Additional material for NIPS submission. Technical details and extended proofs

Abstract

In this additional material we develop the proofs that appear in the paper into further detail. We keep the structure of the paper and the numbers of the propositions so that the reader can easily access the extended proofs.

1 Introduction

2 Overview

2.1 The max-sum algorithm in Pairwise Markov Random Fields

2.2 Neighborhood maximum characterisation of max-sum fixed points

3 Generalizing size and distance optimal bounds

3.1 $\mathcal{C}$-optimal bounds

Although we consider that the proof in the main paper of proposition 1 is clear enough, we enclose the proposition in this document for the sake of completeness, since most of the proofs make reference to it.

Proposition 1. *Let $\mathcal{G} = \langle V, E\rangle$ be a graphical model and $\mathcal{C}$ a region. If $x^{\mathcal{C}}$ is a $\mathcal{C}$-optimum then*

$$\theta(x^{\mathcal{C}}) \geq \frac{cc_*}{|\mathcal{C}| - nc_*}\theta(x^*) \tag{1}$$

where $cc_* = \min_{S\in E} cc(S,\mathcal{C})$, $nc_* = \min_{S\in E} nc(S,\mathcal{C})$, and x^* is the MAP assignment.

3.2 Size-optimal bounds as a particular case of $\mathcal{C}$-optimal

4 Quality guarantees on max-sum fixed-point assignments

4.1 $\mathcal{C}$-optimal bounds based on the SLT region

Proposition 3. *Let $\mathcal{G} = \langle V, E\rangle$ be a graphical model and $\mathcal{C}$ the SLT-region in $\mathcal{G}$. Let $\mathcal{G}' = \langle V', E'\rangle$ be a subgraph of $\mathcal{G}$. Then the bound of equation 1 for $\mathcal{G}$ holds for any SLT-optimal in $\mathcal{G}'$.*

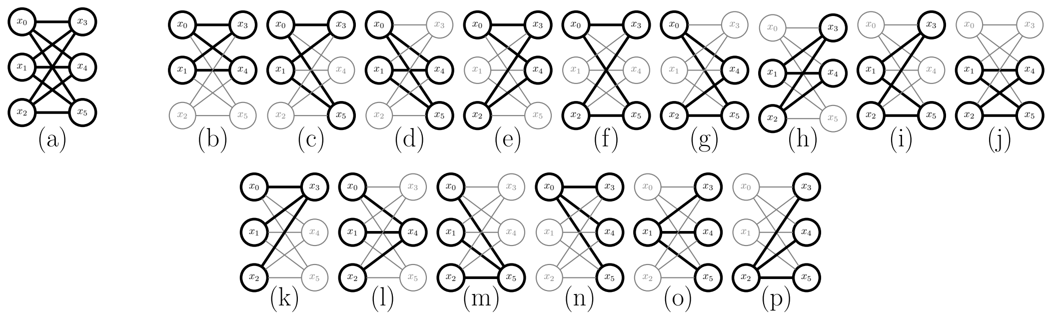

Figure 1: Example of (a) a 3-3 bipartite graph and (b)-(p) sets of variables covered by the SLT-region

Proof. We can compose a region $\mathcal{C}'$ containing the same elements as $\mathcal{C}$ just removing from each element of the region those variables that are not contained in V'. Obviously $|\mathcal{C}| = |\mathcal{C}'|$.

Since we have to deal simultaneously with two regions in the proof, we use $cc_*^{\mathcal{C}} = \min_{S \in E} cc(S, \mathcal{C})$ and $nc_*^{\mathcal{C}} = \min_{S \in E} nc(S, \mathcal{C})$ when referring to $\mathcal{C}$. Likewise, we use $cc_*^{\mathcal{C}'} = \min_{S \in E'} cc(S, \mathcal{C}')$ and $nc_*^{\mathcal{C}'} = \min_{S \in E'} nc(S, \mathcal{C}')$ for $\mathcal{C}'$

It is easy to see that $cc_*^{\mathcal{C}} = \min_{S \in E} cc(S, \mathcal{C}) \leq \min_{S \in E'} cc(S, \mathcal{C}) = \min_{S \in E'} cc(S, \mathcal{C}') = cc_*^{\mathcal{C}'}$ and that $nc_*^{\mathcal{C}} = \min_{S \in E} nc(S, \mathcal{C}) \leq \min_{S \in E'} nc(S, \mathcal{C}) = \min_{S \in E'} cc(S, \mathcal{C}') = nc_*^{\mathcal{C}'}$. Hence, the bound obtained applying equation 1 to $\mathcal{C}'$ is $\frac{cc_*^{\mathcal{C}'}}{|\mathcal{C}'| - nc_*^{\mathcal{C}'}} \geq \frac{cc_*^{\mathcal{C}}}{|\mathcal{C}'| - nc_*^{\mathcal{C}'}} = \frac{cc_*^{\mathcal{C}}}{|\mathcal{C}| - nc_*^{\mathcal{C}'}} \geq \frac{cc_*^{\mathcal{C}}}{|\mathcal{C}| - nc_*^{\mathcal{C}}}$. That is, we can obtain a bound for $\mathcal{G}'$ greater or equal than the bound obtained applying equation 1 to $\mathcal{C}$, as we wanted to prove. $\square$

4.2 SLT-bounds for particular MRF structures and independent of the MRF parameters

4.2.1 Bipartite graphs

Proposition 6. *For any MRF with a n-m bipartite structure where $m \geq n$, and for any max-sum fixed point assignment x^{MS} we have that*

$$\theta(x^{MS}) \geq b(n,m) \cdot \theta(x^*) \qquad\qquad b(n,m) = \begin{cases} \frac{1}{n} & m \geq n+3 \\ \frac{2}{n+m-2} & m < n+3 \end{cases} \tag{2}$$

Proof. Let $\mathcal{C}^A$ be a region including one out of the n variables and all of the m variables (in figure 1, elements (n)-(p)). Since the elements of this region are trees, we can guarantee optimality on them. The number of elements of the region is $|\mathcal{C}^A| = n$. It is clear that each edge in the graph is completely covered by one of the elements of $\mathcal{C}^A$, and hence $cc_* = 1$. Furthermore, every edge is partially covered, since all of the m variables are present in every element, and hence $nc_* = 0$. Applying equation 1 gives the bound $\frac{1}{n}$.

Alternatively, we can define a region $\mathcal{C}^B$ formed by taking sets of four variables, two from each side. Since the elements of $\mathcal{C}^B$ are single-cycle graphs (in figure 1, elements (b)-(j)), we can guarantee optimality on them.

Now, in order to apply proposition 1 to the region $\mathcal{C}^B$ we only need to assess $|\mathcal{C}^B|$, cc_* and nc_*. In a bipartite graph, there are $|\mathcal{C}^B| = \binom{n}{2} \cdot \binom{m}{2}$ different combinations of four vertices when taking two from each side. Hence, the number of elements of $\mathcal{C}^B$ that completely cover any $S \in E$ is assessed as $cc(S, \mathcal{C}^B) = \binom{n-1}{1}\binom{m-1}{1} = (n-1)(m-1)$, because once fixed the two variables in S we have to

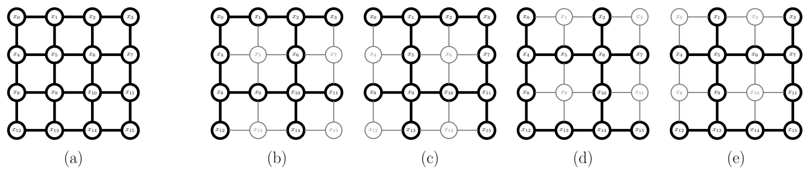

Figure 2: Example of (a) a 4-grid graph and (b)-(e) sets of variables covered by the SLT-region

select an additional variable from each side but excluding the variable in S already included. Clearly, $cc_* = (n-1)(m-1)$ as well. The number of elements of $\mathcal{C}$ that do not cover any $S \in E$ at all can be assessed as $nc(S, \mathcal{C}^B) = \binom{n-1}{2}\binom{m-1}{2}$ because we have to select two elements from each side but excluding the variables in S, which means one variable in each side. Obviously, $nc_* = \binom{n-1}{2}\binom{m-1}{2}$ as well. Applying proposition 1, we get the bound $\frac{2}{n+m-2}$. Observe that $\frac{2}{n+m-2} > \frac{1}{n}$ when $m < n+3$, and so equation 2 holds. $\square$

4.2.2 Two-dimensional (2-D) grids

Proposition 7. *For any MRF with an n grid structure where n is an even number, for any max-sum fixed point assignment x^{MS} we have that*

$$\theta(x^{MS}) \geq \frac{n}{3n-4} \cdot \theta(x^*) \tag{3}$$

Proof. Let $\Gamma(l)$ be a division of l indices into tuples of two elements, such that each index appears exactly in one tuple and any two indexes in the same tuple are non-consecutive. If l is even, the set of tuples $\Gamma(l) = \{(i, \frac{l}{2}+i) | i = 1 \ldots \frac{l}{2}\}$ generates such division. Let $\mathcal{C}$ be a region formed by taking the vertices that result from the combination of any pair of rows $(i,j) \in \Gamma(n)$ and any pair of columns $(k,l) \in \Gamma(n)$. Note that optimality is guaranteed in each $C_i \in \mathcal{C}$ because variables in two non-consecutive rows and two non-consecutive columns create a single-cycle graph. Because $\Gamma(n)$ contains $\frac{n}{2}$ tuples, $|\mathcal{C}| = \left(\frac{n}{2}\right)^2$.

In a grid, any $S \in E$ contains two variables that are either in a column or a row. Let's consider without loss of generality that S contains two variables in a row with index i (the following discussion applies for columns by exchanging rows and columns). First observe that each index i appears exactly in one tuple in the division $\Gamma(n)$ of rows, and each tuple is combined with the $n/2$ tuples of the division $\Gamma(n)$ of columns. Hence, the number of elements of $\mathcal{C}$ that completely cover any $S \in E$ when S is either a row or a column is $cc(S, \mathcal{C}) = n/2$. Clearly, $cc_* = n/2$ as well. Secondly, we assess the number of sets in $\mathcal{C}$ that do not cover S at all. For any $S \in E$, such that the two variables in S are in a row, each variable in S appears in a single column. Notice that the two columns in which each variable in S appears can not be in the same tuple in a division $\Gamma(n)$ because they are consecutive. Thus, there are $|\Gamma(n)| - 1$ rows in a division that does not contain any variable in S and $|\Gamma(n)| - 2$ columns in a division that do not contain any variable in S. Hence, $nc(S, \mathcal{C}) = (\frac{n}{2}-1)(\frac{n}{2}-2)$ if S contains two variables in a row. Following a similar reasoning, $nc(S, \mathcal{C}) = (\frac{n}{2}-2)(\frac{n}{2}-1)$ if S contains two variables in a column. Obviously, $nc_* = (\frac{n}{2}-2)(\frac{m}{2}-1)$. Using these values in equation 1 provides equation 3. $\square$

4.2.3 MRFs which are a union of variable-disjoint cycles

Proposition 8. *For any MRF such that every pair of cycles is variable-disjoint and where there are at most d cycles of size l or larger, and for any max-sum fixed point assignment x^{MS} we have that*

$$\theta(x^{MS}) \geq \left(1 - \frac{2(d-1)}{d \cdot l}\right) \cdot \theta(x^*) = \frac{(l-2) \cdot d + 2}{l \cdot d} \cdot \theta(x^*). \tag{4}$$

Proof. Let $T = \{t_1, \cdots, t_l\}$ be a set of variables such that its vertex induced subraph is a cycle of $l \geq 3$ variables in our MRF. It is evident that by removing any variable from T we break the cycle. Hence, we can define $T_{-k} = T \setminus \{t_k\} = \{t_1, \ldots, t_{k-1}, t_{k+1}, \ldots, t_l\}$, and its induced subgraph is not a cycle. Let U be the variables in the MRF that are not involved in any cycle.

By hypothesis, our graph contains at most d cycles of size $l \leq 4$. We can name them $T^1, \ldots, T^d$. It is clear that if we construct a set including: every variable that is not in any cycle (that is, U), a single cycle (say T^i) and for every remaining cycle T^i all the variables in it but one (say $T^j_{-k_j}$), its induced graph will have SLT-optimality guaranteed. In general for each choice of cycle (i) and each choice of variable to remove in the remaining cycles (k_j) we can create a set

$$C^i_{k_1,\ldots,k_{i-1},k_{i+1},\ldots,k_d} = U \cup T^i \cup \bigcup_{\substack{j \neq i \\ 1 \leq j \leq d}} T^j_{-k_j}.$$

Now, we can define the following region, including each of these sets

$$\mathcal{C} = \{C^i_{k_1,\ldots,k_{i-1},k_{i+1},\ldots,k_d} \mid \forall i \in \{1, \ldots, d\} \;\; \forall j \in \{1, \ldots, i-1, i+1, \ldots, d\} \;\; \forall k_j \in \{1, \ldots, l\}\}$$

The total number of elements in the region is $|\mathcal{C}| = d \cdot l^{d-1}$, since we have to select one out of d cycles and for each of the remaining $d-1$ cycles we have to remove one out of the l variables.

To compute cc_* we have to split our edges into four different types: (i) the ones that have both variables in U (S^1), (ii) the ones that have one variable in a cycle and the other one in U (S^2), (iii) the ones where each variable belongs to a different cycle (S^3), and finally (iv) the ones with both variables in the same cycle (S^4). We will assess cc_* as the minimum of these four values.

Since the variables in U appear in every element of the region, for each S^1 vertex with both variables in U we have that $cc(S^1, \mathcal{C}) = d \cdot l^{d-1}$.

Assume S^2 is an edge with one variable in U and another in one of the cycles (we can fix T^1 without loss of generality). Then S^2 appears in every element of the region that contains T^1 (we have l^{d-1}). Furthermore, for the elements of the region that include T^i completely (with $i > 1$), we can select one out of the $l-1$ variables in T^1 which are not in S^2 and for each of the $d-2$ cycles different from T^1 and T^i we should select one out of the l variables to remove. So, for each i we have $(l-1) \cdot l^{d-2}$ and since we have $d-1$ possible selections for i, we can conclude that $cc(S^2, \mathcal{C}) = l^{d-1} + (d-1)(l^{d-2}(l-1))$.

Assume S^3 is an edge with one variable in a cycle (say T^1) and another variable in a different cycle (say T^2). Including T^1 into our set, we have $(l-1)$ choices for variables to remove from T^2 and l choices for variables to remove from the remaining $d-2$ cycles, for a total of $(l-1)l^{d-2}$. The same happens if we include T^2. If we include one out of the $d-2$ remaining cycles, we have $l-1$ choices for T^1 and $l-1$ choices for T^2 and l choices for each of the remaining $d-3$ cycles. Hence, we have that $cc(S^3, \mathcal{C}) = (l-1)l^{d-2} + (l-1)l^{d-2} + (d-2)(l-1)^2 l^{d-3}$.

Finally, assume S^4 is an edge with both variables in the same cycle (say T^1). Then S^4 appears in every element of the region that contains T^1 (we have l^{d-1}). Furthermore, for the elements of the region that include T^i completely with $i > 1$, we can select one out of the $l-2$ variables in T^1 which are not in S^4, and for each of the $d-2$ cycles different from T^1 and T^i we should select one out of the l variables to remove. So, for each i we have $(l-2) \cdot l^{d-2}$, and since we have $d-1$ possible selections for i, we can conclude that $cc(S^1, \mathcal{C}) = l^{d-1} + (d-1)(l^{d-2}(l-2))$.

It is easy to see that $cc(S^1, \mathcal{C}) \geq cc(S^2, \mathcal{C}) \geq cc(S^3, \mathcal{C}) \geq cc(S^4, \mathcal{C})$, and hence $cc_* = cc(S^4, \mathcal{C}) = l^{d-1} + (d-1)(l^{d-2}(l-2))$.

Since every edge in U is completely covered by every element of the region, we have that $nc_* = 0$.

Using these values into equation 1 and operating provides the desired result. □

4.3 SLT-bounds for arbitrary MRF structures and independent of the MRF parameters

5 Conclusions