On Regularity of Max-CSPs and Min-CSPs

Aleksa Stanković *

KTH Royal Institute of Technology
aleksas@kth.se

August 18, 2021

Abstract

We study the approximability of regular constraint satisfaction problems, i.e., CSPs where each variable in an instance has the same number of occurrences. In particular, we show that for any CSP Λ, the existence of an α approximation algorithm for unweighted regular Max-CSP Λ implies the existence of an $\alpha - o(1)$ approximation algorithm for weighted Max-CSP Λ for which the regularity of instances is not imposed. We also give an analogous result for Min-CSPs, and therefore show that up to an arbitrarily small error it is sufficient to conduct the study of the approximability of CSPs only on regular unweighted instances.

1 Introduction

This article studies the approximability of regular constraint satisfaction problems (CSPs), where we interpret regularity to mean that each variable appears the same number of times in constraints of instances. Since regular CSPs are a subclass of CSPs, approximating their optimal values is not harder than approximating values of general CSPs. Conversely, in this article we show that approximating values of regular CSPs is also *essentially* not easier, i.e., we show that an α approximation algorithm for regular instances of a particular CSP induces $\alpha - o(1)$ approximation algorithm applicable to possibly non-regular instances. Therefore, we show that imposing regularity has *almost* no effect on the approximability of CSPs, and in particular if one is willing to ignore $o(1)$ additive factors in approximation ratios, the study of approximability may be conducted solely on regular instances.

In order to make the result more general, we revisit the previously studied question of *weights vs. no weights* for CSPs [10, 16] in the context of approximation. In particular, we show that it is sufficient to have an α approximation algorithm for regular *unweighted* instances in order to construct an $\alpha - o(1)$ approximation algorithm applicable to possibly *weighted* instances of CSPs without any regularity restriction. In order to do so, we use a result from [16] which shows that weighted versions of CSPs have *essentially* the same (up to $o(1)$ additive error) approximation ratios as their unweighted counterparts. We reprove this result here for the sake of completeness.

1.1 Constraint Satisfaction Problems

Constraint satisfaction problems (CSPs) represent one of the most fundamental classes of problems studied in complexity theory. Each CSP is described by a collection of predicates, which are used in instances of these problems as constraints on tuples of variables. Probably the best known CSP is 3-Sat, in which the constraints are given as disjunctive clauses on at most three literals, where a literal is either a variable or its negation. A basic problem is to determine whether we can satisfy all the constraints of a given CSP instance simultaneously.

*Research supported by the Approximability and Proof Complexity project funded by the Knut and Alice Wallenberg Foundation.

This problem is well understood, due to Schaefer's dichotomy theorem for CSPs on Boolean domains [21] and more recent proofs of a dichotomy theorem on general domains by Bulatov [8] and Zhuk [23].

In this work we focus on optimization variants of CSPs, in which we are interested in finding an assignment which maximizes/minimizes the number of constraints satisfied. Depending on the optimization variant, we refer to these problems as either Max-CSPs or Min-CSPs. A typical problem in this setting is Max-Cut, which has Boolean constraints of the form $x_i \neq x_j$. Many optimization CSPs are intractable, and in that case we typically resort to approximation algorithms in order to estimate their optimal values. The strength of an approximation algorithm is expressed through its approximation ratio α, which measures the quality of a solution produced by the algorithm by comparing[1] it to the optimal one. In studies of approximation algorithms, we are typically interested in finding algorithms with the value of α as close to 1 as possible. We are also interested in studying which values of α are not feasible, in which case we talk about inapproximability.

1.2 Some Important Results on Approximability of CSPs

On the algorithmic side, semidefinite programming (SDP) has been very fruitful tool for approximating optimal values of CSPs. The first approximation algorithm based on SDP dates back to the work of Goemans and Williamson, who devised a ≈ 0.878 approximation algorithm for the Max-Cut problem [11]. Ideas from this work have been very influential for subsequent research of approximation algorithms, and we highlight the 7/8 approximation algorithm for Max-3-Sat [15] and the 0.940 algorithm for Max-2-Sat [5, 18].

On the hardness of approximation side, the celebrated PCP-theorem [3, 4], combined with the usual assumption that P $\neq$ NP, provided a strong starting point used in many results showing impossibility of approximation. The highlight result using this starting point along with parallel repetition of Raz [20] and long codes [7] comes from Håstad, who gave optimal inapproximability results for Max-Ek-Sat and Max-Ek-Lin[2] problems [12]. Recently, Siu On Chan gave optimal (up to a constant factor) inapproximability results for Max-CSPs where the arity k of the predicates is larger than the size of the domain [9].

Even though the PCP theorem was used with great success over the years, researchers still faced seemingly insurmountable difficulties in pursuit of sharp inapproximability results for many fundamental problems such as Max-2-Sat, approximate graph coloring, and minimum vertex cover. More precisely, the starting point of almost all reductions was the Label Cover problem [1], which was constructed by combining the PCP theorem with the parallel repetition of Raz [20]. In order to overcome these difficulties, Khot introduced a modification of Label Cover called *Unique Label Cover* [17] and conjectured it to be NP-Hard. This conjecture is known as the Unique Games Conjecture (UGC), and it quickly became the central problem in the hardness of approximation area, especially since its validity implies optimality of many already known approximation algorithms. Of special importance among UGC-based results is the one from Raghavendra [19], which shows that a certain version of semidefinite programming relaxation is optimal for all constraint satisfaction problems. Therefore, in case UGC is shown to be true, this work would end the quest for optimal approximation algorithms for CSPs.

However, with the validity of the UGC still in question, there is an incentive to derive strong inapproximability results relying on other (weaker) assumptions, most preferably on P $\neq$ NP. Furthermore, while Raghavendra's result shows how to optimally approximate CSPs, it does not give us a suitable way to compute numerical values of optimal approximation ratios; this question remains open for almost all CSPs, even very basic ones.

[1] By convention we assume in this work that approximation algorithms for Max-CSPs always have $\alpha < 1$, while for Min-CSPs $\alpha > 1$.

[2] In Max-Ek-Sat, constraints are clauses of width k, while in Max-Ek-Lin, constraints are linear equations over $\mathbb{Z}_2$. We use abbreviation Ek to denote that each constraint is of width exactly k. Therefore, Max-Ek-Sat allows only clauses of width 3, while Max-3-Sat allows width 1 and 2 as well.

1.3 New Results for Regular CSPs

In order to facilitate further study, it can be valuable to ask whether some additional properties of instances can be assumed when studying the approximability of CSPs. In this work we address this topic by studying regular instances of Max-CSPs and Min-CSPs, i.e. instances in which each variable occurs the same number of times in the constraints. In particular, we prove the following results for poly-time approximation algorithms.

Theorem 1. *If there is an α approximation algorithm for unweighted regular instances of Max-CSP Λ then for every constant $\delta > 0$ there is an $\alpha - \delta$ approximation algorithm for the weighted Max-CSP Λ.*

Theorem 2. *If there is an α approximation algorithm for unweighted regular instances of Min-CSP Λ then for every constant $\delta > 0$ there is an $\alpha + \delta$ approximation algorithm for the weighted Min-CSP Λ.*

The proofs of Theorems 1 and 2 are based on a deterministic reduction introduced in Theorem 14. We also give a randomized reduction for Max-CSPs in order to prove the following theorem.

Theorem 3. *Let us consider a Max-CSP Λ and let $\delta > 0$ be a constant. Then, it is sufficient to have an α approximation algorithm for unweighted regular instances of Max-CSP Λ of degree up to $O(\log(1/\delta)/\delta^2)$ in order to have an $\alpha - \delta$ randomized approximation algorithm for the weighted Max-CSP Λ with success probability of at least $1 - O(2^{-n})$, where n is the number of variables appearing in an instance.*

The details of the randomized reduction can be found in Theorem 12. Randomized reduction also works for Min-CSPs, although with the degree requirement of $\Omega(n^2 \log(n))$, which makes this reduction less efficient than even the deterministic one. For this reason we do not discuss randomized reduction for Min-CSPs.

In Theorem 1 and Theorem 2 instead of a constant δ we can choose $\delta = \Omega(1/poly(n))$, where n is the number of variables, to obtain $\alpha - o(1)$ approximation in poly-time for Max-CSPs (or $\alpha + o(1)$ approximation for Min-CSPs). We also remark that the techniques developed in this work do not allow us to choose δ asymptotically smaller than $1/poly(n)$ since this would make the reductions run in superpolynomial time, as could be inferred by studying the proofs.

1.4 Prior Work

Both the randomized and the deterministic reductions introduced in this paper are based on a construction introduced by Trevisan [22], which was used to show hardness of approximating values of bounded degree instances of the Max-3-Sat problem. The reduction of Trevisan outputs instances in which each variable has degree D in expectation, and therefore, by an argument that relies on Chernoff's bound it is shown that the degree of any variable is with high probability smaller than De^2. Our deterministic reduction comes from derandomization of the aforementioned result, while in the randomized reduction we reuse the mentioned approach of Trevisan [22] and in our argument show that the degrees are with high probability in range $[D - o(1), D + o(1)]$.

In order to make our reductions applicable to the weighted setting, in this work we also show that the approximability of weighted Max-CSPs (or weighted Min-CSPs) is *essentially the same*[3] as the approximability of their unweighted versions. Let us remark that the same result was already proved in [16, Lemma 3.11] by relying on some results that appeared in [10]. We reprove this fact here for the sake of completeness.

2 Preliminaries

We consider constraint satisfaction problems given by the following definition.

[3] If we allow $o(1)$ additive loss in the approximation ratio

Definition 4. A **constraint satisfaction problem** (CSP) over a language $\Sigma = [q], q \in \mathbb{N}$, is a finite collection of predicates $\Lambda \subseteq \{P : [q]^k \to \{0,1\} \mid k \in \mathbb{N}\}$.

For a predicate $P : [q]^k \to \{0,1\}$ we use $\mathrm{ar}(P) = k$ to denote its arity. We are interested in solving instances of CSPs, which are defined as follows.

Definition 5. An **instance** $\mathcal{F}$ of a CSP Λ consists of a set $X = \{x_1, \ldots, x_n\}$ of n variables taking values in Σ and a set $\mathcal{C} = \{C_1, \ldots, C_m\}$ of m constraints, where each constraint C_r is a pair (P_r, S_r), with $P_r \in \Lambda$ being a predicate with arity $k_r := \mathrm{ar}(P_r)$, and S_r being an ordered tuple containing k_r *distinct* variables which we call a **scope**.

In case $\Sigma = [2]$ we call the CSP Boolean. For Boolean CSPs a different definition of an instance is commonly used in the literature, i.e., the definition in which the scopes consist of literals instead of just variables. However, our Definition 5 is more general than the latter, i.e., for every Boolean CSP Λ there is a (Boolean) CSP $\overline{\Lambda}$ such that each instance of Λ with the scopes consisting of literals, can be expressed as an instance of $\overline{\Lambda}$ where the scopes consist only of variables, and vice-versa.

In particular, given $\Lambda = \{P_r\}_r$ we construct a CSP $\overline{\Lambda}$ by taking every P_r of Λ, considering all $I \in \{0,1\}^{\mathrm{ar}(P_r)}$, and adding to $\overline{\Lambda}$ the predicates P_r^I defined as

$$P_r^I(x_1, \ldots, x_n) = P_r(x_1 + I_1, x_2 + I_2, \ldots, x_{\mathrm{ar}(P_r)} + I_{\mathrm{ar}(P_r)}), \tag{1}$$

where I_i is the i-th element of the tuple I, and addition takes place over $\mathbb{Z}_2$. Then each constraint $P_r(x_1 + I_1, x_2 + I_2, \ldots, x_{\mathrm{ar}(P_r)} + I_{\mathrm{ar}(P_r)})$ of Λ can be converted to a constraint $P_r^I(x_1, \ldots, x_n)$ of $\overline{\Lambda}$ which consists only of variables and has the same truth-value, and vice-versa. Hence, throughout the remainder of this work we will consider CSP instances over general domains in the sense of Definition 5, since these encompass Boolean CSPs and also different notions of Boolean CSP instances appearing in the literature.

The degree d_i of a variable x_i is defined as the number of times x_i is mentioned in the constraints, or formally

$$d_i = \sum_{r=1}^{m} \mathbf{1}_{S_r}(x_i), \tag{2}$$

where $\mathbf{1}_{S_r}$ denoted the indicator function. Instances in which all variables have the same degree are called **regular**.

Max/Min-CSP problems frequently appear in a setting in which the constraints of an instance are assigned with non-negative weights, which are typically used to encapsulate the significance of each constraint. Let us now give the definition of these problems.

Definition 6. A **weighted instance** $\mathcal{F}$ of a CSP Λ is an instance of Λ, where each constraint C_r has a weight $w_r \geq 0$, and $\sum_r w_r = 1$.

Obviously, unweighted instances can be seen as weighted where each constraint C_r has a weight $w_r = 1/m$.

Let us denote by a function $\chi \colon X \to \Sigma$ an assignment to variables X of an instance $\mathcal{F}$ of some CSP Λ. We interpret $\chi(S_i)$ as a coordinate-wise action of χ on S_i. Given χ, we define the value $\mathrm{Val}_\chi(\mathcal{F})$ of χ as

$$\mathrm{Val}_\chi(\mathcal{F}) = \sum_{r=1}^{m} w_r P_r(\chi(S_r)). \tag{3}$$

We also define the optimal value of $\mathcal{F}$ in the case of Max-CSP to be

$$\mathrm{Opt}(\mathcal{F}) = \max_\chi \left(\mathrm{Val}_\chi(\mathcal{F})\right). \tag{4}$$

In the minimization version, the correct definition of the optimal value has "min" instead of "max" in the previous expression. Typically, the aim is to find a solution with the value close to the optimal one. In case of a Max-CSP, an α approximation algorithm is an algorithm which in polynomial time finds an assignment χ such that

$$\mathrm{Val}_{\chi}(\mathcal{F}) \geq \alpha \cdot \mathrm{Opt}(\mathcal{F}). \tag{5}$$

For Min-CSPs, the correct definition has "$\leq$" instead of "$\geq$" in the previous inequality.

While introducing weights allows convenient representation of CSPs, the hardness of approximation *essentially* does not change, as shown in [16, Lemma 3.11]. For the sake of completeness we reprove this fact in the appendix. In particular, we prove the following two theorems.

Theorem 7. *Consider a Max-CSP Λ and assume that we can approximate the optimal value of unweighted instances within a multiplicative factor α. Then, for every constant $\delta > 0$, weighted instances of Max-CSP Λ can be approximated within a constant $\alpha - \delta$.*

Theorem 8. *Consider a Min-CSP Λ, and assume that we can approximate the optimal value of unweighted instances within a multiplicative factor α. Then, for every constant $\delta > 0$, weighted instances of the Min-CSP Λ can be approximated within a constant $\alpha + \delta$.*

In this work we will use concentration inequalities which bound the probability that a certain random variable deviates from its mean. While these bounds are widely known, the form in which they appear can vary, and therefore we fix below the versions which are used in this paper.

We use the following variant of Chernoff's inequality.

Lemma 9. *Let $X = \sum_{i=1}^{K} X_i$, where $\{X_i\}_{i=1}^{K}$ are mutually independent random variables with range $\{0, 1\}$. Then*

$$\mathrm{P}\left[|X - \mathbf{E}[X]| \geq \delta\, \mathbf{E}[X]\right] \leq 2e^{-\mathbf{E}[X]\min(\delta/2, \delta^2/4)}, \quad \forall \delta > 0. \tag{6}$$

Proof of this lemma can be found in [2, Corollary A.15]. We also need a concentration bound for the sum of random variables with range $[0, b]$, where $b \in \mathbb{R}$. For that purpose, we use the following variant of Hoeffding's inequality [13].

Lemma 10. *Let $X_1, \ldots, X_K$ be independent variables such that range of each X_i is $[0, b]$, where $b \in \mathbb{R}$. Then for $X = \sum_{i=1}^{k} X_i$ we have*

$$\mathrm{P}[|X - \mathbf{E}[X]| \geq t] \leq 2e^{-\frac{t^2}{Kb^2}}. \tag{7}$$

By plugging in $b = 1, t = \varepsilon K$ in (7) we obtain the following corollary of the previous lemma.

Corollary 11. *Let $X = \sum_{i=1}^{K} X_i$, where $\{X_i\}_{i=1}^{K}$ are mutually independent random variables with range $[0, 1]$. Then*

$$\Pr\left[|X - \mathbf{E}\left[X\right]| \geq \varepsilon K\right] \leq 2 \cdot e^{-\varepsilon^2 K}, \quad \forall \varepsilon \in (0, 1). \tag{8}$$

3 Reduction

We now prove the theorem which shows the existence of the randomized algorithm which will be used for proving Theorem 3. We remark that this theorem uses a reduction that appeared in [22], and that the main difference comes from the fact that we need to create instances in which the degrees of variables are uniform, while bounded degree was sufficient in [22]. Additional complexity lies in the fact that we prove the theorem for any Max-CSP, and therefore we need to account for different arity of predicates, while [22] considered Max-$E3$-Sat with predicates of arity 3. The theorem we prove is given below.

Theorem 12. *Consider an unweighted instance $\mathcal{F}$ of a Max-CSP Λ and let $\varepsilon > 0$ be a constant. Then, there is a randomized algorithm which outputs a regular instance $\mathcal{G}$ of the Max-CSP Λ such that with probability at least $1 - O(2^{-n})$ over the choices made in the randomized algorithm, the following two statements hold:*

(i) *For any assignment ζ to the variables of $\mathcal{G}$, there is an algorithm which runs in polynomial time and finds an assignment χ to the variables of $\mathcal{F}$ such that*

$$\mathrm{Val}_\chi(\mathcal{F}) \geq \mathrm{Val}_\zeta(\mathcal{G}) - \varepsilon. \tag{9}$$

(ii) *The optimal value of $\mathcal{F}$ is upper bounded by $\mathrm{Opt}(\mathcal{G}) + \varepsilon$.*

Furthermore, the runtime of the randomized algorithm is polynomial in terms of the size of $\mathcal{F}$ and $\lceil 1/\varepsilon \rceil$, and the degree of variables in $\mathcal{G}$ is $O(\log(1/\varepsilon)/\varepsilon^2)$.

Throughout this section we will use m to denote the number of constraints in $\mathcal{F}$, and n for the number of its variables. We also use W to denote the average arity of constraints in $\mathcal{F}$ and W_{max} to denote the maximal arity of a constraint in Λ. We use log to denote the natural logarithm.

Before giving the details of the proof, let us give a rough sketch. We start in the same way as [22], by creating d_i *copies* $x_i^j, j = 1, \dots, d_i$, for each variable x_i in the starting instance $\mathcal{F}$ of degree d_i. Then, in order to create a regular instance, we sample constraints of the starting instance, and create a constraint in the new instance by replacing each x_i occurring in the scope by some of its copies x_i^j uniformly at random. Such a procedure outputs an instance $\mathcal{G}'$ in which every variable has the same degree D in expectation. Furthermore, with high probability over the random choices used for construction of $\mathcal{G}'$, every assignment to the variables of $\mathcal{G}'$ can be used to construct an assignment which satisfies a similar fraction of constraints of the starting instance. Let us refer to instances $\mathcal{G}'$ with such property as "nice" in this sketch. The idea of showing that $\mathcal{G}'$ is "nice" is the same as in [22]; namely, we interpret the fraction of variables x_i^j with value 1 as the probability that the variable x_i should have value 1, and this is used in a randomized algorithm which converts the values of x_i^j to values of x_i, and which can be derandomized to show that $\mathcal{G}'$ is "similar" to the starting instance $\mathcal{F}$. We also show that $\mathcal{G}'$ is "close" to being regular with a probability 1/2 . In particular, we show that, with constant probability, most of the variables of $\mathcal{G}'$ have degrees which do not significantly deviate from their expected value D.

The second step of the proof consists in slightly updating $\mathcal{G}'$ to ensure that each variable has the same degree. We first run the first step n times and obtain n instances $\mathcal{G}'_1, \dots, \mathcal{G}'_n$. With probability $1 - O(2^{-n})$ one of these instances will be such that the degree of variables are close to the expected value, which is a property which can be checked by an algorithm running in polynomial time. Furthemore, such an instance will be also "nice" with a very high probability. Let us fix this instance and call it $\mathcal{G}''$. We then update $\mathcal{G}''$ to obtain the regular instance $\mathcal{G}$ as follows. We first fix some small $\beta > 0$ which depends on ε from Theorem 12, and force all variables x_i^j to have degree $(1+\beta)D$. For variables with degree higher than $(1+\beta)D$ we replace their occurrence in some scopes of the constraints of $\mathcal{G}''$ by some new *dummy* variables. If a variable x_i^j appears less then $(1+\beta)D$ times we simply create a new constraint if needed and add the variable x_i^j to it. The final step in our construction consists in making sure that all newly introduced *dummy* variables also have degree $(1+\beta)D$. We then show that these updates changed or added only a small number of constraints, so our regular instance "looks" like $\mathcal{G}''$, and hence like $\mathcal{F}$.

We give the proof of Theorem 12 in two steps. The first step is given in Theorem 13 and it discusses the construction of $\mathcal{G}'$ and its properties, while the second step shows how $\mathcal{G}''$ can be updated to obtain $\mathcal{G}$ and it is given immediately after the Theorem 13 directly in the proof of Theorem 12. Let us now state and prove Theorem 13.

Theorem 13. *Consider an unweighted instance $\mathcal{F}$ of a Max-CSP Λ, let $0 < \beta < 1$, and $\varepsilon > 0$ be constants. Let $D \in \mathbb{N}$ be a constant such that $D \geq \max(\frac{96\log(q)W_{max}}{\varepsilon^2}, \frac{17\log(1/\beta)W_{max}^2}{\beta^2})$, and such that $(1+\beta)D$ is in $\mathbb{N}$. Then, there is a randomized algorithm which outputs an instance $\mathcal{G}'$ of the Max-CSP Λ such that the following statements hold with probabilities over the randomness in creating $\mathcal{G}'$ given in parentheses:*

[a] (With probability $1 - 2^{-n}$): *For any assignment ζ' to the variables of $\mathcal{G}'$, there is an algorithm which runs in polynomial time and finds an assignment χ to the variables of $\mathcal{F}$*

such that

$$\mathrm{Val}_{\chi}(\mathcal{F}) \geq \mathrm{Val}_{\zeta'}(\mathcal{G}') - \frac{\varepsilon}{2}. \tag{10}$$

[b] (With probability $1 - 2^{-n}$): *The optimal value of $\mathcal{F}$ is upper bounded by* $\mathrm{Opt}(\mathcal{G}') + \frac{\varepsilon}{2}$.

[c] (With probability 1/2): *If we use $y_1, \ldots, y_u$ to denote variables of $\mathcal{G}'$, and $r_1, \ldots, r_u$ for their respective degrees, we then have*

$$\sum_{i=1}^{u} \max(0, r_i - (1+\beta)D) \leq 4mW_{max}, \tag{11}$$

and

$$\sum_{i=1}^{u} \max(0, (1-\beta)D - r_i) \leq 4mW_{max} + 2\beta mW_{max}D. \tag{12}$$

Furthermore, the runtime of the randomized algorithm is polynomial in terms of the size of $\mathcal{F}$ and D.

Proof. We begin by formally describing the randomized procedure which creates $\mathcal{G}'$. The instance $\mathcal{F}$ contains m constraints $C_1, \ldots, C_m$, and each constraint is associated to some tuple of variables S_i. First, for each variable x_i from $\mathcal{F}$ with degree d_i, we create d_i new variables[4] $x_i^1, \ldots, x_i^{d_i}$. Then, we create $\mathcal{G}'$ with mD constraints over variables x_i^j by repeating the following procedure mD times:

- Pick a constraint $C_i = (P_i, S_i)$ from $\mathcal{F}$ uniformly at random.
- For each variable x_j appearing in S_i, pick a variable x_j^r from the set $\{x_j^1, \ldots, x_j^{d_j}\}$ uniformly at random.
- Add a constraint $C_i' = (P_i, S_i')$ to $\mathcal{G}'$, which constrains the variables x_j^r picked in the previous step by the predicate P_i of the constraint C_i. Furthermore, each variable x_j^r appears at the same position in the tuple S_i' as the variable x_j in the tuple S_i.

Since each variable x_i^j is picked at each step with probability $1/m$, the degree d_i^j of each variable x_i^j in $\mathcal{G}'$ is in expectation D.

Let us now prove **[a]**. In order to do so, consider ζ' to be any assignment to variables of $\mathcal{G}'$. Then, consider a randomized assignment $\bar{\chi}$ to the variables of $\mathcal{F}$, in which variables x_i get values independently, and the probability of x_i getting the value 1 is proportional to the number of variables x_i^j getting the value 1 under ζ'. Let us denote by $\rho_{\zeta'}$ the expected value of $\mathcal{F}$ under the random $\bar{\chi}$.

Consider now the process of creating $\mathcal{G}'$ from $\mathcal{F}$. At each step, we pick a constraint C_i, which is satisfied by $\bar{\chi}$ with probability $\rho_{\zeta'}$. Furthermore, the respective constraint C_i' in $\mathcal{G}'$ is satisfied by ζ' with probability $\rho_{\zeta'}$ as well. Therefore, our algorithm will create an instance $\mathcal{G}'$ that under ζ' satisfies a fraction $\rho_{\zeta'}$ of constraints in expectation. By Chernoff's bound (Corollary 11), the probability that the fraction of satisfied constraints in $\mathcal{G}'$ is greater than $\rho_{\zeta'} + \varepsilon/2$ is at most $2\exp\left(-\varepsilon^2 mD/4\right)$. Hence, since $D \geq \frac{4\log(2)}{\varepsilon^2} + \frac{8\log(q)W\log(2)}{\varepsilon^2}$ we have that

$$\Pr\left[\mathrm{Val}_{\zeta'}(\mathcal{G}') \geq \rho_{\zeta'} + \varepsilon/2\right] \leq 2^{-2\log(q)mW}. \tag{13}$$

Since there are mW variables x_i^j in $\mathcal{G}'$, there are q^{mW} possible assignments ζ'. Hence, by the union bound the probability that there is an assignment ζ' which satisfies more than $\rho_{\zeta'} + \varepsilon/2$ constraints in $\mathcal{G}'$ is at most $2^{-\log(q)mW}$. Therefore, the property **[a]** holds with probability at least $1 - 2^{-\log(q)mW} \geq 1 - 2^{-n}$.

[4]Variables x_i^j correspond to the variables $y_i, i = 1, \ldots, u$, from the statement of the theorem. We however prefer to use y_i instead of x_i^j in the statement since this allows us to skip superscripts and hence make the statement more readable, which makes subsequent arguments relying on this theorem easier to follow.

Next, in a similar way, we show **[b]**. In particular, let us fix the assignment χ to the variables of $\mathcal{F}$ under which the optimal value $\mathrm{Opt}(\mathcal{F})$ is attained. Then, we construct an assignment ζ' for $\mathcal{G}'$, by setting x_i^j to have the same values as the corresponding x_i under χ. In every trial of the construction of $\mathcal{G}'$, the sampled constraint is satisfied with probability $\mathrm{Opt}(\mathcal{F})$, and therefore the expected fraction of satisfied constraints in $\mathcal{G}'$ is $\mathrm{Opt}(\mathcal{F})$. The probability (over the random choices in the construction of $\mathcal{G}'$) that $\mathrm{Val}_{\zeta'}(\mathcal{G}') \leq \mathrm{Opt}(\mathcal{F}) - \varepsilon/2$ is, by Chernoff's bound (Corollary 11), at most

$$2\exp\left(\frac{-mD\varepsilon^2}{4}\right). \tag{14}$$

Therefore, since $D \geq \frac{4\log(2)(1+W_{max})}{\varepsilon^2}$, this probability will be smaller than 2^{-n}. Therefore, the property **[b]** holds with probability at least $1-2^{-n}$.

Let us now prove **[c]**. For a variable x_i^j let us use R_i^j to denote the random variable which value is its degree, and the randomness comes from the choices in the construction of $\mathcal{G}'$. Let us define $H := \sum_{i=1}^n \sum_{j=1}^{d_i} \max(0, R_i^j - (1+\beta)D)$ and $L = \sum_{i=1}^n \sum_{j=1}^{d_i} \max\left(0, (1-\beta)D - R_i^j\right)$. We can then rephrase our current goal of showing that **[c]** holds with probability 1/2 as

$$\Pr\left[H \leq 4mW_{max} \text{ and } L \leq 4mW_{max} + 2\beta m W_{max} D\right] \geq 1/2. \tag{15}$$

We first show that $\Pr\left[H < 4mW_{max}\right] \geq 3/4$. We begin by calculating the expectation of each summand $\max(0, R_i^j - (1+\beta)D)$ of H, i.e., we calculate $\mathbf{E}[\max(0, R_i^j - (1+\beta)D)]$ for a fixed x_i^j as follows.

$$\begin{aligned}
\mathbf{E}\left[\max(0, R_i^j - (1+\beta)D)\right] &= \sum_{t=(1+\beta)D+1}^{mD} \Pr\left[R_i^j = t\right](t - (1+\beta)D) \\
&= \sum_{t=(1+\beta)D+1}^{mD} \Pr\left[R_i^j = t\right] \sum_{k=(1+\beta)D+1}^{t} 1 \\
&= \sum_{t=(1+\beta)D+1}^{mD} \sum_{k=(1+\beta)D+1}^{t} \Pr\left[R_i^j = t\right] \\
&= \sum_{k=(1+\beta)D+1}^{mD} \sum_{t=k}^{mD} \Pr\left[R_i^j = t\right],
\end{aligned} \tag{16}$$

where the last equality was obtained by changing the order of summation. Since R_i^j is at most mD we have that

$$\sum_{k=(1+\beta)D+1}^{mD} \sum_{t=k}^{mD} \Pr\left[R_i^j = t\right] = \sum_{k=(1+\beta)D+1}^{mD} \Pr\left[R_i^j \geq k\right]. \tag{17}$$

By Chernoff's bound (Lemma 9), the probability that a variable x_i^j appears in more than $(1+\delta)D$ constraints in $\mathcal{G}'$ is given by

$$\mathrm{P}\left[R_i^j \geq (1+\delta)D\right] \leq 2e^{-D\min\left(\frac{\delta}{2}, \frac{\delta^2}{4}\right)}. \tag{18}$$

Hence, $\Pr\left[R_i^j \geq k\right] = \Pr\left[R_i^j \geq (1+\delta)D\right]$ for $\delta = k/D - 1 > 0$, and therefore

$$\Pr\left[R_i^j \geq k\right] \leq 2e^{-D\min\left(\frac{k/D-1}{2}, \frac{(k/D-1)^2}{4}\right)} \tag{19}$$

Having this calculation in mind we can now write

$$\sum_{k=(1+\beta)D+1}^{mD} \Pr\left[R_i^j \geq k\right] \leq \sum_{k=(1+\beta)D+1}^{mD} 2e^{-D\min\left(\frac{k/D-1}{2}, \frac{(k/D-1)^2}{4}\right)}. \tag{20}$$

By introducing $c = k - (1+\beta)D$, we can further simplify this expression as

$$2\sum_{c=1}^{mD-(1+\beta)D} \exp\left(-D\min\left(\frac{c/D+\beta}{2}, \frac{(c/D+\beta)^2}{4}\right)\right) \leq 2\sum_{c=1}^{\infty}\exp\left(-D\frac{c/D+\beta}{2}\right) + 2\sum_{c=1}^{\infty}\exp\left(-D\frac{(c/D+\beta)^2}{4}\right) \tag{21}$$

Let us denote by

$$A := 2\sum_{c=1}^{\infty}\exp\left(-D\frac{c/D+\beta}{2}\right) \tag{22}$$

the first term, and by

$$B := 2\sum_{c=1}^{\infty}\exp\left(-D\frac{(c/D+\beta)^2}{4}\right) \tag{23}$$

the second term. We now show that $A \leq 1/2$ and $B \leq 1/2$ which will imply that

$$\mathbf{E}\left[\max(0, R_i^j - (1+\beta)D)\right] \leq A + B \leq 1. \tag{24}$$

For A we can write

$$\begin{aligned} A = 2\sum_{c=1}^{\infty}\exp\left(-D\frac{c/D+\beta}{2}\right) &= 2\sum_{c=1}^{\infty}\exp\left(-\frac{c+\beta D}{2}\right) = 2\exp\left(\frac{-\beta D}{2}\right)\cdot\sum_{c=1}^{\infty}\exp\left(-\frac{c}{2}\right) \\ &\leq 2\exp\left(\frac{-\beta D}{2}\right)\cdot\int_{c=0}^{\infty}\exp\left(-\frac{c}{2}\right)dc = 2\exp\left(\frac{-\beta D}{2}\right)\cdot 2, \end{aligned} \tag{25}$$

where in the last inequality we used the fact that $e^{-c/2}$ is monotonically decreasing to replace the sum with the integral, while the last equality holds since $\int_0^\infty \exp\left(-\frac{c}{2}\right)dc = 2$. Since D is chosen such that $D \geq \frac{6\log(2)}{\beta}$ we have that

$$A \leq 4\exp\left(\frac{-\beta D}{2}\right) \leq \frac{1}{2}. \tag{26}$$

For B we can write

$$B = 2\sum_{c=1}^{\infty}\exp\left(-D\frac{(c/D+\beta)^2}{4}\right) = 2\sum_{c=1}^{\infty}\exp\left(-\frac{D^2(c/D+\beta)^2}{4D}\right) = 2\sum_{c=1}^{\infty}\exp\left(-\frac{(c+\beta D)^2}{4D}\right). \tag{27}$$

Since the exponent in the last expression is negative and $\frac{c}{\beta} > 0$ we can write

$$\begin{aligned} B \leq 2\sum_{c=1}^{\infty}\exp\left(-\frac{1}{4}\frac{(\beta D+c)^2}{D+\frac{1}{\beta}c}\right) &= 2\sum_{c=1}^{\infty}\exp\left(-\frac{1}{4}\frac{(\beta D+c)^2}{\frac{\beta}{\beta}D+\frac{1}{\beta}c}\right) = 2\sum_{c=1}^{\infty}\exp\left(-\frac{\beta}{4}(\beta D+c)\right) \\ &= 2\exp\left(-\frac{\beta^2 D}{4}\right)\sum_{c=1}^{\infty}\exp\left(-\frac{c\beta}{4}\right) \end{aligned} \tag{28}$$

Since $\exp\left(-\frac{c\beta}{4}\right)$ is monotonically decreasing we can further write

$$B \leq 2\exp\left(-\frac{\beta^2 D}{4}\right)\int_0^\infty \exp\left(-\frac{c\beta}{4}\right)dc = 2\exp\left(-\frac{\beta^2 D}{4}\right)\frac{4}{\beta}, \tag{29}$$

where in the last equality we use the fact that $\int_0^\infty \exp\left(-\frac{c\beta}{4}\right)dc = \frac{4}{\beta}$. Finally, by our choice of $D \geq \frac{16\log(2)}{\beta^2} + \frac{\log(1/\beta)}{\beta^2}$ we have that

$$B \leq 2\exp\left(-\frac{\beta^2 D}{4}\right)\frac{4}{\beta} \leq \frac{1}{2}. \tag{30}$$

Therefore, we have that $\mathbf{E}\left[\max(0, R_i^j - (1+\beta)D)\right] \leq 1$. Observe that H has mW summands $\max(0, R_i^j - (1+\beta)D)$. Hence, $\mathbf{E}[H] \leq mW$. Furthermore, since H takes nonnegative values by Markov's inequality we have

$$\Pr[H \geq 4mW_{max}] \leq \frac{mW}{4mW_{max}} \leq \frac{1}{4}, \tag{31}$$

and therefore $\Pr[H \leq 4mW_{max}] \geq 3/4$. We can now use this property to show **[c]** expressed as (15) with the following argument. First, we study the probability that the average arity of constraints in $\mathcal{G}'$ is smaller than $(1-\beta)W$. In particular, if we denote by $C_i' = (P_i', S_i')$ the constraints of the instance $\mathcal{G}'$, we first show that

$$\Pr\left[H \leq 4mW_{max} \text{ and } \left|\frac{1}{mD}\sum_{C_i' \in \mathcal{G}'} \mathrm{ar}(P_i') - W\right| \leq \beta W\right] \geq \frac{1}{2}. \tag{32}$$

Next, we show that having the two events above occurring simultaneously is sufficient to show that

$$L \leq 4mW_{max} + 2\beta mWD, \tag{33}$$

and therefore the property **[c]**:

$$\Pr\left[H \leq 4mW_{max} \text{ and } L \leq 4mW_{max} + 2\beta mWD\right] \geq 1/2. \tag{34}$$

Hence, let us start by showing (32). An application of Hoeffding's inequality (Lemma 10) gives us

$$\Pr\left[\left|\frac{1}{mD}\sum_{C_i' \in \mathcal{G}'} \mathrm{ar}(P_i') - W\right| \geq \beta W\right] \leq 2\exp\left(-\frac{\beta^2 m^2 D^2 W^2}{mDW_{\max}^2}\right) = 2\exp\left(-\frac{\beta^2 mDW^2}{W_{\max}^2}\right) \leq \frac{1}{4}, \tag{35}$$

where the last inequality holds due to our choice of $D \geq \frac{3\log(2)W_{max}^2}{\beta^2}$. Then (32) holds by the union bound of (31) and (35).

Next, we show that having the average arity of constraints being greater or equal than $(1-\beta)D$ and $H < 4mW_{max}$ is sufficient for $L \leq 4mW_{max} + 2\beta mWD$. It is not hard to observe[5] that the "worst case instance" has one variable x_i^j with the degree $(1+\beta)D + 4mW_{max}$, c_1 variables x_i^j with degrees 0, and the remaining c_2 with degrees $(1+\beta)D$. Let us fix this instance from now on. Since $H \leq 4mW_{max}$, the sum of the degrees of variables in $\mathcal{G}'$ is at most

$$c_1 \cdot 0 + c_2(1+\beta)D + \left((1+\beta)D + 4mW_{max}\right). \tag{36}$$

Since the average degree is at least $(1-\beta)D$ and $\mathcal{G}'$ has mW variables we have that

$$c_2(1+\beta)D + \left((1+\beta)D + 4mW_{max}\right) \geq (1-\beta)DmW. \tag{37}$$

The fact that $\mathcal{G}'$ has mW variables also implies that

$$c_1 + c_2 + 1 = mW. \tag{38}$$

Replacing c_2 from (37) into (38) gives us

$$c_1 + \frac{1-\beta}{1+\beta}mW - \frac{(1+\beta)D + 4mW_{max}}{(1+\beta)D} + 1 \leq mW. \tag{39}$$

The upper bound on H is then given by

$$\begin{aligned} H \leq c_1 \cdot (1-\beta)D \leq c_1 D &\leq mWD - \frac{1-\beta}{1+\beta}mWD + \frac{(1+\beta)D + 4mW_{max}}{(1+\beta)D}D - D \\ &= \frac{2\beta}{1+\beta}mWD + D + \frac{4mW_{max}}{(1+\beta)D}D - D \leq 2\beta mW_{max}D + 4mW_{max}, \end{aligned} \tag{40}$$

where in the last inequality we used the fact that $\beta > 0$. This concludes our proof of the property **[c]** □

[5] One can also show this formally. We omit the discussion here for the sake of brevity.

We next show how the instance $\mathcal{G}'$ output by the previous theorem can be updated to obtain $\mathcal{G}$ with the properties required by the Theorem 12.

Proof of Theorem 12. Let us fix $\varepsilon > 0$ and without loss of generality let us assume that the instance $\mathcal{F}$ from Theorem 13 has the number of constraints $m > 72/\varepsilon + 1$, since the theorem holds trivially for constant sized instances $\mathcal{F}$.
Let us choose $\beta < \varepsilon m/72 - 1$ and $\beta < \varepsilon/(96W_{max})$. Due to our choice of ε we have that $\beta > 0$. Furthermore, we choose $D \in \mathbb{N}$ such that the conditions from Theorem 13 hold and such that $(1+\beta)D$ is coprime with W_{max}, and then run the reduction from Theorem 13 up to n times, stopping if the property **[c]** from Theorem 13 holds (which can be checked by a polynomial time algorithm). If **[c]** did not hold in the first n trials, our algorithm fails, which happens with probability $1 - 2^{-n}$. Otherwise, let us denote by $\mathcal{G}''$ the chosen instance for which **[c]** holds.

Observe that the probability that **[a]** from Theorem 13 holds for $\mathcal{G}''$ is actually equal to the probability that **[a]** holds given that **[c]** holds, and this is at least $1 - 2 \cdot 2^{-n} = 1 - O(2^{-n})$. The same argument shows that the probability that **[b]** holds for $\mathcal{G}''$ is at least $1 - O(2^{-n})$, and hence by the union bound $\mathcal{G}''$ satisfies all three properties **[a]**,**[b]**, and **[c]**, with probability $1 - O(2^{-n})$.

Let us now show how $\mathcal{G}''$ can be altered to obtain the regular instance $\mathcal{G}$ which satisfies the properties of Theorem 12. We first update $\mathcal{G}''$ such that all the variables $y_1, \ldots, y_u$, have degree $(1+\beta)D$. In this process we will introduce some new variables which will be called *dummy* variables. At the beginning we create W_{max} dummy variables and put them into a set $\mathcal{D}$. In the argument that follows each time we say that y_i is replaced by a dummy variable or that a dummy variable is added to a constraint in an instance, we actually use a dummy variable from $\mathcal{D}$ that does not appear in the mentioned constraint. After we pick a dummy variable from $\mathcal{D}$, we increase its degree, and if its degree becomes equal to $(1+\beta)D$ we remove it from $\mathcal{D}$, and add a new dummy variable to $\mathcal{D}$.

Let us now discuss how we update $\mathcal{G}''$ so that each y_i has degree $(1+\beta)D$. Consider first the variables y_i with $\deg(y_i) > (1+\beta)D$. We replace such y_i by dummy variables in $\deg(y_i) - (1+\beta)D$ constraints. By the property **[c]** in total we alter at most $4mW_{max}$ constraints.

Next, consider the variables y_i such that $\deg(y_i) < (1+\beta)D$. For each such variable y_i we create $(1+\beta)D - \deg(y_i)$ new constraints of arity W_{max}, put y_i in each newly created constraint, and fill the remaining $W_{max} - 1$ places of each constraint with dummy variables. Let us now show how many constraints we need to add in this process. In order to make every variable y_i have degree at least $(1-\beta)D$, since **[c]** holds, we add up to $4mW_{max} + 2\beta mWD$ constraints. Then, assuming that each y_i has degree at least $(1-\beta)D$ we need at most $2\beta mWD$ additional constraints to make sure each variable has degree at least $(1+\beta)D$.

Finally, we ensure that the dummy variables have degree $(1+\beta)D$ as well. Observe that after the process of making sure that y_i have degree $(1+\beta)D$ we are left with at most W_{max} dummy variables with the degrees between 0 and $(1+\beta)D$. In order to ensure that they have degree $(1+\beta)D$ we need to use them $s < (1+\beta)DW_{max}$ times. Since D is picked so that $(1+\beta)D$ is coprime with W_{max}, there is some k such that $W_{max} \le k < 2W_{max}$ and $W_{max} | k(1+\beta)D + s$. Let us then introduce k new dummy variables and add $(k(1+\beta)D + s)/W_{max}$ constraints with predicate of arity W_{max}, by assigning the scopes such that the degree of each variable becomes exactly $(1+\beta)D$. In particular, we can assign scopes iteratively, by adding variables with the smallest degree to the scope at each step. Since $k \ge W_{max}$ variables have degree 0 at the start, we can always assign W_{max} distinct variables to the scope. Finally, since $W_{max} | k(1+\beta)D + s$, the iterative procedure will finish.

Let us call the regular instance created by the process above $\mathcal{G}$. When altering $\mathcal{G}''$ to obtain $\mathcal{G}$, we have introduced or changed at most the following number of constraints:

- $4mW_{max}$ constraints by replacing variables y_i by new dummy variables, in order to ensure that no variable y_i has degree greater than $(1+\beta)D$.
- $4\beta mW_{max}D + 4mW_{max}$ constraints in order to ensure that every variable y_i does not have degree smaller than $(1+\beta)D$.

- $3(1+\beta)D$ new constraints introduced to ensure that any additionally added variable has degree $(1+\beta)D$.

Let us use

$$C_r := 4mW_{max}, \quad C_n = 4mW_{max} + 4\beta mW_{max} + 3(1+\beta)D, \tag{41}$$

to denote the upper bounds on the changed and new constraints in $\mathcal{G}$, respectively. Furthermore, define $C := C_r + C_n$. Due to our choice of D and β we have that

$$\frac{4mW_{max}}{mD} \leq \frac{\varepsilon}{12}, \quad \frac{4\beta mW_{max}D}{mD} \leq \frac{\varepsilon}{24}, \quad \text{and} \quad \frac{3(1+\beta)D}{mD} \leq \frac{\varepsilon}{24}. \tag{42}$$

Therefore, we have that $C/mD \leq \varepsilon/4$.

Consider now ζ to be any assignment to the variables of $\mathcal{G}$ and ζ' to be the restriction of ζ to the variables y_i of $\mathcal{G}''$. We then have

$$\mathrm{Val}_\zeta(\mathcal{G}) \leq \frac{\mathrm{Val}_{\zeta'}(\mathcal{G}'')\cdot mD + C}{mD + C_n} \leq \frac{\mathrm{Val}_{\zeta'}(\mathcal{G}'')\cdot mD + C}{mD} = \mathrm{Val}_{\zeta'}(\mathcal{G}'') + \frac{C}{mD} \leq \mathrm{Val}_{\zeta'}(\mathcal{G}'') + \frac{\varepsilon}{4}. \tag{43}$$

Similarly, we have that

$$\begin{aligned}
\mathrm{Val}_\zeta(\mathcal{G}) &\geq \frac{\mathrm{Val}_{\zeta'}(\mathcal{G}'')\cdot mD - C}{mD + C_n} \geq \mathrm{Val}_{\zeta'}(\mathcal{G}'')\frac{mD}{mD + C_n} - \frac{C}{mD + C_n} \\
&= \mathrm{Val}_{\zeta'}(\mathcal{G}'')\frac{mD + C_n - C_n}{mD + C_n} - \frac{C}{mD + C_n} \\
&= \mathrm{Val}_{\zeta'}(\mathcal{G}'') - \mathrm{Val}_{\zeta'}(\mathcal{G}'')\frac{C_n}{mD + C_n} - \frac{C}{mD + C_n} \\
&\geq \mathrm{Val}_{\zeta'}(\mathcal{G}'') - 2\frac{C}{mD} \geq \mathrm{Val}_{\zeta'}(\mathcal{G}'') - \frac{\varepsilon}{2}.
\end{aligned} \tag{44}$$

Hence,

$$\mathrm{Opt}(\mathcal{G}) \geq \mathrm{Opt}(\mathcal{G}'') - \frac{\varepsilon}{2}. \tag{45}$$

Finally, let us show that $\mathcal{G}''$ satisfies properties **[i]** and **[ii]** of Theorem 12. As discussed at the second paragraph of this proof, $\mathcal{G}''$ satisfies **[a]**, **[b]** and **[c]** with probability $1 - O(2^{-n})$. Let us assume this event happened, and show the properties required by Theorem 12. First, let us observe that our instance has degree $(1+\beta)D$ which is $O(\log(1/\varepsilon)/\varepsilon^2)$ thanks to our choice of D. Next, for **[i]**, consider any assignment ζ to the variables of $\mathcal{G}$. There is an algorithm which runs in polynomial time and finds an assignment ζ' to the variables of $\mathcal{G}''$ such that

$$\mathrm{Val}_\zeta(\mathcal{G}) \leq \mathrm{Val}_{\zeta'}(\mathcal{G}'') + \frac{\varepsilon}{2}. \tag{46}$$

The algorithm just takes the restriction ζ' of ζ and the inequality above is due to (43). Finally, due to the property **[a]**, we can apply the polynomial time algorithm to ζ' to find an assignment χ to $\mathcal{F}$ such that

$$\mathrm{Val}_{\zeta'}(\mathcal{G}'') \leq \mathrm{Val}_\chi(\mathcal{F}) + \frac{\varepsilon}{2}. \tag{47}$$

Combining (46) and (47) gives us

$$\mathrm{Val}_\zeta(\mathcal{G}) \leq \mathrm{Val}_\chi(\mathcal{F}) + \varepsilon, \tag{48}$$

which shows **[i]** .

The property **[ii]** holds by combining (45) and the property **[b]**. This concludes the proof of Theorem 12. □

Let us now prove Theorem 3. For that reason, let us suppose that regular instances of Max-CSP Λ can be approximated within some fixed approximation ratio α. Due to Theorem 7 it is sufficient to show that unweighted instances of Max-CSP Λ can be approximated within $\alpha - \varepsilon/2$. Hence, let us consider an arbitrary (possibly not regular) unweighted instance $\mathcal{F}$ of Λ. Then, we construct $\mathcal{G}$ with the probabilistic algorithm given by Theorem 12, apply the α approximation algorithm to find an assignment ζ, and then using the algorithm from Theorem 12, with probability at least $1-O(2^{-n})$, we can find an assignment χ to the instance $\mathcal{F}$ satisfying

$$\begin{aligned} \frac{\mathrm{Val}_\chi(\mathcal{F})}{\mathrm{Opt}(\mathcal{F})} &\geq \frac{\mathrm{Val}_\zeta(\mathcal{G})-\varepsilon}{\mathrm{Opt}(\mathcal{G})+\varepsilon} \geq \frac{\mathrm{Val}_\zeta(\mathcal{G})-\varepsilon}{\mathrm{Opt}(\mathcal{G})}\left(1-\frac{\varepsilon}{\mathrm{Opt}(\mathcal{G})}\right) \\ &\geq \frac{\mathrm{Val}_\zeta(\mathcal{G})}{\mathrm{Opt}(\mathcal{G})} - \frac{\varepsilon\,\mathrm{Val}_\zeta(\mathcal{G})}{\mathrm{Opt}(\mathcal{G})^2} - \frac{\varepsilon}{\mathrm{Opt}(\mathcal{G})} \geq \alpha - 2\frac{\varepsilon}{\mathrm{Opt}(\mathcal{G})}. \end{aligned} \tag{49}$$

Now, we have that $\mathrm{Opt}(\mathcal{G}) \geq \gamma$, for some[6] fixed $\gamma > 0$ which depends only on Λ. Therefore, by choosing $\varepsilon = \delta\gamma/4$, Theorem 3 follows.

Note that using analog of Theorem 12 for Min-CSPs to prove Theorem 2 would require $\varepsilon = O(1/m)$, and therefore the instance $\mathcal{G}$ in the reduction will be of size at least $m^3\log(m)$, with $D = \Omega(m^2\log(m))$. We give a deterministic reduction instead, which works for both Max-CSPs and Min-CSPs, and which creates a regular instance of degree $\lceil \bar{D}/\varepsilon \rceil$ where $\bar{D}$ is the maximal degree of constraints in $\mathcal{F}$. The reduction is given as the following theorem.

Theorem 14. *Consider a Max-CSP (or Min-CSP) Λ and let $\varepsilon > 0$ be a constant. Then, there is a reduction which takes as an input an instance $\mathcal{F}$ of a Max-CSP (or Min-CSP) Λ and outputs a regular instance $\mathcal{G}$ of the Max-CSP (or Min-CSP) Λ such that the following holds:*

[a] $\mathrm{Opt}(\mathcal{F}) \leq \mathrm{Opt}(\mathcal{G})$ *(or* $\mathrm{Opt}(\mathcal{F}) \geq \mathrm{Opt}(\mathcal{G})$ *for Min-CSP)*

[b] *Let ζ be an assignment to the variables of $\mathcal{G}$. There, there is an algorithm which finds an assignment χ to the variables of $\mathcal{F}$ such that*

$$\begin{aligned} \mathrm{Val}_\chi(\mathcal{F}) &\geq \mathrm{Val}_\zeta(\mathcal{G}) - \varepsilon && \text{(Max-CSP case)}, \\ \mathrm{Val}_\chi(\mathcal{F}) &\leq \mathrm{Val}_\zeta(\mathcal{G}) + \varepsilon && \text{(Min-CSP case)}. \end{aligned} \tag{50}$$

Furthermore, the runtime of the reduction and of the algorithm from ***[b]*** *is polynomial in terms of the size of $\mathcal{F}$ and $\lceil 1/\varepsilon \rceil$.*

Proof. We prove this theorem for Min-CSP Λ. The proof for Max-CSP is analogous.

We begin by describing how a regular instance $\mathcal{G}$ is constructed. We start from $\mathcal{F}$ which has n variables $x_1, \ldots, x_n$, with degrees $d_1, \ldots, d_n$, and constraints $\{(P_r, S_r)\}_{r=1}^m$. Let us define $D := \max_{i\in[n]} d_i$ and $N := \lceil D/\varepsilon \rceil$. For each variable x_i from $\mathcal{F}$ we create d_i variables $x_i^1, x_i^2, \ldots, x_i^{d_i}$ in $\mathcal{G}$. The constraints of $\mathcal{G}$ are constructed as follows. We begin by creating N copies of $\mathcal{F}$ which we call blocks. Then, we go through the blocks and at each scope we replace variables x_i by their corresponding copies x_i^j, $j \in [d_i]$. In particular, each x_i can be replaced only by d_i variables $x_i^1, x_i^2, \ldots, x_i^{d_i}$. Since x_i appears d_i times in $\mathcal{F}$, it will get replaced Nd_i times, and in order to impose regularity we replace x_i by each copy x_i^j exactly N times. Therefore, the degree of all variables is N, and the instance $\mathcal{G}$ is regular.

Actually, we have to be a bit more careful when replacing variables x_i by their copies x_i^j. The idea is that each block should resemble $\mathcal{F}$ as much as possible, and therefore we want to avoid replacing x_i by two different copies x_i^j, x_i^k, in the same block. Let us call a block *good* if each variable x_i is replaced by a single copy x_i^j in that block. Our aim is to maximize the number of good blocks, which we do greedily by creating a good block at each step of the algorithm as long as we can. The greedy works iteratively in N steps, and in each step it tries to create a *good* block. At ℓ-th iteration, $\ell \in [N]$, the greedy checks if for each variable x_i we can find x_i^j which was used $N - d_i$ times or less. If this is the case, let us denote with $x_i^{j_i}$ the variables

[6] As in the proof of Theorem 7, w.l.o.g. we assume that the instance $\mathcal{F}$ does not contain predicates which evaluate to 0 under all assignments.

used at most $\leq N - d_i$ times. Then the greedy replaces x_i with $x_i^{j_i}$ to create a good block. In case aforementioned choice of $x_i^{j_i}$ cannot be made, then this and all the following iterations will create a *bad* block; hence the greedy replaces x_i with arbitrary an x_i^j which is not used more than N times.

This finishes our description of $\mathcal{G}$. Before we prove the claims of the theorem, let us find a lower bound on the number of good blocks created. In our greedy algorithm, a good block can be created if for each variable x_i we can find x_i^j which was used $N - d_i$ times or less. Therefore, each variable x_i^j can be used at least $\lfloor N/d_i \rfloor$ times for creating a good block, and therefore the number of good blocks is $\min_{i \in [n]} \lfloor N/d_i \rfloor d_i$. Hence, since $D = \max_{i \in [n]} d_i$, we conclude that there are at least $N - D$ good blocks.

It is straightforward to verify **[a]**, and hence let us now prove the claim **[b]**. For a given assignment ζ of variables in $\mathcal{G}$, let us consider a good block with the smallest value, and let us denote the value of this block by v. We define an assignment χ to be the assignment of ζ on this block. We note that we can do this since the copies of x_i are unique in every good block. We have that $\mathrm{Val}_\chi(\mathcal{F}) = v$, and therefore since v is the minimal value of good blocks we have

$$\mathrm{Val}_\zeta(\mathcal{G}) \geq \frac{1}{N}\left((N - D)v\right) \geq v - \frac{D}{N} = \mathrm{Val}_\chi(\mathcal{F}) - \varepsilon, \tag{51}$$

which finishes the proof of **[b]**. □

Let us now show how this result can be used to prove Theorem 2. Hence, let us fix $0 < \delta < 1$, and consider an instance $\mathcal{F}$ of a Min-CSP Λ. Thanks to Theorem 8 we can assume that $\mathcal{F}$ is unweighted. We apply the algorithm from the previous theorem to $\mathcal{F}$ with $\varepsilon = \delta/m$ to get a regular instance $\mathcal{G}$. Then, we use the α approximation algorithm to get an assignment ζ to variables of $\mathcal{G}$, and then by the algorithm from the point **[b]** of Theorem 14 we obtain an assignment χ for $\mathcal{F}$.

In case $\mathrm{Opt}(\mathcal{F}) = 0$ by claim **[b]** of Theorem 14 we have that $\mathrm{Opt}(\mathcal{G}) = 0$ as well. Therefore, since ζ gives us α approximation of $\mathrm{Opt}(\mathcal{G})$, we have that $\mathrm{Val}_\zeta(\mathcal{G}) = 0$. Finally, by the claim **[a]** of Theorem 14 we have that $\mathrm{Val}_\chi(\mathcal{F}) \leq \delta/m$, which can be only possible if $\mathrm{Val}_\chi(\mathcal{F}) = 0$.

It remains to consider the case when $\mathrm{Opt}(\mathcal{F}) \neq 0$, i.e. $\mathrm{Opt}(\mathcal{F}) \geq 1/m$. In that case we have

$$\begin{aligned} \frac{\mathrm{Val}_\chi(\mathcal{F})}{\mathrm{Opt}(\mathcal{F})} \leq \frac{\mathrm{Val}_\chi(\mathcal{G}) + \varepsilon}{\mathrm{Opt}(\mathcal{F})} \leq \frac{\mathrm{Val}_\chi(\mathcal{G})}{\mathrm{Opt}(\mathcal{F})} + \frac{\varepsilon}{\mathrm{Opt}(\mathcal{F})} \\ \leq \frac{\mathrm{Val}_\chi(\mathcal{G})}{\mathrm{Opt}(\mathcal{G})} + \frac{\delta/m}{1/m} \leq \alpha + \delta, \end{aligned} \tag{52}$$

which finishes the proof of Theorem 2. Theorem 1 can be proved analogously.

4 Conclusion and Some Open Questions

In this article we introduced a reduction which shows how approximation algorithms working on regular unweighted instances of optimization CSPs can be converted (with an arbitrary small loss in the approximation ratio) into approximation algorithms for weighted CSPs in which regularity is not imposed. One interesting question would be to see if we could use this result to obtain better approximation algorithms for different CSPs. Also, the aim of quantifying what makes the problems hard is interesting in its own right, and therefore it would be valuable to analyze whether some additional structure of CSP instances can always be assumed when studying their inapproximability.

It is not uncommon that the reductions showing the hardness of approximation output instances which satisfy some form of regularity. This work shows that we cannot hope to obtain stronger inapproximability results by considering irregular instances of CSPs. However, for many other problems it is still not known whether regular instances might be easier to approximate; answering this question could facilitate the search for the optimal algorithms. One family of problems for which this is an especially interesting topic due to their generality and applicability is defined as "Max Ones" in [16].

On the other side, let us remark that using irregular instances can also be instrumental for showing strong hardness results for certain problems, as recently shown in [6] which treated some cardinality constrained CSPs, i.e., variants of a CSP problem where we also prescribe the cardinality of zeros and ones in admissible assignments. In the aforementioned work the reduction works by first creating a regular instance and showing a hardness for such instance, after which the hardness is "boosted" in certain situations by adding variables of degree[7] 0 . This makes the instances irregular and for some cardinality constraints this procedure gives improved hardness result. On the other side, that work does not show that the regular instances are easier, hence showing this or the converse remains an open question.

With this example in mind, it would be interesting to explore whether we can obtain better hardness results by considering more irregular or asymmetric instances for some problems for which satisfactory understanding of approximability is lacking.

We also mention that the reduction used for proving Theorem 3 does not preserve perfect completeness. Hence, it would be interesting to investigate whether having a *good* approximation algorithm for satisfiable regular instances of a constant degree can be useful for approximating possibly not regular satisfiable instances.

Acknowledgments

I am indebted to Per Austrin for pointing out the reduction in [22] to me. I also thank Johan Håstad and anonymous reviewers for numerous useful comments which significantly improved the quality of presentation of this work. Special thanks goes to Hans Oude Groeniger for proofreading this work.

References

[1] S. Arora, L. Babai, J. Stern, and Z. Sweedyk, *The hardness of approximate optima in lattices, codes, and systems of linear equations*, J. Comput. Syst. Sci., 54 (1997), pp. 317–331.

[2] S. Arora and B. Barak, *Computational Complexity - A Modern Approach*, Cambridge University Press, 2009.

[3] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy, *Proof verification and the hardness of approximation problems*, J. ACM, 45 (1998), pp. 501–555.

[4] S. Arora and S. Safra, *Probabilistic checking of proofs: A new characterization of NP*, J. ACM, 45 (1998), pp. 70–122.

[5] P. Austrin, *Balanced max 2-SAT might not be the hardest*, in Proceedings of the 39th Annual ACM Symposium on Theory of Computing, San Diego, California, USA, June 11-13, 2007, 2007, pp. 189–197.

[6] P. Austrin and A. Stankovic, *Global cardinality constraints make approximating some Max-2-CSPs harder*, in Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2019, September 20-22, 2019, Massachusetts Institute of Technology, Cambridge, MA, USA., 2019, pp. 24:1–24:17.

[7] M. Bellare, O. Goldreich, and M. Sudan, *Free bits, PCPs, and nonapproximability-towards tight results*, SIAM J. Comput., 27 (1998), pp. 804–915.

[8] A. A. Bulatov, *A dichotomy theorem for nonuniform CSPs*, in 58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017, 2017, pp. 319–330.

[7] The "boosting" would still work as long as the degree of these added variables is small.

[9] S. O. Chan, *Approximation resistance from pairwise-independent subgroups*, J. ACM, 63 (2016), pp. 27:1–27:32.

[10] P. Crescenzi, R. Silvestri, and L. Trevisan, *On weighted vs unweighted versions of combinatorial optimization problems*, Inf. Comput., 167 (2001), pp. 10–26.

[11] M. X. Goemans and D. P. Williamson, *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming*, J. ACM, 42 (1995), pp. 1115–1145.

[12] J. Håstad, *Some optimal inapproximability results*, J. ACM, 48 (2001), pp. 798–859.

[13] W. Hoeffding, *Probability inequalities for sums of bounded random variables*, Journal of the American Statistical Association, 58 (1963), pp. 13–30.

[14] O. H. Ibarra and C. E. Kim, *Fast approximation algorithms for the knapsack and sum of subset problems*, J. ACM, 22 (1975), pp. 463–468.

[15] H. J. Karloff and U. Zwick, *A 7/8-approximation algorithm for MAX 3SAT?*, in 38th Annual Symposium on Foundations of Computer Science, FOCS '97, Miami Beach, Florida, USA, October 19-22, 1997, 1997, pp. 406–415.

[16] S. Khanna, M. Sudan, L. Trevisan, and D. P. Williamson, *The approximability of constraint satisfaction problems*, SIAM J. Comput., 30 (2000), pp. 1863–1920.

[17] S. Khot, *On the power of unique 2-prover 1-round games*, in Proceedings on 34th Annual ACM Symposium on Theory of Computing, May 19-21, 2002, Montréal, Québec, Canada, 2002, pp. 767–775.

[18] M. Lewin, D. Livnat, and U. Zwick, *Improved rounding techniques for the MAX 2-SAT and MAX DI-CUT problems*, in Integer Programming and Combinatorial Optimization, 9th International IPCO Conference, Cambridge, MA, USA, May 27-29, 2002, Proceedings, 2002, pp. 67–82.

[19] P. Raghavendra, *Optimal algorithms and inapproximability results for every CSP?*, in Proceedings of the 40th Annual ACM Symposium on Theory of Computing, Victoria, British Columbia, Canada, May 17-20, 2008, 2008, pp. 245–254.

[20] R. Raz, *A parallel repetition theorem*, SIAM J. Comput., 27 (1998), pp. 763–803.

[21] T. J. Schaefer, *The complexity of satisfiability problems*, in Proceedings of the Tenth Annual ACM Symposium on Theory of Computing, STOC '78, New York, NY, USA, 1978, ACM, pp. 216–226.

[22] L. Trevisan, *Non-approximability results for optimization problems on bounded degree instances*, in Proceedings on 33rd Annual ACM Symposium on Theory of Computing, July 6-8, 2001, Heraklion, Crete, Greece, 2001, pp. 453–461.

[23] D. Zhuk, *A proof of the CSP dichotomy conjecture*, J. ACM, 67 (2020), pp. 30:1–30:78.

Appendix

In this section we show that weighted versions of Max/Min-CSPs have *essentially*[8] the same approximation ratios as their unweighted counterparts. We reiterate that this was already shown in [16, Lemma 3.11].

We first consider a reduction which fixes some precision, rounds each weight to the "nearest rational number", and then converts this instance to the unweighted one by repeating the constraints appropriate number of times. In the following lemma we show that the instance obtained by this procedure attains similar values to the starting instance under any assignment to the variables.

[8] We allow $o(1)$ additive error terms.

Lemma 15. *Consider a weighted instance $\mathcal{F}$ of a Max-CSP (or Min-CSP) Λ. Then, for each $\varepsilon > 0$ there is a poly-time algorithm which outputs an unweighted instance $\mathcal{G}$ of the same CSP Λ over the same variables as in $\mathcal{F}$ such that*

$$\mathrm{Val}_\zeta(\mathcal{G}) - \varepsilon \le \mathrm{Val}_\zeta(\mathcal{F}) \le \mathrm{Val}_\zeta(\mathcal{G}) + \varepsilon, \tag{53}$$

where ζ is any assignment to the variables of $\mathcal{F}$ (or $\mathcal{G}$). Furthermore, the size of $\mathcal{G}$ is polynomial in size of $\mathcal{F}$ and $1/\varepsilon$.

Proof. Let $\mathcal{F}$ be an instance over constraints $C_1, \ldots, C_m$, with respective weights $w_1, \ldots, w_m$. We fix $q = \lceil m/\varepsilon \rceil$, and construct $\mathcal{G}$ by creating ℓ_r copies of each constraint C_r, where ℓ_r are chosen such that

$$\sum_r \ell_r = q, \quad \frac{\ell_r}{q} \in \left(w_r - \frac{1}{q}, w_r + \frac{1}{q} \right). \tag{54}$$

We can find such $\{\ell_r\}_{r=1}^m$ by setting first $\ell_r = \lfloor w_r q \rfloor$, and then incrementing some ℓ_i to obtain $\sum_r \ell_r = q$.

For any given assignment ζ to the variables, the contribution towards the value of $\mathcal{F}$ of each constraint C_r in $\mathcal{F}$ is at most $1/q$ different from contributions of replicated constraints in $\mathcal{G}$. Finally, since we have m constraints, the main claim

$$\mathrm{Val}_\zeta(\mathcal{G}) - \varepsilon \le \mathrm{Val}_\zeta(\mathcal{F}) \le \mathrm{Val}_\zeta(\mathcal{G}) + \varepsilon \tag{55}$$

of the theorem follows. □

By relying on this lemma, we can show that weights do not affect the approximability of Max/Min-CSPs, as long as we allow for an additive loss of ε in the approximation ratio. We first prove this claim for Max-CSPs.

Theorem 16 (Theorem 7, restated)**.** *Consider a Max-CSP Λ and assume that we can approximate the optimal value of unweighted instances within a multiplicative factor α. Then, for every $\delta > 0$, weighted instances of Max-CSP Λ can be approximated within a constant $\alpha - \delta$.*

Proof. Without loss of generality let us assume that Λ does not contain a predicate $P \equiv 0$, since we can remove each constraint with a predicate $P \equiv 0$ from an instance and rescale the weights, which does not affect approximability in the discussion that follows since the ratio between the values under any two assignments remains the same.

Now, let us fix a weighted instance $\mathcal{F}$ of a CSP Λ, and consider a random assignment χ in which each variable takes value 0 or 1 with probability 1/2, independently. Then, the expected value of $\mathcal{F}$ under this random assignment is

$$\mathbf{E}_\chi \left[\sum_{r=1}^m w_r P_r(\chi(S_r)) \right] = \sum_{r=1}^m w_r \, \mathbf{E}_\chi \left[P_r(\chi(S_r)) \right]. \tag{56}$$

The value $\mathbf{E}_\chi[P_r(\chi(S_r))]$ depends only on the properties of the predicate P_r. Furthermore, by our assumption $P_r \not\equiv 0$, and therefore $\mathbf{E}_\chi[P_r(\chi(S_r))] > 0$. Thus, since $P_r, r = 1, \ldots, m$ are picked from a finite collection of predicates of Λ, there is a $\gamma > 0$ such that $\mathbf{E}_\chi[P_r(\chi(S_r))] \ge \gamma$, for every $r \in [m]$. Therefore, we have that

$$\mathbf{E}_\chi \left[\sum_{r=1}^m w_r P_r(\chi(S_r)) \right] = \sum_{r=1}^m w_r \, \mathbf{E} \left[P_r(\chi(S_r)) \right] \ge \sum_{r=1}^m w_r \gamma = \gamma. \tag{57}$$

Hence, under the randomized assignment, the instance has a value of at least γ in expectation. By the averaging argument, we have that $\mathrm{Opt}(\mathcal{F}) \ge \gamma$. Now, consider the algorithm from Lemma 15 with parameter $\varepsilon = \delta\gamma/2$, which takes our instance $\mathcal{F}$ and outputs unweighted instance $\mathcal{G}$. We can apply the α approximation algorithm on $\mathcal{G}$ to obtain some assignment ζ for which

$$\frac{\mathrm{Val}_\zeta(\mathcal{G})}{\mathrm{Opt}(\mathcal{G})} \ge \alpha. \tag{58}$$

Then, since $\mathrm{Opt}(\mathcal{G}) \geq \gamma$, for the same assignment ζ we have

$$\begin{aligned} \frac{\mathrm{Val}_\zeta(\mathcal{F})}{\mathrm{Opt}(\mathcal{F})} \geq \frac{\mathrm{Val}_\zeta(\mathcal{G}) - \varepsilon}{\mathrm{Opt}(\mathcal{G}) + \varepsilon} \geq \frac{\mathrm{Val}_\zeta(\mathcal{G}) - \varepsilon}{\mathrm{Opt}(\mathcal{G})}(1 - \varepsilon/\gamma) \\ \geq \frac{\mathrm{Val}_\zeta(\mathcal{G})}{\mathrm{Opt}(\mathcal{G})} - \frac{\varepsilon}{\mathrm{Opt}(\mathcal{G})} - \frac{\varepsilon \,\mathrm{Val}_\zeta(\mathcal{G})}{\gamma \,\mathrm{Opt}(\mathcal{G})} + \frac{\varepsilon^2}{\gamma \,\mathrm{Opt}(\mathcal{G})} \geq \alpha - \frac{\varepsilon}{\gamma} - \frac{\varepsilon}{\gamma} \geq \alpha - \delta, \end{aligned} \tag{59}$$

which proves the statement of the theorem. □

The argument from the previous theorem does not work for Min-CSPs, since in this case $\mathrm{Opt}(\mathcal{G})$ can be arbitrarily small. The analogous claim for Min-CSPs was already proved in [16, Lemma 3.11] by using scaling techniques [14, 10]. For the sake of completeness, we give here somewhat more detailed proof of this claim, using essentially the same techniques.

Theorem 17 (Theorem 8, restated)**.** *Consider a Min-CSP Λ, and assume we can approximate the optimal value of unweighted instances within a multiplicative factor α. Then, for every $\delta \in (0, 1)$, weighted instances of the Min-CSP Λ can be approximated within a constant $\alpha + \delta$.*

Proof. Consider the decision version of CSP Λ, in which we ask whether there is an assignment χ to the variables such that all the constraints of Λ are *not* satisfied. By Schaefer's dichotomy theorem [21], the problem of deciding whether there is an assignment which falsifies all the constraints is either NP-hard or in P. If solving this problem is NP-hard, then both the weighted and the unweighted versions of Min-CSP Λ are obviously NP-hard to approximate within any constant. Therefore, without loss of generality, we assume that deciding whether all constraints can be falsified is in P for Λ.

Hence, given an instance $\mathcal{F}$ of the Min-CSP Λ, we can check in polynomial time whether $\mathrm{Opt}(\mathcal{F}) = 0$. In case $\mathrm{Opt}(\mathcal{F}) = 0$, we have found an optimal assignment, so it only remains to consider $\mathrm{Opt}(\mathcal{F}) > 0$.

Without loss of generality let us assume that the weights of the constraints $w_i, i = 1, \ldots, m$ are sorted in descending order, i.e. $w_1 \geq w_2 \geq \ldots \geq w_m$. We can find in polynomial time the largest $k \geq 1$ such that there is an assignment falsifying each of the constraints $C_1, C_2, \ldots, C_{k-1}$.

For thusly chosen k at least one of $C_1, \ldots, C_k$, will be true in any assignment, so we have that $\mathrm{Opt}(\mathcal{F}) \geq w_k$. Also, since there is an assignment falsifying the first $k-1$ constraints, we have that $\mathrm{Opt}(\mathcal{F}) \leq \sum_{i=k}^m w_i \leq w_k m$.

Let us partition the constraints C_i into the following three groups:

- *light*: constraints C_i with weight $w_i \leq w_k/m^2$.
- *medium*: constraints C_i with weight $w_k/m^2 < w_i < w_k m^2$.
- *heavy*: constraints C_i with weight $w_k m^2 \leq w_i$.

Then, we create an instance $\mathcal{F}'$ by adding medium and heavy constraints C_i from $\mathcal{F}$. Furthermore, we scale down the weights of heavy constraints to $w_k m^2$ in $\mathcal{F}'$. Finally, we normalize the weights by multiplying them by some factor $\sigma > 1$. Note that $\mathrm{Opt}(\mathcal{F}') \geq w_k\sigma$, since $\mathcal{F}'$ still has (although with different weights) constraints $C_1, C_2, \ldots, C_k$. Now, using the algorithm from Lemma 15 with $\varepsilon = \frac{\delta w_k \sigma}{8\alpha}$ we construct a regular instance $\mathcal{G}$, and find an assignment ζ for $\mathcal{G}$ which gives us an α approximation of the optimal value for $\mathcal{G}$ using the assumption that unweighted instances can be approximated within α. Observe that the size of $\mathcal{G}$ is polynomial, since $w_k\sigma \geq m^{-3}$ and therefore $1/\varepsilon \in \mathrm{poly}(n)$. Let us now show that ζ approximates optimal value of $\mathcal{F}'$ within $\alpha + \delta/2$. First, due to (55) we have

$$\frac{\mathrm{Val}_\zeta(\mathcal{F}')}{\mathrm{Opt}(\mathcal{F}')} \leq \frac{\mathrm{Val}_\zeta(\mathcal{G}) + \varepsilon}{\mathrm{Opt}(\mathcal{F}')} = \frac{\mathrm{Val}_\zeta(\mathcal{G})}{\mathrm{Opt}(\mathcal{F}')} + \frac{\varepsilon}{\mathrm{Opt}(\mathcal{F}')} \leq \frac{\mathrm{Val}_\zeta(\mathcal{G})}{\mathrm{Opt}(\mathcal{G}) - \varepsilon} + \frac{\frac{\delta w_k \sigma}{8\alpha}}{\mathrm{Opt}(\mathcal{F}')}. \tag{60}$$

Now, since $\mathrm{Opt}(\mathcal{F}') \geq w_k\sigma$ and $\alpha > 1$ we have that $\frac{\delta w_k \sigma/(8\alpha)}{\mathrm{Opt}(\mathcal{F}')} \leq \delta/4$, and hence it remains to show that $\frac{\mathrm{Val}_\zeta(\mathcal{G})}{\mathrm{Opt}(\mathcal{G}) - \varepsilon} \leq \alpha + \delta/4$. We have

$$\frac{\mathrm{Val}_\zeta(\mathcal{G})}{\mathrm{Opt}(\mathcal{G}) - \varepsilon} = \frac{\mathrm{Val}_\zeta(\mathcal{G})}{\mathrm{Opt}(\mathcal{G})} + \frac{\varepsilon \,\mathrm{Val}(\mathcal{G})}{\mathrm{Opt}(\mathcal{G}) \cdot (Opt(\mathcal{G}) - \varepsilon)} \leq \alpha + \frac{\varepsilon\alpha}{\mathrm{Opt}(\mathcal{G}) - \varepsilon}. \tag{61}$$

Finally, since $\mathrm{Opt}(\mathcal{G}) - \varepsilon \geq \mathrm{Opt}(\mathcal{F}') - 2\varepsilon$ by (55), and since $\mathrm{Opt}(\mathcal{F}') \geq w_k\sigma$, we have that

$$\mathrm{Opt}(\mathcal{G}) - \varepsilon \geq w_k\sigma - 2\varepsilon \geq w_k\sigma - 2\frac{\delta w_k \sigma}{8\alpha} \geq w_k\sigma - \frac{w_k\sigma}{4} \geq \frac{w_k\sigma}{2}. \tag{62}$$

Using this estimate in (61) we get

$$\frac{\mathrm{Val}_\zeta(\mathcal{G})}{\mathrm{Opt}(\mathcal{G}) - \varepsilon} \geq \alpha + \frac{\varepsilon\alpha}{w_k\sigma/2} = \alpha + \frac{\frac{\delta w_k \sigma}{8\alpha}\alpha}{w_k\sigma/2} \geq \alpha + \delta/4. \tag{63}$$

Hence, $\frac{\mathrm{Val}_\zeta(\mathcal{F}')}{\mathrm{Opt}(\mathcal{F}')} \geq \alpha + \delta/2$.

Let us now see how well ζ approximates the optimal value of $\mathcal{F}$. We have that the following two properties hold:

- **Property A:** $\mathrm{Opt}(\mathcal{F}) \geq \frac{1}{\sigma}\mathrm{Opt}(\mathcal{F}')$. This holds since optimal values of both $\mathcal{F}$ and $\mathcal{F}'$ do not satisfy heavy constraints.
- **Property B:** If an assignment χ does not satisfy heavy constraints, then $\mathrm{Val}_\chi(\mathcal{F}) \leq \frac{1}{\sigma}\mathrm{Val}_\chi(\mathcal{F}') + w_k/m$. This statement holds since if we do not satisfy heavy (scaled down) constraints, then the only difference comes from light constraints, which can have a total weight of at most w_k/m.

Finally, consider our $(\alpha+\delta/2)$-approximating assignment ζ to the instance $\mathcal{F}'$. This assignment certainly does not satisfy heavy constraints, since otherwise we have $\mathrm{Val}_\zeta(\mathcal{F}') \geq w_k m^2\sigma$, and $\mathrm{Opt}(\mathcal{F}') \leq w_k m\sigma$, so the approximation ratio would be at least m, which cannot happen since ζ achieves a constant factor approximation. Therefore, by using the properties **A** and **B** for ζ we have

$$\frac{\mathrm{Val}_\zeta(\mathcal{F})}{\mathrm{Opt}(\mathcal{F})} = \frac{\sigma\,\mathrm{Val}_\zeta(\mathcal{F})}{\sigma\,\mathrm{Opt}(\mathcal{F})} \leq \frac{\mathrm{Val}_\zeta(\mathcal{F}') + \sigma w_k/m}{\mathrm{Opt}(\mathcal{F}')} \leq \frac{\mathrm{Val}_\zeta(\mathcal{F}')}{\mathrm{Opt}(\mathcal{F}')} + \frac{\sigma w_k/m}{\sigma w_k} \leq \alpha + \delta/2 + \frac{1}{m}. \tag{64}$$

Since $1/m < \delta/2$ for sufficiently large m, our theorem holds. □