

On Counting t -Cliques Mod 2

Oded Goldreich*

July 12, 2020

Abstract

For a constant $t \in \mathbb{N}$, we consider the problem of counting the number of t -cliques *mod 2* in a given graph. We show that this problem is not easier than determining whether a given graph contains a t -clique, and present a simple worst-case to average-case reduction for it. The reduction runs in linear time when graphs are presented by their adjacency matrices, and average-case is with respect to the uniform distribution over graphs with a given number of vertices.

1 Informal description

For a constant integer $t \geq 3$, finding t -cliques in graphs and determining their mere existence are archetypical computational problems within the frameworks of parameterized complexity and fine grained complexity (see, e.g., [FG] and [W], resp.). The complexity of counting the number of t -cliques has also been studied (see, e.g., [GR, BBB]). In this work, we consider a variant of the latter problem; specifically, the problem of counting the number of t -cliques mod 2.

Determining the number of t -cliques *mod 2* in a given graph is potentially easier than determining the number of t -cliques in the same graph. On the other hand, as shown in Theorem 1, determining the said number mod 2 is not easier (in the worst-case sense) than determining whether or not a graph contains a t -clique. Hence, the worst-case complexity of *counting t -cliques mod 2* lies between the worst-case complexity of *counting t -cliques* and the worst-case complexity of *determining the existence of t -cliques*. Consequently, as far as worst-case complexity is concerned, using the “counting mod 2 problem” as proxy for the “existence problem” is at least as justified as using the “counting problem” as such a proxy.

Our main result (presented in Theorem 2) is an efficient worst-case to average-case reduction for *counting t -cliques mod 2*. The reduction is efficient in the sense that it runs in linear time when graphs are presented by their adjacency matrices. Average-case is with respect to the uniform distribution over graphs with a given number of vertices, and it yields the correct answer (with high probability) whenever the average-case solver is correct on at least a $1 - 2^{-t^2}$ fraction of the instances. In other words, the average-case solver has error rate at most 2^{-t^2} . The question of whether the same result holds with respect to significantly higher error rates, and ultimately with error rate 0.49, is left open.

*Department of Computer Science, Weizmann Institute of Science, Rehovot, ISRAEL. E-mail: oded.goldreich@weizmann.ac.il

Relation and comparison to prior work. Efficient worst-case to average-case reductions were presented before for the related problem of *counting t -cliques* (over the integers). Specifically, Goldreich and Rothblum provided such a reduction with respect to a relatively simple distribution over graphs with a given number of vertices, alas not the uniform distribution [GR]. On the other hand, their reduction works even when the average-case solver has error rate that approaches 1; specifically, its error rate on n -vertex graphs may be as large as $1 - \frac{1}{\text{poly}(\log n)} = 1 - o(1)$. In contrast, Boix-Adsera, Brennan, and Bresler provided an efficient worst-case to average-case reduction with respect to the uniform distribution, but their reduction can only tolerate a vanishing error rate [BBB]; specifically, its error rate on n -vertex graphs is required to be $1/\text{poly}(\log n) = o(1)$.

Hence, our worst-case to average-case reduction, which is for a related (but different) problem, matches the better aspects of the prior works (see Table 1): It refers to the uniform distribution (as [BBB]), and tolerates a constant error rate (which is better than [BBB] but worse than [GR]).

| problem | distribution | error rate | where |
|----------------|-------------------|--|-------|
| counting | relatively simple | $1 - 1/\text{poly}(\log n) = 1 - o(1)$ | [GR] |
| counting | uniform | $1/\text{poly}(\log n) = o(1)$ | [BBB] |
| counting mod 2 | uniform | $\exp(-t^2) = \Omega(1)$ | here |

Table 1: Comparison of different worst-case to average-case reductions for variants of the t -CLIQUE problem, for the constant t , where n denotes the number of vertices. The first column indicates the version being treated, the second indicates the distribution for which average-case is considered, and the third indicates the error rate allowed for the average-case solver.

Techniques. In contrast to [GR, BBB], which relate the t -clique counting problem to the evaluation of lower degree polynomials over large and medium sized fields, we related the counting *mod 2* problem to low degree polynomials over $\text{GF}(2)$. This relation allows us to present reductions that are much simpler than those presented in [GR, BBB].

As noted above, we leave open the problem of improving the error rate that can be tolerated by a worst-case to average-case reduction (for counting t -cliques mod 2). We note that tolerating an error rate that approaches 0.5 presupposes that approximately half of the n -vertex graphs have an odd number of t -cliques (unless finding t -cliques can be done in $\tilde{O}(n^2)$ -time). We were able to prove this combinatorial conjecture only for the case of $t = 3$ (see Proposition 4).

2 Formal statements and proofs

For a fixed integer $t \geq 3$ and a graph G , we denote by $\text{CC}^{(t)}(G)$ the number of t -cliques in G , and let $\text{CC}_2^{(t)}(G) \stackrel{\text{def}}{=} (\text{CC}^{(t)}(G) \bmod 2)$ denote the parity of this number. We often represent n -vertex graphs by their adjacency matrices; hence, $\text{CC}_2^{(t)}(A) = \text{CC}_2^{(t)}(G)$, where A is the adjacency matrix of G , and it follows that

$$\text{CC}_2^{(t)}(A) = \sum_{i_1 < \dots < i_t \in [n]} \prod_{j < k \in [t]} A_{i_j, i_k} \bmod 2, \quad (1)$$

where $A_{u,v}$ is the (u, v) th entry of A (indicating whether or not $\{u, v\}$ is an edge in G).

Theorem 1 (deciding the existence of t -cliques reduces to computing $\text{CC}_2^{(t)}$): *For every integer $t \geq 3$, there is a randomized reduction of determining whether a given n -vertex graph contains a t -clique to computing $\text{CC}_2^{(t)}$ on n -vertex graphs such that the reduction runs in time $O(n^2)$, makes $\exp(t^2)$ queries, and has error probability at most $1/3$.*

Proof: Consider a randomized reduction that, on input $G = ([n], E)$, flips each edge to a non-edge with probability 0.5, leaves non-edges intact, and returns the value of $\text{CC}_2^{(t)}$ on the resulting graph; that is, the reduction generates a random subgraph of G , denoted G' , and returns $\text{CC}_2^{(t)}(G')$.

To analyze the output of this procedure (on input G), consider a (symmetric) n -by- n matrix X such that $x_{i,j}$ is a variable if $\{i, j\} \in E$ and $x_{i,j} = 0$ otherwise. We view $\text{CC}_2^{(t)}(X)$, which is defined as in Eq. (1), as a multivariate polynomial over $\text{GF}(2)$, and observe that it has degree at most $\binom{t}{2}$. The key observation is that $\text{CC}_2^{(t)}(X)$ is a non-zero polynomial if and only if the graph G contains a t -clique (i.e., $\text{CC}^{(t)}(G) > 0$). Hence, the foregoing reduction can be viewed as returning the value of $\text{CC}_2^{(t)}(X)$ on a random (symmetric) assignment to the variables in X . It follows that the reduction always returns 0 if $\text{CC}^{(t)}(G) = 0$, and returns 1 with probability at least $2^{-\binom{t}{2}}$ otherwise (i.e., when $\text{CC}^{(t)}(G) > 0$). The latter assertion is due to the Schwartz-Zippel for small fields (i.e., for $\text{GF}(2)$).¹ Applying the foregoing reduction for $\exp(t^2)$ times, the claim follows. ■

Theorem 2 (worst-case to average-case reduction for $\text{CC}_2^{(t)}$): *For every integer $t \geq 3$, there is a randomized reduction of computing $\text{CC}_2^{(t)}$ on the worst-case n -vertex graph to correctly computing $\text{CC}_2^{(t)}$ on at least a $1 - \exp(-t^2)$ fraction of the n -vertex graphs such that the reduction runs in time $O(n^2)$, makes $\exp(t^2)$ queries, and has error probability at most $1/3$.*

Proof: Setting $d = \binom{t}{2}$, consider the following random self-reduction of $\text{CC}_2^{(t)}$. On input a symmetric and non-reflective n -by- n matrix, A :

1. Select uniformly d random (symmetric and non-reflective) n -by- n matrices, denoted $R^{(1)}, \dots, R^{(d)}$, and let $R^{(0)} = A$.
2. Making adequate queries to $\text{CC}_2^{(t)}$, return $\sum_{I \subseteq \{0,1,\dots,d\}: I \neq \{0\}} \text{CC}_2^{(t)}(R^{(I)}) \bmod 2$, where $R^{(I)} \stackrel{\text{def}}{=} \sum_{i \in I} R^{(i)} \bmod 2$ and $\text{CC}_2^{(t)}(R^{(\emptyset)}) = 0$.

Hence, the foregoing reduction performs $2^{d+1} - 2$ queries, and each of these queries (i.e., each $R^{(I)}$ for $I \notin \{\emptyset, \{0\}\}$) is uniformly distributed over the set of all symmetric and non-reflective n -by- n matrices.

We claim that, for any fixed $R^{(0)}, R^{(1)}, \dots, R^{(d)}$, it holds that $\sum_{I \subseteq \{0,1,\dots,d\}: I \neq \{0\}} \text{CC}_2^{(t)}(R^{(I)})$ equals $\text{CC}_2^{(t)}(R^{(0)}) \bmod 2$. This claim is proved by considering the multivariate polynomial $P(x_0, x_1, \dots, x_d)$ over $\text{GF}(2)$ that is defined to equal $\text{CC}_2^{(t)}(\sum_{i=0}^d x_i R^{(i)})$. Specifically, we use the following facts:

- $P(b_0, b_1, \dots, b_d) = \text{CC}_2^{(t)}(R^{\{\{i:b_i=1\}\}})$; in particular, $P(0, 0, \dots, 0) = 0$ and $P(1, 0, \dots, 0) = \text{CC}_2^{(t)}(R^{(0)})$.
 - P has degree $\binom{t}{2} = d$, since $P(x_0, x_1, \dots, x_d) = \text{CC}_2^{(t)}(L(x_0, x_1, \dots, x_d))$ such that $L(x_0, \dots, x_d)$ is a matrix of linear functions (i.e., the (u, v) th entry of $L(x_0, \dots, x_d)$ equals $\sum_{i=0}^d R_{u,v}^{(i)} x_i$).
- (Indeed, using Eq. (1), it follows that $P = \text{CC}_2^{(t)}(L)$ has degree $\binom{t}{2}$.)

¹See [G, Exer. 5.1].

- for any $(d+1)$ -variate polynomial of degree at most d over $\text{GF}(2)$ it holds that the sum of its evaluation over all 2^{d+1} points is 0.

This general fact can be seen by considering an arbitrary monomial $M(x_0, x_1, \dots, x_d) = \prod_{i \in I} x_i$, where $I \subset \{0, 1, \dots, d\}$. Indeed,

$$\begin{aligned} \sum_{(b_0, b_1, \dots, b_d) \in \text{GF}(2)^{d+1}} M(b_0, b_1, \dots, b_d) &= \sum_{(b_0, b_1, \dots, b_d) \in \text{GF}(2)^{d+1}} \prod_{i \in I} b_i \\ &= 2^{d+1-|I|} \cdot \prod_{i \in I} \sum_{b_i \in \text{GF}(2)} b_i \end{aligned}$$

which equals 0 (mod 2), since $|I| \leq d$.

Combining the foregoing facts, it follows that $\sum_{I \subseteq \{0, 1, \dots, d\}: I \neq \{0\}} \text{CC}_2^{(t)}(R^{(I)})$ equals $\text{CC}_2^{(t)}(R_0)$ (mod 2).

Thus, given oracle access to a program Π such that $\Pr_R[\Pi(R) = \text{CC}_2^{(t)}(R)] \geq 1 - \epsilon$, when making queries to Π rather than to $\text{CC}_2^{(t)}$, the foregoing reduction returns the correct value with probability at least $1 - (2^{d+1} - 2) \cdot \epsilon$ (i.e., whenever all queries are answered correctly). Using $\epsilon = 2^{-t^2}$, we obtain a worst-case to average-case reduction that fails with probability less than $2^{d+1-t^2} = 2^{-(t^2+t-2)/2} < 1/3$ when given access to a procedure that is correct on at least a $1 - 2^{-t^2}$ fraction of the instances.² ■

Remark 3 (the distribution of $\text{CC}_2^{(t)}(R)$ for random R): *The proof of Theorem 2 implies that $2^{-t^2} < \Pr_R[\text{CC}_2^{(t)}(R) = 1] < 1 - 2^{-t^2}$. To see this, using notation as in the proof, suppose towards the contradiction that $\Pr_R[\text{CC}_2^{(t)}(R) = b] \geq 1 - 2^{-t^2}$ for some b . Then, for every R_0 , it holds that*

$$\begin{aligned} &\Pr_{R_1, \dots, R_d} \left[\sum_{I \subseteq \{0, 1, \dots, d\}: I \neq \{0\}} \text{CC}_2^{(t)}(R^{(I)}) \equiv 0 \pmod{2} \right] \\ &\geq \Pr_{R_1, \dots, R_d} \left[(\forall I \subseteq \{0, 1, \dots, d\} \setminus \{\{0\}, \emptyset\}) \text{CC}_2^{(t)}(R^{(I)}) = b \right] \\ &\geq 1 - (2^{d+1} - 2) \cdot 2^{-t^2} > 0 \end{aligned}$$

where the last inequality uses $2^{d+1-t^2} = 2^{-(t^2+t-2)/2} < 1$. But this is impossible when $\text{CC}_2^{(t)}(R_0) = 1$ (e.g., if $\text{CC}_2^{(t)}(R_0) = 1$).

While Remark 3 only asserts that $\mathbb{E}_R[\text{CC}_2^{(t)}(R)]$ is bounded away from both 0 and 1, we conjecture that it is approximately 1/2. We prove this conjecture in the case of $t = 3$.

Proposition 4 (the distribution of $\text{CC}_2^{(3)}(R)$ for random R): *Let R be the adjacency matrix of a random n -vertex graph. Then, $\Pr_R[\text{CC}_2^{(3)}(R) = 1] = 0.5 \pm o(1)$.*

Proof: We consider a two-step process of determining a random graph $G = ([n], E)$. Towards this end, we designate a collection C of $n/4$ disjoint pairs in $[n]$; that is, the collection C covers a set U

²Indeed, we can slightly improve the bound by using any constant $\epsilon < 2^{-d-2} = 2^{-(t^2-t+4)/2}$.

of size $n/2$ (i.e., $U = \{u \in P : P \in C\}$). For each $P \in C$ and $v \in [n] \setminus P$, let $W_{P,v} = \{\{v, u\} : u \in P\}$ (standing for wedge), and note that the $(n/4) \cdot (n-2)$ potential triangles (i.e., $P \cup \{v\}$) are distinct.

The process of selecting a random graph proceeds in two steps. In the first step, for each pair in $\binom{[n]}{2} \setminus C$, we decide at random whether or not to include it as an edge in the graph $G = ([n], E)$; that is, each such pair is included with probability $1/2$ (independent of all others). For $P = \{u_1, u_2\} \in C$ and $v \in [n] \setminus P$, we say that $W_{P,v}$ is *good* if both its pairs (i.e., $\{v, u_1\}$ and $\{v, u_2\}$) were included (i.e., $W_{P,v} \subseteq E$), which happens with probability $1/4$.

We say that $P \in C$ is *good* if an odd number of $W_{P,v}$'s are good, which happens with probability approximately 0.5. Furthermore, by virtue of the pairs $\{u, v\}$ such that $u \in U$ and $v \in [n] \setminus U$, these $n/4$ events are approximately independent (see details below). Hence, letting $C' \subseteq C$ denote the collection of good pairs (i.e., $P \in C$ such that P is good), it holds that, with very high probability, $|C'| \approx |C|/2$.

(To see that the $n/4$ goodness events are independent, consider splitting the first step into two sub-steps such that the pairs in $\binom{U}{2} \setminus C$ are determined in the first sub-step, and the pairs in $U \times ([n] \setminus U)$ are determined in the second sub-step. Fixing the choices made in the first sub-step, we can analyze the size of C' based only on the choices made in the second sub-step, where the choices for different $P \in C$ are disjoint.³ Hence, for any fixing of the first sub-stage, each $P \in C$ is good with probability approximately $1/2$, independently of the goodness of all other pairs in C .)

In the second step, for each pair $P \in C$, we decide at random whether or not to include this pair as an edge. We stress that each decision is independent of all prior choices. Intuitively, assuming that $C' \neq \emptyset$, we observe that the number of triangles that contain $P \in C'$ is odd with probability $1/2$, and show that this implies that the total number of triangles in the graph is odd with probability $1/2$.

In the analysis we fix any choice for the first step (i.e., for pairs $\binom{[n]}{2} \setminus C$) and also fix the decisions for all pairs in $C \setminus C'$. (Hence, we actually split the second step into two sub-steps, the first determining pairs in $C \setminus C'$, and the second determining pairs in C' .) For each $P \in C'$, if the pair P was included as an edge, then it forms a triangle with each v such that $W_{P,v}$ is good, which means that the parity of the number of triangles in the graph is flipped (since the number of good $W_{P,v}$'s is odd for a good P). We stress that the choice made for P only affects potential triangles that include P , and these triangles are exactly the ones counted here (i.e., corresponding to vertices v such that $W_{P,v}$ is good, where the number of such vertices is odd). Furthermore, each triangle in the graph may contain at most one pair $P \in C$. Hence, the foregoing choices are independent. Thus, assuming $|C'| > 0$, the parity of the number of triangles in the resulting graph is distributed as the parity of the number of flips in the current step, which is odd with probability 0.5. The claim follows. ■

Open Problem 5 (stronger worst-case to average-case reduction for $\text{CC}_2^{(t)}$): *For every integer $t \geq 3$ and $\gamma > 0.5$, is there a randomized reduction of computing $\text{CC}_2^{(t)}$ on the worst-case n -vertex graph to correctly computing $\text{CC}_2^{(t)}$ on at least a γ fraction of the n -vertex graphs such that the reduction runs in time $\tilde{O}(n^2)$, and has error probability at most $1/3$.*

This strengthens Theorem 2 by requiring the reduction to tolerate error rate that is arbitrary close to 0.5 rather than error rate $\exp(-t^2)$. The foregoing Proposition 4 may be viewed as a sanity

³In contrast, for $\{u_1, u_2\}, \{u'_1, u'_2\} \in C$, the choice regarding $\{u_1, u'_1\}$ affects the goodness of both pairs.

check for Problem 5, since $|E_R[\mathbb{CC}_2^{(t)}(R)] - 0.5| > \delta$ would have implied that $\mathbb{CC}_2^{(t)}$ can be computed correctly with probability $0.5 + \delta$ in constant time. Additional support is provided by the following.

Proposition 6 (the distribution of $\mathbb{CC}_2^{(t)}(R)$ for random R , revisited): *Let R be the adjacency matrix of a random n -vertex graph. Then, for every $t \geq 3$ and $b \in \{0, 1\}$, it holds that $\Pr_R[\mathbb{CC}_2^{(t)}(R) = b] > 2^{-\lfloor t/2 \rfloor} - o(1)$.*

(This result is not optimal; a more careful analysis of the case of $t = 4$ yields $\Pr_R[\mathbb{CC}_2^{(t)}(R) = b] = 0.5 \pm o(1)$, which yields improvements for all even $t > 4$.)⁴

Proof: We establish the claim by induction on $t \geq 3$, while recalling that the claim holds for $t \in \{2, 3\}$ (by triviality and Proposition 4, respectively). For $t > 3$, we consider a two-step process of determining a random graph $G = ([n], E)$. Towards this end, we designate two vertices $u, v \in [n]$, and let $V = [n] \setminus \{u, v\}$. In the first step we determine at random the adjacency relation between all pairs except for $\{u, v\}$; that is, each pair is placed in E with probability $1/2$. Denoting by $W \stackrel{\text{def}}{=} \{w \in V : \{u, w\}, \{v, w\} \in E\}$ the set of vertices that are adjacent to both u and v , note that $|W| \approx n/4$ with very high probability. In the second step, we add $\{u, v\}$ to E with probability $1/2$.

Considering the situation after the first step, let G_W denote the subgraph of G induced by W , and note that G_W is a random graph over the vertex set W . Hence, by the induction hypothesis, the number of $(t-2)$ -cliques in G_W is odd (i.e., $\mathbb{CC}_2^{(t-2)}(G_W) = 1$) with probability at least $2^{-\lfloor (t-2)/2 \rfloor} - o(1)$. On the other hand, whenever $\mathbb{CC}_2^{(t-2)}(G_W) = 1$, in the second step, the parity of the number of t -cliques in G flips according to whether or not $\{u, v\}$ is included in E . This is the case because t -cliques that contain the pair $\{u, v\}$ appear if and only if $\{u, v\} \in E$, whereas each such t -clique corresponds to a distinct $(t-2)$ -clique in G_W . Hence, $\Pr[\mathbb{CC}_2^{(t)}(G) = b] > (2^{-\lfloor (t-2)/2 \rfloor} - o(1))/2$ for both $b \in \{0, 1\}$. ■

3 Conclusion

Theorem 2 asserts an efficient worst-case to average-case reduction for *counting t -cliques mod 2*, where average-case is with respect to the uniform distribution over graphs with the given number of vertices. Specifically, for any integer $t \geq 3$, computing $\mathbb{CC}_2^{(t)}$ on the worst-case n -vertex graph is reducible (in $O(n^2)$ -time) to computing $\mathbb{CC}_2^{(t)}$ correctly on a $1 - \exp(-t^2)$ fraction of all n -vertex graphs.

We believe that Theorem 2, which has a very simple proof, is as interesting as an analogous result that refers to counting t -cliques (i.e., computing $\mathbb{CC}^{(t)}$), because (as shown in Theorem 1) computing $\mathbb{CC}_2^{(t)}$ is not easier than determining whether a given graph contains a t -clique. The point is that the decisional problem (i.e., t -CLIQUE) is the one that has received most attention in prior work, and results regarding either $\mathbb{CC}^{(t)}$ or $\mathbb{CC}_2^{(t)}$ are mostly proxies for it (i.e., for results

⁴In continuation to the proof of Proposition 6, we use $\ell \in [\omega(1), o(\log n)]$ designated pairs (rather than one), and identify a set I of $\Theta(\log \ell)$ pairs that are not connected by the edges determined among all pairs. For each such pair $\{u_i, v_i\}$ we identify a corresponding W_i , and consider the parity of the number of edges in the subgraph induced by $W'_i = W_i \setminus \bigcup_{j \in I \setminus \{i\}} W_j$. With very high probability, there exists $i \in I$ such that the number of edges in the subgraph induced by W'_i is odd. The key observation is that the edges in W_i are “isomorphic” to 4-cliques that contain $\{u_i, v_i\}$, whereas each of these edges is either internal to W'_i or is internal to $W_i \setminus W'_i$. (In contrast, when $t > 4$, it does not hold that each $(t-2)$ -clique in W_i is either internal to W'_i or has no edge inside W'_i .)

regarding t -CLIQUE). In particular, combining Theorems 1 and 2, it follows that deciding t -CLIQUE on the worst-case n -vertex graph is reducible (in $O(n^2)$ -time) to computing $\text{CC}_2^{(t)}$ correctly on a $1 - \exp(-t^2)$ fraction of all n -vertex graphs.

We note that prior works fall short of establishing results analogous to Theorem 2: The results of [GR] are not for the uniform distribution (but rather for a relatively simple but different distribution), where the results of [BBB] hold for a notion of average-case that allows only a vanishing error rate (i.e., the “average-case algorithm” is required to be correct on at least a $1 - \frac{1}{\text{poly}(\log n)}$ fraction of the n -vertex graphs).

As stated in Problem 5, we leave open the problem of obtaining a result analogous to Theorem 2 for “average-case algorithms” that are correct on a γ fraction of the instances, for every $\gamma > 1/2$.

Acknowledgements

I am grateful to Dana Ron and to Guy Rothblum for useful discussions.

References

- [BBB] Enric Boix-Adsera, Matthew Brennan, and Guy Bresler. The Average-Case Complexity of Counting Cliques in Erdos-Renyi Hypergraphs. In *60th FOCS*, 2019.
- [FG] Jorg Flum and Martin Grohe. *Parameterized Complexity Theory*. Texts in Theoretical Computer Science. An EATCS Series, Springer, 2006.
- [G] Oded Goldreich. *Introduction to Property Testing*. Cambridge University Press, 2017.
- [GR] Oded Goldreich and Guy Rothblum. Counting t -Cliques: Worst-Case to Average-Case Reductions and Direct Interactive Proof Systems. In *59th FOCS*, 2018.
- [W] Virginia Vassilevska Williams. Hardness of Easy Problems: Basing Hardness on Popular Conjectures such as the Strong Exponential Time Hypothesis. In *10th Int. Sym. on Parameterized and Exact Computation*, pages 17–29, 2015.