

# Fractional Edge Cover Number of Model RB

Tian Liu<sup>(✉)</sup>

Key Laboratory of High Confidence Software Technologies, Ministry of Education,  
Peking University, Beijing 100871, China  
lt@pku.edu.cn

**Abstract.** Model RB is a random constraint satisfaction problem with a growing domain size, which exhibits exact phase transition phenomena. Many hard instances with planted solutions can be generated via Model RB, to be used as benchmarks for algorithmic competitions and researches. In the past, some structural parameters of constraint hypergraphs are analyzed to show hardness of Model RB, such as hinge width, decycling number, treewidth, and hypertree width. In this paper, one more structural parameter of constraint hypergraphs of Model RB, namely the fractional edge cover number, is analyzed. We show upper and lower bounds on the fractional edge cover number of Model RB. In particular, the fractional edge cover number of Model RB is shown to be asymptotically linear in the number of variables, like hinge width, decycling number, treewidth and hypertree width. These results together provide further evidences on the hardness of Model RB.

**Keywords:** Model RB · Fractional edge cover · Hardness

## 1 Introduction

Constraint satisfaction problems (CSPs) can model many real world problems, such as  $n$ -queens, Latin squares, etc. A CSP instance consists of a constraint hypergraph on a set of variables and many constraints on the hyperedges. Each constraint gives the compatible assignments of values to the variables in a hyperedge. The task is to decide if the instance is satisfiable, that is, if there is an assignment of values to all the variables, such that it is compatible with all the constraints. The Boolean satisfaction problem 3-SAT is a special case of CSPs, where each variable only takes two different values and each constraint is a conjunction of three variables or their negations.

Since 3-SAT is already an  $\mathcal{NP}$ -complete problem, CSPs are hard to solve in general. Many structural decomposition methods are developed to find tractable classes of CSPs, such as tree decomposition [13], hypertree decomposition [1], and fractional edge cover [11]. For constraint hypergraphs with a structural parameter  $w$  and input size  $\|I\|$ , usually we can solve these instances in time

---

Partially supported by Natural Science Foundation of China (Grant Nos. 61370052 and 61370156).

$O(|I|^{f(w)})$ , where  $f(w)$  is a function of  $w$ , such as a low degree polynomial of  $w$ , or linear in  $w$ . When the structural parameter  $w$  is constantly bounded, we get a tractable class of CSPs. At the moment, the most powerful structural decomposition method is fractional hypertree decomposition [17].

On the other hand, random instances of CSPs can be generated by randomly setting constraint hypergraphs, and then randomly setting compatible assignments for each constraint. A parameter called density, which is the ratio between the number of constraints and the number of variables, can be used to control the number of constraints. With a small density and thus a small number of constraints, the instances are likely to be satisfiable. With a large density and thus a large number of constraints, the instances are unlikely to be satisfiable. When the number of variables goes to infinity, such a change of satisfiability may happen suddenly around a critical value of density. Such phenomena are called the satisfiability phase transition of random CSPs. Moreover, the hardest instances of CSPs are located around the satisfiability thresholds [2, 3, 5, 6, 19, 20]. However, a rigorous link between the hardness of random instances and the satisfiability phase transition is still unknown.

The most common random CSPs are random 3-SAT in theoretical computer science and Model A,B,C,D in artificial intelligence. For random 3-SAT, the exact satisfiability threshold is still unknown, although some upper and lower bounds are shown in the past [4, 18]. Moreover, if planted solutions are used to generate satisfiable random instances with known solutions, the instances usually become much easier to solve. For Model A,B,C,D, when the number of variables goes to infinity, the satisfiability thresholds will go to zero, thus they are useless in generating large hard instances [10]. A random CSP model, called Model RB, is defined by Xu and Li [23], which has an increasing domain size and exhibits exact phase transition phenomena. Many hard instances with planted solutions can be generated via Model RB [24, 26], to be successfully used as benchmarks in various algorithmic competitions and in many research papers, such as the annual CSP solver competitions, the annual Pseudo-Boolean (0-1 Integer Programming) solver competitions, the annual MAX-SAT solver competitions, and the annual SAT solver competitions, etc. [24].

In the past, the hardness of model RB is theoretically shown by exponential lower bounds on the length of resolution [25], and some analysis on the evolution of its solution space [27, 28], besides many experimental results [26]. Some structural parameters of constraint hypergraphs are also analyzed to show hardness of Model RB with respect to the structural decomposition methods, such as hinge width [15], decycling number [12], treewidth [22], and hypertree width [16]. In this paper, one more structural parameter of constraint hypergraphs of Model RB, namely the fractional edge cover number [11], is analyzed. We show upper and lower bounds on the fractional edge cover number of Model RB. In particular, the fractional edge cover number of Model RB is shown to be asymptotically linear in the number of variables, like hinge width, decycling number, treewidth and hypertree width. These results together provide further evidences on the hardness of Model RB.

This paper is structured as follows. After introducing definitions and some facts on Model RB and fractional edge cover respectively, as well as a version of Chernoff bound in Sect. 2, lower and upper bounds on fractional edge cover number of Model RB are shown in Sect. 3. In Sect. 4, we conclude the paper with remarks on our results and open problems.

## 2 Preliminaries

In this section, we give definitions and facts on Model RB and fractional edge cover respectively, as well as a version of the Chernoff bound.

### 2.1 Model RB

An instance  $I$  of *constraint satisfaction problem* (CSP) is a triple  $(V, D, C)$ .  $V$  is a finite set of variables.  $D$  is a finite set of values, called domain.  $C$  is a set of constraints. For each constraint, there is a subset of variables, called the scope of this constraint. For a constraint scope with  $k$  variables, a subset of  $D^k$  is given as compatible assignment of values to the variables in this scope. A solution of  $I$  is an assignment of values to all variables which is compatible to all constraint. If there is at least one solution,  $I$  is called satisfiable, otherwise unsatisfiable. Given an instance, we are asked to decide if it is satisfiable, and in some cases to find a solution if it exists.

A hypergraph is just a set system, which is consist of some subsets (called hyperedges) of a finite set (called vertex set). A hypergraph  $H$  with vertex set  $V$  and hyperedge set  $E$  is denoted by  $H = (V, E)$ . The constraint hypergraph of a CSP is consist of the scopes of the constraints in this CSP.

Model RB is a random CSP defined as follows [23].

1. Given  $n$  variables, each variable takes values in  $\{1, 2, \dots, d\}$ , where  $d = n^\alpha$  and  $\alpha > 0$  is a constant;
2. Select with repetition  $m = rn \ln n$  random constraints, where  $r$  is a constant. For each constraint, select without repetition  $k$  of  $n$  variables, where  $k \geq 2$  is an integer constant;
3. Select uniformly at random without repetition  $(1-p)d^k$  compatible assignments for each constraint, where  $0 < p < 1$  is a constant.

For an instance  $I$  of Model RB, the input size  $||I||$  is about  $O(m(k \log n + D^k))$ , that is, for each of the  $m$  constraints, we list the  $k$  variables involved and at most  $D^k$  compatible assignments.

It is known that in Model RB there exists satisfiability thresholds  $r_{cr} = -\frac{\alpha}{\ln(1-p)}$  [23].

Let  $H_{n,r,k}^{RB}$  denote a random constraint hypergraph in Model RB.

### 2.2 Fractional Edge Cover

Given a hypergraph  $H = (V, E)$ , if there is a mapping

$$\psi : E \rightarrow [0, \infty),$$

such that

$$\sum_{e \in E, v \in e} \psi(e) \geq 1, \text{ for every } v \in V,$$

then  $\psi$  is called a *fractional edge cover* of  $H$  [11].

For a fractional edge cover  $\psi$ , The weight of hyperedge  $e$  under  $\psi$  is  $\psi(e)$ . The weight of  $\psi$  is  $\sum_{e \in E} \psi(e)$ . The optimal fractional edge cover  $\psi^*$  of  $H$  is a fractional edge cover with the minimum weight over all possible fractional edge covers of  $H$ . The *fractional edge cover number* of  $H$ , denoted by  $\rho^*(H)$ , is the weight of an optimal fractional edge cover  $\psi^*$  of  $H$  [11], that is,

$$\rho^*(H) = \min_{\psi} \sum_{e \in E} \psi(e) = \sum_{e \in E} \psi^*(e).$$

It is known that for a CSP instance  $I$ , the number of solutions of  $I$  is at most  $\|I\|^{\rho^*(H_I)}$ , where  $\|I\|$  is the size of  $I$  and  $H_I$  is the constraint hypergraph of  $I$ . It is known that the solutions of  $I$  can be enumerated in time  $\|I\|^{\rho^*(H_I)+O(1)}$  [11].

### 2.3 Chernoff Bound

We say that a random event  $Q$  happens *with high probability* if the probability of this event  $\Pr(Q)$  goes to 1 asymptotically. We will use the following version of Chernoff Bound [21].

**Lemma 1 (Chernoff Bound).** *Given a random variable  $X$ ,  $X$  follows a binomial distribution, i.e.,  $X \sim B(n, \frac{\mu}{n})$ . If  $0 < \epsilon < 1$ , then*

$$\Pr(X \leq (1 - \epsilon)\mu) \leq e^{-\mu\epsilon^2/2}.$$

## 3 Fractional Edge Cover Number of Model RB

In this section, we will give lower and upper bounds on fractional edge cover number of Model RB.

Suppose that  $\rho^*(H_{n,r,k}^{RB})$  is the fractional edge cover number of Model RB, where  $n$  is the total number of vertices in the constraint hypergraph,  $k$  is the number of vertices contained in each hyperedge, and  $rn \ln n$  is the maximum number of hyperedges. We can first give a lower bounds on  $H_{n,r,k}^{RB}$  as follows.

**Theorem 1.**  $\rho^*(H_{n,r,k}^{RB}) \geq \frac{n}{k}$ .

*Proof.* Let  $\psi$  be an arbitrary fractional edge cover of  $H_{n,r,k}^{RB}$ . Then by definition of fractional edge cover,

$$\sum_{e \in E, v \in e} \psi(e) \geq 1, \text{ for all } v \in V.$$

Now we summarize all these  $n$  inequalities over all vertices  $v \in V$ ,

$$\sum_{v \in V} \left( \sum_{e \in E, v \in e} \psi(e) \right) \geq n.$$

Since every hyperedge contains exactly  $k$  vertices, the weight  $\psi(e)$  of every hyperedge  $e$  will appear exactly  $k$  times in the left hand side. Thus,

$$\sum_{v \in V} \left( \sum_{e \in E, v \in e} \psi(e) \right) = k \cdot \sum_{e \in E} \psi(e).$$

From the above two inequalities, we have

$$k \cdot \sum_{e \in E} \psi(e) \geq n,$$

or equivalently,

$$\sum_{e \in E} \psi(e) \geq \frac{n}{k}.$$

Since  $\psi$  is an arbitrary fractional edge cover, the last inequality also holds for the optimal fractional edge cover  $\psi^*$ . Therefore,

$$\rho^*(H_{n,r,k}^{RB}) = \sum_{e \in E} \psi^*(e) \geq \frac{n}{k}.$$

We have finished the proof of this theorem. □

After we get a lower bound on  $\rho^*(H_{n,r,k}^{RB})$  by a counting argument as above, we will give a matching upper bound on  $\rho^*(H_{n,r,k}^{RB})$ , by an explicit construction of a fractional edge cover for  $H_{n,r,k}^{RB}$ . To this end, we need a lower bound on the minimum degree of vertices in  $H_{n,r,k}^{RB}$  by Chernoff bound as follows.

Suppose that the degree of a vertex  $v$  is denoted by  $deg(v)$ , and the minimum degree of  $H_{n,r,k}^{RB}$  is denoted by  $\delta(H_{n,r,k}^{RB})$ .

**Lemma 2.**  $\delta(H_{n,r,k}^{RB}) \geq (1 - \frac{2}{\sqrt{kr}}) \cdot kr \ln n$  with high probability.

*Proof.* Let  $v$  be an arbitrary vertex in  $H_{n,r,k}^{RB}$ . By definition of  $H_{n,r,k}^{RB}$ , we repeat for  $rn \ln n$  times to randomly select hyperedges, and each time independently select  $k$  different vertices from all  $n$  vertices to form an hyperedge. For an arbitrary hyperedge  $e$ , the vertex  $v$  is contained in  $e$  with probability  $\frac{k}{n}$ .

For  $i = 1, \dots, rn \ln n$ , let  $X_i$  be a random variable with probability  $\frac{k}{n}$  to be 1, and with probability  $1 - \frac{k}{n}$  to be 0, respectively. All these  $X_i$ 's are independent and identically distributed 0–1 variables. Then,

$$\deg(v) = \sum_{i=1}^{rn \ln n} X_i.$$

The random variable  $\deg(v)$  has a binomial distribution  $B(rn \ln n, \frac{k}{n})$ . The expectation  $\mu$  of  $\deg(v)$  is

$$\mu = (rn \ln n) \frac{k}{n} = kr \ln n.$$

By the Chernoff bound, for any  $0 < \delta < 1$ ,

$$\Pr(\deg(v) \leq (1 - \delta) \cdot kr \ln n) \leq e^{-(kr \ln n)\delta^2/2}.$$

Let  $\delta = \frac{2}{\sqrt{kr}}$ . If  $kr \leq 4$ ,  $(1 - \delta) \cdot kr \ln n \leq 0$ , this will lead to a trivial case. Otherwise for  $kr > 4$ , we have  $0 < \delta < 1$ . Then

$$\Pr\left(\deg(v) \leq \left(1 - \frac{2}{\sqrt{kr}}\right) \cdot kr \ln n\right) \leq e^{-2 \ln n} = \frac{1}{n^2}.$$

By the Union bound,

$$\Pr\left(\exists v, \deg(v) \leq \left(1 - \frac{2}{\sqrt{kr}}\right) \cdot kr \ln n\right) \leq n \cdot \frac{1}{n^2} = \frac{1}{n}.$$

Thus,

$$\lim_{n \rightarrow \infty} \Pr\left(\delta(H_{n,r,k}^{RB}) \leq \left(1 - \frac{2}{\sqrt{kr}}\right) \cdot kr \ln n\right) = 0.$$

We have finished the proof of this lemma.  $\square$

Once we have a lower bound  $\delta_L$  on the minimum degree of variables in Model RB, we can construct a fractional edge cover of Model RB by putting weight  $\frac{1}{\delta_L}$  on each hyperedge. Since each variable is contained in at least  $\delta_L$  hyperedges, the sum of weight of hyperedges containing this variable is at least 1, which makes sure that it is a fractional edge cover. In this way, we can get a matching upper bound of  $\rho^*(H_{n,r,k}^{RB})$  as follows.

**Theorem 2.**  $\rho^*(H_{n,r,k}^{RB}) \leq \frac{n}{k} \cdot \left(1 - \frac{2}{\sqrt{kr}}\right)^{-1}$ , with high probability.

*Proof.* For a random constraint hypergraph  $H_{n,r,k}^{RB} = (V, E)$  of Model RB, we define a mapping  $\psi_0 : E \rightarrow [0, \infty)$  as follows.

$$\psi_0(e) = \left(1 - \frac{2}{\sqrt{kr}}\right)^{-1} \cdot \frac{1}{kr \ln n}, \text{ for all } e \in E.$$

Recall that with high probability,

$$\delta(H_{n,r,k}^{RB}) \leq \left(1 - \frac{2}{\sqrt{kr}}\right) \cdot kr \ln n.$$

Thus with high probability, for each variable  $v$ ,

$$\sum_{e \in E, v \in e} \psi_0(e) = \deg(v) \cdot \left(1 - \frac{2}{\sqrt{kr}}\right)^{-1} \cdot \frac{1}{kr \ln n} \geq 1.$$

Therefore,  $\psi_0$  is a fractional edge cover of  $H_{n,r,k}^{RB}$  with high probability. The weight of  $\psi_0$  is

$$\sum_{e \in E} \psi_0(e) \leq (rn \ln n) \cdot \left(1 - \frac{2}{\sqrt{kr}}\right)^{-1} \cdot \frac{1}{kr \ln n} = \frac{n}{k} \cdot \left(1 - \frac{2}{\sqrt{kr}}\right)^{-1}.$$

Thus fractional edge cover number of  $H_{n,r,k}^{RB}$  is no larger than  $\frac{n}{k} \cdot \left(1 - \frac{2}{\sqrt{kr}}\right)^{-1}$  with high probability. We have finished the proof.  $\square$

Note that the upper bound is only within a constant ratio  $\left(1 - \frac{2}{\sqrt{kr}}\right)^{-1}$  to the lower bound. For fixed  $r$ , the more larger  $k$  is, the more tighter upper bounds we get.

## 4 Conclusions

In this paper, we show linear lower and upper bounds on fractional edge cover number of Model RB. Since the structural decomposition method based on fractional edge cover runs in time exponential in fractional edge cover number, these results provide evidence for hardness of Model RB.

The fractional edge cover number and hypertree width are incomparable in the sense that, there are hypergraphs with bounded hypertree width and unbounded fractional edge cover number, and vice versa [11]. The most powerful structural parameter is fractional hypertree width, which supersedes both fractional edge cover number and hypertree width [17]. To show lower and upper bounds on fractional hypertree width for Model RB is an open problem.

Among these structural parameters, perhaps tree width is the best understood on the classical random graphs. It is known that there is a threshold where the tree width suddenly jumps from constant to linear in the number of vertices [7–9, 12–14]. Whether there are similar phenomena for the fractional edge cover number and the fractional hypertree width on classical random graphs is also unknown.

**Acknowledgments.** We thank Ms. Yu Song for drafting an earlier version of this paper.

## References

1. Adler, I., Gottlob, G., Grohe, M.: Hypertree width and related hypergraph invariants. *Eur. J. Comb.* **28**, 2167–2181 (2007)
2. Cheeseman, P., Kanefsky, R., Taylor, W.: Where the really hard problems are. In: *Proceedings of IJCAI 1991*, pp. 163–169 (1991)
3. Cook, S.A., Mitchell, D.G.: Finding hard instances of the satisfiability problem: a survey. *DIMACS Ser.* **35**, 1–17 (1997)
4. Dubois, O., Boufkhad, Y., Mandler, J.: Typical random 3-SAT formulae and the satisfiability threshold. In: *Proceedings of SODA 2000*, pp. 126–127 (2000)
5. Fan, Y., Shen, J.: On the phase transitions of random  $k$ -constraint satisfaction problems. *Artif. Intell.* **175**, 914–927 (2011)
6. Fan, Y., Shen, J., Xu, K.: A general model and thresholds for random constraint satisfaction problems. *Artif. Intell.* **193**, 1–17 (2012)
7. Gao, Y.: Phase transition of tractability in constraint satisfaction and Bayesian network inference. In: *Proceedings of UAI*, pp. 265–271 (2003)
8. Gao, Y.: On the threshold of having a linear treewidth in random graphs. In: Chen, D.Z., Lee, D.T. (eds.) *COCOON 2006*. LNCS, vol. 4112, pp. 226–234. Springer, Heidelberg (2006)
9. Gao, Y.: Treewidth of Erdos-Renyi random graphs, random intersection graphs, and scale-free random graphs. *Discrete Appl. Math.* **160**(4–5), 566–578 (2012)
10. Gao, Y., Culberson, J.: Consistency and random constraint satisfaction problems. *J. Artif. Intell. Res.* **28**, 517–557 (2007)
11. Grohe, M., Marx, D.: Constraint solving via fractional edge covers. *ACM Trans. Alg.* **11**(1), Article 4 (2014)
12. Jiang, W., Liu, T., Ren, T., Xu, K.: Two hardness results on feedback vertex sets. In: Atallah, M., Li, X.-Y., Zhu, B. (eds.) *FAW-AAIM 2011*. LNCS, vol. 6681, pp. 233–243. Springer, Heidelberg (2011)
13. Kloks, T.: *Treewidth: Computations and Approximations*, pp. 18–55. Springer, Berlin (1994)
14. Lee, C., Lee, J., Oum, S.: Rank-width of random graphs. *J. Graph. Theor.* **70**(3), 339–347 (2012)
15. Liu, T., Lin, X., Wang, C., Su, K., Xu, K.: Large hinge width on sparse random hypergraphs. In: *Proceedings of IJCAI*, pp. 611–616 (2011)
16. Liu, T., Wang, C., Xu, K.: Large hypertree width for sparse random hypergraphs. *J. Comb. Optim.* **29**, 531–540 (2015)
17. Marx, D.: Tractable hypergraph properties for constraint satisfaction and conjunctive queries. *J. ACM* **60**(6), 42 (2013)
18. Mezard, M., Parisi, G., Zecchina, R.: Analytic and algorithmic solution of random satisfiability problems. *Science* **297**(5582), 812–815 (2002)
19. Mitchell, D.G., Selman, B., Levesque, H.J.: Hard and easy distributions of sat problems. In: *Proceedings of AAAI 1992*, pp. 459–465 (1992)
20. Selman, B., Mitchell, D.G., Levesque, H.J.: Generating hard satisfiability problems. *Artif. Intell.* **81**, 17–29 (1996)
21. Rucinski, A., Janson, S., Luczak, T.: *Random Graphs*. Wiley, New York (2000)
22. Wang, C., Liu, T., Cui, P., Xu, K.: A note on treewidth in random graphs. In: Wang, W., Zhu, X., Du, D.-Z. (eds.) *COCOA 2011*. LNCS, vol. 6831, pp. 491–499. Springer, Heidelberg (2011)
23. Xu, K., Li, W.: Exact phase transitions in random constraint satisfaction problems. *J. Artif. Intell. Res.* **12**, 93–103 (2000)

24. Xu, K.: BHOSLIB: Benchmarks with Hidden Optimum Solutions for Graph Problems. <http://www.nlsde.buaa.edu.cn/kexu/benchmarks/graph-benchmarks.htm>
25. Xu, K., Li, W.: Many hard examples in exact phase transitions. *Theor. Comput. Sci.* **355**, 291–302 (2006)
26. Xu, K., Boussemart, F., Hemery, F., Lecoutre, C.: Random constraint satisfaction: easy generation of hard (satisfiable) instances. *Artif. Intell.* **171**, 514–534 (2007)
27. Xu, W., Zhang, P., Liu, T., Gong, F.: The solution space structure of random constraint satisfaction problems with growing domains. *J. Stat. Mech. Theor. Exp.* **2015**, P12006 (2015)
28. Zhao, C., Zhang, P., Zheng, Z., Xu, K.: Analytical and belief-propagation studies of random constraint satisfaction problems with growing domains. *Phys. Rev. E* **85**, 016106 (2012)