \langle \theta, x \rangle &= \theta_f (1 - 0) + \sum_{e \in \delta(U) \setminus \{f\}} \theta_e (1 - x_e) \\ &\leq -|\theta_f| + \sum_{e \in \delta(U) \cap E^+} \theta_e (1 - x_e) \\ &\leq -|\theta_f| + \sum_{e \in \delta(U) \cap E^+} |\theta_e| \\ &\leq 0. \end{aligned}$$

This concludes the proof. $\square$

Proof of Lemma 3.

Proof. Condition (i) implies that the mapping $p\colon X \to X$ defined by

$$p(x) = \begin{cases} p^y(x) & \text{if } x_{|E_H} = y \\ x & \text{else} \end{cases}$$

is improving. Condition (ii) implies $p(x)_e = \beta$ for all x. $\square$

Proof of Corollary 1.

Proof. We use Lemma 3:

(i) In the case $x_{uw} = 1, x_{uv} = 1, x_{vw} = 0$ apply $p^{\triangle}_{\delta(U)}$. In the case $x_{uw} = 1, x_{uv} = 0, x_{vw} = 1$ apply $p^{\triangle}_{\delta(W)}$. These mappings are improving due to (5) and (6).

(ii) In the case $x_{uw} = 1, x_{uv} = 1, x_{vw} = 0$ apply $p_{\{u,w\}} \circ p_{\delta(U)}$. In the case $x_{uw} = 1, x_{uv} = 0, x_{vw} = 1$ apply $p_{\{u,w\}} \circ p_{\delta(W)}$. These mappings are improving analogously to (i). In the additional case $x_{uw} = 1, x_{uv} = 1, x_{vw} = 1$ apply the mapping $p = p_{\{u,v,w\}} \circ p_{\delta(\{u,v,w\})}$. It is improving, since

$$\begin{aligned}
&\langle \theta, p(x) \rangle - \langle \theta, x \rangle \\
&= \sum_{e \in \delta(\{u,v,w\})} \theta_e(1 - x_e) - \theta_{uv} - \theta_{uw} - \theta_{vw} \\
&\leq \sum_{e \in \delta(\{u,v,w\}) \cap E^+} \theta_e(1 - x_e) - \sum_{e \in \delta(\{u,v,w\}) \cap E^+} \theta_e \\
&\leq 0.
\end{aligned}$$

$\square$

Proof of Theorem 2.

Proof. We use Lemma 3. Let $y \in \mathsf{MC}(H)$ with $y_{uv} = 1$ and suppose $x \in \mathsf{MC}$ with $x_{|E_H} = y$. Then there is a multicut M of H such that $y = \mathbb{1}_M$. Due to (8), every (multi-)cut of H has nonnegative weight. Therefore, there exists some $U \subset V_H$ with $u \in U$ and $v \notin U$ such that $\delta(U, V_H \setminus U) \subseteq M$ and

$$\sum_{e \in E_H} \theta_e x_e = \sum_{e \in M} \theta_e \geq \sum_{e \in \delta(U, V_H \setminus U)} \theta_e. \tag{16}$$

Let $p^y(x) = (p_{V_H} \circ p_{\delta(V_H)})(x)$, then it follows from (9) and (16) that

$$\begin{aligned}
&\langle \theta, p^y(x) \rangle - \langle \theta, x \rangle \\
&= \sum_{e \in \delta(V_H)} \theta_e(1 - x_e) - \sum_{e \in E_H} \theta_e x_e \\
&\leq \sum_{e \in \delta(V_H) \cap E^+} \theta_e(1 - x_e) - \sum_{e \in \delta(V_H) \cap E^+} \theta_e \\
&\leq 0.
\end{aligned}$$

$\square$

Proof of Theorem 3.

Proof. We use Lemma 3. Suppose $y \in \mathsf{CUT}(H)$ with $y_{uv} = 1$. Let $U \subset V_H$ be such that y is the incidence vector of $\delta(U, V_H \setminus U)$ in H and suppose $x \in \mathsf{CUT}$ with $x_{|E_H} = y$. We may assume that

$$\sum_{e \in \delta(U, V_H \setminus U)} \theta_e \geq \sum_{e \in \delta(U, V \setminus V_H)} |\theta_e|,$$

otherwise redefine $U := V_H \setminus U$. Now, let $z = p^y(x) = p^{\triangle}_{\delta(U)}(x)$, then it follows that

$$\begin{aligned}
&\langle \theta, z \rangle - \langle \theta, x \rangle = \sum_{e \in \delta(U)} \theta_e(z_e - x_e) \\
&= \sum_{e \in \delta(U, V_H \setminus U)} \theta_e(0 - 1) + \sum_{e \in \delta(U, V \setminus V_H)} \theta_e(z_e - x_e) \\
&\leq \sum_{e \in \delta(U, V_H \setminus U)} -\theta_e + \sum_{e \in \delta(U, V \setminus V_H)} |\theta_e| \\
&\leq 0.
\end{aligned}$$

$\square$

Proof of Lemma 4.

Proof. As H satisfies Assumption 1, the dual problem (12) evaluates to zero. Thus, since any conflicted cycle contains exactly one edge $e \in E_H$ with $\theta_e < 0$, we must have $\sum_{C : e \in C} \lambda^*_C = |\theta_e|$, which implies $\tilde{\theta}_e = 0$. Furthermore, for any cut $\delta(U)$ of H it holds that

$$\begin{aligned}
&\sum_{e \in \delta(U)} \tilde{\theta}_e \\
&= \sum_{e \in \delta(U) \cap E^+} \tilde{\theta}_e + \sum_{e \in \delta(U) \cap E^-} \tilde{\theta}_e \\
&= \sum_{e \in \delta(U) \cap E^+} \Big(\theta_e - \sum_{C : e \in C} \lambda^*_C \Big) + \sum_{e \in \delta(U) \cap E^-} \Big(\theta_e + \sum_{C : e \in C} \lambda^*_C \Big) \\
&= \sum_{e \in \delta(U)} \theta_e + \sum_{\substack{e \in \delta(U) \cap E^- \\ C : e \in C}} \lambda^*_C - \sum_{\substack{e \in \delta(U) \cap E^+ \\ C : e \in C}} \lambda^*_C \\
&\leq \sum_{e \in \delta(U)} \theta_e.
\end{aligned}$$

The last inequality holds true, because every cycle with precisely one negative edge e, where $e \in \delta(U) \cap E^-$, also contains some positive edge $f \in \delta(U) \cap E^+$, as it crosses $\delta(U)$ at least twice. This concludes the proof. $\square$

B. Running times

In this section we provide additional information on running times for our experiments. We focus here on the time it takes to evaluate our criteria and on the speedup of optimization algorithms when applied to reduced instances instead of original instances.

However, note that the running times of our methods strongly depend on their practical implementation, both on the conceptual level (which persistency criteria are checked) and the technical level (how the algorithms are implemented). Both aspects are highly nontrivial and there is ample potential for further improvement in terms of speed.

In Table 4 we report the average running times of our methods for the experimental results from Table 2 and 3 in the main paper.

Table 4. Average running times for our experimental results in Table 2 (left side) and Table 3 (right side).

Data set	Time	Data set	Time
Image Seg.	0.2s	*Ising Chain*	0.0s
Knott-3D-150	0.2s	*2D Torus*	0.1s
Knott-3D-300	3.5s	*3D Torus*	0.2s
Knott-3D-450	18.7s	*Deconv.*	0.3s
Knott-3D-550	62.8s	*Super Res.*	0.0s
Modularity Clust.	0.1s	*Texture Rest.*	754.2s
CREMI-small	21.2s		
CREMI-large	4h		
Fruit-Fly Level 1–4	18h		
Fruit-Fly Global	69h		

B.1. Overhead and speedup

In this section we compare the running time effort of the different persistency criteria (Table 5) and the running times for optimization algorithms applied to original or reduced instances (Table 6). To this end, we restrict ourselves to medium size instances that can be solved by exact methods but exhibit nontrivial running times. The optimization algorithms we consider include the combination of greedy edge contraction (GAEC) and local search (KLj) [23] as well as an integer linear program solved by branch-and-cut (ILP).

We can see in Table 5 that the overhead of checking edge and triangle subgraphs is minor while the general subgraph criteria require substantially more time.

We can see in Table 6 that both the heuristic solver and the exact solver benefit from the size reduction achieved by our methods. The speedup of the local search algorithm is less than the overhead by our method, which is expected as our method requires primal/dual heuristics and optimization algorithms as subroutines. However, the results from local search on the original instances come without any partial optimality guarantees. The speedup of the exact solver is substantial, which promotes our method as a preprocessing step in an exact optimization pipeline.

C. Improved multicut subgraph criterion

In this section we describe a technical improvement of the MULTICUT subgraph criterion presented in Theorem 2. Here,

Table 5. The table compares the average running time effort of incorporating progressively more expensive persistency criteria. The criteria are listed analogously to Figure 4.

	Time [s]				
Dataset	[29]	Edge	$\triangle$	Greedy	ICP
Knott-3D-450	0.3	0.7	1.6	12.5	18.7
CREMI-small	0.3	0.5	1.4	3.9	21.2

Table 6. The table compares the average running times of optimization methods on original instances and reduced instances. For the reduction all the presented criteria have been applied.

	Time [s] original		Time [s] reduced	
Dataset	GAEC+KL	ILP	GAEC+KL	ILP
Knott-3D-450	5.7	995.5	3.5	252.0
CREMI-small	1.5	497.4	0.5	208.9

improvement means relaxing the inequality (9) such that it applies more often (without compromising the persistency result).

To this end, we need to introduce some more notation. For any set of vertices $U \subseteq V$, let

$$\partial U = \{v \in V \mid \exists uv \in E \text{ with } u \in U\}$$

denote the *boundary* of U in V. The boundary of U consists of those vertices in V that have a neighbor in U but are not in U themselves. For any set $U \subseteq V$, its *closure* $\overline{U}$ is defined as the union of U with its boundary, i.e.

$$\overline{U} = U \cup \partial U.$$

Further, for any connected subgraph $H = (V_H, E_H)$, we define its *positive closure* as the subgraph

$$\overline{H} = (\overline{V_H}, E_H \cup (\delta(V_H) \cap E^+)),$$

which additionally includes all positive edges between V_H and its boundary.

Below we state a more refined version of Theorem 2. The difference in Theorem 4 is that the inner cut is w.r.t. the subgraph $\overline{H}$ instead of H.

Theorem 4 (Multicut Subgraph Criterion). *Let $H = (V_H, E_H)$ be a connected subgraph of G and suppose $uv \in E_H$. If*

$$\min_{y \in \mathsf{MC}(H)} \langle \theta, y \rangle = 0$$

and for all $U \subset \overline{V_H}$ with $u \in U$ and $v \notin U$ it holds that

$$\sum_{e \in \delta(U, \overline{V_H} \setminus U) \cap E_{\overline{H}}} \theta_e \geq \sum_{e \in \delta(V_H) \cap E^+} \theta_e, \tag{17}$$

then $x^*_{uv} = 0$ in some optimal solution x^* of (P_{MC}).

Proof. The proof is largely analogous to the proof of Theorem 2. Suppose $x \in \mathsf{MC}$ with $x_{|E_H} = y \in \mathsf{MC}(H)$ and $y_{uv} = 1$. Apparently, there exists a multicut M of $\overline{H}$ which extends y such that $x_{|E_{\overline{H}}} = \mathbb{1}_M$. Similarly to before, there exists some $U \subset \overline{V_H}$ with $u \in U$ and $v \notin U$ such that $\delta(U, \overline{V_H} \setminus U) \cap E_{\overline{H}} \subseteq M$ and

$$\sum_{e \in \delta(V_H) \cap E^+} \theta_e x_e + \sum_{e \in E_H} \theta_e x_e \geq \sum_{e \in \delta(U, \overline{V_H} \setminus U) \cap E_{\overline{H}}} \theta_e. \tag{18}$$

Eventually, using (17) and (18), we show that the mapping $p^y = (p_{V_H} \circ p_{\delta(V_H)})$ still improves x, as follows:

$$\begin{aligned}
&\langle \theta, p^y(x) \rangle - \langle \theta, x \rangle \\
&= \sum_{e \in \delta(V_H) \cap E^+} \theta_e (1 - x_e) + \sum_{e \in \delta(V_H) \cap E^-} \theta_e (1 - x_e) \\
&\quad - \sum_{e \in E_H} \theta_e x_e \\
&\leq \sum_{e \in \delta(V_H) \cap E^+} \theta_e - \sum_{e \in \delta(V_H) \cap E^+} \theta_e x_e - \sum_{e \in E_H} \theta_e x_e \\
&\leq \sum_{e \in \delta(V_H) \cap E^+} \theta_e - \sum_{e \in \delta(V_H) \cap E^+} \theta_e \\
&= 0.
\end{aligned}$$

□

The inequality (17) is less restrictive than (9) in Theorem 2, because the left-hand side is potentially larger. Indeed, if two neighboring nodes $u, v \in V_H$ are connected by positive edges to some vertex $w \in \partial V_H$ in the boundary (i.e. they form a positive triangle), then the extension of any cut that separates u from v has to cut another edge of the triangle. Thus, the weight of this edge can be subtracted from the right-hand side of the inequality (9) or, equivalently, added to the left-hand side, which is what (17) achieves.

For the special case of a single edge subgraph $H = (\{u, v\}, \{uv\})$ the refined condition is explicitly stated as

$$\theta_{uv} \geq \sum_{e \in \delta(uv) \cap E^+} \theta_e - \sum_{w \neq u, v | uw, vw \in E^+} \min\{\theta_{uw}, \theta_{vw}\}.$$

Note that the refined condition (17) can be checked by computing minimum u-v-cuts in the graph $\overline{H}$ instead of H. Therefore, a minor modification of the algorithm proposed in the main paper suffices.