Using neural networks to infer acceleration for an atom-based accelerometer

A Major Qualifying Project (MQP) Report
Submitted to the Faculty of
WORCESTER POLYTECHNIC INSTITUTE
in partial fulfillment of the requirements
for the Degree of Bachelor of Science in

Physics

By:

Eric Lopes

Project Advisor:

Alex A. Zozulya

Date: August 2021

Abstract

Atom based inertial sensing is widely acknowledged by the atomic physics and precision measurements communities as having enormous potential in applications of both scientific and practical interest. One of the promising approaches to many sensor applications can be addressed by systems based on the physics of optical atomic lattices. An optical lattice is created by interfering laser beams to form a two-dimensional lattice. Such a lattice can be used to trap and manipulate cold atoms. The present project is devoted to the theoretical analysis of one of such interferometric geometries.

Atoms in optical lattices achieve their sensing capability by first preparing them in the quantum-mechanical ground state of the lattice. Subsequently the wavefunction of the atoms are caused to undergo a series of transformations. That series of transformations correspond to those of an interferometer (splitting, propagation, reflection, propagation, and recombination). Each of those wavefunction transformations is achieved by modulating the optical lattice. We introduce and analyze a specific optical lattice modulation protocol that realizes the interferometric cycle. The results can not be interpreted analytically, so the analysis relies on machine learning techniques. Neural networks ware employed to interpret the results, and are shown to be successful for inferring the acceleration of the lattice within multiple ranges of acceleration.

Contents

List of Tables

List of Figures

1 Introduction

1.1 Interferometry

In 1801, Thomas Young experimentally demonstrated the wave nature of light through his famous double-slit experiment. A light beam passing through two slits separated by a small distance produces an interference pattern of dark and light bands on a screen placed behind the two slits.

Consider a light source emitting light of wavelength λ. The two narrow slits separated by a distance d are placed close to the light source so that they receive light in phase. Thus, the two slits act as coherent sources of light [1]. In the far field, the set of angles θ at which constructive interference occurs is given by

$$d \sin\theta = n\lambda \tag{1}$$

where n is an integer, and the angles at which destructive interference occurs are given by

$$d \sin\theta = (2n+1)\lambda/2. \tag{2}$$

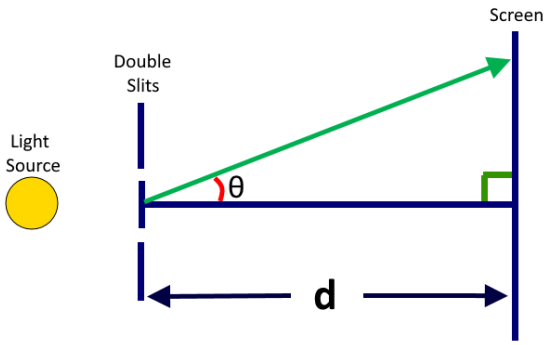

Figure 1: A schematic of the double-slit experiment where the slits and the screen are a distance d apart.

This experiment opened up the branch of optical interferometry. In a typical optical interferometer, a light beam is split by a beamsplitter into two. The split beams travel along different paths, and then are recombined to produce interference. Three very common interferometer types are Mach-Zehnder, Michelson, and Sagnac.

In a Mach-Zehnder interferometer (Fig. 2), the splitting of the original light beam and the recombination of the split beams occur at different points in space. The original beam is split so that each half

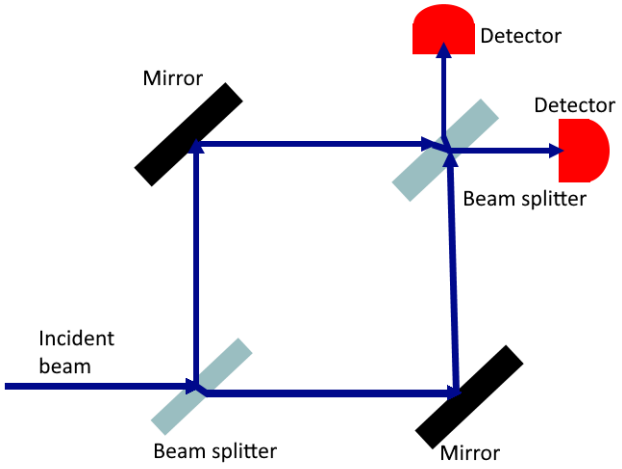

Figure 2: A schematic of a Mach-Zehnder interferometer.

travels at a right angle to the other. They travel for some time and are reflected. After reflection each split beam is moving in the direction the other beam was previously moving. After travelling for some time again, the beams are recombined such that the paths of the split beams trace out a rectangle. In a Michelson interferometer (Fig. 3), the splitting and the recombination happen at the same place. In this case, the retroferlecting mirrors cause each beam to travel in the opposite direction with respect to its original direction. Sagnac interferometer (Fig. 4) has a circular geometry and is used to measure rotations. The two split beams travel along a circular path moving clockwise and counterclockwise, are recombined and interfere. Optical interferometry is used to measure the characteristics of stars [2], the flatness and deformation of a material [3], rotational motion [4], and gravitational waves [5].

In 1924, Louis de Broglie put forward a hypothesis that all matter has wave-like properties, like light. According to de Broglie, every particle has a wave associated with it. The wavelength of this wave (known as the de Broglie wavelength) is given by

$$\lambda = \frac{h}{p} \tag{3}$$

where h is Planck's constant, and p is the momentum of the particle [6]. This hypothesis was proven in the electron diffraction experiment and later the neutron interference experiments of the 1940s. By 1991, interference using more massive particles, like atoms, was demonstrated as well [7].

The thermal de Broglie wavelength of a particle of mass m in thermodynamic equilibrium with a

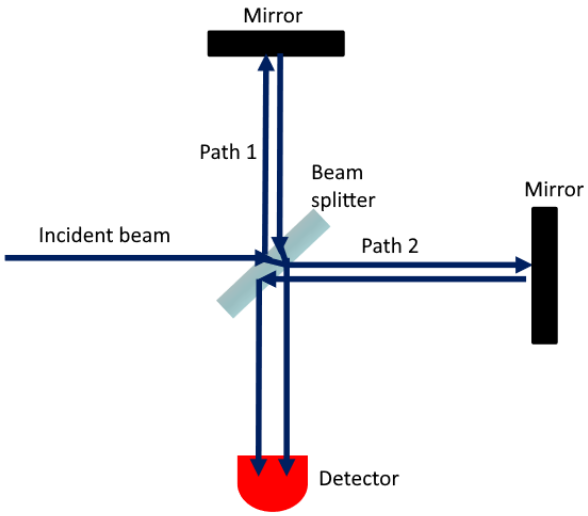

Figure 3: A schematic of a Michelson interferometer.

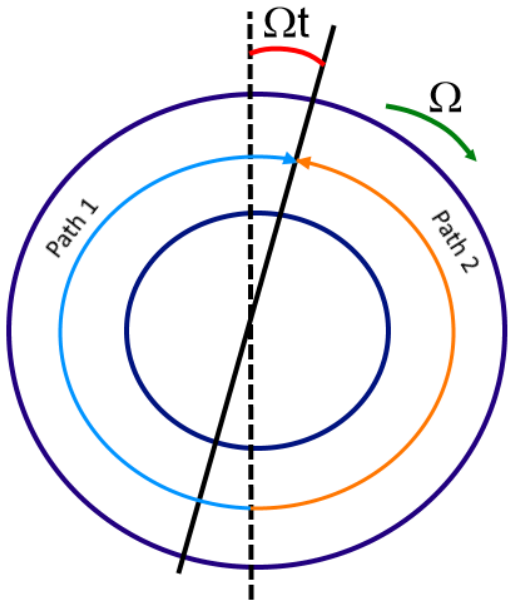

Figure 4: A schematic of a Sagnac interferometer.

reservoir at temperature T is given by

$$\lambda_{th} = \sqrt{\frac{(2\pi\hbar)^2}{mk_BT}} \tag{4}$$

where k_B and $\hbar$ are the Boltzmann and the reduced Planck constants, respectively. The de Broglie wavelength is inversely proportional to the square of temperature and is extremely small unless T is of the order of micro or nanokelvins. Thus, for a Rubidium atom ($m_{Rb} = 1.42 \cdot 10^{-25} kg$) at room temperature ($T = 273K$), $\lambda_{th} = 2.864 \cdot 10^{-11} m$. This wavelength is so small that any interferometry is currently technologically impossible. A reasonable value of the wavelength (say, $1\mu m$) requires $T = 0.2 \cdot 10^{-6} K$. In the realm of ultra-low temperatures, Bose-Einstein condensate (BEC) is the natural candidate for interferometry due to its high density and coherent nature.

Bose-Einstein condensate is one of the most fascinating manifestations of the wave nature of matter. In the 1920s, Satyendra Nath Bose developed statistics to describe quanta of light, and Einstein applied them to a gas of noninteracting atoms. He predicted that all atoms go to the same quantum state below a certain temperature, this state of matter being Bose-Einstein condensate. However, the technology at the time wasn't able to cool atoms down to a low enough temperature. In the 1980s, cooling techniques were developed, and in 1995 BEC was experimentally observed in ^{87}Rb atoms at University of Colorado [8]. In 2001, the Nobel Prize in Physics was awarded to two groups for the achievement of Bose-Einstein condensation in dilute gases of alkali atoms, and for early fundamental studies of the properties of the condensates [9]. In 2017, researchers from MIT developed a technique to make BEC much faster. The researchers were able to cool 2,000 atoms, and from that, generate a condensate of 1,400 atoms, conserving 70 percent of the original amount [10].

The sensitivity of atom interferometers depends on the interferometric cycle time, increasing as cycle time increases. The current cycle times of free-space interferometers are less than one-tenth of a second and on Earth are limited by sagging of the atomic beam due to the gravitational field. One solution to this problem is to use an atomic fountain that increases the physical size of the interferometer, decreasing portability and requiring very sensitive technical details for operation [11]. These limitations lead to the desire for a technique that holds atoms against gravity throughout the whole interferometric cycle without compromising the portability. An example of this type of technique is the use of a confining trap to hold atoms against gravity while the atoms are manipulated by the laser beams. Condensates are perfect to use in this technique because they have very small momentum that allows them to be confined to a small region in space [12].

The first atom interferometers worked with streams of supersonic gases and used mechanical gratings in order to get past the problem of having a small de Broglie wavelength [13]. Following experiments used laser beams that created a periodic potential in place of the material and mechanical gratings to split and recombine streams of gaseous atomic beams [14]. They have been used to measure gravitational constants [15], acceleration [14], electric polarizability [16], and the fine-structure constant [17].

1.2 Atom Interferometers

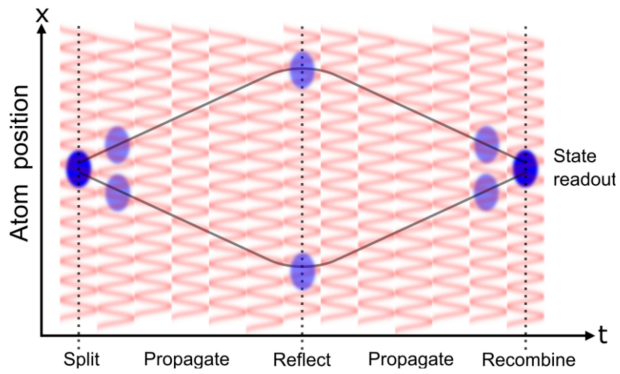

Figure 5: A scheme of a full interferometer sequence, with atom position plotted along the y-axis versus time on the x-axis. Blue clouds represent atom wavepackets interacting with the lattice. (Source: [18])

Atomic interferometers work similarly to light-based interferometers. The inital atomic cloud is split into two, and those two clouds propagate in opposite directions. After some time, the velocity of each cloud is reversed and they propagate back towards each other. When the two clouds are again in the starting location, they are are recombined. This process can be seen in Fig. 5. BEC-based atom interferometers have been realized through various schemes such as trapped-atom schemes and guided wave schemes.

In trapped atom interferometer schemes, a cloud of condensate, which is in the lowest mode of a single well trap and sitting at the center of the trap, is split into two clouds in real space by deforming a single well potential into a double-well potential [19]. Guided-wave atom interferometers use potentials to guide the motion of atomic wave packets by splitting the condensate in momentum space. This can be applied to Michelson and Mach-Zehnder schemes. In these schemes, the splitting of the condensate in momentum space is used to manipulate the condensate in the guide.

In a Michelson atom interferometer [20], the BEC cloud is initially at rest in the waveguide. Splitting pulses consisting of a pair of counterpropagating laser beams detuned from atomic resonance are incident on the cloud. These pulses split the condensate into two harmonics moving in the opposite directions with with velocities $\pm v_0$. In a single-reflection interferometer, the directions of propagation of these harmonics

are reversed at time $T/2$, where T is the total cycle time. The harmonics are allowed to propagate back and are recombined when they overlap again using the same optical pulses that were used to split the original BEC cloud. After the recombination, the condensante is in a superposition of the ground state and two harmonics with the relative amplitudes depending on the amount of the accumulated phase shift between the arms of the interferometer.

In a Mach-Zehnder atom interferometer [21], one of the two counter-propagating waves used to form the splitting pulses is frequency-shifted with respect to the other, resulting in a travelling optical potential. The pulses transform the BEC originally at rest at the center of the trap into two clouds of equal amplitude. One of the clouds remains at rest and the other travels with a velocity v. Another pulse applied mid-cycle stops the moving cloud and brings the stationary cloud into motion. At the end of the cycle, another pulse is used to recombine the clouds.

In recent years a novel approach to atomic sensing based on optical latices has been proposed [18, 22, 23, 24]. An optical lattice is created by interfering laser beams to form a two-dimensional or three-dimensional lattice. Such a lattice can be used to trap and manipulate cold atoms. Atoms in optical lattices achieve their sensing capability by first preparing them in the quantum-mechanical ground state of the lattice. Subsequently the wavefunction of the atoms are caused to undergo a series of transformations. That series of transformations correspond to those of an interferometer (splitting, propagation, reflection, propagation, and recombination). Each of those wavefunction transformation is achieved by modulating the optical lattice appropriately.

1.3 Optical Lattice

The interaction between an atom and electromagnetic field is given in the dipole approximation by

$$H_e = -\vec{d} \cdot \vec{\mathcal{E}} \tag{5}$$

where $\vec{d}$ is the electric dipole moment operator and $\vec{\mathcal{E}}$ is the electric field vector. In a static electric field the change ΔE_g in the ground-state energy of an atom is given to second order by

$$\Delta E_g = -\sum_e \frac{|\langle e| H_e |g\rangle|^2}{E_e - E_g} = -\frac{1}{2}\alpha \mathcal{E}^2, \tag{6}$$

where

$$\alpha = 2\sum_e \frac{|\langle e| \vec{d} \cdot \hat{\epsilon} |g\rangle|^2}{E_e - E_g} \tag{7}$$

is the atomic polarizability and $\hat{\epsilon}$ is the unit vector in the direciton of the electric field.

For a time-dependent electric field with frequency ω, the electric field is

$$\vec{\mathcal{E}} = \vec{\mathcal{E}_\omega} e^{-i\omega t} + \vec{\mathcal{E}_{-\omega}} e^{i\omega t}. \tag{8}$$

The energy shift in this case is then

$$\begin{aligned}
\Delta E_g =& \sum_e \langle g| \vec{d} \cdot \vec{\mathcal{E}_\omega} |e\rangle \frac{1}{E_g - E_e + \hbar\omega} \langle e| \vec{d} \cdot \vec{\mathcal{E}_{-\omega}} |g\rangle \\
&+ \sum_e \langle g| \vec{d} \cdot \vec{\mathcal{E}_{-\omega}} |e\rangle \frac{1}{E_g - E_e - \hbar\omega} \langle e| \vec{d} \cdot \vec{\mathcal{E}_\omega} |g\rangle \\
=& \sum_e |\langle e| \vec{d} \cdot \hat{e} |g\rangle|^2 \left(\frac{1}{E_g - E_e - \hbar\omega} + \frac{1}{E_g - E_e + \hbar\omega} \right) |\mathcal{E}_\omega|^2 \\
=& -\frac{1}{2}\alpha(\omega) < \mathcal{E}(\vec{r}, t) >_t
\end{aligned} \tag{9}$$

where $< ... >_t$ denotes an average over time, and the polarizability is given by

$$\alpha(\omega) = \sum_e \frac{2(E_e - E_g)|\langle e| \vec{d} \cdot \hat{\epsilon} |g\rangle|^2}{(E_e - E_g)^2 - (\hbar\omega)^2}. \tag{10}$$

In many situations frequency of the electric field is close to that of an atomic resonance, and it is then a good approximation to neglect all transitions except the resonant one. The polarizability then reduces to a single term

$$\alpha(\omega) \approx \frac{|\langle e| \vec{d} \cdot \hat{\epsilon} |g\rangle|^2}{E_e - E_g - \hbar\omega}. \tag{11}$$

The analysis so far has assumed the excited state has an infinitely long lifetime, but this is not the case in reality. If the excited state has a lifetime of $1/\Gamma_e$, the polarizability is then

$$\alpha(\omega) \approx \frac{|\langle e| \vec{d} \cdot \hat{\epsilon} |g\rangle|^2}{E_e - i\hbar\Gamma_e/2 - E_g - \hbar\omega}. \tag{12}$$

The energy shift is a complex quantity and can be written as

$$\Delta E_g = V_g - i\hbar\Gamma_g/2 \tag{13}$$

which has the form of an effective potential acting on the atom, the real part corresponding to a shift of the

energy, and the imaginary part to a finite lifetime of the ground state. The shift of the energy level is

$$V_g = -\frac{1}{2}\alpha'(\omega) < \mathcal{E}^2(\vec{r},t) >_t \tag{14}$$

where

$$\alpha'(\omega) \approx \frac{(E_e - E_g - \hbar\omega)|\left\langle e\right| \vec{d}\cdot\hat{\epsilon}\left|g\right\rangle|^2}{(E_e - E_g - \hbar\omega)^2 + (\hbar\Gamma_e/2)^2}. \tag{15}$$

The energy shift can be written as

$$V_g = \frac{\hbar\Omega_R^2\Delta}{\Delta^2 + \Gamma_e^2/4} \tag{16}$$

Here we introduced the detuning Δ, which is the difference between the laser frequency and the frequency $\omega_{eg} = (E_e - E_g)/\hbar$ of the atomic transition,

$$\Delta = \omega - \omega_{eg}. \tag{17}$$

Positive values of Δ are referred to as *blue* detuning and negative Δ as *red* detuning. Parameter Ω_R

$$\Omega_R = |\left\langle e\right| \vec{d}\cdot\vec{\mathcal{E}_\omega}\left|g\right\rangle|/\hbar. \tag{18}$$

is called the Rabi frequency. The rate of loss of atoms from the ground state is given by

$$\Gamma_g = \frac{1}{\hbar}\alpha''(\omega) < \mathcal{E}^2(\vec{r},t) >_t \tag{19}$$

where

$$\alpha''(\omega) \approx \frac{\hbar\Gamma_e/2}{(E_e - E_g - \hbar\omega)^2 + (\hbar\Gamma_e/2)^2}|\left\langle e\right| \vec{d}\cdot\hat{\epsilon}\left|g\right\rangle|^2. \tag{20}$$

In the limit of large detuning

$$\Delta >> \hbar\Gamma_e/2 \tag{21}$$

the real and the imaginary parts of the polarizability behave as

$$\begin{aligned} \alpha'(\omega) &\propto \Delta^{-1} \\ \alpha''(\omega) &\propto \Delta^{-2}, \end{aligned} \tag{22}$$

and so for large detuning, we can neglect the loss of atoms from the ground state [25].

A one-dimensional optical lattice can be created by superimposing two oppositely directed laser beams with the same frequency. Assuming that both beams are linearly polarized with the electric field vector along the z-axis, the total field is

$$E_z = E_0 \cos(k_L x - \omega t) + E_0 \cos(k_L x + \omega t) = 2E_0 \cos(k_L x) \cos(\omega t). \tag{23}$$

Here k_L is the wavevector and ω is the frequency of the laser beams.

As discussed above, the presence of an electric field can be regarded as a potential acting on the atom, and can be given by Eq. (14) which for convenience is presented below:

$$V = -\frac{1}{2}\alpha'(\omega) < E(x,t)^2 >_t . \tag{24}$$

From Eq. (23), we get

$$< E_z^2 >_t = 2E_0^2 \cos^2(k_L x) = E_0^2(\cos(2k_L x) + 1). \tag{25}$$

The associated energy shift is periodic in x with a period π/k_L. When expressed in terms of wavelength $\lambda = 2\pi/k_L$ of the laser, the period is $\lambda/2$.

The constant term in Eq. (23) can be neglected and the potential energy then takes the form [25]

$$V = \frac{V_0}{2} \cos(2k_L x) \tag{26}$$

where

$$V_0 = -\frac{|\langle g|\vec{d} \cdot \hat{\epsilon}|e\rangle|^2 E_0^2}{\hbar\Delta}. \tag{27}$$

All of the calculations done in this paper will be done for Rb^{87} atoms, which have a mass $m = 14.432 \cdot 10^{-23} g$. V_0 will be measured in units of E_R where $E_R = \hbar^2 k_L^2 / 2m$.

1.4 Machine Learning & Neural Networks

Machine learning is a type of artificial intelligence (AI) that trains machines how to learn. Given some set of data, there are various ways that this data can be analyzed and modelled by a computer. A machine learning algorithm creates a model for that data and adapts that model through an iterative process of being given new data. Generally there are two types of machine learning algorithms, supervised and unsupervised. In a supervised machine learning algorithm, the algorithm is given an input set of data

as well as the correct outputs that the algorithm should give. A basic example of this is linear regression algorithms. The algorithm is given both the inputs and the answers at the input nodes. As more data is given, the algorithm is able to create a more accurate model to reflect the data. In an unsupervised machine learning algorithm, the algorithm is not given the correct outputs for its given input set. Instead it looks at all of the input sets and attempts to find some structure among them and analyze that structure. An example of this is separating photos of birds into clusters based on certain characteristics and structures unique to certain types of birds. The algorithm looks for images that are similar and groups them [26].

Neural networks are a specific type of machine learning algorithm that attempt to mimic the functions of a brain (i.e. neurons) to find relationships in data. A basic neural network is made up of three layers of nodes that have connections to the adjacent layers. These layers are the input layer, hidden layer, and output layer. For a typical problem, like linear modelling, both the input layer and output layer are one node. The hidden layer can be given however many nodes are desired to achieve a good model. The input node has a connection to each node in the hidden layer, and each node in the hidden layer has a connection to the output node. These connections are given different weights based on their importance in producing the desired outputs, and they are strengthened through iterations until a desired model accuracy is reached. How the connections are weighted and strengthened, or how the stopping criteria is determined depends on which neural network training function is used. For more complex problems, more hidden layers can be added with various types of connections between them. A neural network with more than one hidden layer is known as a deep neural network [27]. Fig. 6 shows a neural network with one hidden layer. Each node in the input layer is connected to each node in the hidden layer, and each node in the hidden layer is connected to each node in the output layer. In the figure, there are multiple nodes in both the input and output layer, but those numbers don't need to be equal, and they can even be a single node. In the MQP, we use one hidden layer. The input nodes are the relative populations of each plane wave harmonic at the end of the interferometric cycle, and the output layer is a single node which gives the predicted acceleration of the system (see detailed descriptoin later).

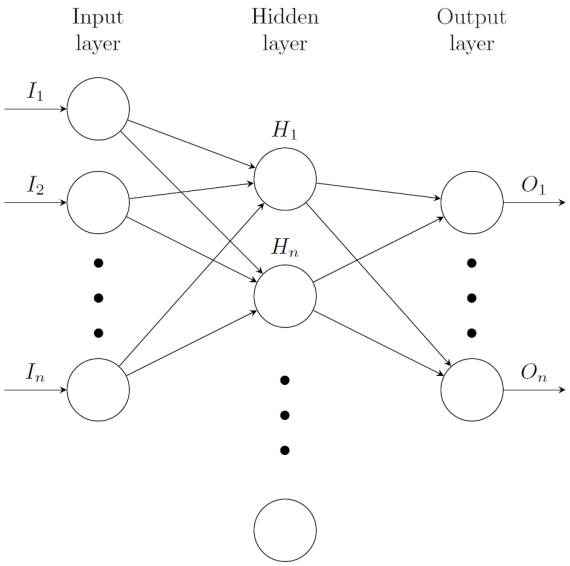

Figure 6: A schematic of a neural network with one hidden layer.

2 Cold atom interferometer with counterpropagating optical potentials

This chapter presents a theoretical analysis of a cold atom interferometer with counterpropagating optical potentials. Section 2.1 gives general description of the proposed interferometric cycle, section 2.2 discusses numerical methods used to solve the time-dependent Schrödinger equation and section 2.3 introduces initial conditions. Section 2.4 describes what happens when the system is in an accelerating frame, and section 2.5 explains the details behind the neural network that we used to infer the value of the acceleration from the interferometric data as well as the results of our analyses.

2.1 Interferometric cycle

Cold atoms that are trapped in an optical potential $V(x,t)$ created by a pair of counterpropagating laser beams with wavevectors k_L given by

$$V(x,t) = \frac{U(t)}{2}\cos(2k_L x + \phi(t)). \tag{28}$$

Here x and t are the coordinate and time respectively, and the phase $\phi(t) = -ak_L t^2$ accounts for the acceleration a of the optical potential with respect to the lab frame.

We consider an amplitude modulation of the optical lattice of the form

$$U(t) = U_0 \cos(2k_L x_p(t)) = U_0 \cos\left(2k_L \int^t v(\tau) d\tau\right) \tag{29}$$

Time dependence of the control parameter $x_p(t)$ is chosen to transform the initial optical lattice into the sum of two counterpropagating optical lattices

$$\begin{aligned} V(x,t) =& \frac{V_0}{2} \cos(2k_L x_p) \cos(2k_L x + \phi(t)) \\ =& \frac{U_0}{4} \cos[2k_L(x + x_p(t)) + \phi(t)] \\ &+ \frac{U_0}{4} \cos[2k_L(x - x_p(t)) + \phi(t)], \end{aligned} \tag{30}$$

moving in opposite directions with velocities $v(t) = dx_p(t)/dt$. A typical time dependence of the velocity $v(t)$ and the displacement $x_p(t)$ is shown in Fig. 7.

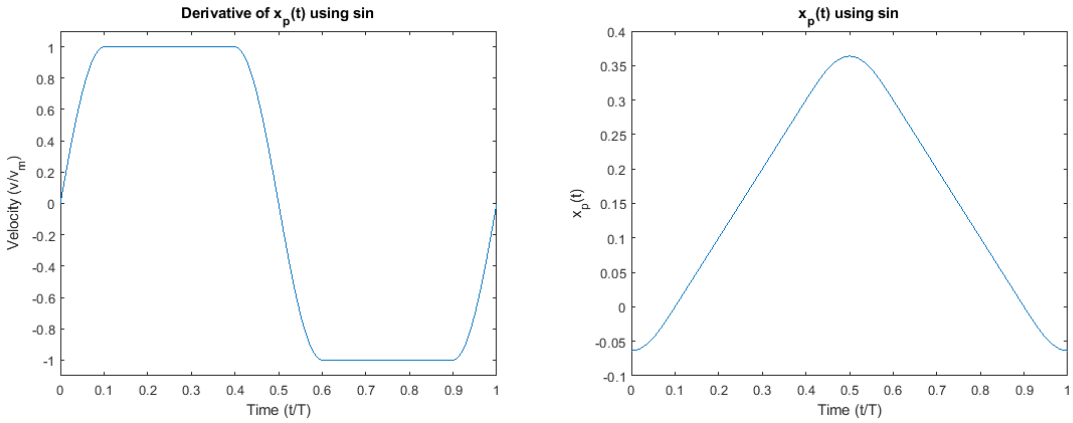

Figure 7: Velocity $v(t)$ (left) and displacement $x_p(t)$ (right).

In the following we will use linear piecewise functions for the velocity $v(t)$. The velocity ramps up from zero to its maximum value v_m during time interval T_1 (see Fig. 8), stays constant during the time interval $T_1 < t < T/2 - T_1$ and then ramps back down to 0 in the same time it took to ramp up T_1. During the first half-cycle

$$v(t) = \begin{cases} v_m \frac{t}{T_1} & 0 \le t \le T_1 \\ v_m & T_1 \le t \le T/2 - T_1 \\ \frac{v_m}{T_1}(T/2 - t) & T/2 - T_1 \le t \le T/2 \end{cases} \tag{31}$$

The second half of the cycle reverses the first half (see Fig. 8). The velocity $v(t)$ goes from zero to

$-v_m$ in the time T_1, keeps that velocity for the time interval $T/2 - 2T_1$, and then ramps back to 0 in the time T_1.

The displacement $x_p(t)$ is given by the integral of $v(t)$. These can be seen in Fig. 8.

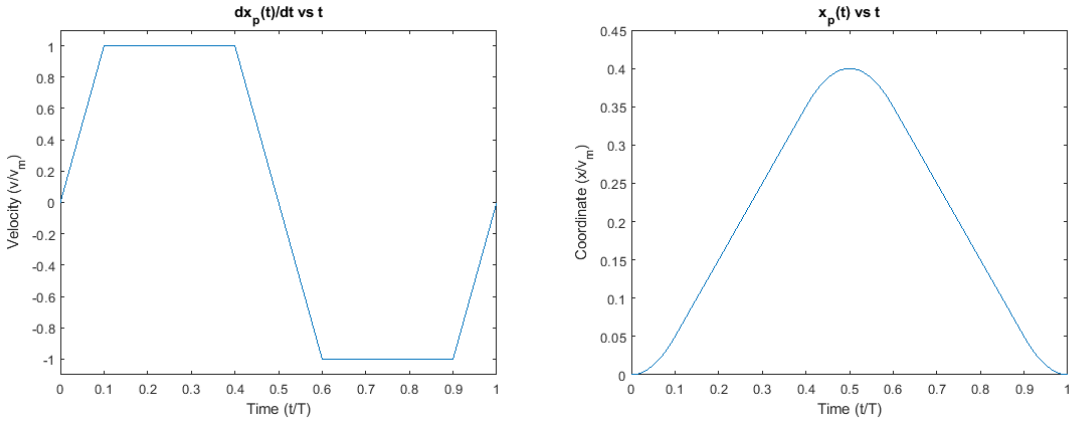

Figure 8: Splitting velocity with linear ramping (left) and the displacement (right).

2.2 Solving the time-dependent Schrödinger equation

Time evolution of the condensate will be described by the one-dimensional Schrödinger equation for the wavefunction $\psi(x,t)$,

$$i\hbar\frac{\partial\psi(x,t)}{\partial t} = -\frac{\hbar^2}{2M}\frac{\partial^2\psi}{\partial x^2} + \frac{U(t)}{2}\cos[2k_L x + \phi(t)]\psi(x,t). \tag{32}$$

In typical experimental conditions the size of the atomic cloud is much larger than the spatial period π/k_L of the optical potential and the displacement of the cloud during the interferometric cycle is much less than its size (see [18]). Therefore we will treat both the optical lattice and the atomic cloud as infinite.

In this approximation the atomic wafefunction can be sought as a product of a plane wave $\exp(iqx)$ and a function $u(x,t)$ that is periodic with the period π/k_L of the optical lattice. In other words, it is a superposition of plane waves separated in Fourier space by an integer number of the wavevectors of the optical lattice:

$$\psi(x,t) = e^{iqx}\sum_{m=-\infty}^{\infty} f_m(t)e^{im2k_L x} \tag{33}$$

Wavevector $q = \text{const.}$ is determined by initial conditions and does not change during the evolution of the wavefunction.

The amplitudes of the Fourier harmonics $f_m(t)$ satisfy the system of coupled equations

$$i\hbar\frac{d}{dt}f_m = -\frac{\hbar^2(2k_Lm+q)^2}{2M}c_m + \frac{U(t)}{4}\Big(f_{m-1}e^{i\phi} + f_{m+1}e^{-i\phi}\Big). \tag{34}$$

2.2.1 Set of ordinary differential equations

In numerics we truncate the set of harmonics to only include terms between some $\pm n_{max}$. The truncated set of equations can be written in matrix form as

$$i\hbar\frac{\partial}{\partial t}\psi = H\psi \tag{35}$$

where H is written as

$$H = \begin{bmatrix}
d_{-n_m} & -Ke^{-i\phi} & & & & & & & \\
-Ke^{i\phi} & d_{-n_m+1} & -Ke^{-i\phi} & & & & & & \\
 & \ddots & \ddots & \ddots & & & & & \\
 & & -Ke^{i\phi} & d_{-1} & -Ke^{-i\phi} & & & & \\
 & & & -Ke^{i\phi} & d_0 & -Ke^{-i\phi} & & & \\
 & & & & -Ke^{i\phi} & d_1 & -Ke^{-i\phi} & & \\
 & & & & & \ddots & \ddots & \ddots & \\
 & & & & & & -Ke^{i\phi} & d_{n_m-1} & -Ke^{-i\phi} \\
 & & & & & & & -Ke^{i\phi} & d_{n_m}
\end{bmatrix} \tag{36}$$

for $d_n = 2\hbar^2k_L^2n^2/m$ and $K = U(t)/4$. From here, numerical solvers, like `ode45` in MATLAB, can be used to solve for the set of coefficients $f_n(t)$.

2.2.2 The split-step method

Another method that can be used to solve Eq. (32) is the split-step method. The split-step method can be used to solve the time-dependent Schrödinger equation by splitting Hamiltonian into the kinetic energy and potential energy operators, and propagating each separately over a time step. For some Hamiltonian of the form $\hat{H} = \hat{K} + \hat{V}$, the propagator for a state $\psi(x,t)$ becomes

$$\psi(x, t+\delta t) = \hat{U}\psi(x,t) = e^{-\frac{i\delta t}{\hbar}\hat{H}}\psi(x,t). \tag{37}$$

The kinetic energy operator is diagonal in momentum space, and the potential operator is diagonal in position space, and so the kinetic energy operator happens in Fourier space, and the potential energy propagation happens in real space. The fast Fourier transform is used for quick computation of this method. Since $\hat{U}$ is unitary, the norm of ψ is preserved. This method assumes periodic boundary conditions, which is ideal for this situation as the Hamiltonian is periodic. Expanding Eq. (37), we get

$$\psi(x, t+\delta t) = e^{-\frac{i\delta t}{\hbar}\hat{H}}\psi(x,t) = e^{-\frac{i\delta t}{\hbar}\hat{K}}e^{-\frac{i\delta t}{\hbar}\hat{V}}\psi(x,t) \tag{38}$$

or, to second order accuracy [28],

$$\psi(x, t+\delta t) = e^{-\frac{i\delta t}{2\hbar}\hat{K}}e^{-\frac{i\delta t}{\hbar}\hat{V}}e^{-\frac{i\delta t}{2\hbar}\hat{K}}\psi(x,t) \tag{39}$$

which is the form used in our simulations. The split-step method in general is much faster than using `ode45`, and so that is the method that we use in our simulations.

2.3 Initial conditions

Consider the Hamiltonian of the form

$$H = -\frac{\hbar^2}{2m}\left(\frac{\partial}{\partial x}\right)^2 - \frac{V_0}{2}\cos\left(2k_L x\right). \tag{40}$$

Delocalized eigenfunctions of this Hamiltonian are called Bloch waves.

For a periodic potential

$$V(x) = V(x+a)$$

where a is the period of the potential, Bloch's theorem [29] states that solutions of the eigenfunction-eigenvalue Schrödinger equation have the form

$$\psi_{nq}(x) = e^{iqx}u_n(x) \tag{41}$$

where q is a constant crystal wave vector, and $u_n(x)$ is a periodic function

$$u_n(x) = u_n(x+a)$$

that has the periodicity of the potential. Integer index n labels solutions with the same value of q but

different values of the eignenergy E_n.

Since $u_n(x)$ is periodic with the period a, it can be written as

$$u_n(x) = \sum_{m=-\infty}^{\infty} f_m e^{i(2\pi/a)mx}. \tag{42}$$

The Hamiltonian (40) is periodic with a period of π/k_L, and so Bloch functions $\psi_{nq}(x)$ can be written as

$$\psi_{nq}(x) = e^{iqx} u_{nq}(x) \tag{43}$$

where

$$u_{nq}(x) = \sum_{m=-\infty}^{\infty} f_m e^{im2k_L x}.$$

Plugging this form into the time-independent Schrödinger equation,

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x) - \frac{V_0}{2}\cos(2k_L x)\psi_{nq}(x) = E_{nq}\psi_{nq}(x) \tag{44}$$

we get a set of equations to solve for f_m

$$\frac{2\hbar^2 k_L^2(q^2+m^2)}{M} f_m - \frac{V_0}{4}(f_{m-1} + f_{m+1}) = E_{nq} f_m. \tag{45}$$

In numerics we work with a finite set of f_m, so we define ψ as the truncated set of coefficients

$$\psi = [f_{-n_m}, ...f_{-1}, f_0, f_1, ...f_{n_m}]^T$$

by introducing the approximation $f_n = 0\ \forall |n| > n_m$. The Hamiltonian H in matrix form has the form

$$H = \begin{bmatrix} d_{-n_m} & -K & & & & & & & \\ -K & d_{-n_m+1} & -K & & & & & & \\ & \ddots & \ddots & \ddots & & & & & \\ & & -K & d_{-1} & -K & & & & \\ & & & -K & d_0 & -K & & & \\ & & & & -K & d_1 & -K & & \\ & & & & & \ddots & \ddots & \ddots & \\ & & & & & & -K & d_{n_m-1} & -K \\ & & & & & & & -K & d_{n_m} \end{bmatrix} \tag{46}$$

where $d_m = 2\hbar^2 k_L^2 (q^2 + m^2)/M$ and $K = V_0/4$. This Hamiltonian is used to numerically solve the eigenfunction-eigenvalue equation. The ground state solution corresponding to $q = 0$ and $n = 1$ is shown in Fig. 9, and the first 5 eigenenergies as a function of q are shown in Fig. 10. In the following we use the

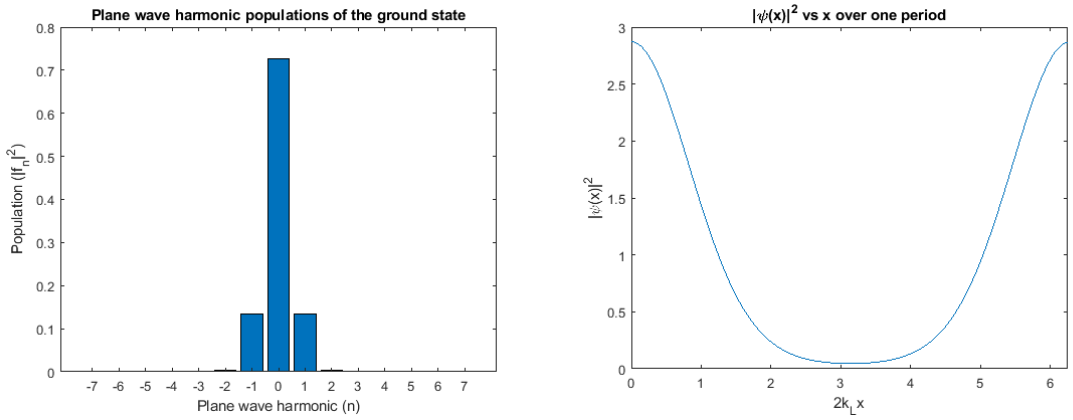

Figure 9: Solution for the ground state of the Hamiltonian for $V_0 = 10E_R$ in momentum (left) and coordinate (right) space.

lowest-energy Bloch eigenfunvtion $n = 1$, $q = 0$ as the initial condition for ψ.

2.4 Output populations of Fourier harmonics for zero and nonzero accelerations

The only information available to experimentalists is the population of Fourier harmonics $|f_m(t)|^2$ (see Eq. (33)) at the end of the interferometric cycle $t = T$. The value of the acceleration a has to be inferred from these data. In this subsection we calculate the populations $|f_m(T)|^2$ by numerically solving

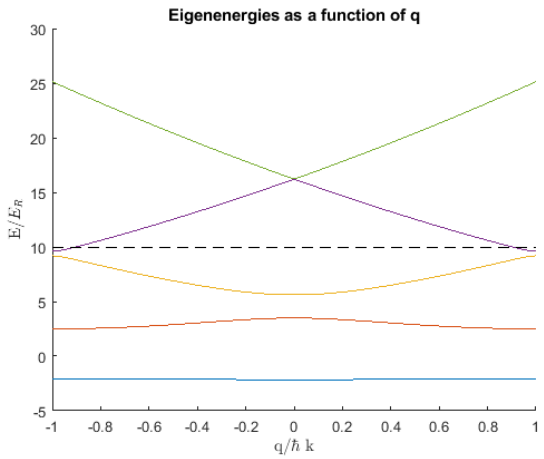

Figure 10: Eigenenergy as a function of q for the ground state (blue) and first four excited states (orange, yellow, purple, and green respectively) for $V_0/E_R = 10$ (the black dotted line).

the time-dependent Schrodinger equation

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} - \frac{V_0}{2}\cos\left(2k_L x_p(t)\right)\cos\left(2k_L x - ak_l t^2\right)\psi. \tag{47}$$

where

$$\psi(x,t) = \sum_{m=-\infty}^{\infty} f_m(t)e^{i(2mk_L x}. \tag{48}$$

with the initial condition being the lowest-energy Bloch function $n = 1$ $q = 0$ discussed above, for zero and nonzero acceleration a.

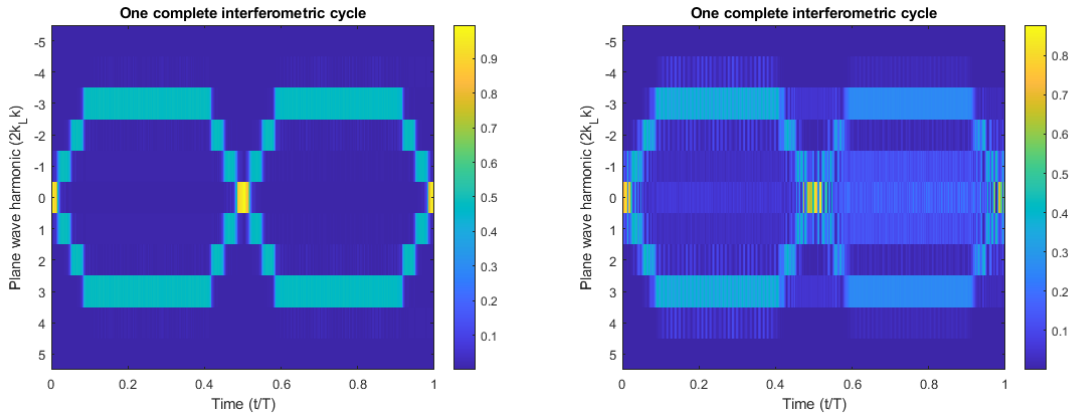

Figure 11: Populations of harmonics versus time for $T = 62.5ms$ and $V_0 = 3E_R$ (left) and $T = 6ms$, $V_0 = 10E_R$ (right). For both cases $v_m = 6k_L\hbar/m$, $a = 0$.

Fig. 11 shows complete interferometric cycle for zero acceleration. It presents populations of Fourier harmonics $|f_m(t)|^2$.

Left graph corresponds to a relative large cycle time $T = 62.5ms$, $v_m = 6k_L\hbar/m$, and $V_0 = 3E_R$. The left graph corresponds to a shorter time $T = 6ms$, $v_m = 6k_L\hbar/m$, and $V_0 = 10E_R$. For both cases we can clearly observe the splitting, propagation, reflection, back propagation, and recombination sequence. The larger cycle time gives better results whereas for a shorter time some loss of the atoms from the moving optical potentials is observed.

Fig. 12 is analogous to Fig. 11 except is is calculated for the nonzero acceleration $a = g/20$. Here $g = 9.8m/\text{s}^2$. Both cases clearly show a bias towards one direction across the entire cycle.

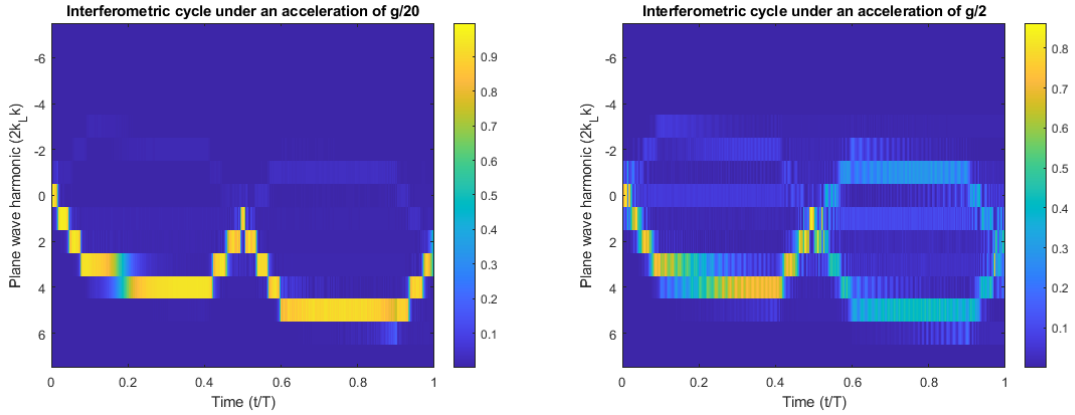

Figure 12: Populations of harmonics versus time for $T = 62.5ms$ and $V_0 = 3E_R$ (left) and $T = 6ms$, $V_0 = 10E_R$ (right). For both cases $v_m = 6k_L\hbar/m$, $a = g/20$.

Fig. 13 shows output populations of the Fourier harmonics versus acceleration a.

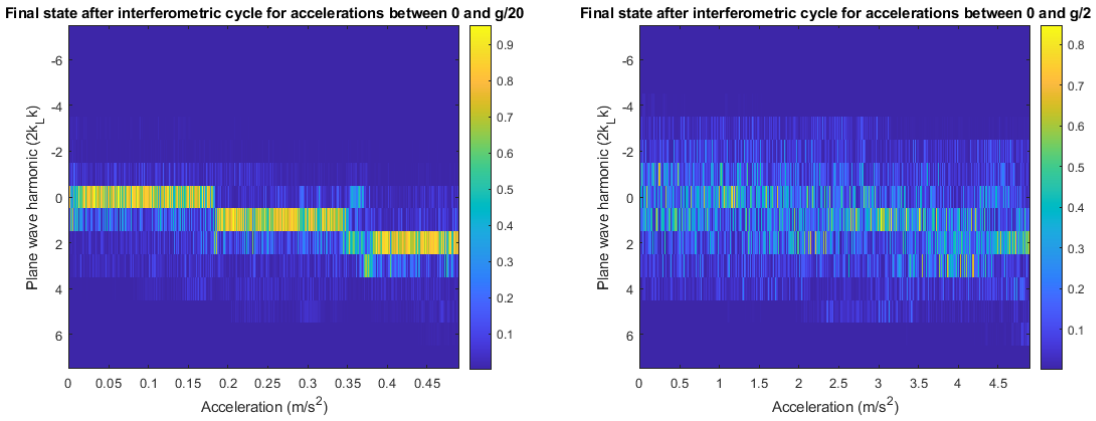

Figure 13: Output populations of Fourier harmonics versus acceleration. For the left image $T = 62.5ms$, $V_0 = 3E_R$, $0 \leq a \leq g/20$. For the right image $T = 6ms$, $V_0 = 10E_R$. $0 \leq a \leq 0.5g$. For both cases $v_m = 6k_L$.

2.5 Inferring the value of acceleration from the output populations

Fig. 13 demonstrates a very rich and complicated structure of the output populations of Fourier harmonics. For their interpretation we use powerful methods of machine learning.

To infer the value of acceleration from the output populations we employ neural networks. The network uses one hidden layer with 32 nodes and the Bayesian Regularization algorithm. Numerical implementation of this network is provided in the MATLAB Neural Fitting toolbox.

For a given range of accelerations we calculate the output populations for a set of uniformly sampled acceleration points. These populations are used as a training set for the neural network. The trained network is tested on a set of output populations calculated for accelerations with values halfway between the values of the accelerations of the training set. For all the results presented below $T = 2.5ms$, $v_m = 4k_L\hbar/m$ and $V_0 = 10E_R$.

We looked at the relative error for four different ranges of acceleration: [0,1g,0.25g], [0.1g,0.5g], [0.1g,0.75g] and [0.1g,1g] for three different numbers of training samples: 50, 500 and 5000. We chose relatively large values of measured acceleration, since our goal was to demonstrate that optical lattices enable operation in highly dynamic environments.

Table 1 shows the mean relative error across the range of the accelerations and the median of the distribution of errors for each of the acceleration ranges and for different number of samples. The data shows that increasing the number of samples results in a decreased error for all cases. Increasing the number of samples from 50 samples to 500 decreases the mean and median by at least an order of magnitude. Transition from 500 to 5000 samples does not decrease the mean and median by nearly as much. From this, we see that it isn't efficient to increase the number of samples above a certain limit, since it takes more time to train the network on them, and the improvement to accuracy slows down.

Relative error of the measurement for different values of the acceleration is shown in Fig. 14 for 500 samples. Some of the points give large errors, but, in fact, the number of such points is relatively small, as is demonstrated by the averaged values of the error presented in Table 1.

Table 2 shows the same data as Table 1 for larger values of accelerations: [1g,1.5g] and [1g,2g]. The errors in Table 2 follow a similar trend as for those in Table 1. However, for the cases where the number of samples is 50, the error is significantly smaller than for those in Table 1. As well, the mean and median of the error in Table 2 are closer to each other in all cases than for those in Table 1. From this, it seems that the accuracy is generally better for ranges farther away from zero acceleration.

Range (g)	# of Samples	Mean Relative Error	Median of Error Dist.
[0.1,0.25]	50	0.144	0.112
	500	7.32e-07	5.79e-07
	5000	5.79e-07	4.94e-07
[0.1,0.5]	50	0.265	0.157
	500	2.11e-3	1.49e-3
	5000	1.13e-3	7.85e-4
[0.1,0.75]	50	0.383	0.207
	500	0.034	0.018
	5000	0.016	9.82e-3
[0.1,1]	50	0.325	0.193
	500	0.081	0.042
	5000	0.036	0.019

Table 1: The mean of the relative error and median of the distribution of errors for neural network trained on different numbers of samples in the ranges from $0.1g$ to $0.25g$, $0.5g$, $0.75g$, and $1g$.

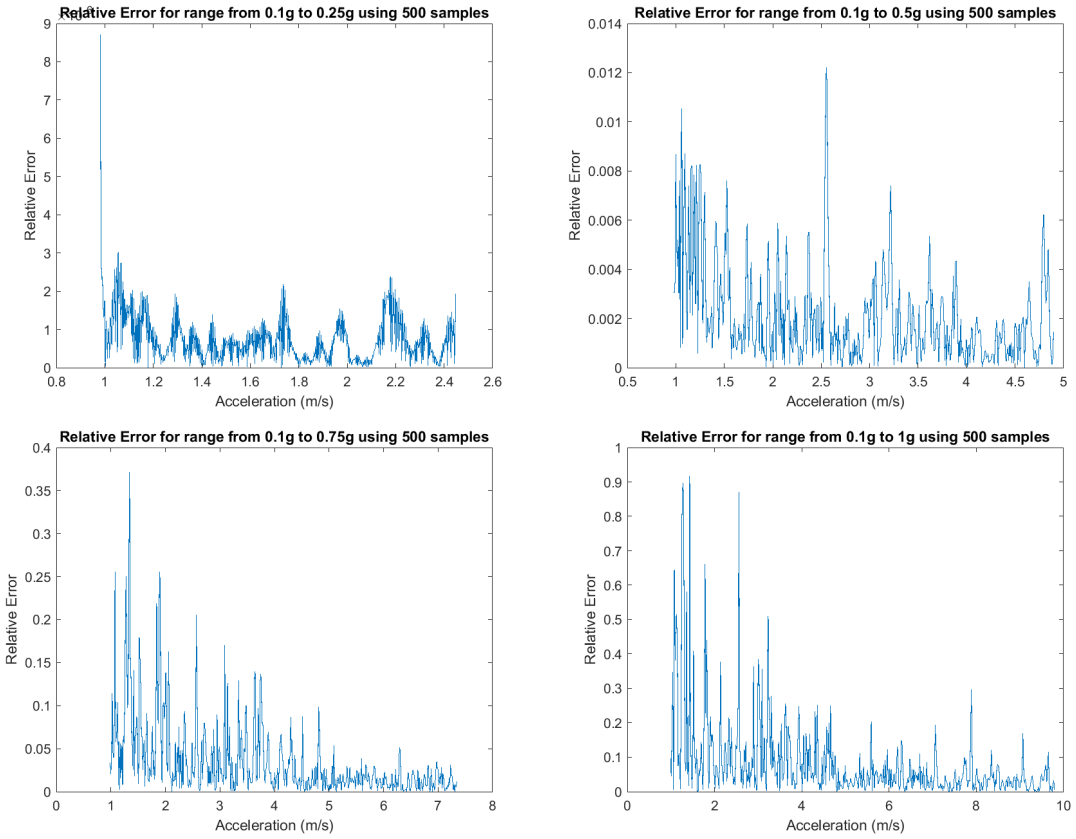

Figure 14: Graphs of relative error in various ranges ($0.1g$ to $0.25g$, $0.5g$, $0.75g$, and g) using 500 samples.

Range (g)	**# of Samples**	**Mean Relative Error**	**Median of Error Dist.**
[1,1.5]	50	0.075	0.074
	500	1.91e-3	1.07e-3
	5000	7.95e-4	6.24e-4
[1,2]	50	0.083	0.065
	500	0.019	0.012
	5000	9.88e-3	7.39e-3

Table 2: The mean of the relative error and median of the distribution of errors for neural network trained on different numbers of samples for the ranges from $1g$ to $1.5g$ and $2g$.

Finally we looked at still larger ranges of acceleration from $1g$ to $5g$ and $10g$, as well as from $2g$ to $6g$ and $10g$. We used 5000, 10000, and 50000 samples for these ranges as they were significantly larger than the previous ranges. Examples of the relative error as a function of acceleration and histograms of the distribution of relative errors are found in Fig. 15. Histograms of the error distribution are included here to give a clearer portrayal of the relative error than would be gleaned from the graphs of relative error as a function of acceleration. Relative error and median of the error distribution for these ranges and numbers of samples are found in Table 3.

From Table 3 we see that neither the mean nor median of the error appreciably or consistently decreases as the number of samples increases, even by an order of magnitude (5000 to 50000). This shows that the accuracy can only be improved to a certain point by increasing the number of samples that the neural network uses to train and further improvements to the accuracy would have to be made by changing some other parameter. As well, increasing the end of the range from $5g$ to $10g$ or $6g$ to $10g$ does not increase the mean relative error by a substantial amount. Comparing the ranges that start at $1g$ to those that start at $2g$, we see that the mean error is smaller by between 0.01 and 0.02 when the range starts at $2g$, even in the cases where the size of the ranges being compared are the same.

Range (g)	# of Samples	Mean Relative Error	Median of Error Dist.
[1,5]	5000	0.047	0.031
	10000	0.044	0.029
	50000	0.048	0.032
[1,10]	5000	0.058	0.033
	10000	0.054	0.029
	50000	0.049	0.025
[2,6]	5000	0.032	0.023
	10000	0.033	0.025
	50000	0.032	0.023
[2,10]	5000	0.038	0.024
	10000	0.035	0.022
	50000	0.036	0.029

Table 3: The mean of the relative error and median of the distribution of errors for neural network trained on different numbers of samples in for the ranges from $1g$ to $5g$ and $10g$, and the ranges from $2g$ to $6g$ and $10g$.

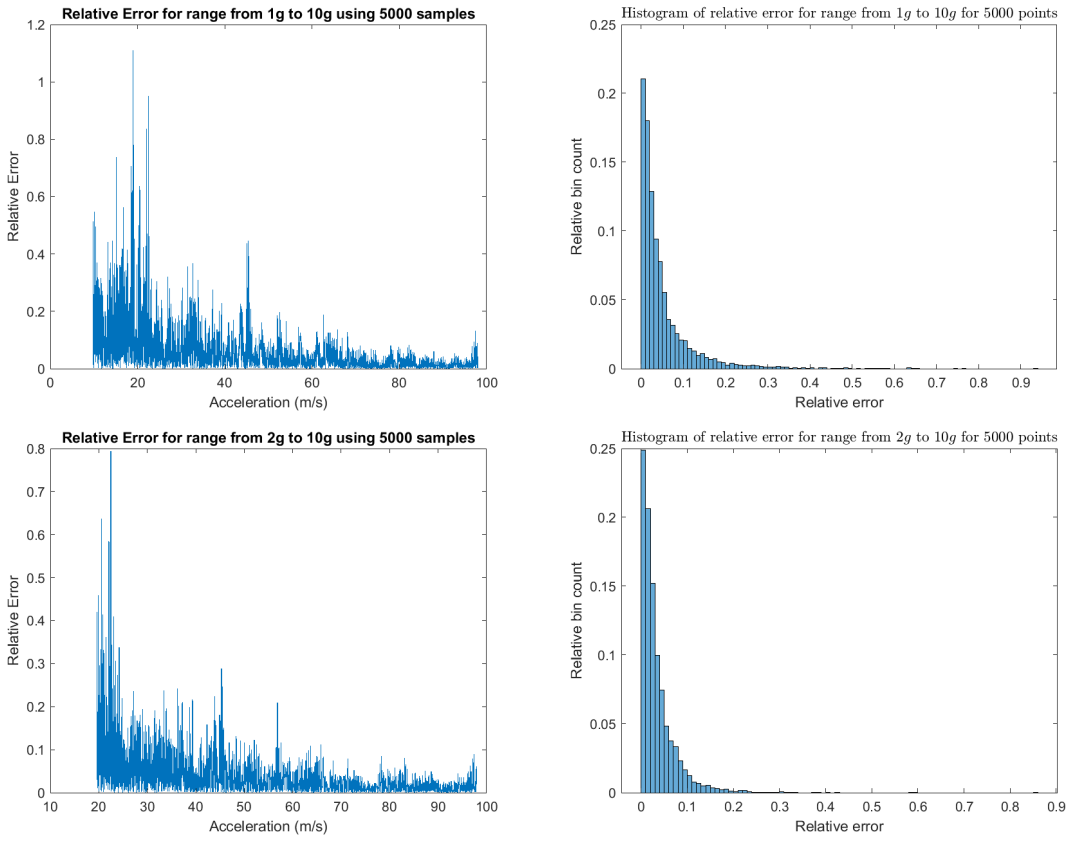

Figure 15: Graphs of relative error for the ranges $1g$ to $10g$ and $2g$ to $10g$ and their respective histograms of error distribution.

3 Conclusions & Future Work

We looked at an optical lattice accelerometer with moving potentials. We calculated the accelerometer's output by numerically solving Schrodinger equation for different value of the acceleration and used neural networks to infer the value of the acceleration from the output data. Our results show that this type of analysis works well for various ranges of acceleration for large numbers of samples. For large ranges, the mean error does not improve appreciably for larger numbers of samples, and so improved accuracy would need to be found by changing some other parameters. As well, as the range was increased through larger multiples of g (i.e. increasing the range $2g - 6g$ to $2g - 10g$) the mean of the error does not change by an substantial amount, and so this method shows promise for increased ranges of acceleration.

The present analysis was done using the theoretical model of infinite plane waves. Future analyses could look at how a wave packet behaves under the potential described in this project or a similar one, if it still follows the expected pattern, and how feasible an analysis using neural networks is in the finite case. As well, our analysis had decreased accuracy when the start of the range was closer to zero, and so for cases where the operational range of accelerations is close to zero, another method may need to be found. Further, for large accelerations, the effects of aliasing in simulations become more prominent unless large numbers of harmonics are used, which then increases computation time beyond the time constraints of this project. Some other way to account for this without having to increase the number of harmonics would help limit increases to computation time.

Appendices

A Dimensionless Variables

When working with computational simulations, it is useful to work in unitless quantities. So we introduce the unitless transformations

$$\begin{aligned} \eta &= 2k_L x \\ \tau &= 8\omega_R t \end{aligned} \tag{49}$$

where $\omega_R = \frac{E_R}{\hbar}$ is the recoil frequency of the lattice and E_R is the recoil energy [18]. The derivative transformations of these are

$$\begin{aligned} \frac{\partial}{\partial x} &= \frac{\partial \eta}{\partial x}\frac{\partial}{\partial \eta} = 2k_L \frac{\partial}{\partial \eta} \\ \frac{\partial}{\partial t} &= \frac{\partial \tau}{\partial x}\frac{\partial}{\partial \tau} = 8\omega_R \frac{\partial}{\partial \tau}. \end{aligned} \tag{50}$$

Additionally, we introduce another unitless variable

$$\alpha = \frac{V_0}{32E_R}. \tag{51}$$

From these transformations, the time-dependent Schrödinger equation can be written in a unitless form

$$\begin{aligned} & i\hbar\frac{\partial}{\partial t}\psi = -\frac{\hbar^2}{2m}\left(\frac{\partial}{\partial x}\right)^2\psi + \frac{V_0}{2}\cos(2k_L x)\psi \\ \implies & i\hbar\frac{\partial \tau}{\partial t}\frac{\partial}{\partial \tau}\psi = -\frac{\hbar^2}{2m}\left(\frac{\partial \eta}{\partial x}\frac{\partial}{\partial \eta}\right)^2\psi + \frac{V_0}{2}\cos(\eta)\psi \\ \implies & i8E_R\frac{\partial \psi}{\partial \tau} = \frac{-4k_L^2\hbar^2}{2m}\frac{\partial^2\psi}{\partial \eta^2} + \frac{V_0}{2}\cos(\eta)\psi \\ \implies & i\frac{\partial \psi}{\partial \tau} = \frac{-k_L^2\hbar^2}{4mE_R}\frac{\partial^2\psi}{\partial \eta^2} + \frac{V_0}{16E_R}\cos(\eta)\psi \\ \implies & i\frac{\partial \psi}{\partial \tau} = \frac{-k_L^2\hbar^2}{4m}\frac{2m}{\hbar^2 k_L^2}\frac{\partial^2\psi}{\partial \eta^2} + 2\alpha\cos(\eta)\psi \\ \implies & i\frac{\partial \psi}{\partial \tau} = \frac{-1}{2}\frac{\partial^2\psi}{\partial \eta^2} + 2\alpha\cos(\eta)\psi. \end{aligned} \tag{52}$$

References

[1] E. Hecht, *Optics*. Pearson Education, Inc., 2002.

[2] J. Monnier, "Optical interferometry in astronomy," *Rep. Prog. Phys.*, vol. 66, no. 5, p. 789, 2003.

[3] R. Youngquist, "Optical distortion evaluation in large area windows using interferometry," *14th Int. Symp. on Nondestructive Characterization of Materials*, 2015.

[4] S. Ezekiel, "Optical gyroscope options: Principles and challenges," *Optical Fiber Sensors*, 2006.

[5] B. A. et al., "Observation of gravitational waves from a binary black hole merger," *Phys. Rev. Lett.*, vol. 116, p. 061102, Feb 2016.

[6] H. Young and R. Freedman, *University Physics*. Pearson Education, 2019.

[7] D. W. Keith, C. R. Ekstrom, Q. A. Turchette, and D. E. Pritchard, "An interferometer for atoms," *Phys. Rev. Lett.*, vol. 66, pp. 2693–2696, May 1991.

[8] M. H. Anderson, J. R. Ensher, M. R. Matthews, C. E. Wieman, and E. A. Cornell, "Observation of bose-einstein condensation in a dilute atomic vapor," *Science*, vol. 269, no. 5221, pp. 198–201, 1995.

[9] "The nobel prize in physics 2001." https://www.nobelprize.org/prizes/physics/2001/press-release/.

[10] J. Hu, A. Urvoy, Z. Vendeiro, V. Crépel, W. Chen, and V. Vuletić, "Creation of a bose-condensed gas of 87rb by laser cooling," *Science*, vol. 358, no. 6366, pp. 1078–1080, 2017.

[11] A. e. a. Peters, "Measurement of gravitational acceleration by dropping atoms.," *Nature*, vol. 400, pp. 849–852, 1999.

[12] Y. e. a. Shin, "Interference of bose-einstein condensates split with an atom chip," *Phys. Rev. A*, vol. 72, no. 2, 2005.

[13] O. Carnal and J. Mlynek, "Young's double-slit experiment with atoms: A simple atom interferometer," *Phys. Rev. Lett.*, vol. 66, pp. 2689–2692, May 1991.

[14] M. Kasevich and S. Chu, "Atomic interferometry using stimulated raman transitions," *Phys. Rev. Lett.*, vol. 67, pp. 181–184, Jul 1991.

[15] J. B. Fixler, G. T. Foster, J. M. McGuirk, and M. A. Kasevich, "Atom interferometer measurement of the newtonian constant of gravity," *Science*, vol. 315, no. 5808, pp. 74–77, 2007.

[16] C. R. Ekstrom, J. Schmiedmayer, M. S. Chapman, T. D. Hammond, and D. E. Pritchard, "Measurement of the electric polarizability of sodium with an atom interferometer," *Phys. Rev. A*, vol. 51, pp. 3883–3888, May 1995.

[17] A. Wicht, J. M. Hensley, E. Sarajlic, and S. Chu, "A preliminary measurement of the fine structure constant based on atom interferometry," *Physica Scripta*, vol. T102, no. 1, p. 82, 2002.

[18] C. Weidner, *Shaken Lattice Interferometry*. The University of Colorado at Boulder, 2018.

[19] G.-B. Jo, Y. Shin, S. Will, T. A. Pasquini, M. Saba, W. Ketterle, D. E. Pritchard, M. Vengalattore, and M. Prentiss, "Long phase coherence time and number squeezing of two bose-einstein condensates on an atom chip," *Phys. Rev. Lett.*, vol. 98, p. 030407, Jan 2007.

[20] Y.-J. Wang, D. Z. Anderson, V. M. Bright, E. A. Cornell, Q. Diot, T. Kishimoto, M. Prentiss, R. A. Saravanan, S. R. Segal, and S. Wu, "Atom michelson interferometer on a chip using a bose-einstein condensate," *Phys. Rev. Lett.*, vol. 94, p. 090405, Mar 2005.

[21] M. Horikoshi and K. Nakagawa, "Dephasing due to atom-atom interaction in a waveguide interferometer using a bose-einstein condensate," *Physical Review A*, vol. 74, p. 031602, 09 2006.

[22] C. A. Weidner, H. Yu, R. Kosloff, and D. Z. Anderson, "Atom interferometry using a shaken optical lattice," *Phys. Rev. A*, vol. 95, p. 043624, Apr 2017.

[23] C. A. Weidner and D. Z. Anderson, "Experimental demonstration of shaken-lattice interferometry," *Phys. Rev. Lett.*, vol. 120, p. 263201, Jun 2018.

[24] C. A. Weidner and D. Z. Anderson, "Simplified landscapes for optimization of shaken lattice interferometry," *New Journal of Physics*, vol. 20, p. 075007, jul 2018.

[25] C. Pethick and H. Smith, *Bose-Einstein Condensation in Dilute Gases*. Cambridge University Press, 2008.

[26] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science*, vol. 349, no. 6245, pp. 255–260, 2015.

[27] C. C. Aggarwal, *An Introduction to Neural Networks*, pp. 1–52. Cham: Springer International Publishing, 2018.

[28] M. McConnell, *On the numerical solution of selected integrable nonlinear wave equations*. The University of Reading, 2002.

[29] N. Ashcroft and N. Mermin, *Solid State Physics*. Saunders College Publishing, 1976.