A PRIORI OPTIMIZATION FOR THE PROBABILISTIC MAXIMUM INDEPENDENT SET PROBLEM

CAHIER N° 166
janvier 2000

Cécile MURAT 1
Vangelis PASCHOS 1

received: October 1999.

1 LAMSADE, Université Paris-Dauphine, Place du Maréchal De Lattre de Tassigny, 75775 Paris Cedex 16, France (murat,paschos@lamsade.dauphine.fr).
# Table of Contents

Résumé ii

Abstract ii

1 Introduction 1

2 The modification strategies and a preliminary result 4
   2.1 Strategy U1 4
   2.2 Strategies U2 and U3 4
   2.3 Strategy U4 5
   2.4 Strategy U5 5
   2.5 A general mathematical formulation for the five functionals 5

3 The complexity of PIS1 6
   3.1 Computing optimal a priori solutions 6
   3.2 Approximating optimal solutions for PIS1 7
   3.3 PIS1 in bipartite graphs 8

4 The complexities of PIS2 and PIS3 8
   4.1 Expressions for $E_{S}^{P2}$ and $E_{S}^{P3}$ 8
   4.2 Bounds for $E_{S}^{P2}$ 11
   4.3 Approximating optimal solutions for PIS2 11
      4.3.1 Using $\text{argmax}\{\sum_{v \in S} p_i\}$ as a priori solution 11
      4.3.2 Polynomial time approximations for PIS2 12
   4.4 PIS2 in bipartite graphs 13

5 The complexity of PIS4 14
   5.1 An expression for $E_{S}^{P4}$ 14
   5.2 Using $S^*$ or $\text{argmax}\{\sum_{v \in S} p_i\}$ as a priori solutions 15
   5.3 PIS4 in bipartite graphs 17

6 The complexity of PIS5 18
   6.1 In general graphs 18
   6.2 In bipartite graphs 19

7 Conclusions 19

A Proof of proposition 1 22
   A.1 Determining $A(z\tau(G))$ 22
   A.2 Determining $A(z\tau(G)-1)$ 22
Optimisation a priori pour le problème du stable maximum probabiliste

Résumé
Nous commençons par donner une définition formelle du concept relativement nouveau des problèmes d'optimisation combinatoire probabilistes (sous la méthode dite a priori développée par Bertsimas, Jaillet et Odoni). Ensuite, nous étudions la complexité de résoudre optimalement le problème du stable maximum probabiliste sous différentes stratégies de modification, ainsi que la complexité de solutions approchées. Pour les diverses stratégies étudiées nous présentons aussi des résultats sur la restriction du stable maximum probabiliste sur les graphes biparties.

Mots-clé : problèmes combinatoires, complexité, stable, graphe
1 Introduction

In probabilistic combinatorial optimization, the probabilities are associated with the data describing an optimization problem (and not with the relations between data as, for example, in random graph theory ([4])). For a particular datum, we can see the probability associated with it as a measure of how this datum is likely to be present in the instance to be optimized, and in this sense, probabilistic elements are explicitly included in problem's formulation. In such a formulation, the objective function is a kind of carefully defined mathematical expectation over
Definition 2. Consider an NPO problem $\Pi = (P, S, v_P, \text{opt})$ and a modification strategy $\mathcal{U}$ modifying $S$ and producing a feasible $\Pi$-solution for $I$. The probabilistic version $\text{P}\Pi$ of $\Pi$ is a quintuple $(\mathcal{I}, S, \mathcal{U}, E_{\text{P}\Pi}(P_S^\mathcal{U}), \text{opt})$ such that:

- $\mathcal{I}$ is the set of instances of $\text{P}\Pi$ defined as $\mathcal{I} = \{\mathcal{P} = (P, \vec{F}_P)\}$, where $P \in \mathcal{P}$ and $\vec{F}_P$ is an $n$-vector of occurrence-probabilities; $\mathcal{I}$ can be recognized in polynomial time;

- given $\mathcal{P} = (P, \vec{F}_P) \in \mathcal{I}$, $S(\mathcal{P}) = S(P)$

- $\mathcal{U}$ is a modification strategy, i.e., an algorithm which, given a solution $S \in S(\mathcal{P})$ and a sub-instance $I$ of $P$, receives $S$ as input and computes a feasible $\Pi$-solution for $I$;

- given $\mathcal{P} \in \mathcal{I}$ and $S \in S(\mathcal{P})$, $E_{\text{P}\Pi}(P_S^\mathcal{U})$ denotes the value of $S$; if $p_i = \Pr[i]$, $i \in P$, is the occurrence probability of datum $i$, $\mathcal{F}(I_S^\mathcal{U})$ and $F(I_S^\mathcal{U})$ are the solution obtained by running $\mathcal{U}(S)$ on $I$ and its value, respectively, $\Pr[I] = \prod_{i \in P} p_i \prod_{i \in P \setminus I} (1 - p_i)$ is the “occurrence” probability of $I$, then $E_{\text{P}\Pi}(P_S^\mathcal{U}) = \sum_{I \subseteq P} \Pr[I] F(I_S^\mathcal{U})$; $E_{\text{P}\Pi}(P_S^\mathcal{U})$ is commonly called functional;

- opt is the same as for $\Pi$.

The complexity of $\text{P}\Pi$ is the complexity of computing $\hat{S} = \text{argopt}\{E_{\text{P}\Pi}(P_S^\mathcal{U}) : S \in S(\mathcal{P})\}$. Dealing with definition 2, the following remark must be underlined.

**Remark 1.** In definition 2, $\mathcal{U}$ is part of the instance and this seems somewhat unusual with respect to standard complexity theory where no algorithm intervenes in the definition of a problem. But note that $\mathcal{U}$ is absolutely not an algorithm for $\text{P}\Pi$, in the sense that it does not compute $S \in S(\mathcal{P})$. It simply fits $S$ (no matter how $S$ has been computed) to $I$.

Moreover, let us note that changing $\mathcal{U}$ one changes the definition of the problem itself. In other words, given a deterministic problem $\Pi$, the probabilistic problems induced by the quintuples
vertices of the input-graph are provided with positive weights, and the objective becomes to maximize the sum of the weights of the vertices in an independent set.

An instance of probabilistic independent set (PIS) is a pair \((G, \overline{F})\) and is obtained by associating with each \(v_i \in V\) an "occurrence" probability \(p_i\) and by considering a modification strategy \(U\) transforming feasible IS-solution \(S\) of \(G\) into an independent set for the sub-graph of \(G\) induced by a set \(I \subseteq V\). As mentioned above, the objective for PIS is to determine the a priori solution \(\hat{S}\) maximizing the functional \(E_{\text{PIS}}(G^S_\delta)\).

Let us note that PIS is quite different from PCOPs already studied ([2, 3, 11, 12]). There, Euclidean versions of minimization routing problems (such as the traveling salesman, or the
In what follows we consider maximal\(^2\) (although not necessarily maximum) a priori independent sets and use five modification strategies, \(U_\ell, \ell = 1, \ldots, 5\). For \(U_1\) and \(U_5\) we express their functionals in a closed form, we prove that they are computed in polynomial time, and we determine the a priori solutions that maximizing them. For \(U_2\) and \(U_3\), the expressions for the functionals are more complicated and it seems that they cannot be computed in polynomial time. Due to the complicated expressions for these functionals, we have not been able to characterize the a priori solutions maximizing them. Finally, for \(U_4\), we prove that the functional associated can be computed in polynomial time, but we are not able to precisely characterize the optimal a priori solution maximizing it. For all the strategies studied we also study the complexity of approximating optimal a priori solutions. Let us, once more, recall here that the strategies studied introduce in fact five distinct PCOPs denoted in the sequel by PIS1, PIS2, PIS3, PIS4 and PIS5, respectively. Finally, we study the probabilistic version of a natural restriction of IS, the one where the input graph is bipartite.

In the sequel, given a graph \(G = (V, E)\) of order \(n\), we sometimes denote by \(V(G)\) the vertex-set of \(G\). We denote by \(S\) a maximal solution of IS of \(G\), by \(\alpha(G)\) its cardinality, by \(S^*\) a maximum independent set set of \(G\), by \(\alpha^*(G)\) its cardinality, by \(\hat{S}\) an optimal PIS-solution (a priori solution) and by \(\hat{\alpha}(G)\) its cardinality. Moreover, by \(\Gamma(v_i), i = 1, \ldots, n\), we denote the set of neighbors of the vertex \(v_i\) and the quantity \(|\Gamma(v_i)|\) is called degree of \(v_i\); by \(\Gamma'(V')\), \(V' \subseteq V\), we denote the set \(\cup_{v_i \in V'} \Gamma(v_i)\); also, \(\delta_G = \min_{v_i \in V} |\Gamma(v_i)|\), \(\Delta_G = \max_{v_i \in V} |\Gamma(v_i)|\) and \(\mu_G\) is the average degree of \(G\); finally, by \(Pr[v_i] = p_i\), we denote the fact that the presence probability of a vertex \(v_i \in V\) equals \(p_i\). Given a set \(I \subseteq V\), we denote by \(G[I] = (I, E_I)\) the subgraph of \(G\) induced by \(I\) (obviously, there are \(2^n\) such graphs). Given a maximal solution \(S\) of IS (the a priori solution) in \(G\), we denote by \(S[I]\) the set \(S \cap I\). For reasons of simplicity, the functional associated with PIS on graph \(G\) is denoted by \(E_S^\psi\) \((= \sum_{I \subseteq V} Pr[I]F(I_S^\psi))\).

2  The modification strategies and a preliminary result

In what follows we denote by GREEDY the classical greedy IS-algorithm. It works as follows: it orders the vertices of \(V\) in increasing degree-order, it includes the minimum-degree vertex in the solution, it deletes it together with its neighbors (as well as all edges incident to these vertices) from \(V\) \((E)\), it reorders the vertices of the surviving graph and so on, until all vertices are removed. Moreover, we denote by SIMGREEDY, a simplified version of GREEDY where after removing a vertex and its neighbors, the algorithm does not reorder the vertices of the surviving graph.

2.1  Strategy \(U_1\)

Given an a priori IS-solution \(S\) and a present subset \(I \subseteq V\), modification strategy \(U_1\) consists in simply moving the absent vertices out of \(S\).

BEGIN (*U1(G,S)*)

\[ \text{OUTPUT } \mathcal{F}(I_S^\psi) \leftarrow S[I] \leftarrow S \cap I; \]

END. (*U1*)

2.2  Strategies \(U_2\) and \(U_3\)

Modification strategy \(U_2\) is a two-step method: it first applies \(U_1\) to obtain \(S[I]\), it next applies GREEDY on the graph \(G[I] = G[I \setminus (S[I] \cup \Gamma(S[I]))]\) and, finally, it retains the union of the two independent sets obtained as final IS-solution for \(G[I]\).

\(^2\)If we add a vertex, the result is not an independent set.
BEGIN (*U2(G,S) /U3(G,S)/*)
S[I] ← U1(G,S);
G[I] ← G[I \ \{S[I] ∪ Γ(S[I])\}];
S[I] ← GREEDY(G[I]);
/\S[I] ← SIMGREEDY(G[I]);/
OUTPUT \mathcal{F}(I_{S}^{U2}) ← S[I] ∪ S[I];
/\OUTPUT \mathcal{F}(I_{S}^{U3}) ← S[I] ∪ S[I];/
END. (*U2 /U3/)

Finally, strategy U3 is identical to U2 modulo the fact that, instead of GREEDY, algorithm SIMGREEDY is executed (instructions between slashes above refer to modification strategy U3).

2.3 Strategy U4

Strategy U4 starts from \( S[I] \) and completes it with the isolated vertices (vertices with no neighbors) of the graph \( G[I] \).

BEGIN (*U4(G,S)*)
S[I] ← U1(G,S);
G[I] ← G[I \ \{S[I] ∪ Γ(S[I])\}];
OUTPUT \mathcal{F}(I_{S}^{U4}) ← S[I] ∪ \{v_i ∈ I : Γ(v_i) = ∅\};
END. (*U4*)

2.4 Strategy U5

Strategy U5 applies the natural relation between a minimal vertex cover\(^3\) and a maximal independent set in a graph, i.e., a minimal vertex cover (resp., maximal independent set) is the complement, with respect to the vertex set of the graph, of a maximal independent set (resp., a minimal vertex cover).

BEGIN (*U5(G,S)*)
(1) \( C ← V \setminus S \);
(2) \( C[I] ← C ∩ I \);
(3) \( R ← \{v_i ∈ C[I] : Γ(v_i) = ∅\} \);
(4) \( C[I] ← C[I] \setminus R \);
(5) OUTPUT \mathcal{F}(I_{S}^{U5}) ← I \setminus C[I];
END. (*U5*)

2.5 A general mathematical formulation for the five functionals

**Theorem 1.** Consider an a priori solution \( S \) of cardinality \( α(G) \) for \( G \); consider strategies \( U_k \), \( k = 1, \ldots, 5 \). With each vertex \( v_i ∈ V \) we associate a probability \( p_i \) and a random variable \( X_i^{U_k,S} \), \( k = 1, \ldots, 5 \), defined, for every \( I ⊆ V \), by

\[
X_i^{U_k,S} = \begin{cases} 1 & v_i ∈ \mathcal{F}(I_{S}^{U_k}) \\ 0 & \text{otherwise} \end{cases}
\]

Then

\[
E_S^{U_k} = \sum_{v_i ∈ S} p_i + \sum_{v_i ∈ (V \setminus S)} \Pr[X_i^{U_k,S} = 1].
\]

In particular, if for each vertex \( v_i ∈ V \) \( v_i = v \ then\)
Proof. By expression (1), \( F(I_{S}^{uk}) = \sum_{i=1}^{n} x_{i}^{uk,S} \). So,

\[
E_{S}^{uk} = \sum_{I \subseteq V} \Pr[I] F(I_{S}^{uk}) = \sum_{I \subseteq V} \Pr[I] \sum_{i=1}^{n} x_{i}^{uk,S} = \sum_{i=1}^{n} \sum_{I \subseteq V} \Pr[I] x_{i}^{uk,S}
\]

\[
= \sum_{i=1}^{n} E \left( x_{i}^{uk,S} \right) = \sum_{i=1}^{n} \Pr \left[ x_{i}^{uk,S} = 1 \right] = \sum_{i=1}^{n} \Pr \left[ x_{i}^{uk,S} = 1 \right] \left( 1_{\{v_{i} \in S\}} + 1_{\{v_{i} \notin S\}} \right)
\]

\[
= \sum_{i=1}^{n} \Pr \left[ x_{i}^{uk,S} = 1 \right] 1_{\{v_{i} \in S\}} + \sum_{i=1}^{n} \Pr \left[ x_{i}^{uk,S} = 1 \right] 1_{\{v_{i} \notin S\}}.
\]

But, if \( v_{i} \in S \), then necessarily \( x_{i}^{uk,S} = 1 \), \( \forall I \) such that \( v_{i} \in I \); so, \( \Pr[\{x_{i}^{uk,S} = 1\}] = p_{i} \), \( v_{i} \in S \), and consequently,

\[
E_{S}^{uk} = \sum_{v_{i} \in S} p_{i} + \sum_{i=1}^{n} \Pr \left[ x_{i}^{uk,S} = 1 \right] 1_{\{v_{i} \notin S\}} = \sum_{v_{i} \in S} p_{i} + \sum_{v_{i} \notin V \setminus S} \Pr \left[ x_{i}^{uk,S} = 1 \right].
\]

If \( p_{i} = p \), \( v_{i} \in V \), then the result of expression (3) is immediately obtained from expression (2).

Let us note that, as it can be easily deduced from the proof of theorem 1, the above result holds for any strategy which first determines \( S[I] \), it next computes an independent set \( S[\bar{I}] \) on \( G[\bar{I}] \), and it finally considers as solution for \( G[I] \) the set \( S[I] \cup S[\bar{I}] \).

3 The complexity of PIS1

3.1 Computing optimal a priori solutions

From expression (1), we have \( x_{i}^{u_{1},S} = 0 \), \( \forall v_{i} \notin S \), and consequently, \( \Pr[\{x_{i}^{u_{1},S} = 1\}] = 0 \), for \( v_{i} \in V \setminus S \); so, the following theorem is immediately derived for strategy \( u_{1} \).

Theorem 2. Given a graph \( G = (V,E) \), an a priori solution \( S \) and the modification strategy \( u_{1} \), then \( E_{S}^{u_{1}} = \sum_{v_{i} \in S} p_{i} \), and is computed in \( O(n) \). Optimal PIS1-solution \( \hat{S} \) is a maximum-weight independent set in a weighted version of \( G \) where vertices are weighted by the corresponding probabilities. If \( p_{i} = p \), \( \forall v_{i} \in V \), then \( E_{S}^{u_{1}} = p \alpha(G) \); in this case, \( \hat{S} = S^{*} \) and \( E_{S}^{u_{1}} = p \alpha^{*}(G) \).

The characterization of \( \hat{S} \) given in theorem 2 immediately introduces the following complexity result for PIS1.

Theorem 3. PIS1 is NP-hard.

We now show that for \( p_{i} = p \), \( v_{i} \in V \), a mathematical expression for \( E_{S}^{u_{1}} \) can be built directly without applying theorem 1 (used in next sections for the analysis of other strategies). Given that \( 0 \leq |S[I]| = F(I_{S}^{u_{1}}) = |S[I]| \leq \alpha(G) \), we get:

\[
\alpha(G)
\]
where in the last summation we count all the sub-graphs $G[I]$ such that $|S[I]| = i$ and we add their probabilities; also, $\sum_{j=0}^{n-\alpha(G)} C_j^n \alpha(G)^j p^j (1-p)^{n-\alpha(G)-j} = 1$. The above proof for $E_{S}^{U1}$ can be generalized in order to compute every moment of any order for $U1$. For instance,

$$E \left[ \left( G_{S}^{U1} \right)^2 \right] = \sum_{I \subseteq V} \Pr[I] F^2(I^{U1}) = \sum_{i=1}^{\alpha(G)} i^2 \binom{\alpha(G)}{i} p^i (1-p)^{\alpha(G)-i} = \alpha(G)p(\alpha(G)+1-p)$$

and, consequently,

$$\text{Var}_{S}^{U1} = E \left[ \left( G_{S}^{U1} \right)^2 \right] - \left( E_{S}^{U1} \right)^2 = \alpha(G)p(1-p).$$

So, for $U1$, the random variable representing the size $\alpha(G)$ of the a priori solution follows a binomial law with parameters $\alpha(G)$ and $p$.

### 3.2 Approximating optimal solutions for PIS1

In this section we show how, even if one cannot compute the optimal a priori solution in polynomial time, one can compute a sub-optimal solution, the value (expectation) of which is always greater than a factor times the value (expectation) of the optimal one. For this, we shall propose in what follows well-known (in the theory of polynomial approximation of NP-complete problems) polynomial algorithms computing "good" sub-optimal solutions, and show that, also in probabilistic case, these algorithms work well.

Let us first recall that given an instance $P$ of an NP-complete problem II, a common way to estimate the capacity of a polynomial approximation algorithm $A$ in finding good sub-optimal solutions for II, is by evaluating its approximation ratio ([8]), i.e., the ratio $A(P) / \text{OPT}(P)$, where $A(P)$ denotes the value of the solution computed by $A$ on $P$ and $\text{OPT}(P)$ denotes the value of the optimal solution of $P$. Then the approximation ratio of $A$ for a maximization problem II is the quantity $\inf_{P} \{ A(P) / \text{OPT}(P) : P \text{ instance of II} \}$; the closer this ratio to 1, the better the approximation algorithm.

Recall that as we have already seen in section 3, PIS1 is equivalent to a weighted IS-problem, where each vertex is weighted by the corresponding probability. Consequently, the following theorem holds immediately.

**Theorem 4.** If there exists a polynomial time approximation algorithm $A$ solving WIS within approximation ratio $\rho$, then $A$ polynomially solves PIS1 within the same approximation ratio $\rho$.

In [6], an algorithm is developed for WIS achieving approximation ratio the minimum value between $\log n / (3(\Delta_G + 1) \log \log n)$ and $O(n^{-4/5})$. Using this algorithm in theorem 4, one gets the following corollary.

**Corollary 1.** PIS1 can be approximated within

$$\min \left\{ \frac{\log n}{3(\Delta_G + 1) \log \log n}, O(n^{-4/5}) \right\}$$

The characterization of PIS1 in terms of a weighted IS-problem draws not only issues for finding reasonable a priori sub-optimal solutions but, unfortunately, limits the capacity of the problem to be "well-approximated" since, via this characterization, all the negative results applying to IS are immediately transferred to PIS1 also. So, PIS1 is hard to approximate within $n^{1-\epsilon}$, for any $\epsilon > 0$ ([10]).
3.3 PIS1 in bipartite graphs

Theorem 5. Consider a bipartite graph \( B = (V_1, V_2, E_B) \). Then, \( E^B_S = \sum_{i \in S} p_i \) and is computed in polynomial time. The optimal a priori solution \( \hat{S} \) is a maximum-weight independent set in \( B \) considering that its vertices are weighted by the corresponding presence probabilities, and can be found in \( O(n^{1/2}|E_B|) \). Consequently, in bipartite graphs, \( PIS1 \in P \).

Proof. Concerning the expression for \( E^B_S \) and the complexity of its computation, the proof is the same as the one of theorem 2.

Determining an optimal IS-solution in a bipartite graph is of polynomial complexity in both weighted and unweighted cases (see [9] for the unweighted case; for the weighted one, we quote here the result of [5] where the polynomiality of weighted IS in a class of graphs including the bipartite ones is proved).

4 The complexities of PIS2 and PIS3

4.1 Expressions for \( E^u_1 \) and \( E^u_2 \)

Let \( A_i = \sum_{i \in V} \sum_{|S \cap |=i} \Pr[I|F(I^u_1)] \). Then, \( E^u_2 \) can be nicely written as follows:

\[
E^u_2 = \sum_{i \in V} \Pr[I|F(I^u_1)] = \sum_{i \in V} \left( \sum_{|S \cap |=i} \alpha(G) \right) \Pr[I|F(I^u_1)] = \sum_{i \in V} \left( \sum_{|S \cap |=i} \Pr[I|F(I^u_1)] \right) = \sum_{i \in V} A_i.
\]

Quantities \( A_i, i = 1, \ldots, \alpha(G) \) \((A_0 = 0)\) are very natural and interesting from both theoretical and practical points of view. For instance, formula for \( E^u_2 \) given by expression above holds for every probability law; also, computing analytical expressions for \( A_i \) seems to be an interesting problem in combinatorial counting of graphs; moreover, thanks to the simple relation between \( E^u_2 \) and \( A_i, i = 1, \ldots, \alpha(G) \), analytical expressions for the latter would produce explicit expressions for the former. Unfortunately, expression above for \( E^u_2 \), even intuitive and smart, does not give any hint allowing precise characterization of \( \hat{S} \).

In proposition 1, the proof of which is given in appendix A, \( A_{\alpha^*(G)} \) and \( A_{\alpha^*(G)-1} \) are explicitly computed, for the case of identical vertex-probabilities. However, the explicit computation for \( A_i \)'s of lower index produces very long and non-intuitive expressions.

Proposition 1. Let \( \Gamma'(v) = \Gamma(v) \setminus \{\Gamma(v) \cap \Gamma(S^* \setminus \{v\})\} \), \( \ell_1 = |\{v \in S^* : \Gamma'(v) = \emptyset\}| \), \( \ell'_1 = \alpha^*(G) - \ell_1 \), and \( p_i = p, \forall v_i \in V \). Then,

\[
A_{\alpha^*(G)} = \alpha^*(G)p^{\alpha^*(G)} \\
A_{\alpha^*(G)-1} = p^{\alpha^*(G)-1}(1-p) \left( \alpha^*(G)\ell_1 - \ell_1 + \alpha^*(G)\ell'_1 - \sum_{v \in S^* \setminus \Gamma'(v)} (1-p)^{|\Gamma'(v)|} \right).
\]

We shall now give an upper bound for the complexity of computing \( E^u_2 \). For this, we will analyze (as an intermediate step) strategy \( U3 \) introduced in section 2.

Let \( G' = G[V \setminus S] = (V', E') \), and let \( V' = \{v_1, \ldots, v_{n-\alpha(G)}\} \) be the list of vertices of \( G' \) sorted in increasing-degree order; let us denote by \( V_i \) the set of the \( i \) first vertices of \( V' \) and let \( G'_i = G'[V_i] \) (of course, for \( G'_i \) the vertices of \( V_i \) are not sorted in increasing-degree order). Let us denote by \( S_i \) the independent set found by \( U3 \) on (the present sub-instance of) \( G'_i \), and by \( \alpha_i \) its cardinality, \( i = 1, \ldots, n - \alpha(G) \). Finally, let \( \alpha(G[I']) \) be the cardinality of the solution provided by \( U3 \) when applied in graph \( G[I'], I' \subseteq V \).
Theorem 6.

\[ E (\alpha_{n-\alpha(G)}) = \sum_{i=1}^{n-\alpha(G)} p_i \Pr [v_i \notin \Gamma(S_{i-1})] \]

\[ E_{S}^{S3} = E_{S}^{S1} + E (\alpha_{n-\alpha(G)}) . \]

If we denote by \( T(E(\alpha_{n-\alpha(G)})) \) and \( T(E_{S}^{S3}) \) the computation times of \( E(\alpha_{n-\alpha(G)}) \) and \( E_{S}^{S3} \), respectively, then \( T(E(\alpha_{n-\alpha(G)})) = O(2^{\alpha(G)}) \) and \( T(E_{S}^{S3}) = O(2^{\alpha(G)}) \).

**Proof.** \( E(\alpha_i) = E(\alpha_i|v_i \text{ present})p_i + E(\alpha_i|v_i \text{ absent})(1 - p_i) \). Moreover, for strategy U3 we have the following relation, setting \( S_0 = \emptyset \), \( \Gamma(S_0) = \emptyset \) and \( \alpha_0 = 0 \):

\[ \alpha_i = \begin{cases} \alpha_{i-1} & \Gamma(v_i) \cap S_{i-1} \neq \emptyset \\ \alpha_{i-1} + 1 & \text{otherwise} \end{cases} \]

So, \( E(\alpha_i|v_i \text{ present}) = E(\alpha_{i-1}) + \Pr[\Gamma(v_i) \cap S_{i-1} = \emptyset] \); consequently,

\[ E(\alpha_i) = p_i E(\alpha_{i-1}) + p_i \Pr [v_i \notin \Gamma(S_{i-1})] + (1 - p_i) E(\alpha_{i-1}) . \]

Since \( E(\alpha_0) = 0 \),

\[ E(\alpha_i) = \sum_{j=1}^{i} p_j \Pr [v_j \notin \Gamma(S_{j-1})] \]

\[ E (\alpha_{n-\alpha(G)}) = \sum_{i=1}^{n-\alpha(G)} p_i \Pr [v_i \notin \Gamma(S_{i-1})] . \]

Now, let \( I' = f(I) = I \setminus (S[I] \cup \Gamma(S[I])) \). This set represents the subset of vertices of \( I \) which are not contained neither in \( S \), nor in the neighbor-set of \( S[I] \). Consequently, U3 will be applied on \( G[I'] \); so, \( F(I_{S3}) = |S[I]| + \alpha(G[I']) \) and consequently

\[ E_{S}^{S3} = \sum_{I \subseteq V} \Pr[I|S[I]] + \sum_{I \subseteq V} \Pr[I] \alpha (G[I']) = E_{S}^{S1} + \sum_{I \subseteq V} \Pr[I] \alpha (G[I']) \]

\[ = E_{S}^{S1} + \sum_{I \subseteq V} \left( \sum_{I \subseteq V} 1(f(I) = r') \Pr[I] \alpha (G[I']) \right) = E_{S}^{S1} + \sum_{I \subseteq V} \sum_{f(I) = r'} \Pr[I] \alpha (G[I']) . \]

Since \( \Pr[I'] = \sum_{I \subseteq V, f(I) = r'} \Pr[I] \), we get

\[ E_{S}^{S3} = E_{S}^{S1} + \sum_{I \subseteq V} \Pr[I'] \alpha (G[I']) = E_{S}^{S1} + E (\alpha_{n-\alpha(G)}) . \]

Let us now introduce the random variable \( C_i \) representing the solution of PIS3 whenever we consider only the present vertices of \( V_i \), i.e., when we apply the greedy algorithm implied by strategy U3 in the graph induced by the present vertices of \( V_i \).

\[ C_i = \begin{cases} C_{i-1} & v_i \text{ is absent} \\ C_{i-1} \cup \{v_i\} & v_i \text{ is present and } \Gamma(v_i) \cap C_{i-1} \neq \emptyset \\ C_{i-1} \cup \{v_i\} & v_i \text{ is present and } \Gamma(v_i) \cap C_{i-1} = \emptyset \end{cases} \]

We then have \( E(\alpha_{n-\alpha(G)}) = \sum_{i=1}^{n-\alpha(G)} p_i \Pr[\Gamma(v_i) \cap C_{i-1} = \emptyset] \).
In order to prove the result of the theorem, we prove that the quantity \( \Pr[\Gamma(v_i) \cap C_{i-1} = \emptyset] \) is not computable in polynomial time.

\[
\Pr[\Gamma(v_i) \cap C_{i-1} = \emptyset] = (1 - p_{i-1}) \Pr[\Gamma(v_i) \cap C_{i-2} = \emptyset]
+ p_{i-1} \Pr[\Gamma(v_i) \cap C_{i-2} = \emptyset] \Pr[\Gamma(v_{i-1}) \cap C_{i-2} \neq \emptyset]
+ p_{i-1} \Pr[\Gamma(v_i) \cap (C_{i-2} \cup \{v_{i-1}\}) = \emptyset] \Pr[\Gamma(v_{i-1}) \cap C_{i-2} = \emptyset]
= (1 - p_{i-1}) \Pr[\Gamma(v_i) \cap C_{i-2} = \emptyset]
+ p_{i-1} \Pr[\Gamma(v_i) \cap C_{i-2} = \emptyset] \Pr[\Gamma(v_{i-1}) \cap C_{i-2} = \emptyset]
+ p_{i-1} \Pr[\Gamma(v_i) \cap (C_{i-2} \cup \{v_{i-1}\}) = \emptyset] \Pr[\Gamma(v_{i-1}) \cap C_{i-2} = \emptyset]

\[
= \Pr[\Gamma(v_i) \cap C_{i-2} = \emptyset]
+ p_{i-1} \Pr[\Gamma(v_i) \cap C_{i-2} = \emptyset] \Pr[\Gamma(v_{i-1}) \cap C_{i-2} = \emptyset]
+ p_{i-1} \Pr[\Gamma(v_i) \cap (C_{i-2} \cup \{v_{i-1}\}) = \emptyset] \Pr[\Gamma(v_{i-1}) \cap C_{i-2} = \emptyset] .
\]

On the other hand,

\[
\Pr[\Gamma(v_i) \cap (C_{i-2} \cup \{v_{i-1}\}) = \emptyset] = \Pr[\Gamma(v_i) \cap C_{i-2} \cup \{v_{i-1}\} = \emptyset]
= \Pr[\Gamma(v_i) \cap C_{i-2} = \emptyset] \cup \{v_{i-1}\} = \emptyset]
= \Pr[\Gamma(v_i) \cap C_{i-2} = \emptyset] 1_{\{\Gamma(v_i) \cap (v_{i-1}) = \emptyset\}} .
\]

Consequently,

\[
\Pr[\Gamma(v_i) \cap C_{i-1} = \emptyset] = \Pr[\Gamma(v_i) \cap C_{i-2} = \emptyset]
- p_{i-1} \Pr[\Gamma(v_i) \cap C_{i-2} = \emptyset] \Pr[\Gamma(v_{i-1}) \cap C_{i-2} = \emptyset]
+ p_{i-1} \Pr[\Gamma(v_i) \cap C_{i-2} = \emptyset] \Pr[\Gamma(v_{i-1}) \cap C_{i-2} = \emptyset] 1_{\{\Gamma(v_i) \cap (v_{i-1}) = \emptyset\}} .
\]

Let \( t_{i-1}(v_i) \) be the computational time of \( \Pr[\Gamma(v_i) \cap C_{i-1} = \emptyset] \). By the above equalities we easily deduce that

\[
t_{i-1}(v_i) = t_{i-2}(v_i) + t_{i-2}(v_{i-1}) .
\]

In order to compute this recurrence relation, at each step we need to know two terms of the precedent step; so for the computation of \( \Pr[\Gamma(v_i) \cap C_{i-1} = \emptyset] \), we need \( 2^{i-1} \) computations and expression for \( T(E)^{\text{IS}} \) immediately follows. ■

Algorithm U3 is, as it has been already noted, a simplified version of algorithm U2. Moreover, there exist graphs where the two algorithms give the same results by performing identical choices and deletions of vertices (for example, consider a graph on \( n \) isolated vertices). Consequently, computation time of U3 is a (worst-case) lower bound for the one of U2 and the following theorem holds.

**Theorem 7.** Let \( T(E)^{\text{IS}} \) be the computational time of \( E^{\text{IS}} \). Then, \( T(E)^{\text{IS}} = \Omega(T(E)^{\text{IS}}) \).

The result of theorem 6 simply gives an upper bound on the complexity of computing \( E^{\text{IS}} \) and does not prove that \( E^{\text{IS}} \) is not computable in polynomial time (if this was true, it would be a very interesting result since, in this case, PIS2 and PIS3 would not belong to NP). In fact, the result of theorem 6 is based upon a particular recursion-formula and a particular way for computing it. In any case, one can easily prove that PIS2 is intractable (following the notation in the appendix of [8], PIS2 is a kind of starry problem).

Indeed, if one can polynomially determine an optimal a priori solution \( \hat{S} \) for PIS2, then one can simply consider an instance of IS as a PIS2-instance with \( p_1 = 1, \forall v_i \in V \). It is easy to see that, in this case, \( \hat{S} = S^{*} \) and the following theorem immediately holds.

**Theorem 8.** Unless \( P=NP \), PIS2 is computationally intractable.
4.2 Bounds for $E_{S}^{12}$

For lack of characterizing the complexity of computing $E_{S}^{12}$, we build in this paragraph upper and lower bounds for it.

**Theorem 9.** Let $\bar{\Delta}$ be the maximum degree of $G[\bar{I}]$. Then, on the hypothesis of distinct vertex-probabilities:
Theorem 10. Approximation of $\tilde{S}$ by $\bar{S} = \text{argmax}\{\sum_{v_i \in S} p_i : S \text{ independent set of } B\}$ guarantees for PIS2 approximation ratio

$$\max \left\{ \frac{p_{\text{min}}}{1 + p_{\text{max}}}, \frac{1}{\Delta_{G} + 1} \right\}.$$ 

Proof. By expression (7) (proof of theorem 9), and since $\alpha(G[I]) \leq \alpha^{*}(G)$, we get

$$\sum_{v_i \in S} p_i \leq E_{\tilde{S}}^{\text{P2}} = \sum_{v_i \in S} p_i + \sum_{I \subseteq V} \Pr[I] \alpha(G[I]) \leq \alpha^{*}(G) \left( p_{\text{max}} + \sum_{I \subseteq V} \Pr[I] \right) = \alpha^{*}(G) (1 + p_{\text{max}}) \quad (8)$$

Since, $\bar{S} = \text{argmax}\{\sum_{v_i \in S} p_i : S \text{ independent set of } G\}$, then

$$E_{\bar{S}}^{\text{P2}} \geq \sum_{v_i \in S} p_i \geq \sum_{v_i \in S} p_i \geq p_{\text{min}} \alpha^{*}(G) \quad (9)$$
Theorem 11. If there exists a polynomial time approximation algorithm \( A \) solving \( IS \) within approximation ratio \( \alpha \), then algorithm \( A \) polynomially solving \( PISP \) within approximation ratio \( \alpha \).
Corollary 2. Suppose $\Pr[v_i] = p_i$, $v_i \in V_1 \cup V_2$, denote by $n_1$ and $n_2$ the sizes of $V_1$ and $V_2$, respectively, and suppose that $n_1 \geq n_2$. Then, $E_{V_1} = pn_1 + p \sum_{v_i \in V_2} (1 - p)^{\Delta(v_i)}$. Naturally, $E_{V_1}$ is computed in $O(n^2)$.

From corollary 2 we can obtain the following framing of $E_{V_1}^{r2}$ by $E_{V_1}^{r1}$ for the case of identical vertex-probabilities: $E_{V_1}^{r1} + n_2 p (1 - p)^{\Delta_B} \leq E_{V_1}^{r2} \leq E_{V_1}^{r2} + n_2 p (1 - p)^{\delta_B}$. So, in regular bipartite graphs (i.e., the ones where $\Delta_B = \delta_B = \Delta$): $E_{V_1}^{r2} = E_{V_1}^{r1} + n_2 p (1 - p)^{\Delta}$.

Consider set $\hat{S} = \operatorname{argmax} \{ \sum_{v_i \in S} p_i : S \text{ independent set of } B \}$ as a priori solution. Then, considering vertex-probabilities as vertex-weights and using the result of [5], $\hat{S}$ can be computed in polynomial time. Moreover,

$$E_{\hat{S}}^{r2} \geq \sum_{v_i \in \hat{S}} p_i \geq \frac{\sum_{v_i \in V} p_i}{n}.$$  

Using the expression above together with expression (11), approximation ratio 1/2 is immediately yielded.

Proposition 2. $\hat{S} = \operatorname{argmax} \{ \sum_{v_i \in S} p_i : S \text{ independent set of } B \}$ is a polynomial time approximation of $\hat{S}$ guaranteeing approximation ratio 1/2 for PIS2 in bipartite graphs.

Obviously, from theorem 12 and the discussion just above, the same approximation ratio can be yielded if one uses $\operatorname{argmax} \{ \sum_{v_i \in V_1} p_i, \sum_{v_i \in V_2} p_i \}$ as a priori solution.

5 The complexity of PIS4

5.1 An expression for $E_{S}^{r4}$

Recall that strategy $U_4$ starts from $S[I]$ and completes it with the isolated vertices of the graph $G[I]$.

Proposition 3. Given a graph $G = (V, E)$, an a priori independent set $S$ and the modification strategy $U_4$, then

$$E(G_{S}^{U_4}) = \sum_{v_i \in S} p_i + \sum_{v_i \in (V \setminus S)} p_i \prod_{v_j \in \Gamma_S(v_i)} (1 - p_j) \times \prod_{v_k \in \Gamma_{V \setminus S}(v_i)} \left(1 - p_k\right) + p_k \left(1 - \prod_{v_l \in \Gamma_{S}(v_k)} (1 - p_l)\right)$$

where $\Gamma_S(v_i) = \Gamma(v_i) \cap S$ and $\Gamma_{V \setminus S}(v_i) = \Gamma(v_i) \cap (V \setminus S)$. $E(G_{S}^{U_4})$ can be computed in polynomial time. If $p_i = p$, for all $v_i \in V$, then

$$E(G_{S}^{U_4}) = p\alpha(G) + \sum_{v_i \in (V \setminus S)} p(1 - p)^{|\Gamma_{S}(v_i)|} \prod_{v_k \in \Gamma_{V \setminus S}(v_i)} (1 - p(1 - p)^{|\Gamma_{S}(v_k)|})$$

Proof. Set $B_i = \Pr[X_i^{U_4,S} = 1]$. Then

$$B_i = \sum_{I \subseteq V \setminus S} \Pr[I] \mathbf{1}_{\{v_i \in \mathcal{F}(I^{U_4})\}}$$

Let $v_i$ be any vertex of $V \setminus S$ and let $I$ be any subset of $V$ containing $v_i$. Obviously,

$v_i \in \mathcal{F}(I^{U_4}) \iff v_i \text{ isolated in } G[I] \iff (v_i \in G[I]) \land (v_i \text{ has lost all its neighbors in } I)$. 

14
Since \( v_i \notin S, v_i \notin \bar{I} \) only if there does not belong to the neighborhood of any vertex in \( S[I] \), i.e.,
\[
v_i \in G[\bar{I}] \iff \Gamma_{S}(v_i) \cap I = \emptyset
\]  
\[(17)\]

On the other hand, all the neighbors of \( v_i \) in \( G[\bar{I}] \) have been removed iff \( \Gamma_{V \setminus S}(v_i) \cap \bar{I} = \emptyset \). This last condition is satisfied only if every vertex of \( \Gamma_{V \setminus S}(v_i) \) is either absent or (being present) has been removed from \( \bar{I} \) because it belonged to the neighborhood of a vertex in \( S[I] \). In all,
\[
\Gamma_{V \setminus S}(v_i) \cap \bar{I} = \emptyset \iff \forall v_j \in \Gamma_{V \setminus S}(v_i) ((v_j \text{ is absent}) \lor ((v_j \text{ is present}) \land (\exists v_k \in \Gamma(v_j) \cap S \text{ such that } v_k \text{ is present})))
\]  
\[(18)\]

From relations (17) and (18) and the discussion above we get \((v_i \in V \setminus S)\):

\[
B_i = \sum_{\bar{I} \in \mathcal{I}} \Pr[\bar{I}] \mathbb{1}_{\{v_i \in \mathcal{I}(\bar{I})\}} = \sum_{\bar{I} \in \mathcal{I}} \Pr[\bar{I}] \mathbb{1}_{\{v_i \in \mathcal{I}(\bar{I})\}}
\]
\[ p\hat{\alpha}(G) + \sum_{v_i \in (V \setminus \hat{S})} p(1-p)^k_i \left( 1 - p(1-p)^{\Delta_G - 1} \right)^{\left| \Gamma(v_i) \right| - k_i} \leq \]

\[ 1 - p \leq p(1-p)^{\Delta_G - 1} \]

\[ \leq p\hat{\alpha}(G) + \sum_{v_i \in (V \setminus \hat{S})} p \left( 1 - p(1-p)^{\Delta_G - 1} \right)^{\left| \Gamma(v_i) \right|} \]

\[ \left| \Gamma(v_i) \right| \geq 1 \]

\[ p\hat{\alpha}(G) + (n - \hat{\alpha}(G)) p(1-p)^{\Delta_G - 1} \leq pn \quad (20) \]

On the other hand, if we consider \( S^* \) as an a priori solution, we get

\[ E(G_{S^*}^{\text{na}}) = p\alpha^*(G) + \sum_{v_i \in (V \setminus S^*)} p(1-p)^{k_i} \prod_{v_k \in \Gamma(v_i)} (1 - p(1-p)^{k_k}) \]

\[ 1 - p(1-p)^{k_k} \geq 1 - p \]

\[ \geq p\alpha^*(G) + \sum_{v_i \in (V \setminus S^*)} p(1-p)^{k_i} \prod_{v_k \in \Gamma(v_i)} (1 - p) \]

\[ = p\alpha^*(G) + \sum_{v_i \in (V \setminus S^*)} p(1-p)^{k_i} (1-p)^{\left| \Gamma(v_i) \right| - k_i} \]

\[ = p\alpha^*(G) + \sum_{v_i \in (V \setminus S^*)} p(1-p)^{\left| \Gamma(v_i) \right|} \]

\[ \geq p\alpha^*(G) + (n - \alpha^*(G)) p(1-p)^{\Delta_G} \]

\[ = p\alpha^*(G)(1 - (1-p)^{\Delta_G}) + pn(1-p)^{\Delta_G} \quad (21) \]

Combining expressions (20) and (21), we have

\[ \frac{E(G_{S^*}^{\text{na}})}{E(G_{S}^{\text{na}})} \geq \frac{\alpha^*(G)}{n} (1 - (1-p)^{\Delta_G}) + (1-p)^{\Delta_G} \geq \frac{\alpha^*(G)}{n} \geq 1 + \frac{1}{\Delta_G + 1}. \]

For instance, if \( G \) is cubic, i.e., \( \Delta_G = 3 \), then \( \alpha^*(G) \geq n/4 \) then, \( E(G_{S^*}^{\text{na}})/E(G_{S}^{\text{na}}) \geq ((1/4)(1 - (1-p)^3)) + (1-p)^3 \geq 1/4 \). If, in addition \( p = 1/2 \), then \( E(G_{S^*}^{\text{na}})/E(G_{S}^{\text{na}}) \geq 11/32 \).

The result of theorem 13 can be easily extended in the case where vertex-probabilities are distinct and \( \hat{S} = \text{argmax}\{\sum_{v_i \in S} p_i : S \text{ independent set of } G\} \) is used as a priori solution. Moreover, without loss of generality, one can suppose that \( \hat{S} \) is maximal. Then, from expression (15) one easily gets:

\[ E(G_{\hat{S}}^{\text{na}}) \leq \sum_{v_i \in \hat{S}} p_i + \sum_{v_i \in (V \setminus \hat{S})} p_i = \sum_{v_i \in V} p_i \quad (22) \]

\[ E(G_{\hat{S}}^{\text{na}}) \geq \sum_{v_i \in \hat{S}} p_i \quad (23) \]

Combining the above expressions we obtain

\[ \frac{E(G_{\hat{S}}^{\text{na}})}{E(G_{S}^{\text{na}})} \geq \frac{\sum_{v_i \in \hat{S}} p_i}{\sum_{v_i \in V} p_i} \geq \frac{1}{\Delta_G + 1} \]

where the last inequality is the weighted version of Turán's theorem.

Finally, let us note that the same approximation ratio can be obtained if one treats vertex probabilities as weights and uses as a priori solution the one computed by the greedy IS-algorithm.
In the weighted case, this algorithm iteratively chooses the vertex maximizing the ratio “vertex-weight over vertex-degree” and eliminates its neighbors. In this way, if $S$ is an independent set
6 The complexity of PIS5

6.1 In general graphs

We recall that strategy U5 considers the restriction $C[I]$ of an a priori vertex cover in the present subgraph $G[I]$ of $G$, it removes the isolated vertices (if any) from $C[I]$ and it finally takes the complement, with respect to $I$, of the resulting set.

**Theorem 15.** Given a graph $G = (V, E)$, an a priori independent set $S$ and the modification strategy U5, then

$$E(G_{S}^{U5}) = \sum_{v_{i} \in S} p_{i} + \sum_{v_{i} \in (V \setminus S), \forall j \in \Gamma(v_{i})} p_{i} \prod_{v_{j} \in \Gamma(v_{i})} (1 - p_{j}). \quad (26)$$

$E(G_{S}^{U5})$ is computable in polynomial time.

**Proof.** Lines (1) to (4) of modification strategy U5 in section 2 constitute a modification strategy, denoted by $U$ in what follows, for probabilistic vertex cover problem. For an a priori vertex cover $C$, the functional associated with $U$ is (see [14] for a detailed computation):

$$E_{C}^{U} = \sum_{v_{i} \in C} p_{i} - \sum_{v_{i} \in C} p_{i} \prod_{v_{j} \in \Gamma(v_{i})} (1 - p_{j}). \quad (27)$$

Using expression (27), we have the following for $E(G_{S}^{U5})$:

$$E(G_{S}^{U5}) = \sum_{I \subseteq V} \Pr[I] F\left(G[I]_{S}^{U5}\right) = \sum_{I \subseteq V} \Pr[I] \left(|I| - F\left(G[I]_{V \setminus S}\right)\right)$$

$$= \sum_{I \subseteq V} \Pr[I] |I| - \sum_{I \subseteq V} \Pr[I] F\left(G[I]_{V \setminus S}\right) = \sum_{I \subseteq V} \Pr[I] \sum_{v_{i} \in V} 1_{(v_{i} \in I)} - E(G_{V \setminus S}^{U5})$$

$$= \sum_{v_{i} \in V} \sum_{I \subseteq V} \Pr[I] 1_{v_{i} \in I} - E(G_{V \setminus S}^{U5}) = \sum_{v_{i} \in V} p_{i} - \sum_{v_{i} \in (V \setminus S)} p_{i} + \sum_{v_{i} \in (V \setminus S)} p_{i} \prod_{v_{j} \in \Gamma(v_{i})} (1 - p_{j})$$

$$= \sum_{v_{i} \in S} p_{i} + \sum_{v_{i} \in (V \setminus S)} p_{i} \prod_{v_{j} \in \Gamma(v_{i})} (1 - p_{j}).$$

Expression (26) can be rewritten as

$$E(G_{S}^{U5}) = \sum_{v_{i} \in V} p_{i} - \sum_{v_{i} \in (V \setminus S)} p_{i} + \sum_{v_{i} \in (V \setminus S)} p_{i} \prod_{v_{j} \in \Gamma(v_{i})} (1 - p_{j})$$

$$= \sum_{v_{i} \in V} p_{i} - \sum_{v_{i} \in (V \setminus S)} p_{i} \left(1 - \prod_{v_{j} \in \Gamma(v_{i})} (1 - p_{j})\right).$$
6.2 In bipartite graphs

Since maximum-weight independent is polynomial in bipartite graphs ([5]), so does minimum-weight vertex covering. So the following theorem immediately holds.

**Theorem 17.** The a priori solution \( \hat{\mathcal{S}} \) maximizing \( E(B^w_{\mathcal{S}}) \) is the complement, with respect to \( V_1 \cup V_2 \), of a minimum-weight vertex cover of \( B \) where every vertex \( v_i \) is weighted by a weight \( p_i(1 - \prod_{v_j \in \Gamma(v_i)} (1 - p_j)) \). Consequently, PIS5 is polynomial for bipartite graphs.

7 Conclusions

We have drawn a formal framework for the study of PCOPs and studied five variants of the probabilistic maximum independent set, defined with respect to natural modification strategies used to adapt an a priori solution to the "present sub-instance".

<table>
<thead>
<tr>
<th>Table 1: Complexities of computing functionals and characterizations and complexities of computing a priori solutions for several variants of probabilistic independent set.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General graphs</strong></td>
</tr>
<tr>
<td>( T(E) )</td>
</tr>
<tr>
<td>( \hat{\mathcal{S}} )</td>
</tr>
<tr>
<td><strong>Complexity</strong></td>
</tr>
<tr>
<td><strong>Bipartite graphs</strong></td>
</tr>
<tr>
<td>( T(E) )</td>
</tr>
<tr>
<td>( \hat{\mathcal{S}} )</td>
</tr>
<tr>
<td><strong>Complexity</strong></td>
</tr>
</tbody>
</table>

(a) Polynomial if \( \arg\max\{|V_1|, |V_2|\} \) is used as an a priori solution.

Dealing with the quality of the solutions obtained we have the following relation for the same a priori solution \( S \) (let us note that, until now, we have not been able to compare \( E(G^{w}_{S}) \) with \( E(G^{w}_{\hat{S}}) \)):

\[
E(G^{w}_{S}) \leq E(G^{w}_{\hat{S}}) \leq E(G^{w}_{S}) \leq E(G^{w}_{\hat{S}})
\]  

(28)

First inequality in expression (28) is obvious and follows from theorem 2 and expression (26) of theorem 15. In order to prove the second inequality, remark that (expression (15))

\[
p_k \times \left(1 - \prod_{v_i \in \Gamma_S(v_k)} (1 - p_i)\right) \geq 0
\]

and consequently,

\[
E(G^{w}_{S}) \geq \sum_{v_i \in S} p_i + \sum_{v_i \in (V \setminus S)} p_i \times \prod_{v_j \in \Gamma_S(v_i)} (1 - p_j) \times \prod_{v_k \in \Gamma_{S \setminus S}(v_i)} (1 - p_k)
\]

\[
\geq \sum_{v_i \in S} p_i + \sum_{v_i \in (V \setminus S)} p_i \times \prod_{v_j \in (\Gamma_S(v_i) \cup \Gamma_{S \setminus S}(v_i))} (1 - p_j)
\]

\[
= \sum_{v_i \in S} p_i + \sum_{v_i \in (V \setminus S)} p_i \times \prod_{v_j \in \Gamma(v_i)} (1 - p_j) \geq E(G^{w}_{\hat{S}}).
\]
Table 2: Approximating a priori solutions in general and bipartite graphs.

<table>
<thead>
<tr>
<th></th>
<th>$S$</th>
<th>Approximation ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>General graphs</td>
</tr>
<tr>
<td>PIS1</td>
<td>The one computed in [6]</td>
<td>$r$</td>
</tr>
<tr>
<td>PIS2</td>
<td>The one computed in [6]</td>
<td>$\max\left{ \frac{p_{\min}}{1+p_{\max}}, r, \frac{1}{\Delta_{G+1}} \right}$ (a)</td>
</tr>
<tr>
<td>PIS4</td>
<td>$\arg\max{\sum_{v_i \in S} p_i}$</td>
<td>$\max\left{ \frac{p_{\min}}{1+p_{\max}}, \frac{1}{\Delta_{G+1}} \right}$ (b)</td>
</tr>
<tr>
<td></td>
<td>$\arg\max{\sum_{v_i \in S} p_i}$</td>
<td>$\frac{1}{\Delta_{G+1}}$</td>
</tr>
<tr>
<td>PIS5</td>
<td>The output of the greedy algorithm</td>
<td>$r$</td>
</tr>
<tr>
<td>PIS2</td>
<td>$\arg\max{\sum_{v_i \in S} p_i}$</td>
<td>$\frac{1}{2}$ (c)</td>
</tr>
<tr>
<td>PIS4</td>
<td>$\arg\max{\sum_{v_i \in S} p_i}$</td>
<td>$\frac{1}{2}$ (d)</td>
</tr>
</tbody>
</table>

(a) $\min\{O(\log n/(3(\Delta_G + 1) \log \log n)), O(n^{-4/5})\}$ if vertex-probabilities independent of $n$.
(b) $\alpha^*(G)/n$ whenever $S = S^*$ and vertex-probabilities are identical.
(c) The same ratio is achieved considering $\arg\max\{\sum_{v_i \in V_1} p_i, \sum_{v_i \in V_2} p_i\}$ as a priori solution.
(d) The same ratio is achieved considering $\arg\max\{\sum_{v_i \in V_1} p_i, \sum_{v_i \in V_2} p_i\}$ as a priori solution.

Last inequality is due to the fact that for every subgraph $G[I]$, the cardinality of the solution computed applying U2 will be greater than the one of the solution computed applying U4 since GREEDY (called by U2) will always add in the solution, at least the isolated vertices of $G[I]$.

Table 1 summarizes the main results of this paper about the complexities of computing the functionals and the ones of computing the a priori solutions maximizing them. In this table we denote by $G^w_p$, a graph $G$ whose vertices are weighted by their corresponding probabilities, by $G^w_p$ a graph whose vertex $v_i$ is weighted by the quantity $p_i(1 - \prod_{v_j \in \Gamma(v_i)}(1 - p_j))$, $1 \leq i \leq n$, by $T(E)$ the time needed for the computation of the functional $E$ and by WIS($G^w_p$) (resp., WIS($G^w_p$)), the fact that the a priori solution maximizing the functional is a maximum-weight independent set in $G^w_p$ (resp., $G^w_p$). Finally, $G$ and $B$ denote general and bipartite graphs, respectively.

In table 2, where we denote by $r$ the quantity $\min\{O(\log n/(3(\Delta_G + 1) \log \log n)), O(n^{-4/5})\}$, a summary of the main approximation results is presented. Let us note that the approximation ratio in the line for PIS5 in general graphs of table 2 is directly obtained with arguments exactly analogous to the ones of corollary 1 in section 3.2, considering $p_i(1 - \prod_{v_j \in \Gamma(v_i)}(1 - p_j))$ as vertex-weight for $v_i$, $i = 1, \ldots, n$. 
References


A Proof of proposition 1

A.1 Determining $A_{\alpha^*}(G)$

Let us first prove a small preliminary lemma.

Lemma 1. If $S \subseteq I$, then $F(I^{R^2}_S) = \alpha(G)$.

Proof. In fact, by definition of $U_2$, $F(I^{R^2}_S) = |S[I]| + |S[\bar{I}]|$. But if $S \subseteq I$, then $S[I] = S$ and this implies $S[I] \cup \Gamma(S[I]) = S \cup \Gamma(S)$. Moreover, the maximality of $S$ implies that $S \cup \Gamma(S) = V$; and consequently, $S[\bar{I}] = \emptyset$. So, $F(I^{R^2}_S) = |S[I]| = \alpha(G)$.

For $A_{\alpha^*}(G)$ we have $|S^* \cap I| = \alpha^*(G)$ and, by lemma 1, $F(I^{R^2}_S) = \alpha^*(G)$; therefore,

$$A_{\alpha^*}(G) = \alpha^*(G) \left( \sum_{|S^* \cap I| \subseteq A_{\alpha^*}(G)} \text{Pr}[I] \right) = \alpha^*(G) p^{\alpha^*(G)} \sum_{i=0}^{n-\alpha^*(G)} \binom{n-\alpha^*(G)}{i} p^i (1-p)^{n-\alpha^*(G)-i} = \alpha^*(G) p^{\alpha^*(G)}$$

where, in the above expression, the term $p^{\alpha^*(G)}$ represents the fact that the $\alpha^*(G)$ vertices of $S^*$ are all present in $I$ ($S^* \cap I = S^*$) and the term $\sum_{i=0}^{n-\alpha^*(G)} C_i^{n-\alpha^*(G)} p^i (1-p)^{n-\alpha^*(G)-i}$ stands for all possible choices for the rest of the elements of $V$ (as we have already seen, this term equals 1).

A.2 Determining $A_{\alpha^*(G)-1}$

Consider an element $v$ of $S^*$ and note that, since $S^*$ is a maximum independent set, $\Gamma'(v)$ is either empty, or a clique on at least one vertex (if not, $S^*$ could be augmented to $(S^* \setminus \{v\}) \cup \Gamma'(v)$). Consequently, by considering that $v$ is the element of $S^*$ which is absent from $I$, we have:

- if $I \cap \Gamma'(v) = \emptyset$, then the independent set $S^* \setminus \{v\}$ remains a maximal one for $G[I]$, so $F(I^{R^2}_S) = \alpha^*(G) - 1$;
- if $I \cap \Gamma'(v) \neq \emptyset$, then the independent set $S^* \setminus \{v\}$ (included in $G[I]$) can be augmented by exactly one element of $\Gamma'(v)$, so $F(I^{R^2}_S) = \alpha^*(G)$.

We now study the two following cases with respect to $\Gamma'(v)$, namely $\Gamma'(v) = \emptyset$ and $\Gamma'(v) \neq \emptyset$.

$\Gamma'(v) = \emptyset$.

$F(I^{R^2}_S) = \alpha^*(G) - 1$ and

$$\sum_{(S^* \setminus \{v\}) \subseteq I} \text{Pr}[I] F(I^{R^2}_S) = (\alpha^*(G) - 1) \sum_{(S^* \setminus \{v\}) \subseteq I} \text{Pr}[I]$$

$$\sum_{(S^* \setminus \{v\}) \subseteq I} \text{Pr}[I] = p^{\alpha^*(G)-1} (1-p) \sum_{i=0}^{n-\alpha^*(G)} \binom{n-\alpha^*(G)}{i} p^i (1-p)^{n-\alpha^*(G)-i} = p^{\alpha^*(G)-1} (1-p)$$

Consequently,

$$\sum_{(S^* \setminus \{v\}) \subseteq I} \text{Pr}[I] F(I^{R^2}_S) = (\alpha^*(G) - 1) p^{\alpha^*(G)-1} (1-p).$$

$\Gamma'(v) \neq \emptyset$.

Here, we study the following two subcases, namely, $I \cap \Gamma'(v) = \emptyset$ and $I \cap \Gamma'(v) \neq \emptyset$.

22
Subcase $I \cap \Gamma'(v) = \emptyset$:
here $F(I_{S^*}^{(2)}) = \alpha^*(G) - 1$ and, moreover, no vertex of $\Gamma'(v)$ is contained in $I$; so, we get
\[ \sum_{(S^* \setminus \{v\}) \subseteq I, \Gamma'(v) \neq \emptyset} \Pr[I] = p^{\alpha^*(G)-1}(1 - p)(1 - p)^{|\Gamma'(v)|} = p^{\alpha^*(G)-1}(1 - p)^{|\Gamma'(v)|+1}. \quad (A.1) \]

Subcase $I \cap \Gamma'(v) \neq \emptyset$:
$F(I_{S^*}^{(2)}) = \alpha^*(G)$ and $I$ contains at least a vertex of $\Gamma'(v)$; so, we get for this case:
\[ \sum_{(S^* \setminus \{v\}) \subseteq I, \Gamma'(v) \neq \emptyset} \Pr[I] = p^{\alpha^*(G)-1}(1 - p) \left[1 - (1 - p)^{|\Gamma'(v)|}\right]; \quad (A.2) \]

Consequently, for case $\Gamma'(v) \neq \emptyset$ we have combining expressions (A.1) and (A.2) together with expressions for $F(I_{S^*}^{(2)})$ of the two subcases:
\[ \sum_{(S^* \setminus \{v\}) \subseteq I} \Pr[I] F(I_{S^*}^{(2)}) = p^{\alpha^*(G)-1}(1 - p) \left[\alpha^*(G) - (1 - p)^{|\Gamma'(v)|}\right]. \]

This concludes the study of case $\Gamma'(v) \neq \emptyset$.

By summing the above expressions over all elements of $S^*$, we get:
\[
A_{\alpha^*(G)-1} = \sum_{v \in S^*} \sum_{I \subseteq V} \Pr[I] F(I_{S^*}^{(2)}) \mathbf{1}_{\{(S^* \setminus \{v\}) \subseteq I\}} \\
= \sum_{v \in S^*} \sum_{I \subseteq V} \Pr[I] F(I_{S^*}^{(2)}) \mathbf{1}_{\{(S^* \setminus \{v\}) \subseteq I\}} \left(1_{\{|\Gamma'(v)|=0\}} + 1_{\{|\Gamma'(v)| \neq 0\}}\right) \\
= \ell_1 \left(\alpha^*(G) - 1\right) p^{\alpha^*(G)-1}(1 - p) + \sum_{v \in S^*} p^{\alpha^*(G)-1}(1 - p) \left(\alpha^*(G) - (1 - p)^{|\Gamma'(v)|}\right) \\
= \ell_1 \left(\alpha^*(G) - 1\right) p^{\alpha^*(G)-1}(1 - p) + \left(\alpha^*(G) - \ell_1\right) p^{\alpha^*(G)-1}(1 - p)\alpha^*(G) \\
- p^{\alpha^*(G)-1}(1 - p) \sum_{\Gamma'(v) \neq \emptyset} (1 - p)^{|\Gamma'(v)|} \\
= p^{\alpha^*(G)-1}(1 - p) \left(\ell_1 \alpha^*(G) - \ell_1 + \alpha^*(G)^2 - \ell_1 \alpha^*(G) - \sum_{\Gamma'(v) \neq \emptyset} (1 - p)^{|\Gamma'(v)|}\right) \\
= p^{\alpha^*(G)-1}(1 - p) \left(-\ell_1 + \alpha^*(G)(\ell_1 + \ell_1') - \sum_{\Gamma'(v) \neq \emptyset} (1 - p)^{|\Gamma'(v)|}\right). 
\]