

1. About Definition 1: In the definition of proper, what is the meaning of the independence of set  $M_j$ . Why is there a concept of a graph? isn't it a set of discrete variables? My understanding is that  $n$  discrete variables  $1, \dots, n$  are combined to form a 2-tuple with  $n(n-1)$  pairs. This  $n(n-1)$  pair has  $m$  proper covers if the 2-tuple in the same set cannot have the same elements. For example  $(1, 2)$  and  $(10, 23)$  can be in one set. But  $(1, 2)$  and  $(1, 12)$  cannot be in one set. Is that right?

Response:  $M_j$  is the independence set such that no two vertices in  $M_j$  are adjacent. Your understanding of independence of set  $M_j$  is correct in the sense that each edge in the graph has at most one end point in  $S$ . So, example you have given is correct. But Since, we are using [1], we do need a concept of a graph because we wanted to mix the concept of chromatic number and proper fractional cover with laplace transform so that hoeffding inequality can be applied. Since the chromatic number and proper cover can be applied only if there is a concept of a graph, hence there is a concept of a graph.

2. About Definition 1: In  $C = (M_j, c_j)$ , what is the meaning of  $w_j$ ? Can it be considered as a subset weight of  $M_j$ ? Is it necessary to define this variable, why not 1? What is the intuitive meaning of the definition of exact fractional cover?

Response: Yes, it is to be considered as the subset weight of  $M_j$ . Yes, it is necessary to define this variable in order to formalize the definition of fractional chromatic number. Intuition behind exact fractional cover is the following: essentially assume you have  $J$  labels, each having a weight. You can put the same label only on non-adjacent vertices. ie. vertices that you put a label on are independent. What is required is that you choose the weights of these labels, and the vertices on which you place them, such that for each vertex, the sum of weights of all its labels exactly equals 1. its integral version would be easier to understand, essentially partition the vertices into  $J$  sets such that each set is an Independent set.

3. About Theorem 1: The expression of  $C(s)$  (The first line in page 4, a equation), what is the meaning of the expectation symbol  $E_{M_j}$ ? What is the expectation under a set  $M_j$ ? The latter summation is to find all the self-measures  $d(z, z)$  belonging to  $M_j$ . Where is the randomness in the summation? Is there a more intuitive explanation about  $C(S)$ ?

Response: The meaning of the expectation symbol  $E_{M_j}$  is mean of dissimilarities of preferred and non-preferred items for that  $M_j$ . We don't understand why we need randomness in the summation. We attach the proof of the theorem with this email so that it is clear how we reach up to this term.  $\mathfrak{C}(S)$  intuitively means the complexity term involving the dissimilarities between the mapping function  $\phi$  of preferred and non-preferred items. In the cases where mapping function has finite range  $k$ , lower values of  $k$  mean lower complexity term  $\mathfrak{C}(S)$  and tighter generalization bound. This observation is also empha-

sized in experimental section when RecNet performs better on lower values of  $k$ .

4. About Theorem 1:  $C(s)$  in the generalization error bound shows metrics of the binary groups in each cover are relatively small, and there is no tag information. The second term of the loss function  $L_p(U, V, S)$ , except for the norm of  $U, V$  is constrained, why the similarity between  $U$  and  $V$  is consistent with the label? Is this proposed isolated or reflected in Theorem 1?

Response: I haven't been able to understand this question. Particularly, I don't understand what is a binary group we have talked about and what does he mean by **The second term of the loss function  $L_p(U, V, S)$ , except for the norm of  $U, V$  is constrained**

## References

- [1] S. Janson. Large Deviations for Sums of Partly Dependent Random Variables. *Random Structures and Algorithms*, 24(3):234–248, 2004.