The Shortest Vector Problem in L_2 is NP-hard for Randomized Reductions.

M. Ajtai
IBM Almaden Research Center

Abstract. We show that the shortest vector problem in lattices with L_2 norm is NP-hard for randomized reductions. Moreover we also show that there is an absolute constant $\epsilon > 0$ so that to find a vector which is longer than the shortest non-zero vector by no more than a factor of $1 + 2^{-n^\epsilon}$ (with respect to the L_2 norm) is also NP-hard for randomized reductions. The corresponding decision problem is NP-complete for randomized reductions.

1. Introduction. A lattice in $\mathbf{R}^n$ is the set of all integer linear combinations of n fixed linearly independent vectors. The question of finding the shortest non-zero vector in a lattice with repsect to the L_∞ was proved to be NP-hard by Van Emde Boas. However the corresponding problem for the L_2 norm (or any other L_p-norms for $1 \leq p < \infty$) remained unsolved. Van Emde Boas conjectured almost twenty years ago (cf. [vEB]) that the L_2 shortest vector problem for lattices in $\mathbf{Z}^n$ is NP-hard and the corresponding decision problem is NP-complete.

The α-approximate version of the problem is the following: find a non-zero vector v in the lattice L so that its length is at most $\alpha\|v_0\|$ where v_0 is a shortest non-zero vector of the lattice. It has been shown by J. Lagarias, H.W Lenstra and, C. P. Schnorr (cf. [LLS]) that if the α-approximate problem is NP-hard for any $\alpha > n^{1.5}$ (where n is the dimension of the lattice) than $NP = co - NP$.

In this paper we show that the shortest vector problem is NP-hard for randomized reductions. That is, there is a probabilistic Turing-machine which in polynomial time reduces any problem in NP to instances of the shortest vector problem. In other words this probabilistic Turing machine can solve in polynomial time any problem in NP, provided that it can use an oracle which returns the solution of the shortest vector problem if an instance of it presented (by giving a basis of the corresponding lattice). We prove the same result about the $1 + 2^{-n^\epsilon}$-approximate problem where $\epsilon > 0$ is a sufficiently small absolute constant and n is the dimension ot the lattice.

Adleman proved in 1995 (see [Adl]) that factoring integers can be reduced to the shortest vector problem in random polynomial time, using some very reasonable but unproved assumptions. The work of the present paper has started as an attempt to give a proof of Adleman's theorem without the unproven assumptions.

Adleman has defined a lattice for his proof, using the logarithms of primes. This lattice plays a crucial role in our proof as well. Actually we use a modified and extended version of the original lattice. Adleman has used his lattice to factor the integer n. By finding a short vector in the lattice it is possible to find supersmooth congruences modulo n, that is, different products of primes (not greater than a polynomial of $\log n$) which are congruent modulo n.

Let $p_1, ..., p_m$ be the first m prime numbers, where m is polynomial in $\log n$. Adleman has defined a lattice V whose basis vectors are rational approximations of the rows of the following matrix:

$$\begin{pmatrix} \sqrt{\log p_1} & \dots & 0 & M \log p_1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & \sqrt{\log p_m} & M \log p_m \\ 0 & \dots & 0 & M \log n \end{pmatrix}$$

The number M is defined so that it balances the contribution of the last component to the norm against the contribution of all of the others. The essential property of the lattice V is that if the length of the shortest non-zero vector v is below a certain bound and its coefficients in the given basis are $\gamma_1, ..., \gamma_{m+1}$ and $P = \prod\{p_i^{\gamma_i} \mid \gamma_i \geq 0\}$, $Q = \prod\{p_i^{\gamma_i} \mid \gamma_i < 0\}$ then $\log n + \sum_{i=1}^{m} \log p_i$ is close to 0, that is $n\frac{P}{Q}$ is close to 1 and therefore $s = nP - Q$ is so small that all of its prime factors are among $p_1, ..., p_m$. Therefore from such a small vector we get a congruence modulo n among the products of small primes. (An unproved assumption, formulated in a different more natural way, guarantees that the shortest vector is really below the given bound.)

For the motivation of the defintion of our extended lattice we note that if n is square free and of the form $n = \prod_{i=1}^{m} p_i^{\gamma_i}$ then (with the right choice of M) $\langle -\gamma_1, ..., -\gamma_m, 1\rangle$ will be a very short vector in the lattice. Of course this does not help much because in this case we can find $\gamma_1, ..., \gamma_m$ easily without using the lattice. We add a new basis vector to the lattice in a way that even if n has larger prime factors, but one of the numbers $n + l\omega$ (ω is a fixed intger $l = \pm 1, ..., \pm[n^\epsilon]$, $\epsilon > 0$) is good, that is, it is squarefree and of the form $n + l\omega = \prod_{i=1}^{m} p_i^{\gamma_i}$ then the vector $\langle -\gamma_1, ..., -\gamma_m, 1, l\rangle$ is a very short vector in the lattice. Actually we will be able to this in a way that the shortest vector must be of this form if there is at least one good number in the given segment of the arithmetic sequence $n + l\omega$. (We will prove that for a random choice of n, with a suitable distribution, with a positive probability there is always at least $2^{n^{\epsilon'}}$ good numbers.)

It may seem surprising that the additive structure of the arithmetic sequence $n + l\omega$ fits in the multiplicative structure of numbers only with small primes. The reason is that $n + l\omega$ can be very well approximated by $n(1 + \frac{\omega}{n})^l$, so the structure is approximately multiplicative. We may consider $1 + \frac{\omega}{n}$ as a “new prime” and add the corresponding row to the lattice.

Since finding the shortest vector in this extended lattice may be helpful in factoring ω (and not n) we change the notation somewhat in the following definition of the lattice L. The dimension of the extended lattice L is larger by one than the dimension of V and we get it by essentially adding a new basis vector to the lattice. More precisely the rows of the following matrix will be the basis of the lattice L:

$$\begin{pmatrix} \sqrt{\log p_1} & \dots & 0 & 0 & B\log p_1 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \dots & \sqrt{\log p_\iota} & 0 & B\log p_\iota \\ 0 & \dots & 0 & 0 & B\log b \\ 0 & \dots & 0 & \omega^{-\kappa} & B\log(1+\frac{\omega}{b}) \end{pmatrix}$$

As we will show, with the right choices for the various parameters, this lattice will have the following property:

Suppose that $v_1, ..., v_{m+2}$ are the rows of the given matrix and $v = \sum_{i=1}^{m+2} \gamma_i v_i$ is a nonzero vector of the lattice generated by $v_1, ..., v_{m+2}$ with $\gamma_{m+1} \geq 0$ and $\|v\| \leq (\log b + \omega^{-1})^{\frac{1}{2}}$. Then $\gamma_{m+1} = 1$, $\gamma_1, ..., \gamma_m \in \{0, -1\}$ and if $g = \prod_{i=1}^{m} p_i^{|\gamma_i|}$ then $g = b + l\omega$ for some $l = 0, \pm 1, ..., \pm[b^c]$. We show that the converse of this statement is also true, that is, if $g = b + l\omega$ for some $l = 0, \pm 1, ..., \pm[b^c]$ and g is of the form described above then the corresponding vector is shorter than $(\log b + \omega^{-1})^{\frac{1}{2}}$. It is also true that the length of every nonzero vector is at least $(\log b)^{\frac{1}{2}}$. (This result is formulated in Lemma 3.2.)

The original motivation for the definition of L was the following. We try to factor the integer ω. Let us pick a random residue class d modulo ω (which is generated as a random a product of random powers of the primes $p_1, ..., p_m$.) After that, we take a representative b of appropriate size (a constant power of ω) from the residue class of d. The important point is that b does not give any information about which representation of d (as a product of small primes) is known to us. Then, by finding a short vector in the lattice, we could get a congruence $b \equiv \prod_{i=1}^{m} p_i^{|\gamma_i|} \pmod{\omega}$. By taking enough different numbers b we could get enough congruences to factor ω. The problematic and still not completed part of this direct reduction of factoring to the shortest vector problem, is to guarantee that for a b generated by this distribution there will be a good number in the arthmetic sequence $b + l\omega$, $l = 0, \pm 1, ..., \pm[b^c]$. (We have some partial results in this direction that we describe at the end of the introduction.) While trying to fill this gap in the proof we noticed that there is another distribution for b where it can be proved easily that with a probability of at least $\frac{1}{2}$ there are always good numbers in the given segment of the arithmetic sequence. Namely b will be the product of h distinct elements of the set $\{p_1, ..., p_m\}$, for a suitably chosen fixed h, with uniform distribtuion on the set of all products with this property. (In this case the number b reveals its known expression as the product of small primes, so we cannot gaurantee that we get enough independent congruences for factoring.) However, for this distribution, we have that with a probability of at least $\frac{1}{2}$, the lattice L has an exponential number of vectors v with $\|v\|^2 < \log b + n^{-1}$.

At this point in the attempted proof we have realized that the lattice L can be used to prove the NP-hardness of the shortest vector problem (even in an approximate sense) for randomized reductions. Namely the lattice has the nice property that it has a basis where all small vectors have $0, 1$ coefficients only (with the exception of the last coefficient), moreover the number of vectors that are small enough to have this nice property is exponential (in the dimension of the lattice.) This creates a possibility to reduce the subset sum problem to the shortest vector problem. (The NP-completeness of the subset sum problem was proved by Karp (see [K])). We will look for a solution of the subset sum problem among the coefficiemt sequences of the short vectors. Although not every $0,1$ sequence will occur as a coefficient sequence, still the number of different 0,1 coefficient sequences is exponential in the size of the subset sum problem. This will make it possible to find the solution of the subset sum problem embedded in some way into a coefficient sequence of a short vector. We will be able to search among these coefficient sequences for a solution of the subset sum problem, by embedding our lattice L into another larger dimensional Euclidean space and therefore change the length of its vectors. According to this new (L_2) norm every short vector will be short in the original sense too, but they also have to satisfy some additional requirements. By defining the embedding in a suitable way this additional requirement can be that the coefficient sequence (in some modified form) is a solution of the subset sum problem. Therefore, by finding the shortest vector in the embedded lattice (or one which is approximately the shortest with an exponentially small error) we will be able to solve the subset sum problem.

As we mentioned above, at some point in the proof, we have to make a transition from an exponential number of $0,1$-sequences (coefficient sequences of short vectors), to all $0,1$ sequences of a certain length, because otherwise we cannot guarantee that the solution of the subset sum vector problem is among them. As a model for such a transition we use a theorem of Sauer about hypergraphs. (This theorem is related to the notion of VC-dimension. We cannot use the theorem itself in the proof because it is not constructive enough, but we give constructive analogue of it, at least in a probabilistic sense.) Sauer's theorem (we give the exact statement in the next section), states that if a set X is a set of subsets of the set S and $|X|$ is above a certain bound depending on $|S|$ and k then there is a $Y \subseteq X$, $|Y| = k$ so that every subset of Y occurs among the sets $Y \cap Z$, $Z \in X$. We may think that S is the basis of the lattice (excluding the last exceptional basis vector) and for each short vector v in the lattice we have $T_v \in X$, where T_v is the set of elements of S where the corresponding coefficient of v is 1. (All of the other coefficients are 0.) Therefore the theorem gaurantees that there is a subset Y of the basis vectors where we get all $0,1$-sequences as coefficient sequences of short vectors. We may look for the solution of the subset sum problem among the $0,1$-sequences defined on Y. (Sauer's theorem also guarantees that $|Y|$ is large enough.) The only problem with this approach is that we do not have any method which could find such a Y in polynomial time. A random choice for Y is not satisfactory because X can be unevenly distributed in S. We formulate an analogue of Sauer's theorem where a a random Y is a good solution with a probability close to 1. Namely instead of taking a single subset Y we will take a sequence of pairwise disjoint subsets $C = \langle C_1, ..., C_k \rangle$ and for each $T \in X$ we define a function $f_T(i) = |T \cap C_i|$ on $\{1, ..., k\}$. We show that if every element of X is of the same appropriate size, $|X|$ contains sufficiently many sets and we take the sequence $C_1, ..., C_k$ at random, then with a probability close to one we get every $0,1$ functions on $\{1, ..., k\}$ in the form of f_T for a suitable $T \in X$. Theorem 2.2 in the next section is an exact formulation of this statement. The proof of this theorem is the technically most difficult part of our paper.

There is a third part of the proof, the embedding of L into another Euclidean space which defines a new norm on it. We may think that we do this by simply adding new columns to the matrix defining the lattice L. E.g. if we would want to guarantee that from a shortest vector $v = \sum_{i=1}^{m} \gamma_i v_i$ we get the solution of a subset sum problem $\sum_{i=1}^{m} a_i x_i = b$ in the form $x_i = \gamma_i$, $i = 1, ..., m$, then we can add a last column to the matrix whose elements are $-a_1 K, ..., -a_n K, bK, 0$. With the right choice of the number K we get that if there is a solution of the subset sum problem of the form $x_i = \gamma_i$, then the shortest vector in the extended lattice L' (that is in the lattice whose basis consists of the new longer rows), must provide such a coefficient sequence γ_i. Of course this is an idealized situation, because in general we do not get every $0, 1$ sequence of length m in the form of $\gamma_1, ..., \gamma_m$ from a short vector v. However combining this idea with the described generalization of Sauer's theorem we can conclude the proof. (See Lemma 2.2 and Corollary 2.2 for an exact formulation of the related results. Corollary 2.1 is only included as an illustration based on the simplified picture described above.)

We have described three parts of the proof, and as we have said, the second one, the constructive analogue of Sauer's theorem, is the most difficult in a technical sense. Still we feel that the most difficult step in the proof was to find the three different parts which together imply our main result and not the proofs of the individual components.

Both the choice of the lattice L and the choice of the sequence $C_1, ..., C_k$ which determines the lattice L' is probabilistic. In the defintion of L we choose the number b at random, with a given distribtuion, and we are able only to prove that the lattice L has the required property with a probability greater than $\frac{1}{2}$. (For the proof that L is good for every choice of b we would need very strong statements about the uniformity of the distribution of numbers with small prime factors.) The second part of the proof, the construction of L' through $C_1, ..., C_k$ is also probabilistic. Still there is a possibility that a single sequence $C_1, ..., C_k$ selected at random can be replaced by a polynomial number of deterministically constructed sequences so that for each fixed subset sum problem at least one of them is good.

Remarks. 1. Adleman has defined the lattice V, so that the coordinates of the basis vectors are rationals, they are approximations of the coordinates of the rows of the first matrix. This way it is possible to perform a computation whose input is the lattice. We define the lattice L as a lattice in $\mathbf{R}^{m+2}$, that is, the coordinates of the basis vectors can be irrational. We will prove all the necessary properties of the lattice L, and then show that L can be approximated by a lattice $\bar{L} \subseteq \mathbf{Q}^{m+2}$ (that is, the basis vectors of $\bar{L}$ has rational coordinates) and show that $\bar{L}$ still has the nice properties of L that are needed for the rest of the proof. Unfortunately both approaches create technical difficulties.

2. The lattice L has been originally defined for a completely different purpose than its final use in this paper. Therefore it may easily happen that other, perhaps in some sense simpler, lattices also have the properties that are required from L to complete the proof.

In the next section we collect these properties in Lemma 2.1. There are different reasons which may motivate the search for such a lattice: to make the proof deterministic; to improve the factor in the approximation result; to make the proof simpler.

3. As we have described above we break down the proof into three different parts. We tried to make the various parts independent, so in each part we are using as little information from the other parts as possible. E.g. we formulate what are the important properties of the lattice L and then use only these properties in the other parts of the proof. This makes the proof more transparent, but we may lose several possibilities for improvements. (Actually the original version of the proof utilized more of the specific properties of the lattice L.)

4. There are many related problems to the shortest vector problem, the nearest vector problem, the nearest codeword problem etc. (for the definitions see e.g. [ABSS]). Van Emde Boas has proved the NP-hardness of the nearest vector problem in all L_p norms $1 \leq p \leq \infty$. Arora, Babai, Stern and Sweedyk (cf. [ABSS]), has shown that even to approximate the nearest vector within a constant factor is NP-hard. A. Vardy has proved recently (cf. [V1] or [V2]) that deciding whether there is a codeword within a given distance is NP-complete. The interested reader may find more detailed information about these and other related problems in [ABSS] or [V2].

5. The motivation for proving that to find short vectors in lattices is hard, in some sense, comes partly from cryptography. (See [Ajt], [AD], [GGH1], [GGH2]). There are cryptosystems whose security is based on the assumtpion that to find short vectors in lattices is computationally infeasible. Since these assumptions imply the hardness of even an approximation of the shortest vector utpo a polynomial factor, we cannot really hope to prove their NP-hardness, since, as we noted earlier, it would imply $NP = co-NP$, at least if the exponent is larger than 1.5, (cf. [LLS]). The fact that the there is no short vector in the lattice can be demonstrated by giving a short basis in its dual (reciprocal) lattice. Still, the fact that closely related lattice problems are NP-hard, makes the hardness assumptions of the cryptographic systems more credible. Therefore it would be particularly important to improve the α-approximate version of our theorem by proving it for greater values of α.

6. Although in this paper we reduce every problem in NP to the shortest vector problem (by random reduction) and so in particular we also reduce factoring to the shortest vector problem, a direct reduction like Adleman's, gives more information about factoring, and also shows that the shortest vector problem is more closely related to factoring than to other problems in NP. Therefore we think that it would be still very important to show (without any unproven assumptions) that Adleman's reduction, or any other direct reduction of factoring, is correct. We have already partly described an attempt of such a proof. Although we are not able to prove that this reduction works for all intger n we can show that it works for almost all n (the number of exceptional integers is exponentially small in $\log N$) that are generated as the products of two primes choosen independently and with uniform distribution from the set of all primes less than N. (Analogue statements can be proved about the products of not two but a constant number of primes. At this point it

is not clear whether there is a limit on the number of primes.) The same technique also gives that the discrete logarithm problem for almost all primes can be directly reduced to the shortest vector problem. We intend to return to these questions in a separate paper. There are signs that indicate that the "almost all" restriction may be avoided. E.g. we can prove that for each number n we get at least one nontrivial congruence among the product of small primes (actually we can prove that we get many, just we cannot prove that there are enough for factoring.) Here we only mention that the proof (of the results concerning factoring) depend on the method of Large Sieve, which was originally developed by Linnik and Rényi and is one of the "elementary" but very powerful methods of number theory. For the result about discrete logarithm the method can be used in its original form which, roughly speaking, says that a large set of integers in a relatively small interval are distributed almost uniformly in the residue classes modulo almost every small primes. For factoring we need a variant given by Montgomery which remains valid if we consider the distributions modulo small composite numbers as well.

2. The outline of the proof.

In this section we break down the proof of our main result (the theorem below) into essentially three different lemmata and show how they imply the theorem. We prove the three lemmata in different sections.

Theorem 2.1. *The shortest vector problem in L_2 is NP-hard for randomized reductions. Moreover there is an absolute constant $\epsilon > 0$ so that the $(1 + 2^{-n^{\epsilon}})$-approximate shortest vector problem is also NP-hard for randomized reductions.*

We prove the NP-hardness (for randomized reductions) of the shortest vector problem. We get the approximate version by minor modifications of the proof.

We will consider probabilistic Turing machines that are using an oracle which returns a solution of the shortest vector problem (in L_2) if a lattice is presented to the oracle (represented by a basis). We will show that the subset sum problem can be solved in polynomial time by such a machine. Since the subset sum problem is known to be NP-complete this will imply that the shortest vector problem is NP-hard for probabilistic reductions. The subset sum problem can be formulated in the following way: assume that $a_1, ..., a_l, b$ are integers, and we are looking for a solution of the equation $\sum_{i=1}^{l} a_i x_i = b$ where $x_i \in \{0,1\}$ for $i = 1, ..., l$. We define the size of the problem as $\log_2(|b| + 1) + l + \sum_{i=1}^{l} \log_2(|a_i| + 1)$. We will not use the subset sum problem directly but we will use a special case of it, which is still NP-complete.

Definition. *The restricted subset sum problem.* Suppose that $a_1, ..., a_l, b$ are integers, $\max\{\log_2(|b|+1), \max_{i=1}^{l} \log_2(|a_i|+1)\} \le l^3$ and we are looking for a solution of the system of equations $\sum_{i=l}^{l} a_i x_i = b$, $\sum_{i=1}^{l} x_i = [\frac{l}{2}]$ where $x_i \in \{0,1\}$ for $i = 1, ..., l$. The size of the problem is $l^3(l+1)$.

The (trivial) proof of the fact that the restricted subset sum problem is NP-complete will be given in another section. (Lemma 4.1)

We will consider problems whose inputs are lattices given in $\mathbf{Q}^n$ where $\mathbf{Q}$ is the field of rational numbers. The lattices will be presented by a basis. So, to define the size of the problem, we have to define the size of a basis.

Definitions. 1. If $a_1, ..., a_n$ are linearly independent vectors in $\mathbf{R}^m$ for some $m \geq n$, then we call the set $\{\sum_{i=1}^n \alpha_i a_i \mid \alpha_1, ..., \alpha_n$ are integers $\}$ a lattice. (Sometimes the defintion is given with the additional requirement $m = n$. Since the real subspace generated by L can be isometrically embedded in $\mathbf{R}^m$ the two definitions are not essentially different.)

2. If $r = \frac{p}{q}$ is a rational number, so that $(p, q) = 1$, then size$(r) = \log_2(|p| + 1) + \log_2(|q| + 1)$, size$(0) = 1$. If $v = \langle r_1, ..., r_n \rangle \in \mathbf{Q}^n$ then size$(v) = \sum_{i=1}^n$ size(r_i). If $v_1, ..., v_t$ is a sequence of vectors then size$(v_1, ..., v_t) = \sum_{i=1}^t$ size(v_i).

3. $\|v\|$ will denote the L_2 norm of the vector v unless it is explicitly stated otherwise.

The motivation for the following lemma is to make some connection between the shortest vector problem and the subset sum problem. The essence of the lemma is that that there exists a lattice in $\mathbf{Q}^m$ where all of the nonzero vectors whose length does not exceed a certain limit has only $0, 1$ coefficients in a fixed given basis, (with the possible exception of the last coefficient for each vector). Moreover the number of vectors below this limit is exponential in the dimension of the lattice. Later we will look for the solution of a subset sum problem among the $0, 1$ sequences which occur as a coefficient sequence of such a short vector. By embedding the lattice (linearly but not isometrically) into $\mathbf{Q}^{\bar{m}}$ for some $\bar{m} > m$, that is defining a new Euclidean norm on it, we will be able to gaurantee that the shortest vector in the new lattice has a coefficient sequence which provides a solution for the subset sum problem.

Lemma 2.1. *There are positive rationals c_1, c_2, c_3, c_4 so that for each sufficiently large positive integer n there is a positive integer $m \in [n^{c_1}, n^{2c_1}]$, a positive rational $\rho < 1$, a lattice $L \subseteq \mathbf{Q}^{m+2}$, and a basis $v_1, ..., v_m, v_{m+1}, v_{m+2}$ of L so that the following holds:*

(1) $v \in L$, $v \neq 0$ *implies* $\|v\| \geq 1$,

(2) *if Z is the set of all $v \in L$, $v = \sum_{i=1}^{m+2} \gamma_i v_i$ with $\sum_{i=1}^m \gamma_i = n$, then $|\{v \in Z \mid \|v\|^2 < 1 + \rho\}| \geq 2^{c_3 n \log n}$,*

(3) *For all $v \in L$, $v \neq 0$, if $\|v\|^2 < 1 + \rho 2^{n^{c_4}}$, $v = \sum_{i=1}^{m+2} \gamma_i v_i$, and $\gamma_{m+1} \geq 0$, then $\gamma_i \in \{0, 1\}$ for $i = 1, ..., m$, $\gamma_{m+1} = 1$.*

(4) *If $u_1 \neq u_2$, $\|u_j\|^2 < 1 + \rho 2^{n^{c_4}}$ and $u_j = \sum_{i=1}^m \gamma_i^{(j)} v_i$ for $j = 1, 2$ then there is a $i = 1, ..., m$ so that $\gamma_i^{(1)} \neq \gamma_i^{(2)}$,*

(5) size$(\rho) \leq n^{c_2}$, $0 < \rho < 2^{-n^{c_4}}$, size$(v_1, ..., v_{m+2}) \leq n^{c_2}$.

(6) $\sqrt{\rho}$ *is rational.*

Moreover there is a probabilistic Turing machine which for each input n gives the following output in time polynomial in n: an integer m, a rational $\rho > 0$ and linearly independent *vectors $v_1, ..., v_{m+2} \in \mathbf{Q}^{m+2}$ so that with a probability of at least $\frac{1}{2}$ if L is the lattice generated by $v_1, ..., v_{m+2}$ then $m, \rho, v_1, ..., v_{m+2}, L$ satisfy conditions (1),(2),(3),(4),(5),(6).*

We will give the proof of this lemma in a separate section. Our next goal is to define an embedding of the lattice L (given by the previous lemma) into a higher dimensional space so that any shortest nonzero vector of the new embedded lattice (which has a different metric since the embedding is only linear but not isometric), has a coefficient sequence (in the embedded image of the basis $\{v_i\}$) which provides a solution of the subset sum problem. Under certain additional assumptions we will be able to do this as our next lemma will show.

Definitions. Assume that $m, \rho, v_1, ..., v_{m+2}, L$ are fixed with the properties listed in Lemma 2.1.

1. We will use the following notation: $S_1 = \{x \in L \mid x \neq 0, \|x\|^2 < 1 + \rho, x \cdot v_{m+1} \geq 0\}$, $S_2 = \{x \in L \mid x \neq 0, \|x\|^2 \leq 1 + \rho 2^{n^{c_4}}, x \cdot v_{m+1} \geq 0\}$

2. For each $x \in \mathbf{R}^{m+2}$ if $x = \sum_{i=1}^{m+2} \gamma_i v_i$ then let $\Lambda(x)$ be the $m+1$ dimensional vector $\langle \gamma_1 \sqrt{\rho}, ..., \gamma_{m+1} \sqrt{\rho} \rangle$. (Note that we do not use the last component of x in this definition.) Suppose that m' is a positive integer and A is a matrix with m' rows and $m+1$ columns. If $x \in R^{m+2}$, then let $\psi_A(x)$ be the $m+2+m'$ dimensional vector $\langle x, A\Lambda(x) \rangle$. $L^{(A)} \subseteq \mathbf{R}^{m+m'+2}$ will be the lattice generated by the vectors $\psi_A(v_i)$, $i = 1, ..., m+2$.

3. $L^{(A)+}$ will denote the set of all $w \in A$ with $w = \phi_A(u)$ for some $u \in L$, with $u \cdot v_{m+1} \geq 0$. Clearly for all $w \in L^{(A)}$, at least one of the vectors $w, -w$ is in $L^{(A)+}$.

Lemma 2.2. *Assume that c_1, c_2, c_3, c_4 are fixed with the properties in Lemma 2.1, $c_5 > 0$, n is sufficiently large, $m, \rho, v_1, ..., v_{m+2}, L$ satisfy the conditions of Lemma 2.1, $m' \leq n^{c_5}$ is a positive integer, A is a matrix with m' rows and $m+1$ columns and all of the entries of A are integers with absolute values no larger than n^{c_5}. If there is a $v \in S_1$, so that $\|A\Lambda(v)\|^2 = \min\{\|A\Lambda(x)\|^2 | x \in S_2\}$ and w is a shortest nonzero vector of $L^{(A)}$ with $w \in L^{(A)+}$, then there is a $u \in S_2$ so that $\|A\Lambda(u)\| = \|A\Lambda(v)\|$ and $w = \psi_A(u)$.*

Definition. Assume that $Y \subseteq \{1, ..., m\}$, $Y = \{y_1, ..., y_l\}$, $|Y| = l$ and $L, v_1, ..., v_{m+2}$ are fixed with the properties given in Lemma 2.1. For each $v \in L$, $v = \sum_{i=1}^{m+2} \gamma_i v_i$, $g_{Y,v}$ will be a function defined by $g_{Y,v}(i) = \gamma_{y_i}$ for all $i \in \{1, ..., l\}$.

In the following Corollary we consider the subset sum problem without the additional restriction $\sum_{i=1}^{m} x_i = [\frac{2}{l}]$. This Corollary and the following Theorem are not parts of our final proof. We use them only to illustrate the main idea of the proof in a much simpler setup.

Corollary 2.1. *Assume that $c_1, ..., c_5$, $m, \rho, v_1, ..., v_{m+2}, L$ are as in Lemma 2.2, l is an integer with $1 \leq l \leq m$ and $Y = \{y_1, ..., y_l\} \subseteq \{1, ..., m\}$, $|Y| = l$. Suppose that $\sum_{i=1}^{l} a_i = b$ is an instance of the (not restricted) subset sum problem so that $x_i = g_{Y,v}(i)$*

$i = 1, ..., l$ is a solution of the problem for some $v \in S_1$, and $K > 0$ is sufficiently large. Let $D = \{d_{i,j}\}_{i=1,j=1,...,m+1}$ be the (1 by $m+1$) matrix defined by $d_{1,y_i} = Ka_i$ for $i = 1, ..., l$, $d_{1,m+1} = -Kb$, and $d_{1,j} = 0$ for all other j. If $w \in L^{(D)+}$ is a shortest nonzero vector in $L^{(D)}$, and $w = \psi_D(u)$, then $x_i = g_{Y,u}(i)$, $i = 1, ..., l$ is a solution of the given instance of the subset sum problem.

We do not prove this Corollary, since it is not part of our proof, but it easily can be proved even without Lemma 2.2. According to this Corollary we are able to solve the non-restricted subset sum problem by solving the shortest vector problem in $L^{(D)}$, provided that we can find an $Y \subseteq \{1, ..., m\}$, $|Y| = l$ so that the subset sum problem has a solution among the evaluations $x_i = g_{Y,v}(i)$, $v \in S_1$. The following theorem of Sauer guarantees the existence of such a set Y provided that $l \leq 2^{n^\delta}$, where $0 < \delta < 1$ and n is sufficiently large with repsect to δ and c_1. (For a proof see [S] or [AS]. The theorem is related to the notion of VC-dimesnion, see [VC].)

Definition. Assume that S is a finite set, X is a set of subsets of S. The pair $\langle S, X\rangle$ is called a hypergraph.

Theorem (Sauer). *If $\langle S, X\rangle$ is a hypergraph and $|X| > \sum_{i=1}^{k} \binom{|S|}{k}$ then there is a $Y \subseteq S$ with k elements such that every subset of Y occurs among the sets $Y \cap Z$, $Z \in X$.*

For each $v \in S_1$, $v = \sum_{i=1}^{m+2} \gamma_i v_i$ let $T_v = \{i \leq m \mid \gamma_i = 1\}$. We apply Sauer's theorem with $S = \{1, ..., m\}$, $X = \{T_v | v \in S_1\}$. We have that $|S| \leq n^{c_1}$ and by (1) and (4) of Lemma 2.1 $|X| \geq 2^{c_3 n \log n}$. Since $l \leq n^\delta$, $0 < \delta < 1$ and n is sufficiently large with respect to c_1 and δ the requirements of Sauer's theorem are met with $k \to l$. (Indeed $\sum_{i=1}^{l} \binom{|S|}{l} \leq l(n^{c_1})^l \leq e^{\delta \log n} e^{n^\delta c_1 \log n} < 2^{c_3 n \log n}$.) Therefore there is a set $Y \subseteq S$, $|Y| = l$ so that *every* $0, 1$ function on Y is of the form $g_{Y,v}$ for some $v \in S_1$. This clearly implies that the solution of the subset sum problem is among them. (The requirement $l \leq n^\delta$ does not cause any problem since if l is given and we pick the smallest n with $l \leq n^\delta$, then the corresponding basis of the lattice L will be still polynomial size in l.) Sauer's theorem therefore guarantees the existence of the required set Y, but it does not provide any way for finding it (finding Y in polynomial time would be necessary to complete the proof this way). Unfortunately the proof of the theorem is not contstructive (from our point of view). To be able to use the described approach we generalize the theorem, namely we replace the set Y with a more complicated structure in a way that makes the proof more constructive.

First we note that it would be sufficient for our proof if a random choice of Y would satisfy the requirements of the theorem with a positive (polynomially large) probability. (In this case by taking many random choices of Y we could get at least one which is good for our purposes.) Unfortunately this is not true if Y is uniformly distributed, since X can be concentrated on a small subset of S. In order to avoid this difficulty, instead of taking a subset Y with l elements, we will take a sequence of disjoint subsets $C_1, ..., C_l$ each with r elements, where r will be a suitably chosen positive integer. Now our goal is to get such a sequence $C_1, ..., C_l$ so that for *every* 0,1-function f on $\{1, ..., l\}$ there is a $T \in X$ with $f(i) = |C_i \cap T|$ for $i = 1, ..., l$. The original setup can be considered as a special case of this, namely when each C_i has exactly 1 element. We will formulate an analogue of Sauer's theorem with this new representation of the $0,1$-functions. We need some additional restrictions on the hypergraph $\langle S, X \rangle$, (which will hold in our case). Among other conditions we will assume that all of the sets in X has the same number of elements. This is not a real restriction since if X is of exponential size (in n) and S is of polynomial size then X always has a subset of exponential size which contains sets of the same cardinality. Fixing this common size helps to find the right choice for the other parameters in the theorem.

Definition. A hypergraph $\langle S, X \rangle$ is called n-uniform if $|T| = n$ for all $T \in X$.

Theorem 2.2. *For all $\alpha_1 > 0$, $\alpha_2 > 0$ there exist $0 < \delta_1 < 1$, $0 < \delta_2 < 1$, $0 < \delta_3 < 1$ so that for all sufficiently large n the following holds:*

Assume that $\langle S, X \rangle$ is an n-uniform hypergraph, $n^2 \leq |S| \leq n^{\alpha_1}$, $|X| \geq 2^{\alpha_2 n \log n}$, $k = [n^{\delta_1}]$ and $C_1, ..., C_k$ is a random sequence of pairwise disjoint subsets each with exactly $[|S| n^{-1-\delta_2}]$ elements, with uniform distribution on the set of all sequences with these properties. Then, with a probability of at least $1 - n^{-\delta_3}$ the following holds:

for each $0,1$-valued function f defined on $\{1, ..., k\}$ there is a $T \in X$ so that $f(j) = |C_j \cap T|$ for all $j = 1, ..., k$.

We will use this theorem to complete our proof. First we formulate an analogue of Corollary 2.1 which is also a corollary of Lemma 2.2 but will be used together with Lemma 2.2 (instead of Sauer's theorem). In this Corollary we will consider the restricted subset sum problem.

Definitions. 1. Assume that $C = \langle C_1, ..., C_l \rangle$ is a sequence of disjoint subsets of $\{1, ..., m\}$ and $L, v_1, ..., v_{m+2}$ are given with the properties described in Lemma 2.1. For each $v \in L$, $v = \sum_{i=1}^{m+2} \gamma_i v_i$, $g_{C,v}$ will denote a function defined by $g_{C,v}(i) = \sum_{j \in C_i} \gamma_j$.

2. Let $D = \{d_{i,j}\}_{i=1,\ldots,l+2,\ j=1,\ldots,m+1}$ be a matrix with $l + 2$ rows and $m + 1$ columns, defined by

(1) $d_{1,j} = a_i l^3$ for all $j \in C_i$, $d_{1,m+1} = -bl^3$ and $d_{1,j} = 0$ for all other values of j

(2) $d_{2,j} = l^3$ for all $j = \bigcup_{i=1}^{l} C_j$, $d_{2,m+1} = -[\frac{l}{2}] l^3$ and $d_{2,j} = 0$ for all other values of j,

(3) For all $i = 1, ..., l$, $d_{i+2,j} = 1$ if $j \in C_i$ and $d_{i+2,j} = 0$ otherwise

If $C_1, ..., C_l$ are consecutive intervals of $\{1, ..., m\}$ then D is the following matrix:

$$\begin{pmatrix} a_1l^3 & \dots & a_1l^3 & \dots & a_il^3 & \dots & a_il^3 & \dots & a_rl^3 & \dots & a_rl^3 & 0 & \dots & 0 & -bl^3 \\ l^3 & \dots & l^3 & \dots & l^3 & \dots & l^3 & \dots & l^3 & \dots & l^3 & 0 & \dots & 0 & -[\frac{l}{2}]l^3 \\ 1 & \dots & 1 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 0 & 0 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & \dots & 1 & \dots & 1 & \dots & 0 & \dots & 0 & 0 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 & \dots & 1 & \dots & 1 & 0 & \dots & 0 & 0 \end{pmatrix}$$

Corollary 2.2. *Assume that $c_1, ..., c_5$, $m, \rho, v_1, ..., v_{m+2}, L$ are as in Lemma 2.2, l is an integer with $1 \le l \le m$ and $C = \langle C_1, ..., C_l \rangle$, is a sequence of pairwise disjoint subsets of $\{1, ..., m\}$. Suppose that $\sum_{i=1}^{l} a_i = b$ is an instance of the restricted subset sum problem so that $x_i = g_{C,v}(i)$, $i = 1, ..., l$ is a solution of the problem for some $v \in S_1$. Let D be the $(l+2$ by $m+1)$ matrix defined above. If $w \in L^{(D)+}$ is a shortest nonzero vector in $L^{(D)}$, and $w = \psi_D(u)$, then $x_i = g_{C,u}(i)$, $i = 1, ..., l$ is a solution of the given instance of the restricted subset sum problem.*

Now, accepting Lemma 2.1, Lemma 2.2, Corollary 2.2 and Theorem 2.2 we can prove our main theorem. We describe a probabilistic algorithm $\mathcal{D}$ using an oracle which gives a shortest vector of the lattice presented to the oracle. Each use of the oracle will be counted as 1 time unit. $\mathcal{D}$ will find a solution of the restricted susbet sum problem $\sum_{i=1}^{l} a_i x_i = b$ (provided that such a solution exists) in time polynomial in l, with a probability exponentially close to one,

Assume now that an instance $\sum_{i=1}^{l} a_i x_i = b$ of the restricted subset sum problem (which has a solution) is given as an input. Let c_1, c_2, c_3, c_4 be the constants whose existence is stated in Lemma 2.1. Let $\delta_1, \delta_2, \delta_3$ be the numbers whose existence is guaranteed by Theorem 2.2 with $\alpha_1 \to 2c_1$, $\alpha_2 \to c_3$. Finally let $n = \lceil l^{\frac{1}{\delta_1}} \rceil$. We apply Lemma 2.1 for this n. According to the lemma $\rho, m, v_1, ..., v_{m+2}$ can be computed in time polynomial in n, and so in l too, with the properties listed in the lemma. (More precisely this happens with a probability of at least $\frac{1}{2}$ only. So we will perform the algorithm described below for the given values of $\rho, m, v_1, ..., v_{m+2}$, if we do not get a solution of the subset sum equation we repeat it with other random values. Performing this a polynomial number of times, with a probability exponentially close to 1 it will happpen at at least once that $\rho, m, v_1, ..., v_{m+2}$ satisfy the conditions of Lemma 2.1). We take a random sequence $C = \langle C_1, ..., C_l \rangle$ of pariwise disjoint subets of $S = \{1, ..., m\}$ each with exactly $[|S|n^{-1-\delta_2}]$ elements, with uniform distribution on the sets of all sequences with these properties. Clearly this randomization can be easily performed by picking the sets C_i recusrsively. $n = \lceil l^{\frac{1}{\delta_1}} \rceil$ implies that $l \le [n^{\delta_1}]$ therefore we may apply Theorem 2.2 with $\alpha_1 \to 2c_1$, $\alpha_2 \to c_3$, $S \to \{1, ..., m\}$, $k \to l$ and

$X \to \{Z \subseteq \{1,...,m\} \mid \exists v \in S_1,\ v = \sum_{i=1}^{m+2} \gamma_i v_i,\ Z = \{i \mid \gamma_i = 1, i = 1,...,m\}\}$.

According to the conclusion of the theorem with a probability of at least $1 - n^{-\delta_3}$ we have that each $0,1$ function $f(j)$ on $\{1,...,m\}$ is of the form $f(i) = |C_i \cap T|$ for some $T \in X$. If this is the case then there is a $T \in X$ so that $x_i = |C_i \cap T|$ is a solution of our subset sum problem. (We have assumed that there is a solution). By the definition of X and $g_{C,v}$ we have that there is a $v \in S_1$ so that $x_i = g_{C,v}$ is a solution of the subset sum problem. We may apply now Corollary 2.2 and get that if $w \in L^{(D)+}$ is a shortest vector in $L^{(D)}$ and $w = \psi_D(u)$, then $x_i = g_{C,u}$, $i = 1,..,l$ is a solution of our susbet sum problem. $\mathcal{D}$ may find such a w by asking the oracle for a shortest vector in $L^{(D)}$. $\mathcal{D}$ presents the lattice $L^{(D)}$ to the oracle by the basis $\psi_D(v_i)$, $i = 1,...,m$. (By the definition of ψ D the size of this basis is polynomial in l.) The oracle gives a shortest vector w' in $L^{(D)}$ since either w' or $-w'$ is in $L^{(D)+}$, $\mathcal{D}$ finds a w with the required property. Writing w as a linear combinations of the vectors v_i and using the fact that ψ_D is a homomorphism, $\mathcal{D}$ finds the vector $u \in L$ and so the solution of the subset sum problem. This way the algorithm has found the solution only with a probability $1 - n^{-\delta_3}$. Repeating this a polynomial number of times independently we may ensure that the probability of succes is exponentially close to one. *Q.E.D.*(Theorem 2.1)

3. Proof of Lemma 2.1. First we reformulate the lemma for lattices in $\mathbf{R}^{m+2}$. We note that it is sufficient to prove the lemma without condition (6) since if $\rho = \frac{p}{q}$ satisfies all of the other conditions and $p \le r^2 \le 2p$, $q \le s^2 \le 2p$ and $\rho' = \frac{r^2}{s^2}$ then the lemma is satisfied by $\rho \to 4\rho'^2$ with slightly modified values of the constants c_4 and c_2.

Lemma 3.1. *There are rationals $\bar c_1, \bar c_2, \bar c_3, \bar c_4$ so that for all $c > 0$ there is a $c' > 0$ so that for each sufficiently large positive integer n there is a positive integer $m \in [n^{\bar c_1}, n^{2\bar c_1}]$, a positive rational $\bar\rho < 1$, a lattice $\bar L \subseteq \mathbf{R}^{m+2}$, and a basis $w_1,...,w_m,w_{m+1},w_{m+2}$ of $\bar L$ so that the following holds:*

$(\bar 1)$ *$v \in \bar L$, $v \neq 0$ implies $\|v\| \ge 1$,*

$(\bar 2)$ *if Z is the set of all $v \in \bar L$, $v = \sum_{i=1}^{m+2} \gamma_i w_i$ with $\sum_{i=1}^{m} \gamma_i = n$, then $|\{v \in Z \mid \|v\|^2 < 1 + \bar\rho\}| \ge 2^{\bar c_3 n \log n}$,*

$(\bar 3)$ *For all $v \in \bar L$, $v \neq 0$, if $\|v\|^2 < 1 + \bar\rho 2^{n^{\bar c_4}}$, $v = \sum_{i=1}^{m+2} \gamma_i w_i$, and $\gamma_{m+1} \ge 0$, then $\gamma_i \in \{0,1\}$ for $i = 1,...,m$, $\gamma_{m+1} = 1$.*

$(\bar 4)$ *If $u_1 \neq u_2$, $\|u_j\|^2 < 1 + \bar\rho 2^{n^{\bar c_4}}$ and $u_j = \sum_{i=1}^{m} \gamma_i^{(j)} w_i$ for $j = 1,2$ then there is a $i = 1,...,m$ so that $\gamma_i^{(1)} \neq \gamma_i^{(2)}$,*

$(\bar 5)$ size$(\bar\rho) \le n^{\bar c_2}$, $0 < \bar\rho < 2^{-n^{\bar c_4}}$,

$(\bar 6)$ $|\det(w_1,...,w_{m+2})| \ge 2^{-n^{\bar c_2}}$

Moreover there is a probabilistic Turing machine $\mathcal{C}$ so that for all $c > 0$, there is a $c' > 0$ so that the following holds. If $\mathcal{C}$ gets n and c as inputs then it returns the following output *in time $n^{c'}$: an integer m, a rational $\bar{\rho} > 0$ and linearly independent vectors $\bar{v}_1, ..., \bar{v}_{m+2} \in \mathbf{Q}^{m+2}$ with* $\mathrm{size}(\bar{v}_1, ..., \bar{v}_{m+2}) \leq n^{c'}$ *so that there exist vectors $w_1, ..., w_{m+2} \in \mathbf{R}^{m+2}$ so that $\|w_i - \bar{v}_i\| \leq 2^{-n^c}$ for $i = 1, ..., m+2$ and with a probability of at least $\frac{1}{2}$ if $\bar{L}$ is the lattice generated by $w_1, ..., w_{m+2}$ then $m, \bar{\rho}, w_1, ..., w_{m+2}, \bar{L}$ satisfy conditions ($\bar{1}$), ($\bar{2}$), ($\bar{3}$), ($\bar{4}$), ($\bar{5}$), ($\bar{6}$).*

Remark. Condition ($\bar{6}$) is not really new, compared to the conditions of Lemma 2.1. There the fact that $v_1, ..., v_{m+2}$ is a basis implies that their determinant is not zero and therefore from $\mathrm{size}(v_1, ..., v_{m+2}) \leq n^{c_2}$ we get a similar lower bound for the absolute value of the determinant.

First we show that Lemma 3.1 imply Lemma 2.1. We claim that if $\bar{c}_1, \bar{c}_2, \bar{c}_3, \bar{c}_4$ are the constants from Lemma 3.1, c is sufficiently large with respect to $\bar{c}_1, \bar{c}_2, \bar{c}_3, \bar{c}_4$, and $\bar{\rho}$, n, m, $\bar{v}_1, ..., \bar{v}_{m+2}$ meet the requirements of Lemma 3.1, then $\rho \rightarrow 8\bar{\rho}$, n m, $v_i \rightarrow (1+\rho)\bar{v}_i$, $c_1 \rightarrow \bar{c}_1$, $c_2 \rightarrow c'$, $c_3 \rightarrow \bar{c}_3$, $c_4 \rightarrow \frac{\bar{c}_4}{2}$ meet the requirements of Lemma 2.1.

($\bar{6}$), $\|w_i - \bar{v}_i\| \leq 2^{-n^c}$ and the assumption that c is sufficiently large with respect to $\bar{c}_2$ implies that if T is a linear transformation with $Tw_i = \bar{v}_i$ for $i = 1, ..., m+2$ then T is invertible and $1 - 2^{-n^{c_5}} \leq\|T\| \leq1 + 2^{-n^{c_5}}$, $1 - 2^{-n^{c_5}} \leq\|T^{-1}\| \leq1 + 2^{-n^{c_5}}$ where c_5 is sufficiently large with respect to $\bar{c}_1, \bar{c}_2, \bar{c}_3, \bar{c}_4$. (We define the norm of a linear transformation A by $\|A\| = \sup\{\|Ax\| \mid \|x\| = 1\}$.) Now we can easliy check that requirements (1),(2),(3),(4),(5) of Lemma 2.1 are met.

(1). $v \in L$, $v \neq 0$ implies that $v = (1+\bar{\rho})Tw$ for some $w \in \bar{L}$, $w \neq 0$ and so $w = (1+\bar{\rho})^{-1}T^{-1}v$, that is $\|w\| \leq (1+\bar{\rho})^{-1}\|T^{-1}\|\|v\|$, Using that by ($\bar{1}$) $\|w\| \geq 1$ we get that $\|v\| \geq (1+\bar{\rho})(1 - 2^{-n^{c_5}}) > 1$. ($\bar{5}$) implies that $\bar{\rho} > 2^{-n^{\bar{c}_2}}$ and so $\|v\| \geq 1$.

(2). Assume that $w \in \bar{L}$, $\|w\|^2 < 1 + \bar{\rho}$, $w = \sum_{i=1}^{m+2} \gamma_i w_i$, $\sum_{i=1}^{m} \gamma_i = n$. Clearly $v = (1+\bar{\rho})Tw \in L$, $v = \sum_{i=1}^{m+2} \gamma_i v_i$ and $\|v\|^2 \leq(1+\bar{\rho})^2(1+2^{-n^{c_5}})^2(1+\bar{\rho})^2 <1+8\bar{\rho} = 1+\rho$.

(3). Using a similar calculation as in the proof of (1) and (2) and using the fact $c_4 = \frac{\bar{c}_4}{2}$, we get that each $v \in L$, $v \neq 0$, $\|v\|^2 \leq 1 + \rho 2^{n^{c_4}}$ can be written in the form $v = (1+\bar{\rho})Tw$ with $w \in \bar{L}$, $w \neq 0$, $\|w\|1 + \bar{\rho}2^{n^{\bar{c}_4}}$. Therefore the defintion of T and ($\bar{3}$) implies (3).

(4). The proof is similar to the proof of (3).

(5). This is an immediate consequence of ($\bar{5}$), $\mathrm{size}(\bar{v}_1, ..., \bar{v}_{m+2}) \leq n^{c'}$, and the definitions of v_i, ρ and c_2.

Proof of Lemma 3.1. For each sufficiently large positive integer we define a lattice $\mathcal{L}_n$. $\bar{L} = \mathcal{L}_n$ will satisfy the conditions of Lemma 3.1. The appropriate choice of $c_1, ..., c_4$ (which do not depend on n) will follow from the properties of $\mathcal{L}_n$. m, $\bar{\rho}$ and the basis $w_1, ..., w_{m+2}$ will be defined together with $\mathcal{L}_n$. The second constructive part of the lemma will be a trivial consequence of these definitions.

First we define a larger class of lattices which depends on the choice of several parameters. We will show later how can we fix the values of these parameters to get the lattice with the appropriate properties.

Definitions. 1. p_i will denote the ith prime number, that is, $p_1 = 2$, $p_2 = 3$, $p_3 = 5$, etc.

2. Assume that b, ω, ι, κ, μ are positive integers. We define a lattice $L = L(b, \omega, \iota, \kappa, \mu)$ in $\mathbf{R}^{\iota+2}$. L will be generated by the vectors $\nu_1, ..., \nu_{\iota+2}$ defined below.

The coordinates of vector ν_i will be denoted by $\chi_{i,j}$ for all $i = 1, ..., \iota+2$, $j = 1, ..., \iota+2$. Let $B = \omega^{\mu+1}$

The definition of ν_i, for $i = 1, ..., \iota$. $\chi_{i,i} = \sqrt{\log p_i}$, $\chi_{i,\iota+2} = B \log p_i$. $\chi_{i,j} = 0$ for all $j \neq i$, $j \neq \iota+2$

The definition of $\nu_{\iota+1}$. $\chi_{\iota+1,\iota+2} = B \log b$. $\chi_{i,j} = 0$ for all $j \neq \iota+2$.

The definition of $\nu_{\iota+2}$. $\chi_{\iota+2,\iota+1} = \omega^{-\kappa}$, $\chi_{\iota+2,\iota+2} = B\log(1+\frac{\omega}{b})$. $\chi_{i,j} = 0$ for all $j \neq \iota+1$, $j \neq \iota+2$.

The vectors $\nu_1, ..., \nu_{\iota+2}$ form the rows of the matrix shown below.

$$\begin{pmatrix} \sqrt{\log p_1} & \dots & 0 & 0 & B\log p_1 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \dots & \sqrt{\log p_\iota} & 0 & B\log p_\iota \\ 0 & \dots & 0 & 0 & B\log b \\ 0 & \dots & 0 & \omega^{-\kappa} & B\log(1+\frac{\omega}{b}) \end{pmatrix}$$

The following lemma is somewhat more general than what we need for the proof of Lemma 3.1 since its original motivation was related to factoring. We will show that if $\sum_{i=1}^{\nu+1} \delta_i \nu_i$ is a small vector in the lattice $L(b, \omega, \iota, \kappa, \mu)$ then $g = \prod_{i=1}^{\iota} p_i^{|\delta_i|}$ is close to b and $g \equiv b \pmod{\omega}$. That is, by searching for a short vector in the lattice, we search for an integer in the arithmetic sequence $b + l\omega$ ($|l|$ is smaller than a small power of b) which has only small prime factors. The last basis vector $\nu_{\iota+2}$ makes it possible that a short vector my indicate any number of this arithmetic sequence, without the last vector the short vector would always indicate b.

Lemma 3.2 . *Suppose that κ, μ are positive real numbers, $c > 0$ is sufficiently large, ω is an integer sufficiently large with respect to c, and ι, b are positive integers so that $(\log \omega)^c < \iota < (\log \omega)^{2c}$, $\omega^\mu < b < 2\omega^\mu$, $\mu > 2\kappa + 6$, $\kappa \geq 2$. If $L = L(b, \omega, \iota, \kappa, \mu)$ then the following holds:*

Assume that $w \in L$, $w = \sum_{i=1}^{\iota+2} \delta_i \nu_i$ and $\delta_{\iota+1} \geq 0$. If $\|w\|^2 \leq \omega^{-1} + \log b$, then

(1) $\delta_{\iota+1} = 1$

(2) $\delta_i \leq 0$ for $i = 1, ..., \iota$

(3) $|\delta_{\iota+2}| \leq \omega^{\kappa+1}$

(4) $\prod_{i=1}^{\iota} p_i^{-\delta_i} \equiv b \pmod{\omega}$.

(5) if $g = \prod_{i=1}^{\iota} p_i^{-\gamma_i} \equiv b \pmod{\omega}$, where $\gamma_i \in \{0, -1\}$ for $i = 1, ..., \iota$, and $|b - g| \leq \min\{\omega^{\frac{\mu}{2}-1}, \omega^{\kappa-\frac{1}{2}}\}$, then there are integers $\gamma_{\iota+1}$, $\gamma_{\iota+2}$, so that $w' = \sum_{i=1}^{\iota+2} \gamma_i \nu_i$ implies $\|w'\|^2 \leq 3\omega^{-2} + \log b \leq \omega^{-1} + \log b$.

(6) For all $v \in L$, $v \neq 0$ we have $\|v\| \geq \log b$.

Remark. Assume that r is an integer with $1 \leq r \leq \omega$. We may define a lattice $L(b, \omega, \iota, \kappa, \mu, r)$ by changing only the value of $\nu_{\iota+2}$ in the definition of $L(b, \omega, \iota, \kappa, \mu)$. Namely let us define the last component of $\nu_{\iota+2}$ as $B \log(1 + \frac{r}{b})$, all of the other components of $\nu_{\iota+2}$ remain the same as before. Lemma 3.2 remains tru for $L = L(b, \omega, \iota, \kappa, \mu, r)$ for any integer r with $1 \leq r \leq \omega$.

First we show that Lemma 3.2 imply Lemma 3.1. Let $\kappa = 2, \mu = 10$. Assume that k is an integer suffciently large with respect to κ and μ and let $\bar{c}_1 = \frac{2}{3}k$, $\bar{c}_2 = k^2$, $\bar{c}_3 = 1$, $\bar{c}_4 = \frac{1}{2}$. We show that Lemma 3.1 holds for these values.

Assume now that n is a sufficiently large integer, we will define $m, \bar{\rho}, w_1, ..., w_m$ as required by the lemma. $\bar{L} = \mathcal{L}_n$ will be defined by $\mathcal{L}_n = (\log b)^{-\frac{1}{2}} L(b, \omega, \iota, \kappa, \mu)$ after we heva selected the values for all of the parameters. (By definition $\alpha L = \{\alpha x \mid x \in L\}$ for a lattice L.) b, ω, ι will be defined in the follwoing way.

Assume that the integer n is sufficiently large with respect to k. Let J be a positive integer so that $n = [\frac{\log J}{k \log \log J}]$. Since n is sufficiently large with respect to k such an integer J clearly exists, its size (the number of bits in its binary representation) is polynomial is n, and J (that is, its binary representation) can be computed in time polynomial in n. Let ι be the number of primes less than $(\log J)^k$. (Note that ι and $p_1, ..., p_\iota$ can be also computed in time polynomial in n.)

We select b at random. Γ will denote the set of integers that are the product of n distinct elements of the set $\{p_1, ..., p_\iota\}$. We pick b with uniform distribution from the set Γ. Then we choose the integer ω with the property $\omega^\mu < b < 2\omega^\mu$. Finally let $\mathcal{L}_n = \frac{1}{\log b} L(b, \omega, \iota, \kappa, \mu)$ with the defined values of parameters. Lemma 3.1 will hold with $L = \mathcal{L}_n$, $m = \iota$, $w_i = (\log b)^{-\frac{1}{2}} \nu_i$ for $i = 1, ..., m + 2$ and $\bar{\rho} = 3\omega^{-2}([\log b])^{-1}$.

The choice of b was the only random step in our construction. We show that if b meets the following requirements then the conditions of Lemma 3.1 are satisfied.

(i) $|b| \geq J^{1 - \frac{1}{k-2}}$

(ii) if $r = \min\{\omega^{\frac{\mu}{2} - 1}, \omega^{\kappa - \frac{1}{2}}\} = \omega^{\frac{3}{2}}$ then in the interval $(b - r, b + r)$ there are at least $2^{n \log n}$ elements of Γ.

First we show that $P((i) \wedge (ii)) \geq \frac{1}{2}$. In the proof of this fact we will use the following lemma which estimates $|\Gamma|$ in terms of J and k.

Lemma 3.3. *Assume that k is a positive integer, J is a sufficiently large positive integer and $h = \frac{\log J}{k \log \log J}$. Then, in the interval $[1, J]$ there are at least $J^{1 - \frac{1}{k-1}}$ squarefree integers which are the product of exactly $[h]$ distinct primes so that each of them is at most $(\log J)^k$.*

Proof. Let D be the set of all squarefree integers in the interval $[1, J]$ that are the product of exactly $[h]$ distinct primes so that each of them is at most $(\log J)^k$. $((\log J)^k)^h = J$ implies that D is the set of all products of $[h]$ different prime numbers less than $(\log J)^k$. Therefore $|D| = \binom{\pi((\log J)^k)}{[h]}$. The Prime Number Theorem implies that if J is sufficiently large then $\pi((\log J)^k) \geq e^{-1}\frac{(\log J)^k}{k\log\log J}$. Consequently $|D| \geq \binom{y}{[h]}$ where $y = e^{-1}\frac{(\log J)^k}{k\log\log J}$. Giving lower resp. upper bounds on $y(y-1)\cdot\ldots\cdot(y-[h]+1)$ resp. $[h]!$ we get the required lower bound on $|D|$. Let $x = \frac{(\log J)^k}{k\log\log J}$. $y(y-1)\cdot\ldots\cdot(y-[h]+1) \geq e^{-h}x(x-1)\cdot\ldots\cdot(x-[h]+1) \geq \exp(-o(\frac{\log J}{k^2}))x(x-1)\cdot\ldots\cdot(x-[h]+1)$.

$x(x-1)\cdot\ldots\cdot(x-[h]+1) \geq (x-h)^{h-1} \geq x^h x^{-1}(1-\frac{h}{x})^h$. We estimate the three factors separately.

$x^h = ((\log J)^k)^h(k\log\log J)^{-h} = Jk^{-h}(\log\log J)^{-h}$ $=\exp(\log J - \frac{\log k\log J}{k\log\log J} - \frac{\log J}{k\log\log J}\log\log\log J) = \exp(\log J - o(\frac{\log J}{k^2}))$.

$x^{-1} = (\log J)^{-k}k\log\log J$ $=\exp(-k\log\log J + \log k + \log\log\log J)$ $\geq\exp(-o(\frac{\log J}{k^2}))$

$(1-\frac{h}{x})^h = (1-\log J(k\log\log J)^{-1}(\log J)^{-k}(k\log\log J))^h$ $=(1-\frac{1}{(\log J)^{k-1}})^h$ $\geq 1-\frac{h}{(\log J)^{k-1}} \geq 1-\frac{1}{k\log\log J(\log J)^{k-2}} \geq \exp(-\frac{2}{k\log\log J(\log J)^{k-2}}) = \exp(-o(\frac{\log J}{k^2}))$.

$[h]! \leq h^h \leq \exp^{h\log h}$ $=\exp(\frac{\log J}{k\log\log J}(\log\log J - \log k - \log\log\log J))$ $\leq\exp(\frac{\log J}{k})$.

The inequalities together yield $\binom{y}{[h]} \geq \exp((1-\frac{1}{k})\log J - o(\frac{\log J}{k^2}))$ which implies the statement of the lemma. *Q.E.D.*(Lemma 3.3)

Now we return to the proof of $P((i)\wedge(ii)) \geq \frac{1}{2}$. According to Lemma 3.3 $|\Gamma| \geq J^{1-\frac{1}{k-1}}$. Clearly $|\Gamma\cap[1, J^{1-\frac{1}{k-2}}]|$ $\leq J^{1-\frac{1}{k-2}}$. Therefore $P(\neg(i)) \leq J^{-\frac{1}{k-2}+\frac{1}{k-1}} \leq J^{-\frac{1}{(k-1)(k-2)}}$, and so

(3.1) $\quad P((i)) \geq 1 - J^{-\frac{1}{(k-1)(k-2)}}$.

Since $n \leq \frac{\log J}{k\log\log J}$ we have that $2^{n\log n} \leq e^{\log 2n\log n} \leq \exp(\log 2\frac{\log J}{k\log\log J}\log\log J) \leq \exp(\frac{\log 2}{k}\log J) = J^{\frac{\log 2}{k}}$. We will prove that

(3.2) $\quad P(|\Gamma| > J^{\frac{\log 2}{k}}) > \frac{2}{3}$.

For the proof of this fact we use the following trivial observation:

(*) *Assume that* $\{Q_1, ..., Q_s\}$ *is a partition of the finite set* Q *and* $A \subseteq Q$. *If we take a random element* a *with uniform distribution from* A, *and* $a \in Q_i$ *then for any* $\lambda > 0$ *we have* $P(\frac{|A\cap Q_i|}{|Q_i|} \leq \lambda\frac{|A|}{|Q|}) \leq \lambda$.

Indeed let G be the set of all $i \in \{1, ..., s\}$ with $\frac{|A\cap Q_i|}{|Q_i|} \leq \lambda\frac{|A|}{|Q|}$. Then $\sum_{i\in G}|A\cap Q_i| \leq \sum_{i\in G}\lambda\frac{|A|}{|Q|}|Q_i| \leq \lambda\frac{|A|}{|Q|}\sum_{i\in G}|Q_i| \leq \lambda\frac{|A|}{|Q|}|Q| \leq \lambda|A|$. *Q.E.D.*(*)

We apply (*) with $Q \to \{1, ..., J\}$, $A \to \Gamma$. The partition Q_i is defined in the following way. First we cut J into intervals with length between r and $\frac{r}{2}$, then we further cut each intervall into residue classes modulo ω. $\{Q_1, ..., Q_s\}$ is the partition that we get this way. According to (*) for any fixed $\lambda > 0$ if we pick b with uniform distribution from Γ and $b \in Q_i$ then $P(\frac{|\Gamma\cap Q_i|}{Q_i} \leq \lambda\frac{|\Gamma|}{J}) \leq \lambda$. Therefore Lemma 3.3 implies that $P(\frac{|\Gamma\cap Q_i|}{Q_i} \leq \lambda J^{-\frac{1}{k-1}}) \leq \lambda$. According to the definition of Q we have $|Q_i| \geq [\frac{1}{2}r\omega^{-1}] \geq \frac{1}{3}\omega^{\frac{3}{2}}\omega^{-1} = \frac{1}{3}\omega^{\frac{1}{2}}$. So for $\lambda = \frac{1}{10}$ we get $P(|\Gamma\cap Q_i| \leq \frac{1}{10}J^{-\frac{1}{k-1}}\omega^{\frac{1}{2}}) \leq \frac{1}{10}$. $\omega \geq b^{\frac{1}{\mu}}$ implies that if (i) holds then $\omega \geq J^{\frac{1}{\mu}(1-\frac{1}{k-1})} \geq J^{\frac{1}{2\mu}}$ and we have in this case $P(|\Gamma\cap Q_i| \leq \frac{1}{10}J^{-\frac{1}{k-1}}J^{\frac{1}{2\mu}}) \leq \frac{1}{10} + P((i))$. Since k is sufficiently large with respect to μ we have that $\frac{1}{10}J^{-\frac{1}{k-1}}J^{\frac{1}{2\mu}} \geq J^{\frac{\log 2}{k}}$. Together with (3.1) this proves (3.2) and so completes the proof of $P((i) \wedge (ii)) \geq \frac{1}{2}$. (Actually this proof with minor modifications gives a much better lower bound.)

Now we assume that (i) and (ii) hold and we show that the requirements of Lemma 3.1 are met.

$(\bar{1})$. This is an immediate consequence of property (6) of Lemma 3.2.

$(\bar{2})$. According to property (5) of Lemma 3.2 we have that for all $a \in \Gamma \cap (b-r, b+r)$ (where r is define in (ii)) there is a $u \in L(b, \omega, \iota, \kappa, \mu)$ so that $\|u\|^2 \leq 3\omega^{-2} + \log b$, $u = \sum_{i=1}^{\iota+2} \gamma_i \nu_i$ and $a = \prod_{i=1}^{\iota} p_i^{-\gamma_i}$. Let Y be the set of all vectors u with these properties and wich also satisfy $\sum_{i=1}^{n} \gamma_i = n$. According to (ii) Y has at least $2^{n \log n}$ elements. Let T be the linear transformation with $T\nu_i = w_i$. We claim that for each $u \in Y$, $v = Tu \in \bar{L}$ with $\|v\|^2 \leq 1 + \bar{\rho}$. $v \in \bar{L}$ follows from $T\nu_i = w_i$. $\|v\|^2 = (\log b)^{-1}\|u\|^2 \leq (\log b)^{-1}(3\omega^{-2} + \log b) = 1 + 3\omega^{-2}(\log b)^{-1} < 1 + 3\omega^{-2}([\log b])^{-1} = 1 + \bar{\rho}$, which completes the proof of $(\bar{2})$.

$(\bar{3})$. Let T be the linear transformation defined above and assume that $v \in \bar{L}$ satisfies the assumptions in $(\bar{3})$. Let $u = T^{-1}v$. Then $u \in L(b, \omega, \iota, \kappa, \mu)$ and $\|u\|^2 < (1 + \bar{\rho}2^{n^{\frac{1}{2}}})\log b = \log b + \bar{\rho} \log b 2^{n^{\frac{1}{2}}}$. (i) implies that $\omega \geq J^{\mu^{-1}(1-\frac{1}{k-2})}$. $2^{n^{\frac{1}{2}}} \leq 2^{(\log J)^{\frac{1}{2}}} < \omega^{\frac{1}{2}}$. Therefore using that $\bar{\rho} = 3\omega^{-2}([\log b])^{-1}$ we get that $\|u\|^2 < \log b + \omega^{-1}$. This and $\gamma_{m+1} \geq 0$ implies that $w \rightarrow u$ satifies the assumptions of Lemma 3.2 and therefore also the conclusions (1), (2), (3), (4). Therefore $w_i = T\nu_i$ implies that $v = Tu$ meets the requirements of $(\bar{3})$.

$(\bar{4})$. Assume that contrary to our statement $\gamma_i^{(1)} = \gamma_i^2$ for all $i = 1, ..., m$. Let $y_i = T^{-1}u_i$, $i = 1, 2$. As in the previous part of the proof we get that $\|y_i\|^2 \leq \log b + \omega^{-1}$. This implies that for $i = 1, 2$ either y_i or $-y_i$ satifies conditions (1), (2), (3), (4). Therefore our indirect assumption implies that $y = y_1 - y_2$ is parallel to $\nu_{\iota+2}$ for $i = 1, 2$ and we also have that $\|y\|^2 = (y_1 + y_2) \cdot (y_1 + y_2) \leq 2(\|y_1\|^2 + \|y_2\|^2) \leq 2(\omega^{-1} + \log b)$. This is however impossible since $\|\nu_{\iota+2}\|^2 \geq B^2(\log(1 + \frac{\omega}{b}))^2 = \omega^{2(\mu+1)}\frac{1}{4}\omega^{-2(\mu-1)} \geq \frac{1}{4}\omega^4 > 2(\omega^{-1} + \log b)$.

$(\bar{5})$. Immediately follows from the defintions of $\bar{\rho}$, J, ω.

$(\bar{6})$. Switching the last two rows in the matrix whose rows are $\nu_1, ..., \nu_{\iota+2}$ we get that $|\det(\nu_1, ..., \nu_{\iota+2})| = n^{-\kappa} B \log B \prod_{i=1}^{\iota+2} \sqrt{\log p_i}$. Therefore $|\det(w_1, ..., w_{m+2})| = (\log b)^{-\frac{m+2}{2}} n^{-\kappa} B \log B \prod_{i=1}^{\iota+2} \sqrt{\log p_i}$. The definitions of m and ι imply that $m \leq n^{2k}$. $B = n^{\mu+1}$, therefore $Bn^{-\kappa} > 1$. $\log b \leq \log(2n^{\mu}) \leq n$. Using this inequalities we get $|\det(w_1, ..., w_{m+2})| \geq 2^{-n^2} = 2^{-n^{\bar{c}_2}}$. *Q.E.D.*(Lemma 3.1).

Proof of Lemma 3.2.

Notation. In the following proof we will denote by $\text{cord}_i(u)$ the ith coordinate of the vector $u \in L$ for $i = 1, ..., \iota + 2$.

In the proof we will use the following trivial inequality.

(3.3) If $\Phi, \chi, \Psi \neq 0, \Gamma$ are real numbers then $|\Phi + \chi\Psi| \leq \Gamma$ implies $|\chi| \geq \frac{|\Phi| - \Gamma}{|\Psi|}$.

Let $g_0 = \prod\{p_i^{\delta_i} | 1 \leq i \leq \iota, \delta_i > 0\}$ and $g_1 = \prod\{p_i^{-\delta_i} | 1 \leq i \leq \iota, \delta_i < 0\}$. The following inequality is an immediate consequence of the definitions of g_0 and g_1:

(3.4) $\|w\|^2 \geq \log g_0 + \log g_1$.

Before we sart the proof we note that $\|\nu_{\iota+2}\|^2 > (\omega^{\mu+1}\log(1+\frac{\omega}{b}))^2 \geq (\omega^{\mu+1}\frac{1}{2}\omega^{-\mu+1})^2 \geq \omega^2 > \omega^{-1} + \log b$. Consequently w is not parallel to $\nu_{\iota+2}$.

(3) $|\mathrm{cord}_{\iota+1}(w)| \leq \|w\| \leq (\omega^{-1} + \log b)^{1/2} \leq (1 + \mu\log\omega)^{\frac{1}{2}} \leq \mu\log\omega$ implies that $|\delta_{\iota+2}\omega^{-\kappa}| \leq \mu\log\omega$ and so $|\delta_{\iota+2}| \leq \omega^{\kappa}\mu\log\omega \leq \omega^{\kappa+1}$.

(1) Assume that contrary to our statement $\delta_{\iota+1} \neq 1$. (We assumed that $\delta_{\iota+2} \geq 0$.)

Case 1. $\delta_{\iota+1} = 0$. Since w is not parallel to $\nu_{\iota+2}$ and $\delta_{\iota+1} = 0$, we have that at least one of the numbers g_0 and g_1 is different from 1 and so, since they have no common prime factors, $g_0 \neq g_1$. Considering only the first ι coordinates of w we get $\|w\|^2 \geq \log g_0 + \log g_1$. Our assumption about the norm of w implies that there is an $i_0 \in \{0,1\}$ so that $\log g_{i_0} \leq \frac{1}{2}(\omega^{-1}+\log b)$ and so $g_{i_0} \leq 2b^{\frac{1}{2}}$, which implies that $|\log g_0 - \log g_1| \geq \log(2b^{\frac{1}{2}}+1) - \log(2b^{\frac{1}{2}})| \geq (2b^{\frac{1}{2}}+1)^{-1} \geq \frac{1}{4}b^{-\frac{1}{2}}$. Therefore $|\sum_{i=1}^{\iota}\delta_i\log p_i| = |\log g_0 - \log g_1| \geq \frac{1}{4}b^{-\frac{1}{2}} \geq \frac{1}{4}\omega^{-\frac{\mu}{2}}$. Since $|\mathrm{cord}_{\iota+2}(w)| \leq \|w\| \leq \omega^{-1} + \log b$ and $\delta_{\iota+1} = 0$ we have that $B^{-1}|\mathrm{cord}_{\iota+2}(w)| = |(\sum_{i=1}^{\iota}\delta_i\log p_i) + \delta_{\iota+2}\log(1+\frac{\omega}{b})| \leq B^{-1}(\omega^{-1}+\log b)$. Applying (3.3) with $\Phi \to \sum_{i=1}^{\iota}\delta_i\log p_i$, $\chi \to \delta_{\iota+2}$, $\Psi \to \log(1+\frac{\omega}{b})$ and $\Gamma \to B^{-1}(\omega^{-1}+\log b)$ we get $\delta_{\iota+2} \geq (\frac{1}{4}\omega^{-\frac{\mu}{2}} - \omega^{-\mu-1}(\omega^{-1}+\log 2+\mu\log\omega))\frac{1}{4}\omega^{\mu-1} \geq \omega^{\frac{\mu}{2}-2}$. Hence $\mathrm{cord}_{\iota+1}(w)$ is at least $\omega^{\frac{\mu}{2}-2}\omega^{-\kappa} \geq \omega$ in contradiction to our upper bound on $\|w\|$.

Case 2. $\delta_{\iota+2} \geq 2$.

$\mathrm{cord}_{\iota+2}(w) = B((\sum_{i=0}^{\iota}\delta_i\log p_i)+\delta_{\iota+1}\log b + \delta_{\iota+2}\log(1+\frac{\omega}{b}))$. (3) implies that $\delta_{\iota+2}\log(1+\frac{\omega}{b})$ is at most 1, so we have $\delta_{\iota+1}\log b - B^{-1}|\mathrm{cord}_{\iota+2}(w)| - 1 \leq |\log g_0 - \log g_1|$. $\mathrm{cord}_{\iota+2}(w) \leq \|w\| \leq (\omega^{-1}+\log b)^{\frac{1}{2}}$, $\delta_{\iota+2} \geq 2$ so $|\log g_0 - \log g_1| \geq 2\log b - B^{-1}(\omega^{-1}+\log b)^{\frac{1}{2}} - 1 \geq \frac{3}{2}\log b$. By (3.4) we have $\|w\|^2 \geq \log g_0 + \log g_1 \geq |\log g_0 - \log g_1| \geq \frac{3}{2}\log b$ in contradiction to our assumption $\|w\|^2 \leq \omega^{-1} + \log b$.

(2) Inequality (3) implies that $\delta_{\iota+2} \leq \omega^{\kappa}\mu\log\omega$ and so $\delta_{\iota+2}\log(1+\frac{\omega}{b}) \leq \omega^{-1}$. $|\mathrm{cord}_{\iota+2}(w)| \leq \|w\| \leq (\omega^{-1}+\log b)^{\frac{1}{2}}$ therefore $|B((\sum_{i=0}^{\iota}\delta_i\log p_i)+\log b+\delta_{\iota+2}\log(1+\frac{\omega}{b}))| \leq (\omega^{-1}+\log b)^{\frac{1}{2}}$, that is $|\log g_0 - \log g_1 + \log b| \leq B^{-1}(\omega^{-1}+\log b)^{\frac{1}{2}} + |\delta_{\iota+2}\log(1+\frac{\omega}{b})| \leq \omega^{-\mu-1}\log b + \omega^{-1} \leq \frac{1}{4}$. If contrary to our assertion there is a $\delta_i > 0$ for some $i = 1, ..., \iota$, then $\log g_0 \geq \log 2$ and so $\log g_1 \geq \log b + \log 2 - \frac{1}{4} \geq \log b + \frac{1}{4}$. By (3.4) this would would imply $\|w\|^2 \geq \log b + \frac{1}{2}$ in contradiction to our assumed upper bound.

(4) According to (2) we have $g_0 = 1$ and so our estimate on $|\mathrm{cord}_{\iota+2}(w)| \leq \|w\| \leq (\omega^{-1}+\log b)^{\frac{1}{2}}$ and (2) implies that $|-\log g_1+\log b+\delta_{\iota+2}\log(1+\frac{\omega}{b})| \leq B^{-1}(\omega^{-1}+\log b)^{\frac{1}{2}} < \omega^{-\mu-\frac{1}{2}}$. By (3) this implies that $|-\log g_1 + \log b| < \log 2$ and therefore, $\frac{b}{2} < g_1 < 2b$. Let $y = b(1+\frac{\omega}{b})^{\delta_{\iota+2}}$. We have

$$\textbf{(3.5)} \qquad |-\log g_1 + \log y| \leq \omega^{-\mu-\frac{1}{2}}.$$

We claim that g_1 is the closest integer to y. Indeed this is an immediate consequence of (3.5) and the inequalities $|\log g_1 - \log(g_1+\frac{1}{2})| \geq \frac{1}{2}\frac{1}{g_1+\frac{1}{2}} \geq \frac{1}{16}\omega^{-\mu}$, and $|\log g_1 - \log(g_1-\frac{1}{2})| \geq \frac{1}{2}\frac{1}{g_1-\frac{1}{2}} \geq \frac{1}{16}\omega^{-\mu}$. If $l = \delta_{\iota+2}$ then $b(1+\frac{\omega}{b})^l = b + l\omega + b\binom{l}{2}(\frac{\omega}{b})^2 + ... = b + l\omega + R$, where $|R| \leq 2l^2\frac{\omega^2}{b} \leq 2\omega^{2\kappa+2}\omega^2\omega^{-\mu} = 2\omega^{-\mu+2\kappa+4} \leq \omega^{-1}$. This implies that the closest integer to $b + l\omega + R$ is g_1 that is $g_1 = b + l\omega$ which together with $g_0 = 1$ implies (4).

(5) Let $g = b + l\omega$, $\gamma_{\iota+2} = l$ and $\gamma_{\iota+1} = 1$. $\|w'\|^2 = (\sum_{j=1}^{\iota} \gamma_j^2 \log p_j) + l^2\omega^{-2\kappa} + B^2((\sum_{i=1}^{\iota} \gamma_i \log p_i) + \log b + l \log(1 + \frac{\omega}{b}))^2$. We estimate the various parts of this expression separately.

$(\sum_{j=1}^{\iota} \gamma_j^2 \log p_j) = \log g = \log(b + l\omega) = \log b + \log(1 + \frac{l\omega}{b}) = \log b + R_1$, where $|R_1| < \frac{l\omega}{b}^2$

$(\sum_{i=1}^{\iota} \gamma_i \log p_i) + \log b + l \log(1 + \frac{\omega}{b}) = -\log g + \log b + l\frac{\omega}{b} + R_3 = -\log(b + l\omega) + \log b + l\frac{\omega}{b} + R_3 = -\log b - \frac{l\omega}{b} - R_2 + \log b + l\frac{\omega}{b} + R_3 = -R_2 + R_3$, where $|R_2| \leq (\frac{l\omega}{b})^2$, $|R_3| \leq |l|(\frac{\omega}{b})^2$.

We got that $\|w\|^2 \leq \log b + |R_1| + l^2\omega^{-2\kappa} + B^2(|R_2| + |R_3|)^2 \leq \log b + (\frac{l\omega}{b})^2 + l^2\omega^{-2\kappa} + \omega^{2\mu+2}4(\frac{l\omega}{b})^4$. Our assumption about $|g - b|$ implies that $|l| \leq \min\{\omega^{\frac{\mu}{2}-2}, \omega^{\kappa-\frac{1}{2}}\}$. Using this we get that each of the last three terms in the sum is at most ω^{-2}, that is, $\|w'\|^2 \leq \log b + 3\omega^{-2} \leq \log b + \omega^{-1}$.

(6) Assume that v is a nonzero vector in L with minimal length. If $\|v\| > \omega^{-1} + \log b$ then our statement obviously holds. Assume that $\|v\| \leq \omega^{-1} + \log b$. Clearly either $w = v$ or $w = -v$ satisfies the assumptions of the lemma, and in both cases $\|w\| = \|v\|$. According to the already proven parts of the lemma (1), (2), (3), (4) hold. Let $g = \prod_{i=1}^{\iota} p^{-\delta_i}$. If $g \geq b$ then $\|w\|^2 \geq \log g \geq \log b$. Assume now that $g < b$. Suppose that contrary to our assertion $\|w\|^2 < \log b$. If $l = \delta_{\iota+2}$, then $\|w\|^2 \geq \log g + l^2\omega^{-2\kappa}$ Therefore $\|w\|^2 < \log b$ implies $l^2\omega^{-2\kappa} < \log b$ and so $l \leq \omega^{\kappa}(\log b)^{\frac{1}{2}}$. At the end of the proof of (4) we have concluded that $g = b + l\omega$. Using these facts we get the following: $(\log b - \log g) \leq \frac{1}{g}(b - g) \leq (b + l\omega)^{-1}|l|\omega < 2b^{-1}|l|\omega \leq 4\omega^{-\mu+1}|l|$. Finally $\|w\|^2 \geq \log g + l^2\omega^{-2\kappa} = \log b + l^2\omega^{-2\kappa} - (\log b - \log g) > \log b + l^2\omega^{-2\kappa} - 4|l|\omega^{-\mu+1} = \log b + |l|(|l|\omega^{-2\kappa} - 4\omega^{-\mu}) \geq \log b$ a contradiction. *Q.E.D.*(Lemma 3.2).

4. Proof of Lemma 2.2.

In this section we prove Lemma 2.2, its Corollary and some related results.

Lemma 4.1. *The restricted subset sum problem is NP-complete.*

Proof. We reduce the subset sum problem to the restricted sum problem. Let $\sum_{i=1}^{m} a_i x_i = b$ be an instance of the subset sum problem whose size is n. We choose l so that l is polynomial in n and $2q + \max\{\log_2(|b| + 1), \max_{i=1}^{l} \log_2(|a_i| + 1)\} \leq l$. For each fixed $r = 1, ..., l - q$ we consider the instance I_r of the restricted subset sum problem: $\sum_{i=l}^{q} a_i x_i = b + r2^{l^2}$, where $a_1, ..., a_q, b$ are from the original subset sum problems $a_{q+1} = \ldots = a_{q+r} = 2^r$, $a_{q+r+1} = \ldots = a_l$.

Assume now that $x_i = \delta_i$, $i = 1, .., q$ is a solution of the original subset sum problem and $\sum_{i=1}^{q} \delta_i = s$. Let $r = [\frac{r}{2}] - s$. Then $x_i = \delta_i$, $i = 1, ..., q$, $x_j = 1$, $j = q + 1, ..., q + r$, $x_j = 0$ for $j = q + r + 1$ is a solution of I_r.

Suppose now that $x_i = \epsilon_i$ is an arbitrary solution of I_r. $\sum_{i=1}^{q} a_i\epsilon_i = b + 2^{l^2}(r - \sum_{i=q+1}^{l} a_i\epsilon_i 2^{-l^2})$

The expression in the parenthesis is an integer $|a_i| \leq 2^l$ for $i = 1, ...q$ and $q < l$, therefore we have $\sum_{i=1}^{q} \epsilon_i x_i = b$. This shows that by solving all of the instances I_r,

$r = 1, ..., l - q$ of the restricted subset sum problem we get a solution of the subset sum problem. $Q.E.D.$(Lemma 4.1)

Proof of Lemma 2.2. Assume that v and w are given with the properties described in the lemma. By the definition of $L^{(A)+}$ and ψ_A there is a $u \in L$ so that $w = \psi_A(u)$. Suppose that $u = \sum_{i=1}^{m+2} \gamma_i v_i$. $w \in L^{(A)+}$ implies that $\gamma_{m+1} \geq 0$. Assume that contrary to our assertion $u \notin S_2$. Then $\|u\|^2 > 1 + \rho 2^{n(c_4)}$. Therefore $\|w\|^2 = \|\psi_A(u)\|^2 = \|u\|^2 + \|A\Lambda(u)\|^2 > 1 + \rho 2^{n(c_4)}$.

On the other hand $\|\psi_A(v)\|^2 = \|v\|^2 + \|A\Lambda(v)\|^2$.

$v \in S_1$ therefore $\|v\|^2 \leq 1 + \rho$. A is an m' by $m+1$ matrix, $m \leq n^{2c_1}$, $m' \leq n^{c_5}$ and each entry of A is at most n^{c_5}. By the definition of S_1 and Λ, $\Lambda(v)$ is a vector whose each entry is 0 or $\sqrt{\rho}$. therefore $\|\Lambda(v)\|^2 \leq n^{c_6}\rho$ where ρ is an absolute constant. Thus we get that $\|\psi_A v\|^2 \leq 1 + \rho + \rho n^{c_6} 1 + \rho 2^{n(c_4)} < \|w\|^2$ in contradiction to the assumption that w is a shortest nonzero vector. Therefore $u \in S_2$.

We have to show that $\|A\Lambda(u)\|^2 = \|A\Lambda(v)\|^2$. By our assumption about the minimality of $\|A\Lambda(v)\|^2$ in S_2 we have $\|A\Lambda(u)\|^2 \geq \|A\Lambda(v)\|^2$. Assume that contrary to our statement $\|A\Lambda(u)\|^2 > \|A\Lambda(v)\|^2$. Since the entries of A are integers and the coefficients of u and v in the basis v_i are also integers the definition of Λ implies that $\|A\Lambda(u)\|^2 - \|A\Lambda(v)\|^2 \geq \rho$.

$v \in S_1$ implies $\|v\|^2 < 1 + \rho$ and (1) of Lemma 2.1 implies that $\|u\|^2 \geq 1$. Therefore $\|u\|^2 - \|v\|^2 > -\rho$. This yields

$\|u\|^2 + \|A\Lambda(u)\|^2 - (\|v\|^2 + \|A\Lambda(v)\|^2) > 0$

that is $\|v\|^2 = \|\psi_A(u)\|^2 > \|\psi_A(v)\|^2$ in contradiction to the assumtpion the w is a shortest vector. $Q.E.D.$(Lemma 2.2)

Proof of Corollary 2.2. We claim that $\|D\Lambda(v)\|^2 = \min\{\|D\Lambda(x)\|^2 | x \in S_2\}$. According to our assumption $v \in S_1$ so every component of the vector $\Lambda(v)$ is either 0 or $\sqrt{\rho}$. Since $x_i = g_{C,v}(i)$ is a solution of the subset sum problem the first component of the vector $D\Lambda(v)$ is 0. Since it is a solution of the restricted subset sum problem (that is, the number of 1's is $[\frac{l}{2}]$) the second component of $D\Lambda(v)$ is also 0. The same property of $g_{C,v}$ implies that all of the other components are $\sqrt{\rho}$. Therefore we have that $\|D\Lambda(v)\|^2 = \rho l$.

Assume now that $x \in S_2$ if $x_i = g_{C,x}(i)$ is a solution of the restricted subset sum problem then we get the same way as for v that $\|D\Lambda(x)\| = \rho l$. Assume that it is not a solution of the restricted subset sum problem.

If $\sum_{i=1}^{l} x_i a_i \neq b$ or $\sum_{i=1}^{l} x_i \neq [\frac{l}{2}]$ then either the first or the second component of $D\Lambda(x)$ is not 0. Since they coordinates of $(\rho)^{-\frac{1}{2}} D\Lambda(x)$ are integers divisible by l^3, we get in this case that $\|D\Lambda\|^2 \geq l^6 \rho$ and therefore $\|D\Lambda(x)\|^2 > \|D\Lambda(v)\|^2$.

Assume now that $\sum_{i=1}^{l} x_i a_i = b$ and $\sum_{i=1}^{l} x_i = [\frac{l}{2}]$ but $x_i = g_{C,x}$ is still not a solution of the restricted subset sum problem. This is possible only if the is an $i = 1, ..., l$ so that $x_i \notin \{0, 1\}$. Since $x \in S_2$ we have that $x_i \geq 0$ and therefore $^{(\rho)^{-\frac{1}{2}}} D\Lambda(x)$ is a vector with integer coordinates so that the sum of its coordinates is $[\frac{l}{2}]$ and there is at least one of them which is at least 2. This implies that $\|D\Lambda(x)\|^2 ge\rho(4+l-1) = \rho(l+3)$ and so $\|D\Lambda(x)\|^2 > \|D\Lambda(v)^2\|$.

We can apply now Lemma (2.2) and get that $u \in S_2$ and $\|D\Lambda(u)\|^2 = \|D\Lambda^v\|^2$. As we have proved above this implies that $x_i = g_{C,u}(i)$ is a solution of the restricted subset sum porblem. *Q.E.D.*(Corollary 2.2)

5. Proof of Theorem 2.2. In this section we prove Theorem 2.2. We will prove the large deviation theroems used here in the next section.

Definition. Assume that $\langle S, X\rangle$ is a hypergraph. $\langle S, X\rangle$ will be called (M, α, k)-dispersed, if $|\{T \in X | B \subseteq T\}| \le \alpha^{-i} M$ holds for all $i = 0, 1, ..., k$ and for all $B \subseteq S$, $|B| = i$. If $Z \subseteq S$ then we will say that $\langle S, X\rangle$ is (M, α, k)-dispersed on Z if $|\{T \in X | B \subseteq T\}| \le \alpha^{-i} M$ holds for all $i = 0, 1, ..., k$ and for all $B \subseteq Z$, $|B| = i$.

Remark. If S is the set of primes $p_1, ..., p_\iota$ (as defined in Lemma 3.2) and X consists of those subsets $T \subseteq S$ with $\prod_{p \in T} p \in I$ where I is a fixed interval, then X will have very strong dispersion properties in the sense defined above. Namely if $B \subseteq S$ then we have (approximately) that $|\{T \in X \mid B \subseteq T\}| \le |I| \prod_{p \in T} p^{-1}$. We do not utilize this property in our proof of the NP-hardness result (although it was used in an earlier version of the proof). Theorem 2.2 does not assume any dispersion property of the set X, we have assumptions only about its cardinality. (We will show that there is large a subset of X with nice dispersion properties.) Taking into account the dispersion properties of the hypergraph defined from the primes and adding the corresponding condition to the assumptions of Theorem 2.2 is a possibility for improvments. (It may result only in better constants.)

Sketch of the proof of Theorem 2.2. We will randomize the sets $C_1, ..., C_k$ sequentially. The elements of the individual sets C_i will be also randomized sequentially, more precisely in bigger blocks. After each randomization we count the number of sets $T \in X$ which intersect the already randomized (or partially radnomized) sets C_i in a given number of elements. The "given number" will be only 0 or 1. In other words assume that f is a fixed 0, 1-function defined on k, and suppose that $C_1, ..., C_{i-1}$ has been already randomized and C_i partially randomized ($C_i^{\prime}$ will denote the part of C_i which has been already selected). We count the number of $T \in X$ with $f(j) = |C_j \cap T|$ for $j = 1, ..., i-1$ and $f(i) = |C_i^{\prime} \cap T|$. We prove the theorem by showing that with a high probability this number g is always close to its expected value. More precisely this will not be necessarily true for our original set X but we show that X has a subset X' with nice dispersion properties which still has sufficiently many elements (Lemma 5.2) and for such an X, g and its expected value will be always close.

Without some dispersion property we cannot expect that such a statement is true. Indeed assume that all of the sets in X contains a single point $a \in S$. If a is selected as a memeber of a C_j then automatically $|T \cap C_j| \neq 0$ for all $\in X$ although the expected number of such sets $T \in X$ can be quite large. Therefore we need an assumtpion that the number of sets in X containing any fixed point of S is not too large compared to $|X|$ (less than a polynomial fraction). Actually we will use a similar assumption about the cardinality of sets containing two fixed points. This is equivalent to the statement that X is $(|X|, n^{-\alpha}, 2)$ dispersed. Even if we start with such an X our proof may break down, because after selecting $C_1, ..., C_i$ we will need that if we replace X by X_i the set of all $T \in X$ with $f(j) = |T \cap C_j|$ for $j = 1, ..., i$, then X_i has the same dispersion property. Unfortunately the $(|X|, n^{-\alpha}, 2)$ dispersion property is not inherited this way. (We will not have a similar problem while selecting the elements a single set C_i because during this process we measure the error relative to the size of X at the beginning of the selection process.) Assume now that X is $(|X|, n^{-\alpha}, l)$ dispersed, for some positive integer l and let's see what happens during the choice of C_1. If we are able to prove that the number g is close to its expected value using that X is s $(|X|, n^{-\alpha}, 2)$ dispersed, then we may apply the same proof to the subset of X containing $l - 2$ fixed points. (This subset is $(|X|, n^{-\alpha}, 2)$ dispersed.) This way we get (approximatley) that X_1 is $(|X_1|, n^{-\alpha}, l-2)$ dispersed. Continueing selections of the sets C_i we get that X_i is $(|X_i|, n^{-\alpha}, l-2i)$ dispersed. Lemma 5.2 guarantees the dispersrion property of the selected $X' \subseteq X$ with a large l, so during the selections of the sets C_i, l will decrease but it will remain always larger than 2.

The subset X' selected in Lemma 5.2 has a special structure. namely all of it sets contain a common subset B with at most $(1-\beta)n$ elemetns for some constant $0 < \beta < 1$. In the remaining part of the proof we will consider only sets C_i which has an empty intersection with B. Since the probability that $B \cap \bigcup C_i = \emptyset$ is at least $1 - n_{-\delta}$ for some constant $\delta > 0$ it does not cause any problem. However if we could state the theorem

End of sketch.

We formulate now another weaker version of Theorem 2.2 where the statements are made only about hypergraphs $\langle |S|, X \rangle$ that are $(|X|, n^{-\alpha}, n)$ dispersed for some constant α. We also give some indication about the choice of the numbers δ_i, $i = 1, 2, 3$. Later we show that this waeker version implies the original theorem.

Notation. $x \prec y_1, ..., y_k$ will mean that x is sufficently small with respect to $y_1, ..., y_k$.

Lemma 5.1. *For all $\alpha_1 > 0$, $\alpha_2 > 0, \alpha_3 > 0$ and $\delta_1 > 0$ $\delta_2 > 0$, $\delta_3 > 0$ if $\delta_3 \prec \delta_1 \prec \delta_2 \prec \alpha_1, \alpha_2, \alpha_3$, and n is sufficiently large then the following holds:*

Assume that $\langle S, X \rangle$ is an n-uniform $(|X|, n^{-\alpha_3}, n)$-dispersed hypergraph, $n^2 \leq |S| \leq n^{\alpha_1}$, $|X| \geq 2^{\alpha_2 n \log n}$, $k = [n^{\delta_1}]$ and $C_1, ..., C_k$ is a random sequence of pairwise disjoint subsets each with exactly $[|S| n^{-1-\delta_2}]$ elements, with uniform distribution on the set of all sequences with these properties. Then with a probability of at least $1 - 2^{-n^{\delta_3}}$ the following holds:

for each $0,1$-valued function f defined on $\{1, ..., k\}$ there is a $T \in X$ so that $f(j) = |C_j \cap T|$ for all $j = 1, ..., k$.

We will show that this lemma implies theorem 2.2. The following lemma is needed for this proof.

Lemma 5.2. *For all $\alpha_1 > 0$, $\alpha_2 > 0$ there exist a $\beta > 0$ so that if n is sufficiently large then the following holds. Assume that $\langle S, X\rangle$ is an n-uniform hypergraph, $|S| \le n^{\alpha_1}$ and $|X| \ge 2^{\alpha_2 n \log n}$. Then there is a $B \subseteq S$, $|B| \le (1-\beta)n$, so that if $S' = S - B$, $X' = \{T - B \mid T \in X, B \subseteq T\}$ then $|X'| \ge 2^{\beta n \log n}$ and $\langle S', X'\rangle$ is an $n - |B|$-uniform $(|X'|, n^{-\beta}, n - |B|)$-dispersed hypergraph.*

Proof. Assume that $\beta > 0$, is sufficiently small with respect to α_1 and α_2. Let $B \in X$ be maximal with the following property:

(5.1) $\quad |\{T \in X \mid B \subseteq T\}| \ge n^{-\beta|B|}|X|.$

We claim that $|B| \le (1-\beta)n$. Indeed, otherwise the number of elements of X containing B is at most $|S|^{\beta n} \le (n^{\alpha_1})^{\beta n} = 2^{\beta\alpha_1 n \log n}$. The definition of B implies that this number is at least $|X|n^{-\beta n} \ge 2^{\alpha_2 n \log n - \beta n \log n} = 2^{(\alpha_2-\beta)n \log n}$. We got $2^{(\alpha_2-\beta)n\log n} \le 2^{\beta\alpha_1 n \log n}$. This is a contradiction since β is sufficiently small with respect to α_1, α_2.

Now we show that B satisfies the conditions of the lemma. According to (5.1) $|X'| > n^{-\beta|B|} \ge n^{-\beta n} \ge 2^{-\beta n \log n}$. Suppose that contrary to our assertion $\langle S', X'\rangle$ is not $(|X'|, n^{-\beta}, n - |B|)$-dispersed. Then there is a $D \in |X'|$ with $|D| = j$, $1 \le j \le n - |B|$ so that $|\{T' \in X' \mid D \subseteq T'\}| > |X'|n^{-\beta j}$. For each set T' in the last inequality we have that $B \cup T' \in X$. Therefore if $B' = B \cup D$ then $|\{T \in X \mid B' \subseteq T\}| \ge |X'|n^{-\beta j} \ge n^{-\beta|B| - \beta j} = n^{-\beta|B'|}$ in contradiction to the maximality of B. *Q.E.D.*(Lemma 5.2)

Now we show that Lemma 5.1 implies Theorem 2.2.

We apply Lemma 5.2 to the hypergraph $\langle S, X\rangle$ of theorem 2.2. Let $\langle S', X'\rangle$ be the hypergraph defined in Lemma 5.2. We have that it is $(|X'|, n^{-\beta}, n - |B|)$-dispersed and $|X'| \ge 2^{\beta n \log n}$. Let $m = n - |B|$. We may formulate some consequences of these properties in terms of m. We get that $|X'| \ge 2^{\beta m \log m}$ and $\langle S', X'\rangle$ is $(|X'|, n^{-\frac{\beta}{2}}, m)$ dispersed (since $m \ge \beta n \ge n^{\frac{1}{2}}$).

Now we may apply Lemma 5.1 to the hypergraph $\langle S, X\rangle$, with $n \to m$, $\alpha_2 \to \beta$, $\alpha_3 \to \frac{\beta}{2}$. We pick the numbers $\delta_i = 1, 2, 3$ so that they meet the requirement of Lemma 5.1. We claim that the requirements of Lemma 2.2 are also met with the same δ_1, δ_2 and $\delta_3 \to \frac{\delta_3}{2}$. Assume that we pick $C_1, \ldots, C_k$ at random with the distribution given in Lemma 2.2. $F = \bigcup_{i=1}^{k} C_i$ is a random subset of S with exactly $[n^{\delta_1}]|S|n^{-1-\delta_2} \le |S|n^{-1-\delta_2+\delta_1} \le |S|n^{-1-\frac{\delta_2}{2}}$ elements. (δ_1 was sufficiently small with repsect to δ_2). Since $|B| \le n$, we have that $P(F \cap B = \emptyset) \le nn^{-1-\frac{\delta_2}{2}} = n^{-\frac{\delta_2}{2}}$. The distribution of $C_1, \ldots, C_k$ with the condition $F \cap B = \emptyset$ is the same as the distribution of $C_1, \ldots, C_k$ according to Lemma 5.1. Therefore the probability that $C_1, \ldots, C_k$ satisfies the conclusion of Lemma 5.1 and so the conclusion of Theorem 2.2 as well is at least $1 - (n^{-\frac{\delta}{2}} - (1 - n^{-\frac{\delta}{2}})n^{-\delta_3}) \ge 1 - n^{-\frac{\delta_3}{2}}$. *Q.E.D.*(Theorem 2.2)

Proof of Lemma 5.1. We will pick the sets $C_1, ..., C_k$ sequentially and describe the distribution of the number of intersections $T \cap C_i$, $T \in X$ with one or zero elements. The following lemma describes what happens at the choice of a single C_i (C in the lemma below). The hypergaph $\langle S, Z \rangle$ in the lemma will play both the role of $\langle S', X \rangle$ (for some $S' \subseteq S$) and $\langle S', X_D \rangle$, where $D \subseteq S$ and $X_D = \{T \in X \mid D \subseteq T\}$.

Lemma 5.3. *For all $\beta_1 > 0, \beta_2 > 0$ and $\gamma_1 > 0, \gamma_2 > 0$, if $\gamma_2 \prec \gamma_1 \prec \beta_1, \beta_2$ then the following holds. Assume that $M > 0$ and $\langle S, Z \rangle$ is an n-uniform $(M, n^{-\beta_2}, 2)$-dispersed hypergraph, $\frac{1}{2}n^2 \leq |S| \leq n^{\beta_1}$ and C is chosen at random with uniform distribution from the set of subsets of $|S|$ with exactly t elements where $|S|n^{-1-2\gamma_2} \leq t \leq |S|n^{-1-\frac{1}{2}\gamma_2}$. Suppose further that the numbers λ_i, $i = 0, 1$ are given so that for any $W \subseteq S$ with $|W| = n$, $\lambda_i = P(|C \cap W| = i)$ for $i = 0, 1$.*

Then with a probability of at least $1 - 2^{-n^{\gamma_1}}$ we have that for $i = 0, 1$
if $|Z| \leq M$ then, $|\{T \in Z \mid |T \cap C| = i\}| = \lambda_i |Z| + R_i$ where $|R_i| \leq n^{-\gamma_1} M$

We continue the proof of Lemma 5.1, accepting Lemma 5.3. Assume that $\alpha_1, \alpha_2, \alpha_3$ are fixed, and $\delta_1, \delta_2, \delta_3$ were picked as described in the statement of the lemma.

Let f be a fixed $0, 1$-valued function defined on $\{1, ..., k\}$. Since the number of choices for f is at most $2^{n^{\delta_1}}$ it is sufficient to prove that for each fixed f the probability probability of the event A_f is at least 1-$2^{-n^{2\delta_1}}$, where A_f holds iff there is a $T \in X$ with $f(j) = |C_j \cap T|$ for $j = 1, ..., k$.

We estimate now the probability of A_f for a fixed f.

We will use the following notation. If $Y \subseteq S$ then $Y_f^{(i)}$ will denote the set of all $T \in X$, $Y \subseteq T$ with $f(j) = |C_j \cap T|$ for $j = 1, ..., i$. Let $b_i = \sum_{j=1}^{i-1} f(j)$. Suppose that $T \subseteq S$, $|T| = n$ and $|T \cap C_j| = f(j)$ for $j = 1, ..., i-1$. For fixed $C_1, ..., C_{i-1}$ and for the randomization of C_i let $\mu_i = P(|B \cap C_i| = f(i))$. (Clearly μ_i does not depend on the choice of $T, C_1, ..., C_i$.) Let $\kappa_i = \prod_{j=1}^{i} \mu_i$. Finally let $\Phi_i = \bigcup_{j=1}^{i} C_j$.

For later use we give a lower bound on κ_k. If $f(i) = 0$ then $|C_i| = t \leq |S|n^{-1-2\gamma_2}$, and $|S - \Phi_{i-1}| \geq \frac{1}{2}n^2 - n^{\delta_i}$ implies that $\mu_i \geq \frac{1}{2}$. If $f(i) = 1$ then it is easy to see that $\mu_i \geq n^{-3\gamma_2}$. Therefore $\kappa_k \geq \prod_{i=1}^{k} \mu_k \geq n^{-3k\gamma_2} \geq n^{-3\gamma_2 n^{\delta_1}}$

We have $\delta_3 \prec \delta_1 \prec \delta_2 \prec \alpha_1, \alpha_2, \alpha_3$. Let $\delta_4 > 0$ so that $\delta_2 \prec \delta_4 \prec \alpha_1, \alpha_2, \alpha_3$. Suppose these numbers are fixed and n is sufficiently large.

Claim 5.1. *For all $i = 1, ..., k$ with a probability of at least $1 - 2^{-n^{\delta_4} + in^{\frac{1}{2}\delta_4}}$ the following holds:*
for all $D \subseteq S$ with at most $2k - 2i$ elements if $X_D = \{T \in X \mid D \subseteq X\}$, then we have that either $D \cap \Phi_i \neq \emptyset$ or $|D_f^{(i)}| \leq (1 + in^{-\delta_4})\kappa_i |X| n^{-\alpha_3 |D|}$.

We prove the statement by induction on i. Assume that our statement holds for $i - 1$ and we prove it for i.

First we estimate the probability of the event that the Claim holds for a fixed set $D \subseteq S$. Assume therefore that $D \subseteq S$ is fixed $|D| \le 2k - 2i$. We formulate a consequence of the inductive assumption. After the randomization of $C_1, ..., C_{i-1}$ with a probability of at least $1 - 2^{-n^{\delta_4} + (i-1) n^{\frac{1}{2}\delta_4}}$: for all $D' \supseteq X$ with at most $|D| + 2$ elements if $X_{D'} = \{T \in X \mid D' \subseteq T\}$, then we have that either $D' \cap \Phi_{i-1} \neq \emptyset$ or $|(D')_f^{(i-1)}| \le (1 + (i-1)n^{-\delta_4})\kappa_{i-1}|X|n^{-\alpha_3|D'|}$. (In other words we need the inductive hypothesis only for sets containing D.) Assume that $D \cap \Phi_{i-1} = \emptyset$. The last inequality implies that if $S' = S - \Phi_{i-1} - D$ and $X' = \{S' \cap T \mid T \in D_f^{(i-1)}\}$ then the hypergraph $\langle S', X' \rangle$ is $(M, n^{-\alpha_3}, 2)$ dispersed where $M = (1 + (i-1)n^{-\delta_4})\kappa_{i-1}|X|$. Now we randomize C_i. If $C_i \cap D \neq \emptyset$ then the conclusion of the lemma holds. Assume that $C_i \cap D = \emptyset$, that is, we randomize C_{i-1} with this condition. This means that we pick C_{i-1} with uniform distribution from the set of subsets of $S - \Phi_{i-1} - D$ with exactly $|S|n^{-1-\delta_2}$ elements. Therefore we may apply the upper bound of Lemma 5.3 with $S \to S'$, $Z \to X'$, $C \to C_i$, $M \to (1 + (i-1)n^{-\delta_4})\kappa_{i-1}|X|n^{-\alpha_3|D|}$, $i \to f(i)$, $\beta_1 \to \alpha_1$, $\beta_2 \to \alpha_3$, $t \to k$, $\gamma_1 \to 4\delta_4$, $\gamma_2 \to \delta_2$.

We get that with a probability of at least $1 - 2^{-4\delta_4}$

$$|D_f^{(i)}| \le \lambda_{f(i)}(1 + (i-1)n^{-\delta_4})(1 + n^{-4\delta_4})\kappa_{i-1}|X|n^{-\alpha_3|D|}.$$

The definitions of λ_j and μ_j and $|S| \ge n^2$ imply that $|\lambda_{f(i)} - \mu_{f(i)}| < \frac{1}{n}$ therefore $|D_f^{(i)}| \le (1 + in^{-\delta_4})\kappa_{i-1}|X|n^{-\alpha_3|D|}$. We got that for every fixed D with a probability of at least $1 - 2^{-4\delta_4}$ either $D \cap \Phi_i \neq \emptyset$ or the upper bound on $|D_f^{(i)}|$ (described in the claim) holds. This is true for every fixed $D \subseteq S$ with at most $2k - 2i$ elements. The number of possible sets D is at most $|S|^k \le n^{\alpha_1 n^{\delta_1}} \le 2^{\alpha_1 \log n n^{\delta_1}} < 2^{n^{\frac{1}{4}\delta_4}}$, therefore by taking the sum of the exceptional probabilities for each D we get that the assertion of the Claim is true.

Now we may repeat the proof of the claim for $D = \emptyset$ but both the upper and lower bounds of Lemma 5.3. (The already proven Claim will guarantee that the requirements of Lemma 5.3 are met.) The application of (1) yields that with a probability of at least $1 - 2^{-n^{\frac{1}{2}\delta_4}}$ the lower bound $|\emptyset_f^i| \ge (1 - in^{-\delta_4})\kappa_i|X|$ holds for $i = 1, ..., k$. This, $\kappa_k > n^{-3\gamma_2 n^{\delta_1}} \ge 2^{-3\gamma_2 n^{\delta_1}} \log n$, $|X| \ge 2^{\alpha_2 \log n}$ implies that with a probability of at least $1 - 2^{-n^{\frac{1}{2}\delta_4}}$ the set $\emptyset_f^k$ is not empty and therefore A_f holds. $\delta_1 \prec \delta_4$ implies that we have proved the required lower bound on the probability of A_f. *Q.E.D.*(Lemma 5.1).

Proof of Lemma 5.3. Let $\delta > 0$ with $\gamma_2 \prec \gamma_1 \prec \delta \prec \beta_1, \beta_2$ and let $r = \lceil n^{\delta} \rceil$. We randomize a C by selecting at random a sequence of pairwise disjoint subsets $\mathcal{C}_1, ..., \mathcal{C}_r$ of S whose union will be C. For the description of this randomization let q_i be a fixed positive integer for $i = 1, .., r$ so that $\frac{1}{2}tr^{-1} < q_i < 2tr^{-1}$ and $\sum_{i=1}^{r} q_i = t$. $\mathcal{C}_1, ..., \mathcal{C}_r$ will be a random sequence of pairwise disjoint subsets of S so that $|\mathcal{C}_i| = q_i$ for $i = 1, ..., r-1$, moreover we take the sequence $\mathcal{C}_1, ..., \mathcal{C}_r$ with uniform distribution form the set of all sequences with the described properties. Let $C = \bigcup_{i=1}^{r} \mathcal{C}_i$. Clearly the distribution of C meets the requirements of the lemma. We will randomize the sets $\mathcal{C}_1, ..., \mathcal{C}_r$ sequentially. Let $\Phi_i = \bigcup_{j=1}^{i} \mathcal{C}_i$ and $Z_{\iota}^{(i)} = \{T \in Z \mid |T \cap \Phi_i| = \iota\}$ for $i = 0, 1, ..., r$, $\iota = 0, 1$, $(\Phi_0 = \emptyset,$

$Z_0^{(0)} = Z$, $Z_0^{(1)} = \emptyset$). We note that if $\mathcal{C}_1,...,\mathcal{C}_{i-1}$ are fixed then the distribution of $Z_\iota^{(i)}$ depends only on the sets $Z_0^{(i-1)}$, $Z_1^{(i-1)}$ and the remaining part of Z is irrelevant. (These other elements of Z have already an intersection with Φ_{i-1} containing at least two elements, so they cannot contribute the sets $Z_\iota^{(i)}$, $\iota = 0, 1$.) Therefore we will follow the evolution of the sets $Z_\iota^{(i)}$, $\iota = 0, 1$ as i grows from 0 to r. Actually we will need only the numbers $|Z_\iota^{(i)}|$, $\iota = 0, 1$, since, as we will prove, the numbers $|Z_\iota^{(i-1)}|$, $\iota = 0, 1$ approximately determine $|Z_\iota^{(i)}|$, $\iota = 0, 1$. To show this we need a lemma about the choice of a single set $\mathcal{C}_i$. This lemma is very similar to Lemma 5.3, only the random set is of smaller size. The advantage of the smaller size is that the probability that $\mathcal{C}_i$ intersects a fixed $T \in Z$ in more than one point is so small that everything that we get this way can be included in the error term.

Lemma 5.4. *For all $\beta_1 > 0, \beta_2 > 0$ and $\gamma_1 > 0$, if $\gamma_1 \prec \beta_1, \beta_2$ then the following holds. Assume that $M > 0$ and $\langle S, Z\rangle$ is an n-uniform $(M, n^{-\beta_2}, 2)$-dispersed hypergraph, $\frac{1}{2}n^2 \le |S| \le n^{\beta_1}$ and $\mathcal{C}$ is chosen at random with uniform distribution from the set of subsets of $|S|$ with exactly r elements where $r = |S|n^{-1-\gamma_2}$, $\gamma_1 \le \gamma_2 \le 2\gamma_1$. Suppose further that the numbers σ_i, $i = 0, 1$ are given so that for any $W \subseteq S$ with $|W| = n$, $\sigma_i = P(|\mathcal{C} \cap W| = i)$ for $i = 0, 1$.*

Then with a probability of at least $1 - 2^{-n^{\gamma_1}}$ we have that for $i = 0, 1$ if $|Z| \le M$ then, $|\{T \in Z \mid |T \cap \mathcal{C}| = i\}| = \sigma_i|Z| + R_i$ where $|R_i| \le 10n^{-2\gamma_2}M$.

Proof. First we estimate $\{T \in Z \mid |T \cap \mathcal{C}| \ge 2\}$. For this we use the following lemma.

Lemma 5.5. *For all sufficiently small $\delta > 0$ and for all sufficiently large positive integers s and t the following holds: assume that A is a finite set, $w(x, y)$ is a real-valued function of two variables defined on A and $K > 0$ is a real number with the following properties:*
(1) $0 \le w(x, y)$ for all $x \in A$,
(2) $\sum_{x\in A} w(x, y) \le K$ for all $y \in A$,
(3) $w(x, y) = w(y, x)$ for all $x, y \in A$,
(4) $s^2|A|^{-2}(\sum_{x\in A}\sum_{y\in A} w(x, y)) \le Kt$,
(5) $|A| > st^2$
Assume further that Y is a random variable uniformly distributed on the set of subsets of A with exactly s elements. Then with a probability of at least $1 - 2^{-t^\delta}$, we have that

$$\sum_{x\in Y}\sum_{y\in Y} w(x, y) \le 2Kt$$

We apply Lemma 5.5 with $A \to S$. Let $w(x, x) = 0$ for all $x \in S$ and for all $x \ne y$ let $w(x, y) = |\{T \in Z \mid x, y \in Z\}|$, $K = Mn^{-\beta_2}$. Since $\langle S, Z\rangle$ is $(M, n^{-\beta_2}, 1)$ dispersed we have that for each fixed $y \in A$, $\sum_{x\in S} w(x, y) \le |\{T \in Z \mid x \in Z\}| \le K$. $\sum_{x\in S}\sum_{y\in S} w(x, y) = \sum_{T\in Z} |Z|(n^2 - n)| \le Mn^2 \le Kn^2n^{\beta_2}$. Therefore

$r^2|S|^2 \sum_{x\in S}\sum_{y\in S} w(x,y) \le n^{-2-2\gamma_2}Kn^{\beta_2} = Kn^{\beta_2-2\gamma_2}$. Therefore we apply Lemma 5.5 with $t \to n^{\beta_2-2\gamma_2}$. We pick $\delta > 0$ from Lemma 5.5 with $\gamma_1 \prec \delta \prec 1$. The conclusion of Lemma 5.5 is that with a probability of at least $1 - 2^{-n^{\delta(\beta_2-2\gamma_2)}}$ $\sum_{x\in\mathcal{C}}\sum_{y\in\mathcal{C}} w(x,y) \le 2Mn^{-\beta_2}n^{\beta-2\gamma_2} = 2Mn^{-2\gamma_2}$. Since $\gamma_2 \prec \delta(\beta_2 - \gamma_2)$ this implies that with a probability of at least $1 - 2^{-n^{2\gamma_1}}$, $|\{T \in Z \mid |T \cap \mathcal{C}| \ge 2\}| \le n^{-2\gamma_2}M$.

Using this we get that with a probability of at least $1 - 2^{-n^{2\gamma_2}}$

(5.2) $|\{T \in Z \mid |T\cap\mathcal{C}| = 1\}| = \sum_{x\in\mathcal{C}} |\{T \in Z \mid x \in T\}| + R_3$ where $|R_3| \le n^{-2\gamma_2}$.

We estimate the sum $\sum_{x\in\mathcal{C}} |\{T \in Z \mid x \in T\}|$ using the following lemma:

Lemma 5.6. *For all sufficiently small $\delta > 0$ and for all sufficiently large positive integers s and t the following holds: assume that A is a finite set, w is a real-valued function on A and $K > 0$ is a real number with the following properties:*

(1) $0 \le w(x) \le K$ for all $x \in A$,

(2) $s|A|^{-1}\sum_{x\in A} w(x) \le Kt$,

(3) $|A| > st^2$.

Assume further that Y is a random variable uniformly distributed on the set of subsets of A with exactly s elements. Then with a probability of at least $1 - 2^{-t^{\delta}}$, we have that

$$|(\sum_{y\in Y} w(y)) - s|A|^{-1}\sum_{x\in A} w(x)| \le Kt^{1-\delta}$$

We apply this lemma with $A \to S$, $Y \to \mathcal{C}$. We define the function w by $w(x) = |\{T \in Z \mid x \in T\}|$. Since $\langle S, Z\rangle$ is $(M, n^{-\beta_2}, 1)$-dispersed we have that $w(x) \le Mn^{-\beta_2}$ for all $x \in S$, therefore we put $K = Mn^{-\beta_2}$. $s|A|^{-1}\sum_{x\in A} w(x) \le n^{-1-\gamma_2}|M|n = Mn^{-\gamma_2} = Kn^{\beta_2-\gamma_2}$. Therefore we apply Lemma 5.6 with $t \to n^{\beta_2-\gamma_2}$. Let $\gamma_1 \prec \delta \prec 1$. Lemma 5.6 implies that with a probability of at least $1 - 2^{-n^{2\gamma_1}}$ we have that

(5.3) $|(\sum_{x\in\mathcal{C}} |\{T \in Z \mid x \in T\}|) - r|S|^{-1}\sum_{x\in S} |\{T \in Z \mid x \in T\}|| \le n^{-2\gamma_2}$.

We will also need an upper bound on the probability that a fixed $W \subseteq S$, $|W| = n$ intersects $\mathcal{C}$ in at least two points. We claim that

For any $W \subseteq S$, $|W| = n$ we have that

(5.4) $P(|\mathcal{C} \cap W| \ge 2) \le n^{-\gamma_2}$ for any $W \subseteq S$, $|W| = n$.

Indeed $P(|\mathcal{C} \cap W| \ge 2) \le \binom{n}{2}(n^{-1-\gamma_2})^2 \le n^{-2\gamma_2}$.

We estimate the second term in (5.3). $r|S|^{-1}\sum_{x\in S} |\{T \in Z \mid x \in T\}| = r|S|^{-1}|Z|n$. According to the definition of σ_1 and (5.4) if $q = r|S|^{-1}$ then $\sigma_1 = nq + R_4$, where $|R_4| \le n^{-2\gamma_2}$. Therefore we get that

$r|S|^{-1}\sum_{x\in S} |\{T \in Z \mid x \in T\}| = r|S|^{-1}|Z|n = \sigma_1|Z| + R_5$ where $|R_5| \le 4Mn^{-2\gamma_2}$. This, (5.3), (5.2) implies that with a probability of at least $1 - 2^{-\gamma_1}$ we have $|\{T \in Z \mid |T \cap \mathcal{C}| = 1\}| = \sigma_1|Z| + R_1$, where $|R_1| \le 4n^{-2\gamma_1}M$, that is we have proved the statement of the Lemma for $i = 1$.

(5.4) implies that $\sigma_0 + \sigma_1 = 1 + R_6$, where $|R_6| \le n^{-2\gamma_2}$. This, the already proven part of the Lemma for $i = 1$ and (5.2) and (5.4) implies that with a probability of at least $1 - 2^{-\gamma_1}$ we also have $|\{T \in Z \mid |T \cap \mathcal{C}| = 1\}| = \sigma_0|Z| + R_0$, where $|R_0| \le 10n^{-2\gamma_1}M$. *Q.E.D.*(Lemma 5.4)

Now we continue the proof of lemma 5.3. Let $\chi_{i,\iota} = |Z^i_\iota|$, $i = 0, 1, ..., r$, $\iota = 0, 1$. Let $\xi_{i,\iota}$ be the expected number of elements of Z which intersect Φ_i in exactly ι points for $i = 0, 1, ..., r$, $\iota = 0, 1$. First we show that the sequence of 2-dimensional vectors $\langle \xi_{i,0}, \xi_{i,1} \rangle$, $i = 1, ..., r$ satisfies a linear recursion, then we show that the sequence $\langle \chi_{i,0}, \chi_{i,1} \rangle$ satisfies the same recursion approximately. From this we will get an upper bound on the distance of the corresponding elements of the two sequences.

Let $\mu_{i,\iota}$ be the probability that any fixed element T of Z intersects Φ_i in exactly ι points. (Since Z is μ-uniform the value of $\mu_{i,\iota}$ does not depend on the choice of T.) We have that $\xi_{i,\iota} = |Z|\mu_{i,\iota}$. Let T be any fixed element Z and we consider the conditional probabilities: $\gamma_{i,\iota,\kappa} = P(|T \cap \mathcal{C}_i| = \iota \mid |T \cap \mathcal{C}_{i-1}| = \kappa)$ for $i = 1, ..., r$, $\iota = 0, 1$, $\kappa = 0, 1$. (Again $\gamma_{i,\iota,\kappa}$ does note depend on the choice of T. Also note that $\gamma_{i,0,1} = 0$) Let Γ_i be the matrix

$$\begin{pmatrix} \gamma_{i,0,0} & \gamma_{i,0,1} \\ \gamma_{i,1,0} & \gamma_{i,1,1} \end{pmatrix}$$

If $\bar{\mu}_i = \begin{pmatrix} \mu_{i,0} \\ \mu_{i,1} \end{pmatrix}$, then by the defintions of $\gamma_{i,\iota,\kappa}$ and $\mu_{i,\iota}$ we have $\Gamma_i\bar{\mu}_{i-1} = \bar{\mu}_i$ for $i = 1, ..., r$. Let $\bar{\xi}_i = Z\bar{\mu}_i$. We also have $\Gamma_i\bar{\xi}_{i-1} = \bar{\xi}_i$, $i = 1, ..., r$. Let $\bar{\chi}_i = \begin{pmatrix} \chi_{i,0} \\ \chi_{i,1} \end{pmatrix}$. We claim that with a probability of at least $1 - 2^{-2\gamma_1}$ we have

(5.5) $\quad \Gamma_i\bar{\chi}_{i-1} = \bar{\chi}_i + R_i$ where $\|R_i\|_{L_1} \le n^{-2\delta}M$.

(Note that the error term R_i is a two dimensional vector and we use the L_1 norm to measure it.) Indeed this is an immediate consequence of Lemma 5.4 if we apply it separately with $Z \to Z^{(i-1)}_0$, $n \to n$, $S \to S - \Phi_i$ and $Z \to Z^{(i-1)}_1$, $n \to n - 1$, $S \to S - \Phi_i$.

Therefore, if (5.5) holds, we have two sequences of vectors $\bar{\xi}_0, ..., \bar{\xi}_r$ and $\bar{\chi}_0, ..., \bar{\chi}_r$ so that $\bar{\xi}_0 = \bar{\chi}_0$, $\bar{\xi}$ satisfies the recursion $\Gamma_i\bar{\xi}_{i-1} = \bar{\xi}_i$ and $\bar{\chi}_i$ satisfy the same recursion approximately. Any 2 by 2 matrix A induces a linear transformation on $\mathbf{R}^2$ so we may define its norm as the norm of the corresponding linear transformation of $\mathbf{R}^2$ with respect to the L_1 norm namely $\|A\| = \sup\{\|Ax\|_{L_1} \mid x \in \mathbf{R}^2, \|x\|_{L_1} \le 1\}$. Since the entries of Γ_i are nonnegative numbers and their sum is at most 1 in each column (they are the probabilities of disjoint events with the same condition) we have that $\|\Gamma_i\| \le 1$. So we get, using the exact and approximate recursions that

$\|\bar{\xi}_i - \bar{\chi}_i\| \le \|\Gamma_i(\bar{\xi}_{i-1} - \bar{\chi}_{i-1} + R_i)\| \le \|\Gamma_i\|(\|\bar{\xi}_{i-1} - \bar{\chi}_{i-1}\|_{L_1} + \|R_i\|\text{Ł}_1) \le \|\bar{\xi} - \bar{\chi}_i\|_{L_1} + n^{-2\delta}M$. Therefore if $D_i = \|\bar{\xi}_i - \bar{\chi}_i\|$ then we got that $D_i \le D_{i-1} + n^{-2\delta}M$ and therefore $D_r = \|\bar{\chi}_r - \bar{\xi}_r\|_{L_1} \le rn^{-2\delta M} \le n^\delta n^{-2\delta}M = n^{-\delta}M \le n^{-\gamma_1}M$, that is

(5.6) $\quad \|\bar{\chi}_r - \bar{\xi}_r\|_{L_1} \le n^{-\gamma_1}M$.

Since $C = \Phi_r$, we have $\lambda_\iota = \mu_{r,\iota}$ and so $\bar{\xi}_r = \begin{pmatrix} \lambda_0 |Z| \\ \lambda_1 |Z| \end{pmatrix}$. Therefore the definitions of $\bar{\chi}_r$ $\chi_{r,\iota}$ and $Z_\iota^{(r)}$ and (5.6) imply the conclusion of the lemma. *Q.E.D.*(Lemma 5.3).

6. Large deviation theorems. In this section we prove Lemma 5.6 and Lemma 5.5. In their proofs we use the following well-known large deviation theorem about independent random variables with identical distributions.

Theorem 6.1. *For all sufficiently small $\delta > 0$ if t is sufficiently large, $K > 0$, n is a positive integer then the following holds. Assume that $X_1, ..., X_n$ are mutually independent random variables with identical distributions which take their values in the interval $[0, K]$ so that $E(\sum_{i=1}^n X_i) \le Kt$. Then with a probability of at least $1 - 2^{-t^\delta}$ we have*

$$|(\sum_{i=1}^n X_i) - E(\sum_{i=1}^n X_i)| \le Kt^{1-\delta}$$

Remarks. 1. The theorem remains valid if the distributions of the random variables $X_1, ..., X_n$ are not necessarily identical. However we use the theorem only in the case of identical distributions.

2. Although the theorem is "well-known" we haven't located yet a reference which would imply it without much additional work. (Any suggestion to an appropriate reference is welcome.) The theorem can be proved by Chernoff's method of bounding large deviations (see [Ch]). In the special case when the random variables X_i take only the two extreme values 0 and K the theorem follows immediately from Corollary A.14. (p. 93) of [AS]. The proof given there (based on Chernoff's method) can be modified for the general case. If we do not want to get the optimal constants then the proof can be simplified by proving only the one-sided inequality

$(\sum_{i=1}^n X_i) - E(\sum_{i=1}^n X_i) \le Kt^{1-\delta}$.

This implies the inequality in the other direction if the random variables satisfy the additional condition, $E(X_i) \ge \frac{K}{2}$. In this case we can apply the already proven part to the random variables $K - X_i$. (Because of our additional assumption $E(\sum_{i=1}^n (1 - X_i)) \le E(\sum_{i=1}^n X_i)$ and therefore the requirements of the theroem are met.) The general case can be reduced to this special case by dividing the interval $[1, n]$ into $\sqrt{t}$ subintervals of equal size (with one possible exception), and for each subinterval I consider $Y_i = \sum_{i \in I}^n X_i$ as a single random variable. (We change Y_i slightly by excluding all of its values which are higher then $\frac{4}{3}$ times its expected value. By the already proven part of the theorem this causes only an exponentially small change in the distribution.) With the appropriate change in K the new random variables satsify the additional condition so we may apply the theorem to them.

Proof of Lemma 5.6. We pick the set in the following way. We choose a sequence of points $a_1, a_2, \ldots$ independently and with uniform distribution from A. Let q be the smallest integer so that $|\{a_1, \ldots, a_q\}| = s$ and let $Y = \{a_1, \ldots, a_q\}$.

We claim that if δ is sufficiently small than with a probability of at least $1 - 2^{-2t^{\delta}}$,

(a) $|(\sum_{y \in Y} w(y)) - \sum_{j=1}^{q} w(a_i)| \le K t^{1-2t^{\delta}}$.

(b) $|s - q| \le t^{\delta}$

Theorem 6.1 implies that with a probability of at least $1 - 2^{-2t^{\delta}}$ we have $|(\sum_{i=1}^{q} w(a_i)) - q|A|^{-1} \sum_{x \in A} w(x)| \le K t^{1-4\delta}$. (a) and (b) imply the statement of the lemma.

The inequality $|A| > st^{-1}$ imply that for any fixed $j = 1, \ldots, r$, $P(a_j \in \{a_1, \ldots, a_{j-1}) < t^{-1}$. Therefore by Theorem 6.1 (b) holds. This also implies that the upper bound of Lemma 5.6 is true.

For the proof of (a) for each integer i let

$Y_i = \{a \in A \mid a \text{ occurs with multiplicity } i \text{ in the sequence } a_1, \ldots, a_q\}$,

Let $|Y_i| = y_i$, and $p_i = P(y_i > t^{-\frac{i}{2}} q)$. Theorem 6.1 imply that $\sum_{i=1}^{\infty} p_i < 2^{-n^{3\delta}}$ therefore with probability of at least $1 - 2^{-n^{2\delta}}$ we have that $|Y_i| < t^{-\frac{i}{2}q}$ for all $i = 1, \ldots, 2$. Each Y_i cam be exteded into a random subset with exactly $[t^{-\frac{i}{2}q}]$ elements. For $i \le t$ we estimate $\sum_{j \in Y_i} w(a_j)$ by the the already proven upper bound of Lemma pontok. For all $i > t$ we use Markov's inequality to estimate the probability that $\sum\{w(a_j) | j \in \bigcup_{i > t} Y_i$ is at least twice as much as its expected value. These bounds together imply (b). *Q.E.D.*(Lemma 5.6)

Proof of Lemma 5.5. Let $r = [t^{\frac{1}{3}}]$ and let $q_1, \ldots, q_r$ be integers with $\frac{1}{2} t^{-\frac{1}{3}} s \le q_i \le 2t^{-\frac{1}{3}} s$. (If $s < t^{\frac{1}{2}}$ then the conclusion of the Lemma is a trivial consequence of (2)). Let $Y_1, \ldots, Y_r$ be a sequence of pairwise disjoint random subsets of A so that $|Y_i| = q_i$ for $i = 1, \ldots, r$ with uniform distribution on the set of all sequences with these properties. We claim that with a probability of at least $1 - 2^{-t^{2\delta}}$ we have

(a) for each $1 \le i < j \le q$, $\sum_{x \in Y_i} \sum_{y \in Y_j} w(x, y) \le \frac{q_i q_j}{s^2} 2Kt$

(b) for each $i = 1, .., q$, $\sum_{x \in Y_i} \sum_{y \in Y_i} w(x, y) \le \frac{q_i^2}{s^2} 2K t^{1+\frac{1}{3}}$.

Clearly (a) and (b) imply the conclusion of the lemma.

Proof of (b). Assume that i is fixed and we randomize Y_i by choosing a sequence of distinct random points $y_1, \ldots, y_{q_i}$ with uniform distribution from A. Assume that $y_1, \ldots, y_m$ has been already selected. We claim that

(c) $P(\sum_{j=1}^{q_i} \sum_{k=m+1}^{q_i} w(y_j, y_k) > \frac{q_i^2}{s^2} 2tK) < \frac{1}{2}$.

This is an immediate consequence of Markov's inequality. Using (c) we may conclude the proof of (b) in the following way: As we randomize the points $y_1, \ldots, y_{q_i}$ sequenntially, we also select a sequnce of positive integers $d_0 = 0 < d_1 < \ldots$ so that d_j is the smallest integer with the property

$\sum_{j=1}^{q_i} \sum_{k=d_{j-1}+1}^{d_j} w(y_j, y_k) > \frac{q_i^2}{s^2} 2tK$.

Because of (c) the probability that after selecting d_j we will not be able to select a d_{j+1} is at least $\frac{1}{2}$. Therefore the probability that d_j cannot be selected for already some $j < t^{\frac{1}{3}}$ is at least $1 - t^{\frac{1}{3}}$ which implies (b).

Proof of (a). Assume that we randomize first the set Y_i and then the set Y_j. Suppose that the first randomization has been already completed, that is, Y_i is fixed. For each $y \in Y_1$ let $w_i = \sum_i$ For all $a \in A - Y_i$ let $w'(a) = \sum_{x \in Y_i} w(x, a)$. (2) implies that $w'(a) \leq 1$ for all $a \in A - Y_1$. We apply Lemma 5.6 with $A \to A - Y_1$, $w \to w'$. *Q.E.D.*(Lemma 5.5)

REFERENCES

[ABSS] S. Arora, L. Babai, J. Stern, Z. Sweedyk, The hardness of approximate optima in lattices, codes and systems of linear equations, Proc. 34-th Annual Symp. Found. Computer Science, 1993, pp. 724-733.

[AD] M. Ajtai, C. Dwork, A Public Key Cryptosystem with Worst-Case/Average-Case Equivalence. Proc. 29-th Annual ACM Symp. on the Theory of Computing, 1997, pp. 284-293, and Electronic Colloquium on Computational Complexity, 1996, http://www.eccc.uni-trier.de/eccc/

[Adl] L. Adleman, Factoring and Lattice Reduction, Manuscript, 1995.

[Ajt] M. Ajtai, Generating Hard Instances of Lattice Problems, Proc. 28-th Annual ACM Symp. on the Theory of Computing, 1996, pp. 99-108, and Electronic Colloquium on Computational Complexity, 1996, http://www.eccc.uni-trier.de/eccc/

[AS] N. Alon, J. Spencer, The Probabilisitc Method, John Wiley & Sons, 1992.

[Ch] H. Chernoff, A Measure of Asymptotic Efficiency for Tests of a Hypothesis Based on the Sum of Observations, Annals of Mathematical Statistics 1952, (23), pp 493-507.

[vEB] P. Van Emde Boas, Another NP-complete partition problem and the complexity of computing short vectors in a lattice, Tech. Report 81-04, Dept. of Mathematics, Univ. of Amsterdam, 1980.

[GGH1] O. Goldreich, S. Goldwasser, S. Halevi, Collision-free hashing from lattice problems, Electronic Colloquium, 1996, on Computational Complexity, http://www.eccc.uni-trier.de/eccc/

[GGH2] O. Goldreich, S. Goldwasser, S. Halevi, Public-key cryptosystems from lattice reduction problems, Electronic Colloquium on Computational Complexity, 1996, http://www.eccc.uni-trier.de/eccc/

[K] R. M. Karp, Reducibility among combinatorial problems, in Complexity of Computer Computation, e.d. R. E. Miller and J. W. Thatcher, New York: Plenum Press, 1972

[LLS] J. Lagarias, H.W Lenstra and, C. P. Schnorr, Korkine-Zolotarev bases and successive minima of a lattice and its reciprocal lattice. Combinatorica **10** (1990), 333-348

[S] N. Sauer On the density of families of sets, Journal of Combinatorial Theory, Series A13, 1972, 145-147

[V1] A. Vardy, The Intractability of Computing the Minimum Distance Code, Preprint

[V2] A. Vardy, Algorithmic Complexity in Coding Theory and the Minimum Distance Problem. Proceeding of the Twenty-Ninth Annual Symposium on Theory of Computing, May 4-6, El Paso, Texas, pp 92-109

[VC] V.N. Vapnik and Ya. Chervonenkis, On the uniform convergence of relative frequencies of events to their probabilities, Theory of Probability Applications 1971, (16), 264-280.