Quantum algorithm for a generalized hidden shift problem*

Andrew M. Childs† Wim van Dam‡

Abstract

Consider the following generalized hidden shift problem: given a function f on $\{0,\ldots,M-1\}\times\mathbb{Z}_N$ promised to be injective for fixed b and satisfying $f(b,x)=f(b+1,x+s)$ for $b=0,1,\ldots,M-2$, find the unknown shift $s\in\mathbb{Z}_N$. For $M=N$, this problem is an instance of the abelian hidden subgroup problem, which can be solved efficiently on a quantum computer, whereas for $M=2$, it is equivalent to the dihedral hidden subgroup problem, for which no efficient algorithm is known. For any fixed positive ϵ, we give an efficient (i.e., $\mathrm{poly}(\log N)$) quantum algorithm for this problem provided $M\geq N^{\epsilon}$. The algorithm is based on the "pretty good measurement" and uses H. Lenstra's (classical) algorithm for integer programming as a subroutine.

1 Introduction.

Quantum mechanical computers can solve certain problems asymptotically faster than classical computers, but the extent of this advantage is not well understood. The most significant example of quantum computational speedup, Shor's algorithm for factoring and calculating discrete logarithms [29], is essentially based on an efficient quantum algorithm for the abelian hidden subgroup problem. This naturally leads to the question of whether the general *nonabelian* hidden subgroup problem can be solved efficiently on a quantum computer. Although efficient algorithms are known for a number of special cases of this problem [3, 10–12, 17, 18, 24], the two cases known to have significant applications, the dihedral group and the symmetric group, remain unsolved. In particular, an efficient quantum algorithm for the hidden subgroup problem (HSP) over the symmetric group would lead to an efficient quantum algorithm for graph isomorphism [4, 8]; and an efficient quantum algorithm for the dihedral HSP would lead to efficient quantum algorithms for certain lattice problems [27].

Although no polynomial-time algorithm is known for the dihedral HSP, Kuperberg discovered a subexponential-time quantum algorithm [21]. Kuperberg's algorithm uses a superpolynomial amount of time, space, and queries; Regev subsequently improved the space requirement to be only polynomial [28]. These algorithms are closely related to a connection between the dihedral HSP and an average case subset sum problem observed by Regev [27].

Recently, together with Bacon, we have developed an approach to the hidden subgroup problem based on the "pretty good measurement" (PGM) [2, 3]. In this approach, the PGM is used to distinguish the members of an ensemble of quantum states corresponding to the various possible hidden subgroups. For a variety of groups that can be written as the semidirect product of an abelian group and a cyclic group of prime order, we found that this measurement is closely related to a certain kind of average case algebraic problem. In particular, the measurement succeeds when the algebraic problem is likely to have a solution, and can be implemented if the solutions to that problem can be found. For the dihedral group, this problem is simply the average case subset sum problem [2]; more generally, we refer to it as the *matrix sum problem*. In some cases, the matrix sum problem can be solved, giving an efficient quantum algorithm for the corresponding hidden subgroup problem [3]. However, since the average case subset sum problem appears to be difficult, this approach has not yielded an improved algorithm for the dihedral HSP.

In this article, we show how the PGM approach provides an efficient quantum algorithm for a problem that interpolates between the abelian and dihedral hidden subgroup problems. The dihedral HSP is equivalent to the *hidden shift problem*, in which the goal is to determine a hidden shift $s\in\mathbb{Z}_N$ given two injective functions f_0, f_1 satisfying $f_0(x)=f_1(x+s)$. Instead of only two such functions, we consider M such functions, each one shifted from the previous by a fixed hidden shift s. If $M=N$, this problem is an instance of the abelian HSP on $\mathbb{Z}_N\times\mathbb{Z}_N$ with the hidden subgroup $\langle(1,s)\rangle$, which can be solved efficiently using abelian Fourier sampling. Even the case $M=N$ is classically intractable, and the

*AMC was supported in part by the National Science Foundation under Grant No. PHY-456720, and by the Army Research Office under Grant No. W911NF-05-1-0294. WvD was supported in part by the Disruptive Technology Office (DTO) under Army Research Office (ARO) contract number W911NF-04-R-0009.

†Institute for Quantum Information, California Institute of Technology, Pasadena, CA 91125, USA; amchilds@caltech.edu

‡Departments of Computer Science and Physics, University of California, Santa Barbara, Santa Barbara, CA 93106, USA; vandam@cs.ucsb.edu

problem only becomes more difficult for smaller M. In particular, for $M \ll N$, abelian Fourier sampling fails to determine the hidden shift. Using the PGM approach, we give, for any fixed integer $k \geq 3$, an efficient quantum algorithm that solves this problem for $M = \lfloor N^{1/k} \rfloor$. The algorithm works by implementing a joint measurement on k copies of certain quantum states that encode the hidden shift. Because for each $M \geq M'$ the generalized hidden shift problem on $\{0, \ldots, M-1\} \times \mathbb{Z}_N$ can be reduced to the generalized hidden shift problem on $\{0, \ldots, M'-1\} \times \mathbb{Z}_N$, for any fixed $\epsilon > 0$, this gives an efficient quantum algorithm for all $M \geq N^{\epsilon}$.

By applying the general PGM approach developed in [3], we find that the matrix sum problem corresponding to the generalized hidden shift problem is the following: given uniformly random $x \in \mathbb{Z}_N^k$ and $w \in \mathbb{Z}_N$, find $b \in \{0, \ldots, M-1\}^k$ such that $b \cdot x \bmod N = w$. We show how to express this problem as an instance of integer programming, so that it can be solved using Hendrik Lenstra's algorithm for that problem [23], which is efficient as long as the dimension k is constant. Thus our algorithm for the generalized hidden shift problem reiterates a theme of [3]: by combining abelian quantum Fourier transforms with nontrivial classical (or quantum) algorithms, one can find efficient quantum algorithms for HSP-like problems via the implementation of entangled quantum measurements. This result is encouraging since entangled measurements are known to be necessary for some hidden subgroup problems—in particular, for the symmetric group [13, 25].

The remainder of this article is organized as follows. In §2, we describe the generalized hidden shift problem in detail and explain how it can be viewed as a quantum state distinguishability problem. In §3, we review the pretty good measurement approach to such problems, prove that this approach solves the generalized hidden shift problem when the number of states is $k \geq \lceil 1/\epsilon \rceil$, and explain how it can be implemented by solving an appropriate matrix sum problem. In §4, we explain how the matrix sum problem can be solved efficiently (for constant k) using Lenstra's algorithm for integer programming, thereby giving an implementation of the PGM, and consequently, an algorithm for the hidden shift problem. Finally, we conclude in §5 with a discussion of the results and some open questions.

2 The Generalized Hidden Shift Problem.

It is well known that the dihedral HSP is equivalent to the *hidden shift problem*, which is defined as follows. Given two injective functions $f_0 : \mathbb{Z}_N \to S$ and $f_1 : \mathbb{Z}_N \to S$ (where S is some finite set) satisfying $f_0(x) = f_1(x+s)$ for some unknown $s \in \mathbb{Z}_N$, find s. For a proof of this equivalence, see Theorem 2 of [9] and Proposition 6.1 of [21]. For certain explicit functions of interest, such as the Legendre symbol, the hidden shift problem can be solved efficiently on a quantum computer [7]. However, for arbitrary black box functions, no efficient algorithm for the hidden shift problem is known.

A natural generalization of this problem, which we call the *generalized hidden shift problem*, is as follows. Consider a single function $f : \{0, \ldots, M-1\} \times \mathbb{Z}_N \to S$ satisfying two conditions: for fixed b, $f(b, x) : \mathbb{Z}_N \to S$ is injective; and $f(b, x) = f(b+1, x+s)$ for $b = 0, 1, \ldots, M-2$ for some fixed $s \in \mathbb{Z}_N$. Given such a function, our goal is again to find the hidden shift s. For $M = 2$, this problem is simply the usual hidden shift problem (with $f_b(x) = f(b, x)$ for $b = 0, 1$), and hence is equivalent to the dihedral HSP. For $M = N$, this problem is an instance of the abelian hidden subgroup problem (where the hidden subgroup is $\langle (1, s) \rangle \leq \mathbb{Z}_N \times \mathbb{Z}_N$). As a step toward understanding the dihedral HSP, we would like to investigate the difficulty of the problem for intermediate values of M. (Note that for intermediate values of M, the generalized hidden shift problem appears not to be an instance of the HSP for any group.)

On a quantum computer, this problem can be turned into a state distinguishability problem in the same manner as the standard approach to the hidden subgroup problem. Prepare a uniform superposition over all values of $b \in \{0, \ldots, M-1\}$ and $x \in \mathbb{Z}_N$ and then compute the value of $f(b, x)$, giving the state

$$\frac{1}{\sqrt{MN}} \sum_{b=0}^{M-1} \sum_{x \in \mathbb{Z}_N} |b, x, f(b, x)\rangle \,. \tag{2.1}$$

Then measure the third register, giving the state

$$|\phi_{x,s}\rangle := \frac{1}{\sqrt{M}} \sum_{b=0}^{M-1} |b, x + bs\rangle \tag{2.2}$$

for some unknown $x \in \mathbb{Z}_N$. Equivalently, the result is the mixed state described by the density matrix

$$\rho_s := \frac{1}{N} \sum_{x \in \mathbb{Z}_N} |\phi_{x,s}\rangle\langle\phi_{x,s}| \,. \tag{2.3}$$

Using a single copy of the state, we can identify s by the standard period finding algorithm (e.g., as in Shor's algorithm) only when M is very large (i.e., a reasonable fraction of N). Given the state $|\phi_{x,s}\rangle$, we can try to find s by applying the Fourier transform over $\mathbb{Z}_N \times \mathbb{Z}_N$ to the two registers, which yields the state

$$\frac{1}{N\sqrt{M}} \sum_{y,z \in \mathbb{Z}_N} \omega^{xz} \sum_{b=0}^{M-1} \omega^{b(y+sz)} |y, z\rangle \,, \tag{2.4}$$

where $\omega := \exp(2\pi \mathrm{i}/N)$. In the case $M = N$, this state equals

$$\frac{1}{\sqrt{N}} \sum_{z \in \mathbb{Z}_N} \omega^{xz} |-sz, z\rangle \,. \tag{2.5}$$

Measuring in the computational basis, we will observe $(-sz, z)$ for a uniformly random $z \in \mathbb{Z}_N$. If z is invertible modulo N, which happens with probability $\Omega(1/\log\log N)$, then we can deduce s immediately from the values $-sz$ and z. However, in general, for $M \leq N$, the outcome will only be of the form $(-sz, z)$ with probability M/N. If $M \leq N^\epsilon$ with $\epsilon < 1$, this probability is exponentially small in $\log N$, and the approach fails. A similar argument shows that an analogous approach using a Fourier transform over $\mathbb{Z}_M \times \mathbb{Z}_N$ followed by a computational basis measurement also fails for $M \ll N$. (Note that $\mathrm{poly}(\log N)$ such classical samples information theoretically determine the answer even for $M = 2$ [9], but it is not known how to process this data efficiently.)

Instead, we will use $k > 1$ states and apply the pretty good measurement. To see the connection to the matrix sum problem, it is helpful to write these states in a different basis. Fourier transforming the second register over $\mathbb{Z}_N$, we find

$$\tilde{\rho}_s^{\otimes k} = \frac{1}{(MN)^k} \sum_{\substack{x \in \mathbb{Z}_N^k \\ b,c \in \{0,\ldots,M-1\}^k}} \omega^{(b\cdot x - c\cdot x)s} |b, x\rangle\langle c, x| \tag{2.6}$$

$$= \frac{1}{(MN)^k} \sum_{\substack{x \in \mathbb{Z}_N^k \\ w,v \in \mathbb{Z}_N}} \omega^{(w-v)s} \sqrt{\eta_w^x \eta_v^x} |S_w^x, x\rangle\langle S_v^x, x| \tag{2.7}$$

where we have introduced the states

$$|S_w^x\rangle := \frac{1}{\sqrt{\eta_w^x}} \sum_{b \in S_w^x} |b\rangle \tag{2.8}$$

which are uniform superpositions over the solutions of the matrix sum problem,

$$S_w^x := \big\{ b \in \{0, \ldots, M-1\}^k : b \cdot x = w \bmod N \big\} \,. \tag{2.9}$$

Here the number of solutions is $\eta_w^x := |S_w^x|$. If there are no solutions (i.e., if $\eta_w^x = 0$), then we define $|S_w^x\rangle := 0$. Given the state $\tilde{\rho}_s^{\otimes k}$, we would like to determine the value of s.

3 Pretty Good Measurement Approach.

In this section, we review the PGM approach to distinguishing hidden subgroup states [2,3] as applied to the generalized hidden shift states (2.7).

The *pretty good measurement* (also known as the square root measurement or least squares measurement) is a measurement that often does well at distinguishing the members of an ensemble of quantum states [14–16] (and in fact is sometimes optimal in a certain sense). For an ensemble of states $\{\sigma_j\}$ with equal a priori probabilities, the pretty good measurement is the POVM with elements

$$E_j := \Sigma^{-1/2} \sigma_j \Sigma^{-1/2} \tag{3.10}$$

where

$$\Sigma := \sum_j \sigma_j \tag{3.11}$$

and where the inverse is taken over the support of the ensemble.

For the states (2.7), the PGM normalization matrix is

$$\Sigma = \frac{N}{(MN)^k} \sum_{x \in \mathbb{Z}_N^k} \sum_{w \in \mathbb{Z}_N} \eta_w^x |S_w^x, x\rangle\langle S_w^x, x| \,, \tag{3.12}$$

giving POVM elements

$$E_j = \frac{1}{N} \sum_{x \in \mathbb{Z}_N^k} \sum_{w,v \in \mathbb{Z}_N} \omega^{(w-v)j} |S_w^x, x\rangle\langle S_v^x, x| \,. \tag{3.13}$$

The probability of successfully identifying the hidden shift s is independent of s, and is given by

$$\Pr(\text{success}) := \operatorname{tr} E_s \tilde{\rho}_s^{\otimes k} \tag{3.14}$$

$$= \frac{1}{M^k N^{k+1}} \sum_{x \in \mathbb{Z}_N^k} \Bigg(\sum_{w \in \mathbb{Z}_N} \sqrt{\eta_w^x} \Bigg)^2 . \tag{3.15}$$

Using this expression, we can show that the success probability is appreciable when the matrix sum problem is likely to have a solution. Specifically, we have

LEMMA 3.1. (LEMMA 2 OF [3]) *If* $\Pr(\eta_w^x \geq \alpha) \geq \beta$ *for uniformly random* $x \in \mathbb{Z}_N^k$ *and* $w \in \mathbb{Z}_N$ *(i.e., if most instances of the matrix sum problem have many solutions), then* $\alpha\beta^2 N/M^k \leq \Pr(\text{success}) \leq M^k/N$.

Proof. For the upper bound, we have

$$\Pr(\text{success}) \leq \frac{1}{M^k N^{k+1}} \sum_{x \in \mathbb{Z}_N^k} \Bigg(\sum_{w \in \mathbb{Z}_N} \eta_w^x \Bigg)^2 \tag{3.16}$$

$$= \frac{M^k}{N} \tag{3.17}$$

since the η's are integers and $\sum_{w\in\mathbb{Z}_N} \eta_w^x = M^k$ for any x. For the lower bound, we have

$$\Pr(\text{success}) \geq \frac{N}{M^k}\left(\frac{1}{N^{k+1}}\sum_{x\in\mathbb{Z}_N^k}\sum_{w\in\mathbb{Z}_N}\sqrt{\eta_w^x}\right)^2 \tag{3.18}$$

by Cauchy's inequality applied to (3.15). Now

$$\frac{1}{N^{k+1}}\sum_{x\in\mathbb{Z}_N^k}\sum_{w\in\mathbb{Z}_N}\sqrt{\eta_w^x} \geq \sqrt{\alpha}\Pr(\eta_w^x \geq \alpha)\,, \tag{3.19}$$

so by the hypothesis, $\Pr(\text{success}) \geq \alpha\beta^2 N/M^k$ as claimed. This completes the proof.

For uniformly random $x \in \mathbb{Z}_N^k$ and $w \in \mathbb{Z}_N$, the expected number of matrix sum solutions is

$$\mu := \mathop{\mathbf{E}}_{x,w}[\eta_w^x] = \frac{M^k}{N}\,, \tag{3.20}$$

where we have again used the fact that $\sum_w \eta_w^x = M^k$ for any x. Thus, we expect the matrix sum problem to typically have no solutions for $k \ll \log N/\log M$, many solutions for $k \gg \log N/\log M$, and a constant number of solutions for $k \approx \log N/\log M$. For the specific case of the generalized hidden shift problem, this intuition can be formalized as follows:

LEMMA 3.2. *For the generalized hidden shift problem with $M = \lfloor N^{1/k} \rfloor$ with $k \geq 3$ and N sufficiently large, $\Pr(1 \leq \eta_w^x \leq 4)$ is lower bounded by a constant.*

A proof is given in the appendix.

Together, Lemmas 3.1 and 3.2 show that the PGM has at least a constant probability of successfully identifying the hidden shift. In fact, it turns out that the PGM is the POVM that maximizes the probability of successfully determining s given the states (2.7). For more details, we refer the reader to §4 of [2] and §4.4 of [3].

To give an efficient algorithm based on the PGM, we must show how the measurement can be implemented efficiently on a quantum computer. Such an implementation can be achieved using Neumark's theorem [26], which states that any POVM can be realized by a unitary transformation U on the system together with an ancilla followed by a measurement in the computational basis. For a POVM consisting of N rank one operators $E_j = |e_j\rangle\langle e_j|$ in a D-dimensional Hilbert space, U has the block form

$$U = \begin{pmatrix} V & X \\ Y & Z \end{pmatrix} \tag{3.21}$$

where the rows of the $N \times D$ matrix V are the D-vectors $\langle e_j|$, i.e., $V = \sum_{j=1}^N |j\rangle\langle e_j|$, and where X, Y, Z are arbitrary matrices subject to the constraint that U is unitary.

Recall from (3.13) that the POVM operators for the PGM on hidden subgroup states can be written

$$E_j = \sum_{x\in\mathbb{Z}_N^k} E_j^x \otimes |x\rangle\langle x| \tag{3.22}$$

where $E_j^x := |e_j^x\rangle\langle e_j^x|$ with

$$|e_j^x\rangle := \frac{1}{\sqrt{N}}\sum_{w\in\mathbb{Z}_N}\omega^{wj}|S_w^x\rangle\,. \tag{3.23}$$

In other words, each E_j is block diagonal, with blocks labeled by $x \in \mathbb{Z}_N^k$, and where each block is rank one. Thus, the measurement can be implemented in a straightforward way by first measuring the block label x and then performing the POVM $\{E_j^x\}_{j\in\mathbb{Z}_N}$ conditional on the first measurement result.

To implement the POVM $\{E_j^x\}_{j\in\mathbb{Z}_N}$ using Neumark's theorem, we would like to implement the unitary transformation U^x with the upper left submatrix

$$V^x = \frac{1}{\sqrt{N}}\sum_{j,w\in\mathbb{Z}_N}\omega^{-wj}\,|j\rangle\langle S_w^x|\,. \tag{3.24}$$

It is convenient to perform a Fourier transform (over $\mathbb{Z}_N$) on the left (i.e., on the index j), giving a unitary operator $\tilde{U}^x$ with upper left submatrix

$$\tilde{V}^x = \frac{1}{N}\sum_{j,w,v\in\mathbb{Z}_N}\omega^{(v-w)j}\,|v\rangle\langle S_v^x| \tag{3.25}$$

$$= \sum_{w\in\mathbb{Z}_N}|w\rangle\langle S_w^x|\,. \tag{3.26}$$

Therefore, the PGM can be implemented efficiently if we can efficiently perform a unitary transformation satisfying

$$|w, x\rangle \mapsto |S_w^x, x\rangle \tag{3.27}$$

for all matrix sum problem instances (x, w) with $\eta_w^x > 0$. Since the state $|S_w^x\rangle$ is a uniform superposition of the solutions of the matrix sum problem instance (x, w), we refer to (3.27) as *quantum sampling* of solutions to the matrix sum problem. If we can efficiently quantum sample from matrix sum solutions, then by running the circuit in reverse, we can efficiently implement $\tilde{U}^x$, and hence the desired measurement.

By applying these unitary transformations directly to the state (2.7), we can obtain a description of the algorithm without reference to generalized measurement.

Performing the inverse of the quantum sampling transformation (3.27) followed by a Fourier transform, we obtain the state

$$\rho' := \frac{1}{N^k} \sum_{x \in \mathbb{Z}_N^k} |\rho'_x, x\rangle\langle\rho'_x, x| \tag{3.28}$$

where

$$|\rho'_x\rangle := \frac{1}{\sqrt{M^k N}} \sum_{j,w \in \mathbb{Z}_N} \omega^{w(s-j)} \sqrt{\eta_w^x} |j\rangle \,. \tag{3.29}$$

Roughly speaking, if the distribution of η_w^x is close to uniform, then the sum over w in (3.29) is close to zero unless $j = s$, so that a measurement of the first register is likely to yield the hidden shift s.

In general, it may be difficult to implement (3.27) exactly. Instead, we may only be able to perform an approximate quantum sampling transformation satisfying

$$|w, x\rangle \mapsto \begin{cases} |S_w^x, x\rangle & (x,w) \in Z_{\text{good}} \\ |\mu_w^x, x\rangle & (x,w) \in Z_{\text{bad}} \end{cases} \tag{3.30}$$

for some states $|\mu_w^x\rangle$, where $Z_{\text{good}}, Z_{\text{bad}}$ form a partition of the matrix sum instances (x, w) for which $\eta_w^x > 0$. The good matrix sum problem instances $(x, w) \in Z_{\text{good}}$ are those for which the quantum sampling can be done correctly. Assuming the bad matrix sum instances $(x, w) \in Z_{\text{bad}}$ can be efficiently recognized, we can include a label in the states $|\mu_w^x\rangle$ to ensure that $\langle S_w^x | \mu_{w'}^x \rangle = 0$ for all $x \in \mathbb{Z}_n^k$, $w, w' \in \mathbb{Z}_N$ with $(x, w') \in Z_{\text{bad}}$, even if $(x, w) \in Z_{\text{bad}}$ (and indeed, even if $w = w'$). Applying the approximate quantum sampling transformation followed by the Fourier transform gives the state

$$\rho'_{\text{apx}} = \frac{1}{N^k} \sum_{x \in \mathbb{Z}_N^k} |\rho'_{x,\text{apx}}, x\rangle\langle\rho'_{x,\text{apx}}, x| \tag{3.31}$$

where

$$|\rho'_{x,\text{apx}}\rangle := \frac{1}{\sqrt{M^k N}} \sum_{j \in \mathbb{Z}_N} \left(\sum_{(x,w) \in Z_{\text{good}}} \omega^{w(s-j)} \sqrt{\eta_w^x} |j\rangle + \sum_{(x,w) \in Z_{\text{bad}}} \omega^{w(s-j)} \sqrt{\eta_w^x} |\nu_j^x\rangle \right) \tag{3.32}$$

for some states $|\nu_j^x\rangle$ with $\langle j | \nu_{j'}^x \rangle = 0$ for all $x \in \mathbb{Z}_N^k$, $j, j' \in \mathbb{Z}_N$ (because of the promise on $\langle S_w^x | \mu_{w'}^x \rangle$ and the fact that (3.30) is unitary). The fidelity between the ideal final state ρ' and the actual final state ρ'_{apx} resulting from approximate quantum sampling is thus

$$\frac{1}{(MN)^k} \sum_{(x,w) \in Z_{\text{good}}} \eta_w^x \,. \tag{3.33}$$

Now $\eta_w^x > 1$ for all $(x, w) \in Z_{\text{good}}$, so if $|Z_{\text{good}}|$ is sufficiently large, the actual final state is close to the ideal final state, and hence a measurement of the first register yields the hidden shift s with reasonable probability. As we will show in the next section, the instances with $1 \le \eta_w^x \le 4$ can be quantum sampled efficiently (i.e., these instances are good). Then, letting $M = \lfloor N^{1/k} \rfloor$, Lemma 2 shows that $|Z_{\text{good}}|/N^{k+1}$ is lower bounded by a constant, and thus the fidelity between ρ' and ρ'_{apx} is lower bounded by a constant, so that the probability of successfully determining s is lower bounded by a constant.

4 Solution of the Matrix Sum Problem.

Recall that the matrix sum problem for the generalized hidden shift problem is the following: given $x \in \mathbb{Z}_N^k$ and $w \in \mathbb{Z}_N$ chosen uniformly at random, find $b \in \{0, \ldots, M-1\}^k$ such that $b \cdot x = w \bmod N$. This is a linear equation over $\mathbb{Z}_N$ in k variables, where the solutions are required to come from $\{0, \ldots, M-1\}$. Such solutions can be found using integer programming, which has an efficient algorithm if M is sufficiently large. We assume $M = \lfloor N^{1/c} \rfloor$ for some positive integer $c \ge 3$. Since we can always decrease M by only considering a subset of the inputs to the first argument of the hiding function f, this will not constitute a loss of generality. According to Lemma 3.2, we take $k = c$ so that there are between 1 and 4 solutions with probability at least some constant.

To see the connection to integer programming, we note that the solutions form an integer lattice. We begin by considering the equation $b \cdot x = w \bmod N$ as a $(k+1)$-variable linear equation over all the integers $\mathbb{Z}$. Define an extension of x by $\bar{x} := (x_1, \ldots, x_k, N)$ and consider the solutions $\bar{b} \in \mathbb{Z}^{k+1}$ of the equation $\bar{b} \cdot \bar{x} = w$. For any $b \in \mathbb{Z}^k$ that solves the equation $b \cdot x = w \bmod N$, there is a unique $\lambda \in \mathbb{Z}$ such that $\bar{b} = (b, \lambda)$ is a solution to $\bar{b} \cdot \bar{x} = w$; and conversely, for any $\bar{b} \in \mathbb{Z}^{k+1}$ that solves $\bar{b} \cdot \bar{x} = w$, there is a unique $b \in \mathbb{Z}^k$ (namely, the first k components of $\bar{b}$) that solves $b \cdot x = w \bmod N$. Hence there is a bijection between the solutions $\bar{b} \in \mathbb{Z}^{k+1}$ to the equation $\bar{b} \cdot \bar{x} = w$ and the solutions $b \in \mathbb{Z}^k$ to the equation $b \cdot x = w \bmod N$.

By Lemma A.1 in the appendix, we see that the linear Diophantine equation $\bar{b} \cdot \bar{x} = w$ will have no solutions if $\gcd(x_1, \ldots, x_k, N)$ does not divide w. If $\bar{b} \cdot \bar{x} = w$ does have a solution, then the solutions $\bar{b}$ comprise a shifted k-dimensional lattice $\bar{b}^{(0)} + L$ with some particular solution $\bar{b}^{(0)}$ satisfying $\bar{b}^{(0)} \cdot \bar{x} = w$ and the elements of $L \subset \mathbb{Z}^{k+1}$ the solutions of the equation $\bar{b} \cdot \bar{x} = 0$. By omitting the last coordinate of these solutions, we obtain all solutions $b \in \mathbb{Z}^k$, which comprise

a shifted k-dimensional lattice in $\mathbb{Z}^k$:

$$b = b^{(0)} + \sum_{j=1}^{k} \beta_j b^{(j)} \tag{4.34}$$

for all $\beta_1, \ldots, \beta_k \in \mathbb{Z}$. Due to the aforementioned bijection, each solution $b \in \mathbb{Z}^k$ has a unique set of coordinates $\beta_1, \ldots, \beta_k \in \mathbb{Z}$. The vectors $b^{(0)}, \ldots, b^{(k)} \in \mathbb{Z}^k$ can be found efficiently by applying the extended Euclidean algorithm to the equation $\bar{b} \cdot \bar{x} = w$ (see for example Algorithm 1.3.6 in [6]).

To solve the matrix sum problem, we would like to find the solutions b that lie in $\{0, \ldots, M-1\}^k$, which is the set of integer points in the convex region described by the inequalities

$$0 \leq b_i \leq M - 1\,, \quad i = 1, \ldots, k\,. \tag{4.35}$$

The problem of finding such points (or more precisely, deciding whether such a point exists) is simply an instance of integer programming in k dimensions, which can be solved efficiently if k is a constant. In general, the integer programming problem is the following. Given a rational matrix $A \in \mathbb{Q}^{m \times k}$ and a rational vector $\gamma \in \mathbb{Q}^m$, does there exist an integral vector $\beta \in \mathbb{Z}^k$ such that $A\beta \leq \gamma$? Although this general problem is NP-complete [5,20], if the dimension k is held constant, then the problem can be solved in time polynomial in the input size [23] using an algorithm based on lattice basis reduction [22].

By rewriting the convex constraints (4.35) in terms of the lattice of solutions (4.34), we see that solutions of the matrix sum problem correspond precisely to vectors $\beta \in \mathbb{Z}^k$ satisfying the constraints

$$\sum_{j=1}^{k} \beta_j b_i^{(j)} \leq (M-1) - b_i^{(0)}\,, \quad i = 1, \ldots, k \tag{4.36}$$

$$-\sum_{j=1}^{k} \beta_j b_i^{(j)} \leq b_i^{(0)}\,, \qquad i = 1, \ldots, k\,. \tag{4.37}$$

But this is precisely an instance of integer programming in k dimensions with $m = 2k$ constraints, with

$$A_{ij} = \begin{cases} b_i^{(j)} & i = 1, \ldots, k \\ -b_{i-k}^{(j)} & i = k+1, \ldots, 2k \end{cases} \tag{4.38}$$

$$\gamma_i = \begin{cases} (M-1) - b_i^{(0)} & i = 1, \ldots, k \\ b_{i-k}^{(0)} & i = k+1, \ldots, 2k\,. \end{cases} \tag{4.39}$$

Therefore, it can be solved efficiently whenever k is a constant. Note that integer programming as described above is a decision problem, whereas we would like to find the actual solutions. However, this is easily accomplished using bisection, recursively dividing the set $\{0, \ldots, M-1\}^k$ into halves, to find all of the solutions efficiently (for the cases in which there are few solutions—in particular, for those for which there are between 1 and 4 solutions).

Overall, we see that we can efficiently solve the matrix sum problem (in a regime where the pretty good measurement solves the generalized hidden shift problem with constant probability) whenever $M \geq N^\epsilon$ for some fixed $\epsilon > 0$. For the cases in which the number of solutions is small, they can be explicitly enumerated, and hence we can efficiently perform the approximate quantum sampling transformation (3.30) (see for example footnote 2 of [3]). Therefore, we find the following result.

THEOREM 4.1. *The generalized hidden shift problem with $M \geq N^\epsilon$ for any fixed $\epsilon > 0$ can be solved in time* $\mathrm{poly}(\log N)$ *on a quantum computer.*

Proof. Given ϵ, we will use $k = \max\{\lceil 1/\epsilon \rceil, 3\}$ copies of the unknown quantum state (2.3). Because the generalized hidden shift problem on the full domain $\{0, \ldots, M-1\} \times \mathbb{Z}_N$ can be solved by solving the same problem on a reduced domain $\{0, \ldots, M'-1\} \times \mathbb{Z}_N$ with $M' \leq M$, it is sufficient to prove the theorem for a specific $M \leq N^\epsilon$. We will do this for $M = \lfloor N^{1/k} \rfloor$.

The algorithm is as follows:

1. Create k copies of the hidden shift state (2.3):

$$\left(\frac{1}{N} \sum_{x \in \mathbb{Z}_N} |\phi_{x,s}\rangle\langle\phi_{x,s}| \right)^{\otimes k}\,. \tag{4.40}$$

2. Perform the Fourier transform over $\mathbb{Z}_N$ on the second register of each copy. After reordering the registers this gives

$$\begin{aligned} &\frac{1}{(MN)^k} \sum_{\substack{x \in \mathbb{Z}_N^k \\ b,c \in \{0,\ldots,M-1\}^k}} \omega^{(b \cdot x - c \cdot x)s} |b, x\rangle\langle c, x| \\ &\quad = \frac{1}{(MN)^k} \sum_{\substack{x \in \mathbb{Z}_N^k \\ w,v \in \mathbb{Z}_N}} \omega^{(w-v)s} \sqrt{\eta_w^x \eta_v^x} |S_w^x, x\rangle\langle S_v^x, x|\,. \end{aligned} \tag{4.41}$$

3. Using Lenstra's algorithm to solve the integer program defined by (4.38–4.39), perform the inverse of the approximate quantum sampling transformation (3.30). This means that we apply the transformation $|S_w^x, x\rangle \mapsto |w, x\rangle$ for a significant fraction of the values $w \in \mathbb{Z}_N$. Combined with the statistics

of the η_w^x values (Lemma 3.2), this gives an approximation of the state

$$\frac{1}{(MN)^k} \sum_{\substack{x \in \mathbb{Z}_N^k \\ w,v \in \mathbb{Z}_N}} \omega^{(w-v)s} |w,x\rangle\langle v,x| \,. \tag{4.42}$$

Note that the first register now takes values in $\mathbb{Z}_N$.

4. Perform the inverse Fourier transform over $\mathbb{Z}_N$ on the first register, which will lead to an approximation of the state $|s\rangle\langle s|$ for that register.
5. Measure the first register and return the outcome.

As described in the second half of §3, this procedure gives an approximate implementation of the PGM, and succeeds with constant probability provided the PGM has constant success probability. By Lemma 3.2, there is a constant probability of having between 1 and 4 solutions to the random matrix sum equation $b \cdot x = w \bmod N$, and hence the success probability of the PGM is a constant (by Lemma 3.1). This completes the proof.

In fact, the algorithm remains efficient even if ϵ decreases (very) slowly with N. Lenstra's algorithm for integer programming in dimension k takes time $2^{O(k^3)}$ [23], so the generalized hidden shift problem can be solved efficiently for $M = N^{O(1/(\log\log N)^{1/3})}$. Indeed, a subsequent improvement by Kannan solves k-dimensional integer programming in time $2^{O(k \log k)}$ [19], which can be used to decrease M slightly further.

5 Discussion.

We have applied the PGM approach to the generalized hidden shift problem, which interpolates from the dihedral HSP to the abelian HSP as M varies from 2 to N. We found an efficient quantum algorithm for this problem for any $M \geq N^\epsilon$ with ϵ fixed (or decreasing very slowly with N). The algorithm works by solving the matrix sum problem using Lenstra's algorithm for integer programming in constant dimensions, thereby illustrating (as in [3]) that nontrivial classical algorithms can be useful for implementing entangled measurements to distinguish states obtained by weak Fourier sampling.

Our original motivation for studying this problem was the observation that a solution to the generalized hidden shift problem for sufficiently small M could lead to new algorithms for the unique shortest vector in a lattice problem, just as Regev showed for the case $M = 2$ [27]. Unfortunately, $M \geq N^\epsilon$ does not appear to be sufficiently small to yield interesting lattice algorithms. Nevertheless, attempting to solve the generalized hidden shift problem for yet smaller M may be a promising path toward improved quantum algorithms for lattice problems. Indeed, for the case $M = 2$, Kuperberg's subexponential-time algorithm outperforms the algorithm given in this paper, so it seems likely that an improved algorithm could be found for values of M intermediate between 2 and N^ϵ. If the strategy for such an algorithm is to implement the optimal measurement to distinguish the hidden shift states, then this would seem to require an improved algorithm for (average-case) integer programming. However, it is also possible that an entirely different strategy could extend the accessible values of M.

It should be pointed out that the problem of distinguishing the states (2.2) has the following variant, which also has an efficient solution. For given N and M we define an s-periodic state with an unknown offset $x \in \mathbb{Z}_N$, namely

$$|\varphi_{x,s}\rangle := \frac{1}{\sqrt{M}} \sum_{b=0}^{M-1} |x + bs \bmod N\rangle \tag{5.43}$$

where we assume that the parameters are such that the state is properly normalized. Just as with the state $|\phi_{x,s}\rangle$ of (2.2), the task is to determine the hidden shift s from a polynomial number of copies of $|\varphi_{x,s}\rangle$, where each state has its own random offset $x \in \mathbb{Z}_N$. This version of the problem, where the states do not contain a register for the label b, has a much simpler solution than the problem of distinguishing the labeled states arising from the generalized hidden shift problem. If we apply the quantum Fourier transform over $\mathbb{Z}_N$ to $\varrho_s := \sum_{x \in \mathbb{Z}_N} |\varphi_{x,s}\rangle\langle\varphi_{x,s}|/N$, we obtain a mixed state $\tilde{\varrho}_s$ that is diagonal in the computational basis and hence is effectively a classical probability distribution. This distribution has period N/s, and the width of the distribution around its maximum values at $0, N/s, 2N/s, \ldots$ depends on the value $\varepsilon := \log M / \log N$. For fixed ε, one can show that there is an efficient algorithm using $k = \lceil 1/\varepsilon \rceil$ samples of ϱ_s that determines the shift s reliably in time poly$(\log N)$. This algorithm is again based on Lenstra's algorithm for integer programming.

Another problem suggested by this work is the following generalization of graph isomorphism. Suppose that we are given a list of n-vertex graphs $G_0, \ldots, G_{M-1}$, and are promised that either no two graphs are isomorphic, or $G_b = \pi(G_{b+1})$ for some fixed permutation $\pi \in S_n$ for $b = 0, 1, \ldots, M-2$. It would be interesting to show that this problem can be solved efficiently even for very large M (where the graphs can be specified by a black box in the case where M is superpolynomial in n).

Acknowledgments

We thank Oded Regev for discussions about the relationship between lattice problems and the generalized

hidden shift problem, and for pointing out that one can classically find the period of a noisy, periodic probability distribution using Lenstra's algorithm.

References

[1] N. Alon and J. H. Spencer, *The Probabilistic Method*, 2nd ed., Wiley Interscience, New York, 2000.

[2] D. Bacon, A. M. Childs, and W. van Dam, *Optimal measurements for the dihedral hidden subgroup problem*, to appear in Chicago Journal of Theoretical Computer Science. arXiv:quant-ph/0501044.

[3] ———, *From optimal measurement to efficient quantum algorithms for the hidden subgroup problem over semidirect product groups*, Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science, 2005, pp. 469–478. arXiv:quant-ph/0504083.

[4] R. Boneh and R. Lipton, *Quantum cryptanalysis of hidden linear functions*, Advances in Cryptology – Crypto'95, 1995, pp. 424–437.

[5] I. Borosh and L. B. Treybig, *Bounds on positive integral solutions of linear diophantine equations*, Proceedings of the American Mathematical Society **55** (1976), 299–304.

[6] H. Cohen, *A Course in Computational Algebraic Number Theory*, Springer-Verlag, Berlin, 1993.

[7] W. van Dam, S. Hallgren, and L. Ip, *Quantum algorithms for some hidden shift problems*, Proceedings of the ACM-SIAM Symposium on Discrete Algorithms, 2003, pp. 489–498. arXiv:quant-ph/0211140.

[8] M. Ettinger and P. Høyer, *A quantum observable for the graph isomorphism problem*. arXiv:quant-ph/9901029.

[9] ———, *On quantum algorithms for noncommutative hidden subgroups*, Advances in Applied Mathematics **25** (2000), no. 3, 239–251. arXiv:quant-ph/9807029.

[10] K. Friedl, G. Ivanyos, F. Magniez, M. Santha, and P. Sen, *Hidden translation and orbit coset in quantum computing*, Proceedings of the 35th Annual ACM Symposium on Theory of Computing, 2003, pp. 1–9. arXiv:quant-ph/0211091.

[11] D. Gavinsky, *Quantum solution to the hidden subgroup problem for poly-near-Hamiltonian groups*, Quantum Information and Computation **4** (2004), no. 3, 229–235.

[12] M. Grigni, L. Schulman, M. Vazirani, and U. Vazirani, *Quantum mechanical algorithms for the nonabelian hidden subgroup problem*, Combinatorica **24** (2004), no. 1, 137–154.

[13] S. Hallgren, C. Moore, M. Rötteler, A. Russell, and P. Sen, *Limitations of quantum coset states for graph isomorphism*, Proceedings of the 38th Annual ACM Symposium on Theory of Computing, 2006, pp. 604–617. arXiv:quant-ph/0511148, arXiv:quant-ph/0511149.

[14] P. Hausladen and W. K. Wootters, *A 'pretty good' measurement for distinguishing quantum states*, Journal of Modern Optics **41** (1994), no. 12, 2385–2390.

[15] C. W. Helstrom, *Quantum Detection and Estimation Theory*, Academic Press, New York, 1976.

[16] A. S. Holevo (Kholevo), *On asymptotically optimal hypothesis testing in quantum statistics*, Theory of Probability and its Applications **23** (1979), no. 2, 411–415. English translation of Teoriya Veroyatnosteĭ i ee Primeneniya **23** (1978), no. 2, 429–432.

[17] Y. Inui and F. Le Gall, *An efficient algorithm for the hidden subgroup problem over a class of semi-direct product groups*. arXiv:quant-ph/0412033.

[18] G. Ivanyos, F. Magniez, and M. Santha, *Efficient quantum algorithms for some instances of the non-abelian hidden subgroup problem*, International Journal of Foundations of Computer Science **14** (2003), no. 5, 723–739. arXiv:quant-ph/0102014.

[19] R. Kannan, *Minkowski's convex body theorem and integer programming*, Mathematics of Operations Research **12** (1987), no. 3, 415–440.

[20] R. M. Karp, *Reducibility among computational problems*, Complexity of Computer Computations, 1972, pp. 85–103.

[21] G. Kuperberg, *A subexponential-time quantum algorithm for the dihedral hidden subgroup problem*, SIAM Journal on Computing **35** (2005), no. 1, 170–188. arXiv:quant-ph/0302112.

[22] A. K. Lenstra, H. W. Lenstra, Jr., and L. Lovász, *Factoring polynomials with rational coefficients*, Mathematische Annalen **261** (1982), no. 4, 515–534.

[23] H. W. Lenstra, Jr., *Integer programming with a fixed number of variables*, Mathematics of Operations Research **8** (1983), no. 4, 538–548.

[24] C. Moore, D. N. Rockmore, A. Russell, and L. J. Schulman, *The hidden subgroup problem in affine groups: Basis selection in Fourier sampling*, Proceedings of the 15th Annual ACM-SIAM Symposium on Discrete Algorithms, 2004, pp. 1113–1122. arXiv:quant-ph/0211124, extended version available at arXiv:quant-ph/0503095.

[25] C. Moore, A. Russell, and L. J. Schulman, *The symmetric group defies strong Fourier sampling*, Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science, 2005, pp. 479–490. arXiv:quant-ph/0501056.

[26] M. A. Neumark, *On a representation of additive operator set functions*, Comptes Rendus de l'Académie des Sciences de l'URSS (Doklady Akademii Nauk SSSR) **41** (1943), 359–361.

[27] O. Regev, *Quantum computation and lattice problems*, Proceedings of the 43rd Annual Symposium on Foundations of Computer Science, 2002, pp. 520–529. arXiv:cs.DS/0304005.

[28] ———, *A subexponential time algorithm for the dihedral hidden subgroup problem with polynomial space*. arXiv:quant-ph/0406151.

[29] P. W. Shor, *Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer*, SIAM Journal on Computing **26** (1997), no. 5, 1484–1509.

A Number of Solutions of the Matrix Sum Problem.

In this appendix, we prove Lemma 3.2. Before giving the proof, we need the following fact:

LEMMA A.1. *For any fixed b, the number of solutions $x \in \mathbb{Z}_N^k$ to the equation $b \cdot x = 0 \bmod N$ is $N^{k-1} \gcd(b_1, \ldots, b_k, N)$.*

Proof. First, consider the case where $N = p^r$ is a prime power. Then $\gcd(b_1, \ldots, b_k, p^r) = p^s$ for some $s \in \{0, 1, \ldots, r\}$. In particular, there must be an index i such that $\gcd(b_i, p^r) = p^s$, and hence $b_i = cp^s$ for some $c \in \{1, \ldots, p-1\}$. Without loss of generality, assume $i = 1$. Now we can rewrite the equation $b \cdot x = 0$ as $cp^s x_1 + \sum_{j=2}^k b_j x_j = 0 \bmod p^r$, or equivalently, since p^s is a common divisor of all b_j, $cx_1 = -\sum_{j=2}^k b'_j x_j \bmod p^{r-s}$ where $b'_j = b_j/p^s$. Because $c \in \mathbb{Z}_{p^r}^\times$, for any fixed $(x_2, \ldots, x_k) \in \mathbb{Z}_{p^r}^{k-1}$, there are p^s solutions $x_1 = (\sum_{j=2}^k b'_j x_j)/c + \lambda p^{r-s} \bmod p^r$ (one solution for each $\lambda \in \{0, \ldots, p^s - 1\}$). Hence the total number of solutions $(x_1, \ldots, x_k)$ is $N^{k-1}p^s$, proving the lemma for the case $N = p^r$.

Now if N is not a prime power, let $N = p_1^{r_1} \cdots p_t^{r_t}$ be the factorization of N into powers of distinct primes, and let $\tau : \mathbb{Z}_N \to \mathbb{Z}_{p_1^{r_1}} \times \cdots \times \mathbb{Z}_{p_t^{r_t}}$ be the ring isomorphism provided by the Chinese remainder theorem: for $x \in \mathbb{Z}_N$, $\tau(x) = (x \bmod p_1^{r_1}, \ldots, x \bmod p_t^{r_t})$. Since τ is a ring isomorphism, $b \cdot x = 0 \bmod N$ if and only if $b \cdot \tau(x)_i = 0 \bmod p_i^{r_i}$ for all $i = 1, \ldots, t$. By the special case of the lemma for N a prime power, the number of solutions to the ith equation is $p_i^{r_i(k-1)} \gcd(b_1, \ldots, b_k, p_i^{r_i})$; hence the total number of solutions is $\prod_{i=1}^t p_i^{r_i(k-1)} \gcd(b_1, \ldots, b_k, p_i^{r_i}) = N^{k-1} \gcd(b_1, \ldots, b_k, N)$ as claimed.

Now we are ready to give the proof of Lemma 3.2:

LEMMA 3.2. *For the generalized hidden shift problem with $M = \lfloor N^{1/k} \rfloor$ with $k \geq 3$ and N sufficiently large, $\Pr(1 \leq \eta_w^x \leq 4)$ is lower bounded by a constant.*

Proof. The main idea of the proof is the same as in the proof of Lemma 5 of [3]: we show that the variance of η_w^x is small, so that the number of solutions of the matrix sum problem is typically close to its mean. Because we have $M = \lfloor N^{1/k} \rfloor$, the mean value of η_w^x is $\mu := \mathbf{E}_{x,w}[\eta_w^x] = M^k/N = 1 + O(1/N)$ as N grows.

The variance of the number of solutions $b \in \{0, \ldots, M-1\}^k$ of the equation $b \cdot x = w \bmod N$ for uniformly random $x \in \mathbb{Z}_N^k, w \in \mathbb{Z}_N$ is $\sigma^2 := \mathbf{E}_{x,w}[(\eta_w^x)^2] - \mu^2$, and

$$\begin{aligned}
&\mathop{\mathbf{E}}_{x \in \mathbb{Z}_N^k, w \in \mathbb{Z}_N} [(\eta_w^x)^2] && \text{(A.44)}\\
&\quad = \frac{1}{N^{k+1}} \sum_{x \in \mathbb{Z}_N^k, w \in \mathbb{Z}_N} (\eta_w^x)^2 \\
&\quad = \frac{1}{N^{k+1}} \sum_{x,w} \Big(\sum_b \delta_{b \cdot x, w} \Big) \Big(\sum_c \delta_{c \cdot x, w} \Big) && \text{(A.45)}\\
&\quad = \frac{1}{N^{k+1}} \sum_{x,w} \Big(\sum_b \delta_{b \cdot x, w} + \sum_{b \neq c} \delta_{b \cdot x, w}\, \delta_{b \cdot x, c \cdot x} \Big) && \text{(A.46)}
\end{aligned}$$

(with the b, c summations over $\{0, \ldots, M-1\}^k$). The first (diagonal) term is just the mean. To handle the second (off-diagonal) term, we can write

$$\begin{aligned}
&\mathop{\mathbf{E}}_{x,w} [(\eta_w^x)^2] && \text{(A.47)}\\
&\quad = \mu + \frac{1}{N^{k+1}} \sum_{b \neq c} \sum_{x \in \mathbb{Z}_N^k} \delta_{b \cdot x, c \cdot x} \sum_{w \in \mathbb{Z}_N} \delta_{b \cdot x, w} \\
&\quad = \mu + \frac{1}{N^{k+1}} \sum_{b \neq c} \sum_{x \in \mathbb{Z}_N^k} \delta_{b \cdot x, c \cdot x} && \text{(A.48)}\\
&\quad = \mu + \frac{1}{N^2} \sum_{b \neq c} \gcd(b_1 - c_1, \ldots, b_k - c_k, N) && \text{(A.49)}\\
&\quad \leq \mu + \frac{1}{N^2} \sum_{b \neq c} \gcd(b_1 - c_1, \ldots, b_k - c_k) && \text{(A.50)}
\end{aligned}$$

where the next to last step follows from Lemma A.1. Now for any $q \in \{0, \ldots, M-1\}$, for a fixed value of c_i, there are $1 + \lfloor (M - c_i - 1)/q \rfloor + \lfloor c_i/q \rfloor \leq 1 + M/q$ choices of $b_i \in \{0, \ldots, M-1\}$ that are divisible by q, and hence the number of b, c such that $\gcd(b_1 - c_1, \ldots, b_k - c_k) = q$ is upper bounded by $M^k[(M+q)/q]^k$. Therefore, for fixed $k \geq 3$ and $M^k/N = 1 + O(1/N)$, we have

$$\begin{aligned}
&\mathop{\mathbf{E}}_{x,w} [(\eta_w^x)^2] - \mu && \text{(A.51)}\\
&\quad \leq \frac{M^k}{N^2} \sum_{q=1}^{M-1} q \Big(\frac{M+q}{q} \Big)^k \\
&\quad = \frac{M^k}{N^2} \sum_{q=1}^{M-1} \sum_{j=0}^{k} \binom{k}{j} \frac{M^j}{q^{j-1}} && \text{(A.52)}\\
&\quad \leq \frac{M^k}{N^2} \Big[O(M^2 \log M) + \sum_{j=3}^{k} \binom{k}{j} M^j \sum_{q=1}^{\infty} \frac{1}{q^{j-1}} \Big] && \text{(A.53)}\\
&\quad \leq \frac{M^k}{N^2} \Big[O(M^2 \log M) + \frac{\pi^2}{6} \sum_{j=3}^{k} \binom{k}{j} M^j \Big] && \text{(A.54)}\\
&\quad = \frac{\pi^2}{6} + o(1)\,, && \text{(A.55)}
\end{aligned}$$

where in the next to last step we have used the fact that $\sum_{q=1}^\infty q^{-(j-1)} \leq \pi^2/6$ for any $j \geq 3$. As $\mu = M^k/N = 1 + o(1)$, we find $\sigma^2 = \mathbf{E}_{x,w}[(\eta_w^x)^2] - \mu^2 \leq \pi^2/6 + o(1)$.

Since the variance is small, Chebyshev's inequality shows that the probability of deviating far from the mean number of solutions is small:

$$\Pr(|\eta_w^x - \mu| \geq \Delta) \leq \frac{\sigma^2}{\Delta^2}\,. \qquad \text{(A.56)}$$

Putting $\Delta = 4$ and using the fact that η_w^x must be an integer, we find $\Pr(\eta_w^x \geq 5) \leq \pi^2/96 + o(1)$.

To see that we are unlikely to have no solutions, we need a slightly stronger bound than the Chebyshev

inequality. Since $\eta_w^x \in \mathbb{N}$, we have $\Pr(\eta_w^x = 0) \leq \sigma^2/(\mu^2 + \sigma^2)$ [1, p. 58]. Now, noting that the gcd in (A.49) is at least 1, we have

$$\mathop{\mathbf{E}}_{x,w}[(\eta_w^x)^2] \geq \mu + \frac{M^k(M^k - 1)}{N^2} = 2 + o(1) \tag{A.57}$$

so that $\sigma^2 \geq 1 + o(1)$. Therefore, we find $\Pr(\eta_w^x = 0) \leq \pi^2/12 + o(1)$. Combining these results, we see that $\Pr(1 \leq \eta_w^x \leq 4) \geq 1 - 3\pi^2/32 + o(1) \geq 0.0747 + o(1)$, so that the probability is lower bounded by a constant for sufficiently large N.

While the above bounds apply to arbitrary values of N, they are not tight, and better bounds can be obtained using knowledge of the factorization of N. For example, if N is prime, $\sigma^2 \sim 1$. For $k = 2$, the above argument is not sufficient except for special values of N (such as N prime); indeed, if N has an unbounded number of distinct prime factors, then it appears that the variance of η_w^x might be unbounded. However, for this case, one can simply decrease M and use $k = 3$ copies, as mentioned previously.