Information Processing Letters 145 (2019) 48–52

Contents lists available at ScienceDirect

Information Processing Letters

www.elsevier.com/locate/ipl

# Reductions in **PPP**

Frank Ban [a], Kamal Jain [b], Christos H. Papadimitriou [c], Christos-Alexandros Psomas [d,*], Aviad Rubinstein [e]

[a] *UC Berkeley, USA*
[b] *Faira.com, USA*
[c] *Columbia University, USA*
[d] *Carnegie Mellon University, USA*
[e] *Stanford University, USA*

ARTICLE INFO

*Article history:*
Received 17 May 2018
Accepted 29 December 2018
Available online 17 January 2019
Communicated by Ryuhei Uehara

*Keywords:*
Computational complexity
Theory of computation
Combinatorial problems

ABSTRACT

We show several reductions between problems in the complexity class **PPP** related to the pigeonhole principle, and harboring several intriguing problems relevant to Cryptography. We define a problem related to Minkowski's theorem and another related to Dirichlet's theorem, and we show them to belong to this class. We also show that MINKOWSKI is very expressive, in the sense that all other non-generic problems in **PPP** considered here can be reduced to it. We conjecture that MINKOWSKI is **PPP**-complete.



## 1. Introduction

Total search problems [7,9] still constitute an exotic domain of complexity theory — exotic in both senses "strange, unusual" and "remote, unexplored." Take the class **PPP**, for example. It is known to include all of **PPAD**, and is defined in terms of the generic complete problem

PIGEON: Given a Boolean circuit $C$ with $n$ inputs and $n$ outputs, find $x \neq y \in \{0,1\}^n$ such that either $C(x) = 0^n$ or $C(x) = C(y)$.

The class **PPP** consists of all search problems reducible to PIGEON. As for other problems known to be in **PPP**, [9] only mentions

EQUAL SUMS: Given positive integers $a_1, \ldots, a_n$ such that $\sum_i a_i < 2^n - 1$, find two subsets $S \neq T \subseteq \{1, \ldots, n\}$ such that $\sum_{i \in S} a_i = \sum_{i \in T} a_i$.

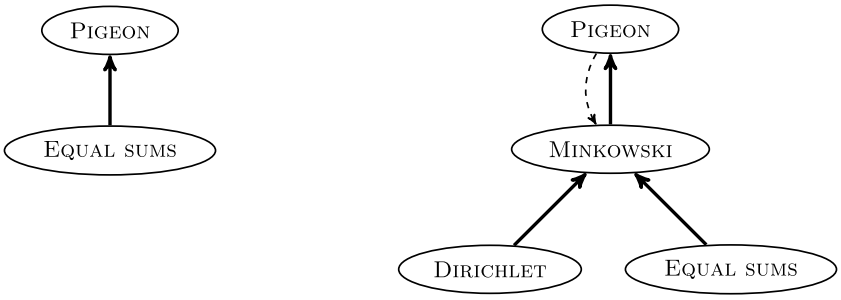
and these two problems, PIGEON and EQUAL SUMS, are to our knowledge[1] the only problems discussed in the literature that are known to be in **PPP** and not known to be in included classes, e.g. NASH.

Over the past decade, the class **PPAD** and its complete problems have substantially informed and advanced our algorithmic understanding of Game Theory [8]. Can **PPP** serve the same role for Cryptography? The generic problem PIGEON is loosely about collisions of hash functions; furthermore, it was recently pointed out that FACTORING belongs to **PWPP** (the class of problems whose totality is proved through a weak pigeonhole principle), a subclass of **PPP**, via randomized reductions [6]. The two new problems that we introduce here, MINKOWSKI and DIRICHLET, are also motivated by Cryptography, since they are both the complexity renderings of mathematical results which have been used in the foundations of lattice-based crypto systems [1]. There may be important dividends in understanding better this class and its complete problems.

* Corresponding author.
*E-mail addresses:* fban@berkeley.edu (F. Ban), kamaljain@gmail.com (K. Jain), christos@cs.columbia.edu (C.H. Papadimitriou), cpsomas@cs.cmu.edu (C.-A. Psomas), aviad@cs.stanford.edu (A. Rubinstein).

https://doi.org/10.1016/j.ipl.2018.12.009
0020-0190/

[1] See Section 1.2 for a discussion on results published after writing this note.




(a) The class **PPP** before this paper (b) The class **PPP** with our contributions

**Fig. 1.** Problems in **PPP**.

Incidentally, other classes of total problems from [9] have recently been connected to Cryptography: FACTORING was shown to also belong in the class **PPA**, again via randomized reductions, and also in the class **PWPP** [6]. Note that the only other problems known to be in **PPP** ∩ **PPA** are the problems in **PPAD**, such as NASH [3,4]. Determining whether FACTORING is in **PPAD** (via randomized reductions) is a most important open question. Recall also the recent insight that **PPAD** is intractable under standard cryptographic assumptions [2] (standard in the sense that they are adopted in parts of the mainstream literature).

### 1.1. Our contributions

As we mentioned, PIGEON and EQUAL SUMS are, as far as we know, the only problems in **PPP** discussed in the literature.[2] In other words, the picture of **PPP** prior to this note has been as shown in Fig. 1(a). In this note we introduce two new problems:

MINKOWSKI: Given an $n \times n$ matrix $A$ with $|\det(A)| < 1$, find a nontrivial integer combination of its rows with $l_\infty$ norm less than one.

DIRICHLET: Given $n$ rational numbers $a_1, \ldots, a_n$ and an integer $N$, find integers $q, p_1, \ldots, p_n$ such that $|a_i - \frac{p_i}{q}| < \frac{1}{qN}$ for all $i$, and $1 \leq q \leq N^n$.

We show that these two problems are in **PPP**. In fact, we show that MINKOWSKI is remarkably expressive, in the sense that the two currently known non-generic problems in **PPP** (EQUAL SUMS and DIRICHLET, leaving out problems, such as NASH, which belong to subclasses) are reducible to MINKOWSKI.

**Theorem 1.** MINKOWSKI *and* DIRICHLET *are reducible to* PIGEON.

**Theorem 2.** EQUAL SUMS *and* DIRICHLET *are reducible to* MINKOWSKI.

None of these results is trivial, but the reduction from MINKOWSKI to PIGEON is, surprisingly, the hardest to prove. In other words, the new picture of problems in **PPP** is as shown in Fig. 1(b). These results suggest that MINKOWSKI is a natural candidate for a non-generic complete problem for **PPP**. Thus, the following important open problem is the main message of this work:

**Conjecture 1.** PIGEON *is reducible to* MINKOWSKI.

### 1.2. Additional related work

When writing this note the authors were not aware of the following two, very relevant works (or the works didn't exist in the case of the second paper by Sotiraki et al. [10]).

**Hoberg et al. [5]**

In a recent work, Hoberg et al. [5] show (independently of this work) that the problem NUMBER BALANCING is equivalent to polynomial approximations of Minkowski's theorem. The result most relevant to this note is the following. Given an algorithm that takes as input a lattice $\Lambda \subseteq \mathbb{R}^n$ with $|\det(\Lambda)| < 1$ and finds a non-zero vector $\mathbf{x} \in \Lambda$ such that $\|\mathbf{x}\|_\infty \leq \rho$, there exists a $\delta$-approximation algorithm for the number balancing problem, where $\delta = 2^{-n^{\Theta(1/\rho)}}$. Their proof is similar to the reduction of EQUAL SUMS to MINKOWSKI presented in this note.

**Sotiraki et al. [10]**

In a recent unpublished manuscript, Sotiraki et al. [10] identify two new problems in **PPP**: a computational problem associated with Blichfeldt's fundamental theorem, and a generalized version of the Short Integer Solution problem from lattice based cryptography. They prove that both problems are in **PPP**, and furthermore that both problems are **PPP**-hard. This breakthrough therefore gives us the first natural **PPP**-complete problems. The authors provide a new proof that MINKOWSKI is in **PPP**, via a reduction to BLICHFELDT. Their reduction of BLICHFELDT to PIGEON is similar in spirit to the reduction of MINKOWSKI to PIGEON presented in this note.

## 2. Preliminaries

We are interested in *search problems*, that is, problems in which we are given the input to a problem in **NP**, and we are asked for a witness of this input (if none exists, we

[2] We refer the reader to Section 1.2 for a discussion on recent results.

return "no"). This class of problems is often called **FNP** (for function problems in **NP**).

Within **FNP**, we are interested in *total* search problems, that is, problems in which "no" is never a legitimate answer, for all inputs. The class of all total search problems is denoted **TFNP**. Evidently, every total search problem in **TFNP** must possess a mathematical proof of its totality. Consequently, total search problems can be further categorized in terms of the corresponding existence proofs. There are certain simple combinatorial facts, such as the pigeonhole principle and the parity argument, asserting the existence of an object with a certain property from among a population of objects, which are used often in such existence proofs. Syntactic subclasses of **TFNP**, such as **PPP, PPA, PLS, PPAD**, etc., are defined in terms of such facts. (Incidentally, it is remarkable and intriguing that we know of only five existence proofs that are not immediately constructive.) In particular, the class **PPP** is defined in terms of the pigeonhole principle as explained in the Introduction.

## Minkowski, Dirichlet and the Pigeonhole principle

We now state and prove the existence theorems of interest in this note.

**Theorem 3** *(Minkowski's convex body theorem). Let $S$ be a closed convex body in $\mathbb{R}^n$, symmetric with respect to the origin* 0 *and having volume* $\mathrm{vol}(S)$*. Then every lattice $L$ of determinant* $\det(L)$ *such that* $\mathrm{vol}(S) > 2^n|\det(L)|$ *has a point in $S$ distinct from the origin.*

**Proof.** Let $P$ denote the fundamental parallelepiped spanned by the basis of lattice $L$; we also denote $2P \triangleq \{2x : x \in P\}$. Define $f : S \to 2P$ as a "modulo-$2P$ operator", where for every $x \in S$, $f(x) = x + 2v \in 2P$ for some lattice vector $v \in L$. Since $\mathrm{vol}(S) > 2^n|\det(L)| = \mathrm{vol}(2P)$ and $f$ is locally area-preserving, then $f$ is not injective.

Thus, we can find two distinct points $p$ and $q$ in $S$ where $f(p) = f(q)$. By the definition of $f$, there exists a lattice point $w \neq 0$ such that $q = p + 2w$. Since $S$ is symmetric with respect to the origin, then $-p \in S$. Since $S$ is convex, then $w = (q - p)/2$ is in $S \cap L$. □

For our computational problem, MINKOWSKI, we take $S$ to be the hypercube $[-1, 1]^n$.

**Theorem 4** *(Dirichlet's approximation theorem). Given real numbers $\alpha_1, ..., \alpha_k$ and a natural number $N$, then there are integers $p_1, ..., p_k, q \in \mathbb{Z}$, $1 \leq q \leq N^k$, such that* $\left|\alpha_i - \frac{p_i}{q}\right| \leq \frac{1}{qN}$.

**Proof.** We define a map $\phi : \{0, 1, 2, ..., N^k\} \to [0, 1]^k$ taking $q$ to $(\{q\alpha_1\}, \{q\alpha_2\}, \ldots, \{q\alpha_k\})$, where $\{a\}$ denotes the fractional part of $a$. Now subdivide the hypercube $[0, 1]^k$ into hypercubelets of side length $1/N$. There are $N^k$ such cubelets and $N^k + 1$ elements in the domain of $\phi$ so by the pigeonhole principle, there exist $q' \neq q''$ such that $\phi(q')$ and $\phi(q'')$ are in the same cubelet.

Let $q = |q'' - q'|$. Then each component of $\phi(q)$ has a fractional part smaller than $1/N$. Thus, there exist integers $p_i$ such that $|q\alpha_i - p_i| \leq 1/N$ for all $i$. The result follows. □

We note that one can prove Dirichlet's approximation theorem using Minkowski's convex body theorem. We also comment that in the case of $k = 1$, the computational problem DIRICHLET can be solved in polynomial time. It is enough to compute the continued fraction expansion of $\alpha_1$ for sufficiently many terms.

## 3. Reducing to PIGEON

**Lemma 1.** MINKOWSKI *is reducible to* PIGEON.

**Proof.** Given an $n$ by $n$ matrix $A$ with determinant strictly smaller than 1, let $P$ be the fundamental parallelepiped of the vectors spanned by the rows of $A$. We will construct a circuit $C$ that computes a "modulo lattice" function similar to the function $f$ in the proof of Minkowski's theorem, mapping the hypercube $[0, 1)^n$ to $P$. A collision in this circuit (two inputs mapped to the same output) will give us a short lattice vector, i.e. an integer combination of rows of $A$ with $l_\infty$ smaller than 1.

## The challenge: how to succinctly represent a point in a parallelepiped

At first look, the reduction from MINKOWSKI to PIGEON seems like a trivial adaptation of the proof of Minkowski's Theorem. For the ultimate step which invokes the Pigeonhole Principle, we used $A$'s determinant to argue that the lattice (the range of the "modulo lattice" function) is smaller than the unit hypercube (the domain). In PIGEON, the circuit must have an equal number of input and output bits; this is how the range is guaranteed to be smaller than the domain. Hence, we must represent the output of the "modulo lattice" function with optimal succinctness. Even if we only waste one bit in every dimension, the reduction would only work if the parallelepiped is exponentially smaller than the cube. For example, the naive representation of a point in the parallelepiped as a linear combination of lattice vectors is far too wasteful. (By number of "necessary bits" we mean the precision we use for the input in the unit hypercube.) How, then, do you represent a point in a parallelepiped without wasting even a single bit?

Wlog, we can assume the entries of $A$ are all in $\delta\mathbb{Z}$ for a small $\delta$. Thus, the lattice points of $P$ all lie on a hypergrid that has unit length $\delta$. We want to be able to efficiently encode all the grid points in $P$. Note that if $P$ were a hyperrectangle rather than a general parallelipiped, it would be easy to express its grid points. This is because if the hyperrectangle was $\bigtimes_{i=1}^n [0, k_i)$, then the set of grid points would be exactly the set of all $(c_1\delta, c_2\delta, \ldots, c_n\delta)$ where the $c_i$ are integers and $0 \leq c_i < k_i/\delta$ for all $i$.

We first transform our matrix $A$ so that it is in upper triangular form with positive pivot elements, all entries are still integer multiples of $\delta$, and its volume is preserved (e.g. Hermite Normal Form).

Let $\hat{A}$ denote the resulting upper triangular matrix, and let $P$ denote the parallelepiped spanned by its rows. Given

a grid point in $P$, we embed it into the equal-volume hyperrectangle $H := \bigtimes_{i=1}^{n} \left[0, \hat{A}_{i,i}\right)$ by simply taking the $i$-th coordinate modulo $\hat{A}_{i,i}$:

$$f_i(x) = x_i \pmod{\hat{A}_{i,i}},$$

where $f_i$ denotes the $i$-th entry of our embedding $f : P \to H$.

We want to show that the embedding is one-to-one, i.e. if $fx = fy$, then $x = y$ where $x$ and $y$ are grid points in $P$. We can deduce by induction that the first $n-1$ coordinates of $x$ and $y$ are equal. Below, we argue that the respective $n$-th entries of $x$ and $y$ differ by less than $\hat{A}_{n,n}$. Since by definition they differ by a multiple of $\hat{A}_{n,n}$, they must indeed be equal.

We now argue that $x_n$ and $y_n$ differ by less than $\hat{A}_{n,n}$. Both $x$ and $y$ are linear combinations of the rows of $\hat{A}$ with coefficients in $[0, 1)$. In particular, their difference is a linear combination (with coefficients in $(-1, 1)$) of the rows of $\hat{A}$. Notice that: (i) by our inductive hypothesis, $x$ and $y$ are equal on the first $n-1$ coordinates; (ii) because of the upper-triangular form, the upper left submatrix $\hat{A}_{[n-1],[n-1]}$ has full rank; and (iii) the $n$-th row does not contribute to the first $n-1$ coordinates (again by the upper triangular structure). Therefore it follows that the difference has 0 coefficient on all but the last row. Hence we have that indeed $|x_n - y_n| < \hat{A}_{n,n}$.

**Finishing the proof of Lemma 1**

Since we have an embedding from $P$ to $H$, we can encode the grid points of $P$ in our output strings of length at most

$$\lceil \log_2 \prod_{i=1}^{n} \left(\hat{A}_{i,i}/\delta\right) \rceil = \lceil \log_2 |\det A| + n \log_2 (1/\delta) \rceil \leq \lceil n \log_2 (1/\delta) \rceil,$$

where we used the fact that $\prod_{i=1}^{n} \hat{A}_{i,i} = |\det A|$. Thus we can construct a PPP circuit with input and output length equal to $\lceil n \log_2 (1/\delta) \rceil$. Any input values greater than $1/\delta^n$ will be mapped to themselves to avoid spurious collisions. Therefore the circuit never outputs any values in $\left(|\det A|/\delta^n, 1/\delta^n\right]$. □

**Lemma 2.** DIRICHLET *is reducible to* PIGEON.

**Proof.** To develop some intuition, we first present the reduction for $n = 1$, i.e. our DIRICHLET input is a rational $a$ and an integer $N$, and we are looking for two integers $q$ and $p$, with $q \leq N$, such that $|a - \frac{p}{q}| < \frac{1}{qN}$.

Let $C_a(x)$ be the circuit that takes as input an ($\log N$ bit) integer $x$ from 1 to $N$ and outputs an ($\log N$ bit) integer $t$ such that $xa \mod 1 \in \left[t\frac{1}{N}, (t+1)\frac{1}{N}\right)$, where $t$ takes values from 0 to $N-1$. This circuit is a valid input for PIGEON. Let $x_1$ and $x_2$ be integers such that $C_a(x_1) = C_a(x_2)$, and let $x_2 > x_1$ without loss of generality (we deal with the case these don't exist later). Also, let $b_1$ and $b_2$ be numbers such that $x_1 a = \lfloor x_1 a \rfloor + b_1$ and $x_2 a = \lfloor x_2 a \rfloor + b_2$. Notice that

$$|(x_2 - x_1)a - (\lfloor x_2 a \rfloor - \lfloor x_1 a \rfloor)| = |b_2 - b_1| < \frac{1}{N},$$

thus $p = \lfloor x_2 a \rfloor - \lfloor x_1 a \rfloor$ and $q = x_2 - x_1$. If no such $x_1, x_2$ exist, and the solution to PIGEON is a number $x'$ such that $C_a(x') = 0$, then simply set $q = x'$ and $p = \lfloor x'a \rfloor$

For a general $n$, our circuit will map integers not to intervals of $[0, 1)$, but on hypercubelets of $[0, 1)^n$. Each hypercubelet will have sides of length $\frac{1}{N}$.

Let $C(x)$ be the circuit that takes as input an integer $x$ from 1 to $N^n$, and outputs the hypercubelet (under some valid encoding) where the vector $(xa_1 \mod 1, xa_2 \mod 1, \ldots, xa_n \mod 1)$ lies. There are $N^n$ possible outputs, i.e. $C$ is a valid input for PIGEON.

Let $x_1$, $x_2$, with $x_2 > x_1$, be a solution to PIGEON. Setting $q = x_2 - x_1$, and $p_i = \lfloor x_2 a_i \rfloor - \lfloor x_1 a_i \rfloor$ approximates all $a_i$ simultaneously. □

## 4. Reducing to MINKOWSKI

**Lemma 3.** EQUAL SUMS *reduces to* MINKOWSKI

**Proof.** We will construct an $(n+1) \times (n+1)$ lower diagonal matrix, with determinant strictly less than 1. The first column of the matrix is $(2^n - 1, a_1, a_2, \ldots a_n)$. The diagonal is $(2^n - 1, 1/2, 1/2, \ldots, 1/2)$, and all other entries are zero:

$$\begin{bmatrix} 2^n - 1 & 0 & 0 & \ldots & 0 \\ a_1 & \frac{1}{2} & 0 & \ldots & 0 \\ a_2 & 0 & \frac{1}{2} & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_n & 0 & 0 & \ldots & \frac{1}{2} \end{bmatrix}$$

The determinant of the matrix is $\frac{1}{2^n}\left(2^n - 1\right) < 1$. By Minkowski's theorem, there must be a non-trivial integer combination of rows so that each coordinate is less than 1 in magnitude. Moreover, this integer combination must satisfy the following:

- The coefficient of rows 2 to $n+1$, can only be $-1$, 0, or 1: a coefficient greater than 1 or less than $-1$ will get multiplied with the diagonal element, which can never be canceled (since it's the only element of the column).
- The elements of the first column are integers, thus the coefficient of rows 2 to $n+1$ must be chosen in such a way that the $a_i$'s sums up to zero.
- The coefficient of the first row is 0: the other rows with coefficient $-1$, 0, 1 can't cancel the $2^n - 1$.

The subsets with the same sum can be recovered as follows: one of the sets consists of the rows that were picked with coefficient 1, and the other subset consists of the ones with coefficient $-1$. This completes the reduction from EQUAL SUMS to MINKOWSKI. □

**Lemma 4.** DIRICHLET *reduces to* MINKOWSKI

**Proof.** The reduction from DIRICHLET to MINKOWSKI is very similar. Given $n$ numbers $a_1, \ldots, a_n$ and an integer $N$ we will construct an $(n+1) \times (n+1)$ upper diagonal matrix,

with determinant strictly less than 1 as follows: the first row of the matrix is $(N^n + 1/2)^{-1}, a_1 N, a_2 N, \ldots, a_n N$. For rows 2 through $n+1$, the $i$-th element is $N$, and everything else is zero:

$$\begin{bmatrix} (N^n + 1/2)^{-1} & a_1 N & a_2 N & \ldots & a_n N \\ 0 & N & 0 & \ldots & 0 \\ 0 & 0 & N & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & N \end{bmatrix}$$

The determinant of this matrix is $\frac{N^n}{N^n+1/2} < 1$, thus there exists an integer combination $q, p_1, \ldots, p_n$ of the rows with norm infinity smaller than 1. We will show that $q$ and $p_1, \ldots, p_n$ are exactly the integers that satisfy the conditions of Dirichlet's approximation theorem.

Since $(N^n+1)^{-1}$ is the only element of the first column, we get that $-\left(N^n + 1/2\right) < q < N^n + 1/2$. Since $q$ is an integer, then we get $-N^n \leq q \leq N^n$. Moreover, for every column $i$ from 2 through $n+1$, we get that $-1 < a_i q N + p_i N < 1$, which implies that $|a_i - \frac{p_i}{q}| < \frac{1}{qN}$. □

## Acknowledgements

We are grateful to anonymous reviewers for a major simplification of the proof of Lemma 1 and other helpful comments.

## References

[1] Miklós Ajtai, Generating hard instances of lattice problems, in: Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing, ACM, 1996, pp. 99–108.
[2] Nir Bitansky, Omer Paneth, Alon Rosen, On the cryptographic hardness of finding a nash equilibrium, Technical report, Cryptology ePrint Archive, Report 2014/1029, 2 014. http://eprint.iacr.org, 2014.
[3] Xi Chen, Xiaotie Deng, Settling the complexity of two-player nash equilibrium, in: FOCS, vol. 6, 2006, pp. 261–272.
[4] Constantinos Daskalakis, Paul W. Goldberg, Christos H. Papadimitriou, The complexity of computing a nash equilibrium, SIAM J. Comput. 39 (1) (2009) 195–259.
[5] Rebecca Hoberg, Harishchandra Ramadas, Thomas Rothvoss, Xin Yang, Number balancing is as hard as Minkowski's theorem and shortest vector, in: International Conference on Integer Programming and Combinatorial Optimization, Springer, 2017, pp. 254–266.
[6] Emil Jeřábek, Integer factoring and modular square roots, arXiv preprint, arXiv:1207.5220, 2012.
[7] Nimrod Megiddo, Christos H. Papadimitriou, On total functions, existence theorems and computational complexity, Theor. Comput. Sci. 81 (2) (1991) 317–324.
[8] Noam Nisan, Tim Roughgarden, Eva Tardos, Vijay V. Vazirani, Algorithmic Game Theory, Vol. 1, Cambridge University Press, Cambridge, 2007.
[9] Christos H. Papadimitriou, On the complexity of the parity argument and other inefficient proofs of existence, J. Comput. Syst. Sci. 48 (3) (1994) 498–532.
[10] Katerina Sotiraki, Manolis Zampetakis, Giorgos Zirdelis, PPP-completeness with connections to cryptography, arXiv preprint, arXiv:1808.06407, 2018.