Optimizing persistent currents in a ring-shaped Bose-Einstein condensate using machine learning

Simeon Simjanovski , Guillaume Gauthier , Matthew J. Davis, Halina Rubinsztein-Dunlop, and Tyler W. Neely

ARC Centre of Excellence for Engineered Quantum Systems, School of Mathematics and Physics, University of Queensland, Saint Lucia, QLD 4072, Australia

We demonstrate a method for generating persistent currents in Bose-Einstein condensates by using a Gaussian process learner to experimentally control the stirring of the superfluid. The learner optimizes four different outcomes of the stirring process: (O.I) targeting and (O.II) maximization of the persistent current winding number and (O.III) targeting and (O.IV) maximization with time constraints. The learner optimizations are determined based on the achieved winding number and the number of spurious vortices introduced by stirring. We find that the learner is successful in optimizing the stirring protocols, although the optimal stirring profiles vary significantly depending strongly on the choice of cost function and scenario. These results suggest that stirring is robust and persistent currents can be reliably generated through a variety of stirring approaches.

I. INTRODUCTION AND BACKGROUND

In a Bose-Einstein condensate (BEC) superfluid, a persistent current (PC) is the quantized circulation of the bulk fluid in a multiply connected geometry, such as a ring [1,2]. As the phase singularity associated with the PC is located at the center of the ring, where there is no density, the flow is stabilized by topology against dissipation [1,3]. While PCs are of fundamental interest in BEC research, they are also relevant to atomtronics [4,5], where they can form the basis of compact matter-wave interferometers [6–8]. PCs are also the basis of fluxon systems [9] that have potential applications in atomtronics [10], for example as qubits which can be physically relocated [11].

Multiple methods for experimentally generating PCs have been developed. One can imprint the desired currents directly to the wave-function phase using an optical potential [12,13], inducing superfluid flow proportional to the phase gradient [1]. However, in doing so the phase profile of the condensate acquires a sharp $2N\pi$ phase jump for a charge-N current at some point along the ring leading to significant density excitations [13,14]. Alternatively, orbital angular momentum can be directly transferred from Laguerre-Gaussian laser modes [15,16], to the condensate, although the typical efficiency is limited by mode matching to around 50% [17]. A more straightforward method is to stir the condensate using an optical barrier, analogous to stirring a classical fluid to generate flow [18,19]. In this paper, we explore the parameter space of a stirring method, with an aim to determine optimal stirring profiles with a machine learning approach. Determining the best-case stirring scenarios will highlight the advantages and disadvantages of this approach compared to the other schemes.

Machine learning has increasingly been used for the experimental control of quantum systems [20–22]. The key advantage of this approach is the lack of prior knowledge a machine learner has on the system of interest, allowing for unique or counterintuitive solutions to a given problem

to be found [21–24]. These solutions can be found even when there is an incomplete physical model of the system available, since the learner requires only a restricted set of data points [25]. Furthermore, computers have the ability to handle large amounts of data efficiently as compared to human operators. These factors have established machine learning as a potentially superior method of experimental control [26].

The stirring of a BEC to generate a PC is complicated by a variety of experimental realities that can be challenging or inefficient to model. For example, there may be density variations due to trap imperfections or roughness that can cause the PC to degrade or decay [27]. The BEC also has a finite lifetime within which stirring, system dynamics, and data extraction must occur, introducing time restrictions to the stirring process [28]. Aggressive stirring of the system will likely result in the generation of undesirable excitations such as quantum vortices and phonons [29]. These aspects suggest the stirring can be a complex process covering a large parameter space that may be challenging to accurately describe and control. Thus, under a set of desired constraints, optimal stirring parameters for this complex system are not *a priori* obvious. For example, various approaches have moved the barrier at constant speed [29,30] or accelerating [31] speed profiles, but it is unclear which approach is superior. The stirring barrier must also be removed after inducing flow at some optimal rate. Finally, one might be interested in creating the current in the shortest time possible, to minimize the effect of atom loss due to a finite BEC lifetime. This large parameter space motivates the use of machine learning approaches to optimize the stirring process, subject to user-chosen sets of constraints. Recently, machine learning optimization of persistent currents in a ring lattice has been explored in numerical simulations [32], but such approaches have not yet been applied to experiments.

Here, we use reinforcement learning through the open-source Machine-Learning Online Optimization Package (M-LOOP) [33] to explore the stirring parameter space. Using the
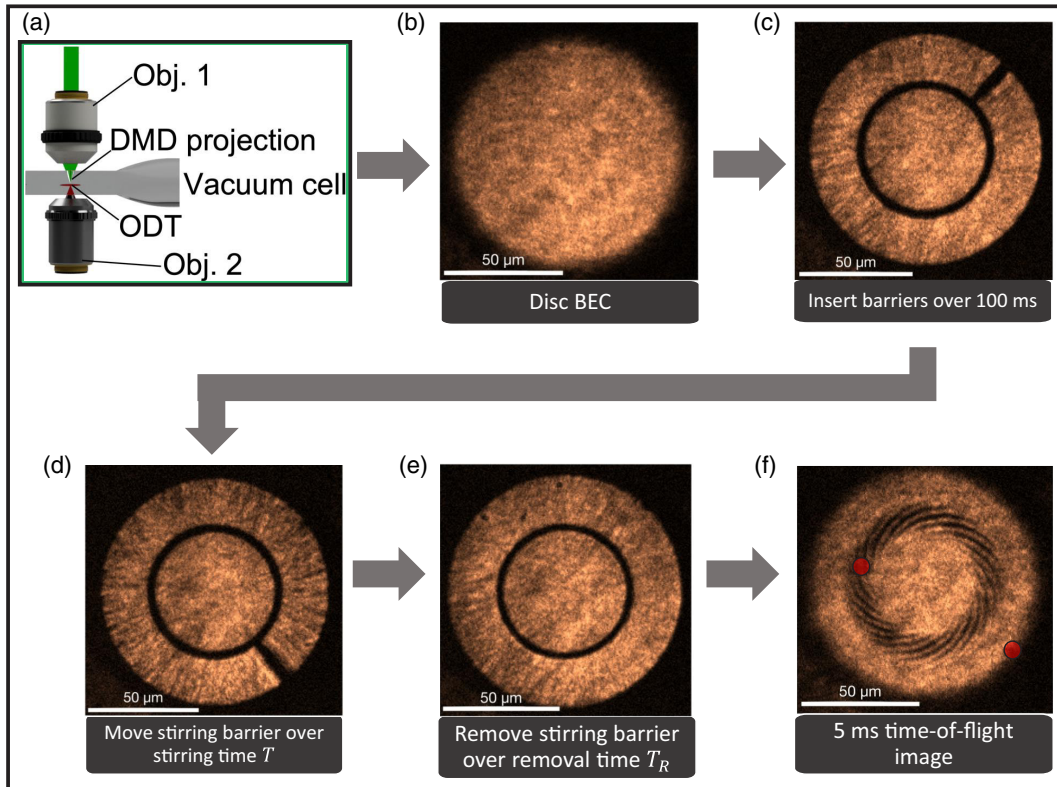
FIG. 1. (a) A simplified diagram of the main apparatus used to generate and control BECs. The light from the digital micromirror device (DMD) is shown by the green beam along the projection (Obj. 1) and imaging (Obj. 2) objectives. The light used for Faraday imaging propagates along the same direction. (b–e) Experimental images demonstrating the stirring process using the DMD projected light. The ring trap shown here has a larger width than the subsequent optimizations for illustrative clarity. (f) A typical time-of-flight experiment using stirring as described in Sec. II, with red dots indicating spurious vortices.

Gaussian process learner in M-LOOP and applying it to four separate optimization conditions, we find that the learner is able to successfully optimize the stirring process, generating the desired PCs while minimizing spurious vortices. However, we find that the optimized parameters vary widely depending on the chosen cost function and target parameters. The results suggest that stirring is robust against variations in the particular details of the stirring profile.

II. EXPERIMENTAL IMPLEMENTATION

Our experiment implements configured optical potentials with blue-detuned repulsive light to both trap a ^{87}Rb BEC in a ring and simultaneously stir it, as shown in Fig. 1. The optical potential for the ring geometry and stirring is generated using a digital micromirror device (DMD). The DMD is an array composed of 2560×1600 mirrors that can be individually switched "on" and "off" corresponding to $\pm 12°$ tilts, respectively. This enables the creation of a "binary" hard-wall trapping potential [34]. A detailed description of the process for creating and configuring BECs using this apparatus can be found in our prior work [34,35].

The experimental sequence is shown in Fig. 1. The atoms are initially trapped in a 100-μm diameter disk. A separation barrier is then inserted, growing from zero to 7-μm width over 100 ms symmetrically about a mean radius of 34 μm.

This insertion method was observed to be sufficiently slow to avoid density excitations. This process results in two separated condensates: an outer ring BEC where the PC is created, and an inner reference BEC that has a uniform phase, and can be interfered with the outer BEC resulting in a pattern from which the PC winding number can be inferred [36]. Simultaneously, a 3-μm-thick stirring barrier is inserted into the ring BEC by linearly increasing the barrier thickness from 0 μm. This barrier is used for the stirring of the outer ring, subject to machine learning controlled parameters (see Sec. III) which aim to generate the desired persistent current. After the stirring process is completed, the stirring barrier is removed by linearly decreasing its thickness over the removal time T_R while maintaining the same angular velocity. The stirring region and reference BEC are then interfered by releasing the system from the trap, and imaged in time of flight (TOF) [13,19,37].

The number of spiral fringes resulting from the interference of the ring and reference condensates identifies the phase winding number of the ring, which is directly proportional to the PC flow speed [37]. Spuriously generated quantum vortices can form due to the stirring and are resolvable as density dips in TOF, as seen in Fig. 1. A single TOF image can thus be used to determine both the winding and vortex number. A detailed description of the image analysis is given in Appendix A.

III. MACHINE LEARNING IMPLEMENTATION

A. Stirrer parametrization

Machine learning control of the stirring is implemented with Gaussian process regression via the freely available package M-LOOP [33,38]. As a reinforcement machine learning method, M-LOOP requires an external measure of the error produced by the parameter model chosen by the learner [25,39]. This error can be quantified using a cost function. The choice of cost function can vary specific to the type of optimization required, but there is one underlying property across all choices: the cost should be minimized when the parameter set is optimized [20,24].

We begin by specifying the angular stirring function of the barrier:

$$\theta(t) = \alpha \left(\frac{t}{A} \right)^P, \tag{1}$$

where the constant $A = 400$ ms is introduced to ensure the stirring is slow enough to prevent destroying the BEC. The parameters controlled by the machine learner are as follows: (1) *stirring time* of the barrier after insertion, T; (2) *removal time* of the barrier after the stirring time, T_R; (3) *stirring exponent*, P; and (4) *angular coefficient*, α. Each parameter is given a lower and upper bound: (1) $100 \leqslant T \leqslant 1500$ ms, (2) $10 \leqslant T_R \leqslant 450$ ms, (3) $1 \leqslant P \leqslant 4.5$ (dimensionless), and (4) 1.75×10^{-2} rad $\leqslant \alpha \leqslant \alpha_{max}$, where $\alpha_{max} = 3.49$ rad in the case of optimizations considered in Sec. III and $\alpha_{max} = 10.5$ rad for those considered in Sec. V. This was done to restrict the parameter space initially (Sec. IV) for a simpler optimizations while increasing complexity with a larger parameter space in the later optimizations (Sec. V) to test the limitations of the learner employed. The lower bound on T is selected so that even for the minimum stirring time sufficient fringes are generated for the detection algorithm to operate accurately [40] (see also Appendix A). The upper bound is imposed due to the frame-rate constraints on the DMD, ensuring that the sequence projected onto the BEC is smooth when the stirring occurs, avoiding the stirring barrier moving around the ring in large discrete jumps. Bounds on T_R are also chosen such that the removal time is comparable to the insertion time. The ranges for α and P were chosen by observation to ensure parameters that both were sufficient to consistently generate observable spiral fringes and did not result in aggressive stirring such that the BEC was destroyed.

B. Cost functions

Two main optimizations are considered: targeting a specific winding number, with the cost function

$$\mathcal{C}_{\mathrm{T}} = |N_T - N_W| + N_V + 10 \frac{T}{T_{max}}, \tag{2}$$

and maximizing the winding number, with the cost function

$$\mathcal{C}_{\mathrm{M}} = -N_W + N_V + 10 \frac{T}{T_{max}}. \tag{3}$$

Here, the measured winding number N_W and the measured spurious vortex number N_V are obtained from the image analysis (Appendix A). The target winding number N_T is chosen prior to optimization. The maximum allowed stirring time T_{max} is introduced for optimizations O.III and O.IV where we impose a time restriction for the stirring process (Sec. V). This time restriction on stirring is made so that any experiments which require both stirring and observation of subsequent dynamics of the BEC system can be made within the BEC lifetime, ≈ 20 s for our system. By minimizing the stirring time, we make more time available for these subsequent dynamics to be observed. Otherwise, we set $T_{max} \to \infty$ such that the last term in the cost functions vanishes. Cost is computed using an average measure of cost for five instances of the experiment. The standard error of this cost measurement is used as the uncertainty estimate for the cost. This uncertainty is not incorporated into the cost itself, but is fed directly to M-LOOP alongside the cost. M-LOOP is then able to estimate errors on the parameter landscapes [e.g., Fig. 3(b)] using these input uncertainties.

While vortices are the primary spurious effects we are interested in minimizing, density excitations (sound) may also result depending on the stirring profile [41]. This information is not readily extracted from the TOF images, although it may affect the regularity of the fringe pattern. Minimizing sound excitation may be an additional refinement in future implementations of this approach.

Equation (2) is minimized when the winding number is exactly N_T and no vortices are present in the system, resulting in optimization to the target. Equation (3) is minimized for the largest possible winding number given a minimized vortex number, thus maximizing the PC. The weights on each term were chosen such that any spurious vortices are weighted equally to the error in the winding for Eq. (2) or to how large the winding becomes in Eq. (3). This ensures that the best "score" is assigned not only in achieving the target or maximizing the current, but also when spurious vortices are removed. When including the stirring time penalty, the factor of 10 ensures that that longer stirring times are heavily punished, to the same level of having ten spurious vortices for every stirring time T near the upper limit T_{max}.

IV. OPTIMIZATIONS O.I AND O.II: VARIABLE STIRRING TIME

We begin our investigation by fixing the stirring exponent $P = 2$ (constant acceleration) and considering $T_{max} \to \infty$ to remove the stirring time penalty from Eqs. (2) and (3), meaning the stirring time is limited only by its upper bound.

A. Optimization O.I: Targeting

Using Eq. (2) as a cost function aims to realize a target winding number. This case is desirable for experiments investigating various initial conditions of a superfluid shear layer such as with the superfluid Kelvin-Helmholtz instability [42,43] or when fluxons [9,10] are experimentally initialized by stirring. A target winding number of $N_T = 20$ is chosen. This choice was made since a winding of 20 was already known to be achievable in the experiment though manually optimized stirring. The general process for optimization is
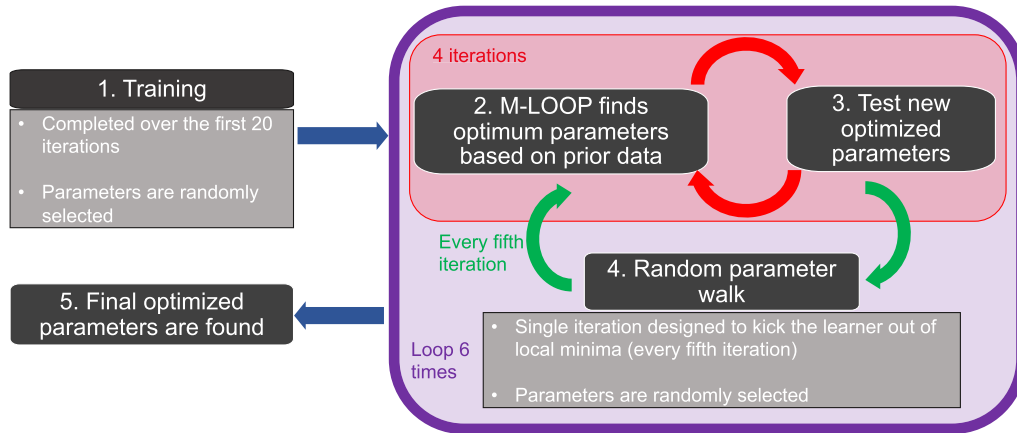
FIG. 2. Process flow chart showing the general stirring parameter optimization process using M-LOOP over 50 iterations. Each iteration is in reality composed of five experimental runs, taken with the same parameter settings. The cost of each iteration is computed as the average cost for these five runs. This results in a mean cost and uncertainty (standard error) returned to M-LOOP for each parameter setting.

described in Fig. 2. After 50 iterations, the learner was terminated, and the results of the procedure can be seen in Fig. 3.

The cost versus run number is shown in Fig. 3(a) and illustrates how the learner evolved in choosing parameters and their effectiveness in experimentally acquiring a circulation of $N_T = 20$. We note that in this scheme it is difficult to discern between large cost values due to the measured winding number being vastly different to the target or whether it is due to many spurious vortices. Generally, spurious vortices vary greatly shot to shot and so large cost values with larger uncertainties can be associated with large spurious vortex numbers. In contrast, the winding will be relatively stable shot to shot and data points associated with these defects will have lower uncertainty. The blue curve superimposed on the data tracks the running minimum cost. As can be seen, convergence is quickly achieved, justifying the cutoff run number of 50 since we see no significant improvement in the cost or parameters. With these parameters, the cost landscape forms a four-dimensional region. Landscape curves, shown in Fig. 3(b), are computed as the cross sections of the cost landscape for each of the three parameters, chosen such that the other two parameters coincide with experimentally minimal costs. The optimum stirring parameters can be read directly from the global minima of the curves.

Figure 3(c) shows the results of stirring under the learner-proposed optimum parameters. For this particular image, no spurious vortices were observed while the winding count is $N_W = 19$. Averaging over five runs under these optimum parameters, it was found that the mean winding was $\overline{N}_W = 19.6 \pm 0.3$ with $\overline{N}_V = 0.2 \pm 0.2$ spurious vortices. The optimal stirring time T is 760.6 ms, near the middle of its bound range. The optimal removal time $T_R = 176.2$ ms is relatively short, near its lower bound. The coefficient parameter α exhibits multiple minima but has one clear global minimum at 0.786 rad. Using these, it is possible to also deduce the final speed of the barrier from Eq. (1) to be $v_B = 8.9 \times 10^{-3}$ rad/ms or equivalently $v_B = 0.38$ mm/s at the inner radius of 32.25 μm. The optimum parameters, along with the parameters for the subsequent optimizations, are summarized in Table I.

B. Optimization O.II: Maximization

Optimization O.II seeks to maximize the winding number while suppressing spurious vortices. This can be achieved by running M-LOOP with the cost function in Eq. (3). Figure 4 shows the outcome of this optimization. There is immediately a clear and steady decrease in cost after the final training run, leading to a plateau in the cost for the final runs indicated by the minimum cost curve in Fig. 4(a). This behavior suggests convergence to an optimal parameter set.

The parameter landscapes show the optimized parameter choices as clear global minima of the curves. Unlike the targeting optimization, all curves contain a single minima. The optimal stirring protocol applies aggressive movement of the barrier, indicated by the large $\alpha = 2.36$ rad, but for a relatively short stirring time at $T = 335.3$ ms and removal time $T_R = 15.2$ ms. These parameters correspond to a final barrier speed of $v_B = 1.0 \times 10^{-2}$ rad/ms or $v_B = 0.44$ mm/s at a radius of 32.25 μm. An example TOF image using the optimized parameters is shown in Fig. 4(c), which contains 25 windings. The parameters result in $\overline{N}_W = 25.8 \pm 0.4$, with $\overline{N}_V = 0.2 \pm 0.2$ over five experimental runs.

C. Remarks on the optimizations

Considering first Figs. 3(a) and 4(a), we observe that during training runs the computed cost and uncertainty vary significantly from run to run. The concept behind this random

TABLE I. Summary of optimized parameter values for the optimization cases O.I (targeting), O.II (maximization), O.III (time-limited targeting), and O.IV (time-limited maximization). v is the PC flow speed in the middle of the ring, while $c_s \approx 1300$ μm/s is the speed of sound. The final speed of the barrier is indicated by v_B.

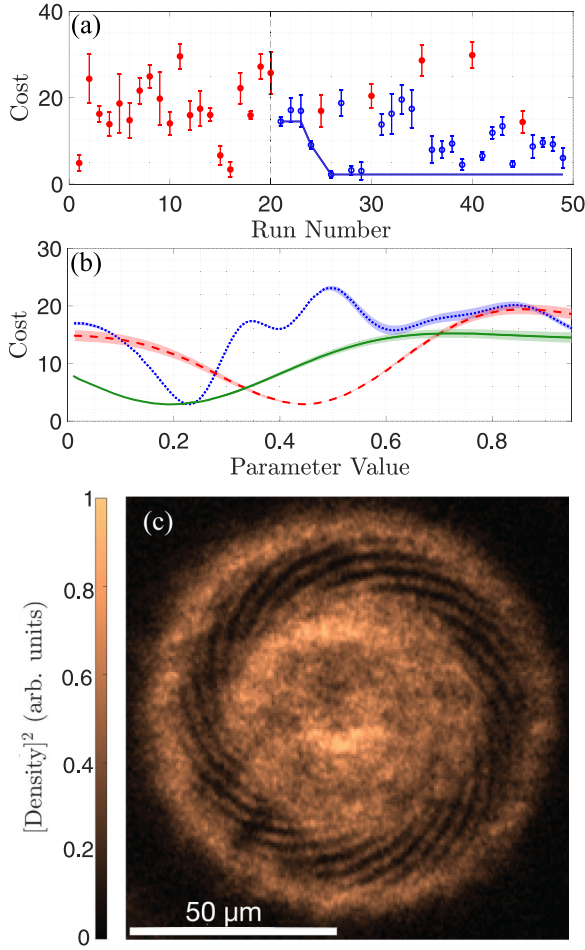| Optimization	T (ms)	T_R (ms)	α (rad)	P	v/c_s	v_B/c_s
O.I	721	176	0.786	2	0.27	0.38
O.II	335	15.2	2.36	2	0.35	0.34
O.III	240	100	2.64	1	0.25	0.22
O.IV	164	100	8.22	3.5	0.25	0.84

FIG. 3. M-LOOP output graphics for optimization to a target of $N_T = 20$ windings [optimization O.I: cost function given by Eq. (2)]. (a) Cost vs run number over 50 runs. The initial 20 random runs are shown before the vertical dashed line. Filled (red) points represent the randomly chosen parameter runs, and open (blue) points represent the runs with machine learner chosen parameters. Filled (red) points following the training can also be observed, which also represent random walks designed to drive the learner out of potential local minima while testing the optimum parameters. The solid (blue) line tracks the minimum cost achieved as a function of run number. (b) The predicted landscape for each of the three parameters computed using the costs. The dashed (red) line represents the stirring time parameter T, the solid (green) line represents the removal time T_R, and the dotted (blue) line represents the angular coefficient α. Parameter values are normalized to lower and upper (P_{min} and P_{max}) bounds using $(P - P_{min})/(P_{max} - P_{min})$. Uncertainties, shown by the shaded regions for each curve, are computed through M-LOOP by feeding the learner an error estimate for each run. (c) The -ms TOF image resulting from using the optimum parameters in stirring. This particular image contains 19 fringes.

training is to walk through as large a portion of the parameter space as possible before the optimization effectively begins. Doing so ensures that the Gaussian process regression is more accurately able to produce meaningful landscape curves [33]. Following the training, we observe a rapid stagnation of the cost in both optimization cases. This is indicated by the minimum cost curves superimposed on Figs. 3(a) and 4(a), which
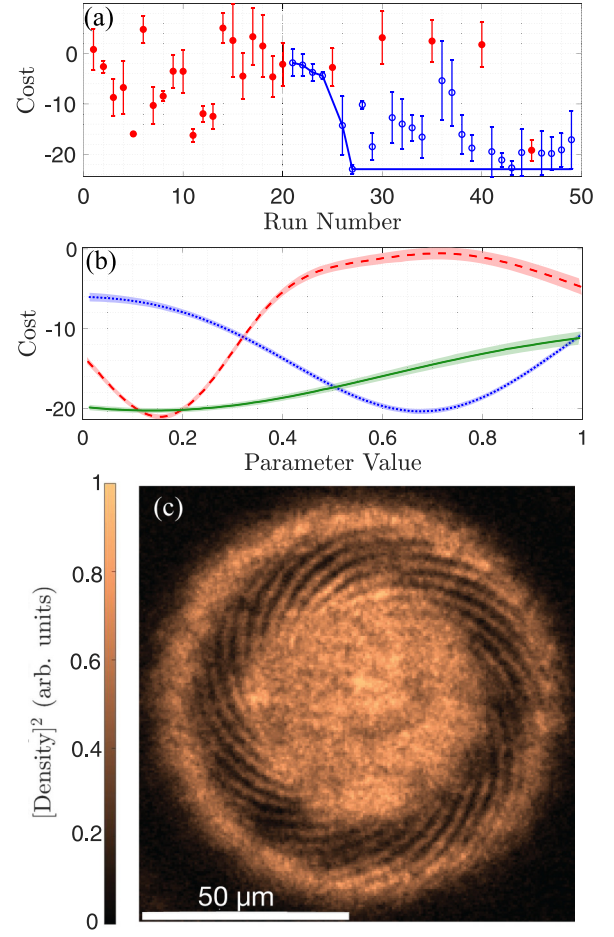

FIG. 4. M-LOOP output graphics for maximization of the winding number [O.II: cost function given by Eq. (3)]. (a) The cost vs the run number over 50 runs. (b) The predicted landscape showing the optimum parameter choices. The dashed (red) line represents the stirring time parameter T, the solid (green) line represents the removal time T_R, and the dotted (blue) line represents the angular coefficient α. Uncertainties are indicated by the shaded regions. (c) 5-ms TOF image under optimized parameters. This particular image shows 25 fringes.

show a drop after training to a sudden plateau suggesting convergence.

For maximization O.II, the optimal parameters can be interpreted as those which reach the upper limit for the winding or flow speed of the persistent current before the shedding of vortices from the stirring barrier cannot be avoided, i.e., the critical velocity has been reached [44]. Applying Eq. (B2) (Appendix B) using the final winding numbers for both the targeting and maximization case gives a flow speed of $v_T(R) \approx 347\ \mu m/s$ and $v_M(R) \approx 457\ \mu m/s$ respectively (at the inner ring radius $R = 32.25\ \mu m$). Equivalently, this gives the critical velocity of the experimental system to be $v_M \approx 0.35 c_s$ where $c_s \approx 1.3\ mm/s$ is the two-dimensional (2D) speed of sound for the system [see Appendix B, Eq. (B3)]. This critical velocity measurement is consistent with previous estimates of vortex shedding due to flow past an obstacle [45–47], suggesting the current is limited by features or roughness in the optical potential and that the machine learner has optimized the stirring process to this limit. We also note that the final

barrier and fluid speeds for O.I do not coincide, as shown in Table I, indicating that the fluid is out of equilibrium with the barrier for this stirring protocol. This behavior is discussed in more detail in Sec. VI.

Overall, the machine learner is able to optimize two different stirring protocols for different final PC conditions. Given the constant angular acceleration used in the stirring profile here, an experimenter might opt to maximize the current via stirring using the same profile as O.I for longer times, given that this is the most straightforward approach. Remarkably, the machine learner finds a different and more efficient approach to stirring: aggressively stirring for a shorter period of time.

V. OPTIMIZATIONS O.III AND O.IV: MINIMUM STIRRING TIME

Motivated by the above results, we next repeat the previous optimizations while attempting to minimize the stirring time. The learner now has to satisfy the desired winding conditions with a cost penalty for long stirring times T. We no longer hold $P = 2$, in Eqs. (3) and (2), instead allowing it to be a parameter optimized by the learner. This extra flexibility is allowed in order to compensate for the added complexity in the cost function, giving the learner more control over the barrier motion. $T_{max} = 1500$ ms is now the maximum allowed stirring time, which is consistent with the upper bound on the stirring time T. Since the cost was observed to vary weakly with the barrier removal time (see Figs. 3 and 4) the removal time is fixed to $T_R = 100$ ms, simplifying the cost functions and aiding convergence.

A. Optimization O.III: Targeting

Figure 5 shows the outcomes from target optimization subject to a time cost imposed by Eq. (2). The stirring time curve in Fig. 5(b) shows a global minimum near 240 ms, which is close to its lower boundary of 100 ms, indicating that only short stirring times are required to achieve the desired winding target. The power parameter also has its global minimum at its lower bound value $P = 1$ and the associated coefficient $\alpha = 2.64$ rad. This results in an instantaneous and constant angular velocity of 6.6×10^{-3} rad/ms, corresponding to 0.28 mm/s at the inner radius of 32.25 μm. Spurious vortices were more prevalent in this optimization, as shown in the example image, Fig. 5(c). Here we also observe 18 windings. Averaging over five separate images, a winding number of $N_W = 18.4 \pm 0.4$ is observed with $N_V = 1.0 \pm 0.3$ spurious vortices.

B. Optimization O.IV: Time-limited maximization

Figure 6 shows the outcome of trying to maximize the winding by employing Eq. (3) with a time cost for M-LOOP. The learner was able to find optimized parameters as indicated by the clear minima in the landscape curves of Fig. 6(b). All global minima lie away from the boundaries of the parameters in this optimization. The stirring is highly nonlinear, with a value of $P = 3.5$. The stirring time was optimized to $T = 163.5$ ms suggesting that the learner correctly identified the cost function and time restriction. Accompanying the
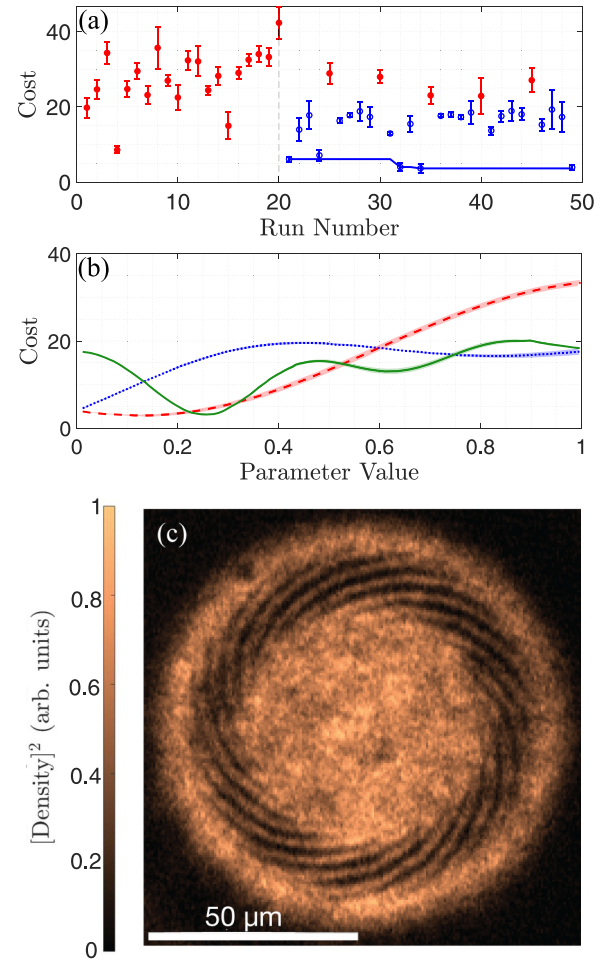

FIG. 5. M-LOOP output graphics for time-limited targeting a winding number of $N_T = 20$ windings [O.III: cost function given by Eq. (2)]. (a) The cost vs the run number over 50 runs. The color convention is consistent with Fig. 3. (b) The predicted landscape showing the optimum parameter choices. The dashed (red) line represents the stirring time T landscape while the dotted (blue) line represents the acceleration coefficient α. Here, the solid (green) line represents the exponent parameter P. Uncertainties are indicated by the shaded regions. (c) 5-ms TOF image under optimized parameters. This particular image shows 18 fringes.

short, nonlinear stirring is a large, $\alpha = 8.22$ rad, coefficient contributing to the highly aggressive stirring. In this case, the final barrier speed is $v_B = 2.5 \times 10^{-2}$ rad/ms or 1.1 mm/s at a radius of 32.25 μm. Under this proposed optimal stirring, the resulting BEC looks similar to Fig. 6(c) which contains five spurious vortices but only 18 windings. An average winding number of $\bar{N}_W = 18.4 \pm 0.4$ with $\bar{N}_V = 4.5 \pm 0.5$ spurious vortices is observed overall. Likely this smaller winding, relative to O.II, is due to the tradeoff placed in the cost function weights between the winding and the stirring time, which are necessary to enable the fast creation of the PC.

C. Remarks on the time-limited optimizations

The time series of the cost versus run number for O.III in Fig. 5(a) shows a small dip followed by stagnation within ten machine learner controlled runs. Both cases exhibit the
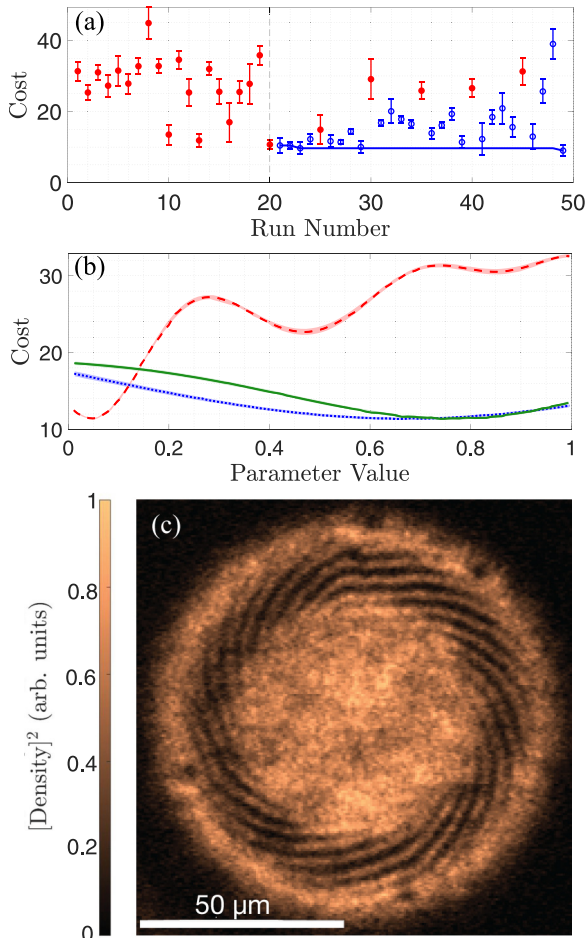
FIG. 6. M-LOOP output graphics for time-limited maximization of the windings number [Optimization IV: cost function given by Eq. (3)]. (a) The cost vs the run number over 50 runs. (b) The predicted landscape shows the optimum parameter choices. The dashed (red) line represents the stirring time T landscape while the dotted (blue) line represents the acceleration coefficient α. Here the solid (green) line represents the exponent parameter P. Uncertainties are indicated by the shaded regions. (c) 5-ms TOF image under optimized parameters. This particular image shows 18 fringes.

characteristic random cost distribution during training, and the effectiveness of said training is highlighted by the smoothness of the parameter landscape curves in Figs. 5(b) and 6(b). In both cases, the parameter landscape curves have minima near the lower boundary of the stirring time T. This suggests that the learner correctly includes the added restriction on the stirring time.

Observing the time series for both O.III and O.IV, Figs. 5(a) and 6(a) respectively, we see the cost versus run number in both cases stagnates rapidly, suggesting fast convergence. However, the convergence in winding number is exactly the same between the two cases. Since both cases share $N_W = 18.4 \pm 0.4$, both have a flow speed of $v(R) \approx 325$ μm/s or $v \approx 0.25c_s$ (at $R = 32.25$ μm). For O.III, this is close to the target. For O.IV however, this flow rate is far from the critical velocity measured in O.II. The inability of the learner to optimize the stirring as desired arises from the fact that the cost function is designed with a

new weighted punishment term for the stirring time, introducing a tradeoff between maximum winding and shortest stirring.

VI. CONCLUSION AND OUTLOOK

The first aim of this paper was to realize optimal stirring parameters for creating PCs with a stirring barrier. The machine learning algorithm was successful in this endeavor, creating PCs for each of the optimizations explored, while minimizing spurious vortices [48]. While our initial hypothesis was that there would be a unique approach to optimal stirring, the results depended greatly on the choice of cost function and the constraints chosen. In the case of variable stirring time without penalty (O.I and O.II), the learner chose dramatically different stirring times, stirring gently for a long time (O.I) and stirring aggressively for a short time (O.II). For the variable time optimizations (O.III and O.IV), the choice of parameters is even more distinct. O.III found the best protocol was to instantaneously accelerate the barrier to move with constant angular velocity, while O.IV found that aggressive stirring with a highly nonlinear profile, $P = 3.5$, was optimal.

These results suggest that while the learner can successfully control stirring for desired results and minimum cost, the problem of determining an optimal stirring profile is underconstrained by the chosen parameters. We take this result to imply that stirring to create a PC is a robust process, and can be successfully achieved with a variety of stirring profiles. From a practical perspective, the relatively counterintuitive approach of aggressive stirring for a short time at constant speed (O.III) may be the best approach, since this minimizes atom loss due to the finite lifetime of the BEC. Another interesting feature of these optimizations is that there are some out-of-equilibrium dynamics between the fluid and barrier. This is seen through the difference in final barrier speeds relative to the superfluid velocity. For example, O.IV resulted in a final barrier speed of $v_B = 0.84c_s$ for a final superfluid speed of $v = 0.25c_s$, but relied on a short stirring time of 164 ms. These results suggest counteradiabatic stirring protocols may be advantageous, provided the barrier beam is removed at the right time. Previous work has explored the possibility of spatial rapid adiabatic passage for engineering angular momentum states in a short time [49,50]. Future extension of similar approaches to stirring of persistent currents may be a fruitful area of both theoretical and experimental future work.

Overall, the successful optimizations demonstrate the suitability of using machine learners for the creation and control of PCs through stirring, albeit with no universal stirring protocol. These results suggest useful future applications of the algorithm. Since PCs establish a basis for atom interferometry [51–53], having high levels of control over PCs using machine learning can be desirable for engineering atom interferometers [54,55]. Beyond interferometry, it is possible to apply this algorithm to prepare PC states with specified winding numbers for studies in quantum turbulence, such as determining the dynamics of the Kelvin-Helmholtz instability [42,43]. Atom inertial sensors have used PCs as a source of synthetic rotation for calibration [56], and this method presents a precise way of creating these references. The general structure of the feedback loop
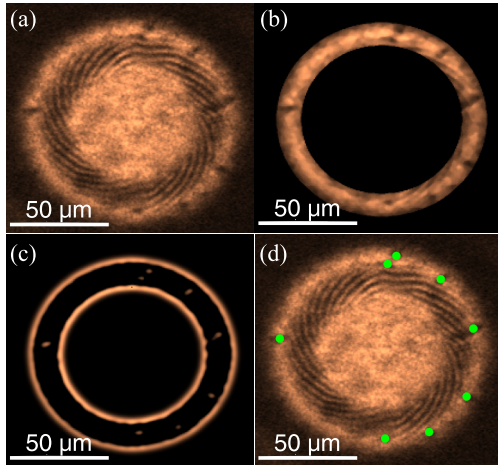
FIG. 7. Processing of a TOF image using the Gaussian blob algorithm. (a) The original image. (b) Masked image to remove the fringe region. The processed image is also smoothed using a Gaussian filter. (c) Laplacian of the masked and filtered image. Thresholding of the image is now possible and the sharp ring edges can be masked out. (d) Detected vortex positions (green solid circles) superimposed on the original image. Additional details can be found in Refs. [59–61]. Note that due to the masking of the fringe region, vortices in this region are not detected despite being clearly present as evident in the warping and forking of the fringes.

described could also be adapted to optimize alternative methods for generating PCs, such as phase imprinting [13].

APPENDIX A: IMAGE PROCESSING

Images of the BEC are produced using Faraday imaging [57,58]. Automatic image processing is used to extract the winding number and vortex number from the Faraday images. Vortices become directly resolvable in TOF as clusters of dark pixels or density dips, which means that detection of such sites can be treated analogously to blob detection in general image processing [59–61]. The use of this algorithm ensures successful detection of vortices outside of the fringe region but tends to count dark fringes as vortex objects as well. To avoid this issue, the algorithm is restricted to operate only outside the fringe region through masking while the fringe region is analyzed using a Fourier signal method discussed in more detail below. Specifically, we restrict the vortex detection to the outer ring where stirring occurs.
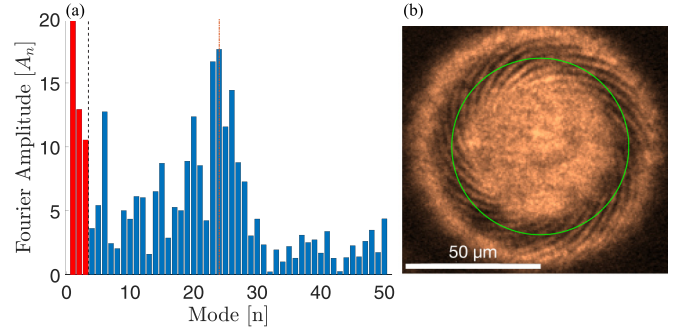

FIG. 8. Example of the processing of a TOF image using the winding counting algorithm. (a) Fourier amplitude for mode numbers 1–50. The largest amplitude mode is taken to be the winding number. Mode numbers at or under the cutoff mode of $n = 3$ are indicated by the black dashed vertical line and are highlighted in red. This cutoff is introduced since these modes have a large amplitude but do not reflect the fringe number. The investigation does not require modes close to $n = 3$ to be detected. The largest amplitude has a mode number of 24, indicated by the orange dot-dashed line. (b) Color coded image highlighting the region of interest (green ring) considered in the example, superimposed on the original image. Note that only one small region of interest is shown here, while in practice 11 similar rings of varying radii are considered. These 11 rings overlap slightly since their inner radii are separated by 0.25 μm while their thickness remains fixed at 1 μm. Over the 11 rings for the example image the mean fringe number is 24.8 ± 0.4 where the uncertainty is estimated as the standard error of the 11 ring measurements. The observed count in this particular image is 24 fringes.

Vortices may still be present in the fringe region and can be observed as either a warping or a fork in a fringe which separates into two fringes, most clearly observable in Fig. 7(a). This results in an increased standard deviation when estimating the winding number as described below. The fringe number standard deviation is simply added to the vortex number N_V when computing the average cost for the set of five runs. Although this potentially overestimates the number of vortices in the fringe region by mixing the standard deviation in achieved winding number into N_V, we find that this approach successfully minimizes spurious vortices in both the masked and interference fringe regions. Note that any spurious vortices in the interior of the reference condensate are ignored by the algorithm since they are not caused by the stirring process.

A Fourier signal algorithm can be used for the counting winding number and its standard deviation for the aforementioned estimate of N_V. We begin by considering the fringe region as indicated by the solid green line in Fig. 8(b). The azimuthal Fourier transform is determined and the mode with the largest Fourier amplitude is identified as corresponding to the winding number for the selected region. A cutoff mode of $n = 3$ is introduced since the signal of these modes is generally large, leading to erroneous detection of winding numbers that do not correspond to the observed spiral pattern. We found this cutoff value sufficient for reducing this effect.

The spiraling nature of the fringes requires some care in considering how to best extract the winding number from the

discrete Fourier transform. The solution is to subdivide the region into 11 overlapping rings of 1-µm thickness. These smaller rings are set apart via their inner and outer radii being translated by 0.25 µm. This choice allows the entire fringe region to be spanned by a series of thin rings over which the accuracy of the detection is improved. The standard deviation of the winding number across the rings is used to estimate the number of vortices in the fringe region. Finally, in order to extract a useful measure of winding number, the mean of the mode number with the largest amplitude is selected which need not be an integer. Figure 8(a) shows the Fourier amplitudes for the considered modes for the small ring region considered in the image in Fig. 8(b).

APPENDIX B: ESTIMATING FLOW SPEED FROM WINDING NUMBERS

It is possible to estimate the speed of the PCs generated by these different stirring processes via the velocity-phase relationship in a supercurrent:

$$\vec{v}(\vec{r}, t) = \frac{\hbar}{m} \nabla \phi(\vec{r}, t). \tag{B1}$$

Assuming that the winding is in the azimuthal direction, $\hat{\theta}$, due to the ring geometry, any radial derivatives of the phase

are zero. The PC should not decay making the speed independent of time. Therefore, the flow speed can be computed via

$$\vec{v}(r) = \frac{\hbar}{m} \frac{1}{r} N_W \hat{\theta}. \tag{B2}$$

The speed of sound in the quasi-2D BEC system can be estimated for comparison using the relation

$$c_s = \sqrt{\frac{\mu_{2D}}{m}}, \tag{B3}$$

where μ_{2D} is the effective 2D chemical potential of the persistent current ring region. μ_{2D} can be deduced from the condensed atom number in the ring $N_c \approx 9.6 \times 10^5$ along with the inner and outer radii of the ring, $r_I = 36$ µm and $r_O = 50$ µm respectively, using the following relation [62]:

$$\mu_{2D} = \frac{2}{5} \left(\frac{3}{2} \frac{g N_c \sqrt{m \omega_z^2}}{\pi \left(r_O^2 - r_I^2 \right)} \right)^{2/3}. \tag{B4}$$

Using a vertical trapping frequency of $\omega_z = 2\pi \times 108$ Hz and the s-wave scattering length $a_s = 98.13 a_0$ to compute the effective 2D interaction parameter $g = 4\pi \hbar^2 a_s/m$, we compute that $\mu_{2D} \approx 2.5 \times 10^{-31}$ J/kg and therefore $c_s \approx 1.3$ mm/s.

[1] A. J. Leggett, Bose-Einstein condensation in the alkali gases: Some fundamental concepts, Rev. Mod. Phys. **73**, 307 (2001).

[2] C. Ryu, M. F. Andersen, P. Clade, V. Natarajan, K. Helmerson, and W. D. Phillips, Observation of persistent flow of a Bose-Einstein condensate in a toroidal trap, Phys. Rev. Lett. **99**, 260401 (2007).

[3] T. W. Neely, A. S. Bradley, E. C. Samson, S. J. Rooney, E. M. Wright, K. J. H. Law, R. Carretero-González, P. G. Kevrekidis, M. J. Davis, and B. P. Anderson, Characteristics of two-dimensional quantum turbulence in a compressible superfluid, Phys. Rev. Lett. **111**, 235301 (2013).

[4] L. Amico, M. Boshier, G. Birkl, A. Minguzzi, C. Miniatura, L.-C. Kwek, D. Aghamalyan, V. Ahufinger, D. Anderson, N. Andrei *et al.*, Roadmap on atomtronics: State of the art and perspective, AVS Quantum Science **3**, 039201 (2021).

[5] L. Amico, D. Anderson, M. Boshier, J.-P. Brantut, L.-C. Kwek, A. Minguzzi, and W. von Klitzing, Colloquium: Atomtronic circuits: From many-body physics to quantum technologies, Rev. Mod. Phys. **94**, 041001 (2022).

[6] C. Ryu, P. W. Blackburn, A. A. Blinova, and M. G. Boshier, Experimental realization of Josephson junctions for an atom SQUID, Phys. Rev. Lett. **111**, 205301 (2013).

[7] P. L. Halkyard, M. P. A. Jones, and S. A. Gardiner, Rotational response of two-component Bose-Einstein condensates in ring traps, Phys. Rev. A **81**, 061602(R) (2010).

[8] J. L. Helm, T. P. Billam, A. Rakonjac, S. L. Cornish, and S. A. Gardiner, Spin-orbit-coupled interferometry with ring-trapped Bose-Einstein condensates, Phys. Rev. Lett. **120**, 063201 (2018).

[9] A. Oliinyk, A. Yakimenko, and B. Malomed, Tunneling of persistent currents in coupled ring-shaped Bose-Einstein condensates, J. Phys. B **52**, 225301 (2019).

[10] A. Oliinyk, B. Malomed, and A. Yakimenko, Nonlinear dynamics of Josephson vortices in merging superfluid rings, Commun. Nonlinear Sci. Numer. Simul. **83**, 105113 (2020).

[11] V. M. Kaurov and A. B. Kuklov, Josephson vortex between two atomic Bose-Einstein condensates, Phys. Rev. A **71**, 011601(R) (2005).

[12] J. Denschlag, J. E. Simsarian, D. L. Feder, C. W. Clark, L. A. Collins, J. Cubizolles, L. Deng, E. W. Hagley, K. Helmerson, W. P. Reinhardt, S. L. Rolston, B. I. Schneider, and W. D. Phillips, Generating solitons by phase engineering of a Bose-Einstein condensate, Science **287**, 97 (2000).

[13] G. Del Pace, K. Xhani, A. Muzi Falconi, M. Fedrizzi, N. Grani, D. Hernandez Rajkov, M. Inguscio, F. Scazza, W. J. Kwon, and G. Roati, Imprinting persistent currents in tunable fermionic rings, Phys. Rev. X **12**, 041037 (2022).

[14] A. Kumar, R. Dubessy, T. Badr, C. De Rossi, M. de Goër de Herve, L. Longchambon, and H. Perrin, Producing superfluid circulation states using phase imprinting, Phys. Rev. A **97**, 043615 (2018).

[15] L. Allen, M. W. Beijersbergen, R. J. C. Spreeuw, and J. P. Woerdman, Orbital angular momentum of light and the transformation of Laguerre-Gaussian laser modes, Phys. Rev. A **45**, 8185 (1992).

[16] S. Beattie, S. Moulder, R. J. Fletcher, and Z. Hadzibabic, Persistent currents in spinor condensates, Phys. Rev. Lett. **110**, 025301 (2013).

[17] M. F. Andersen, C. Ryu, P. Cladé, V. Natarajan, A. Vaziri, K. Helmerson, and W. D. Phillips, Quantized rotation of atoms from photons with orbital angular momentum, Phys. Rev. Lett. **97**, 170406 (2006).

[18] K. C. Wright, R. B. Blakestad, C. J. Lobb, W. D. Phillips, and G. K. Campbell, Driving phase slips in a superfluid atom

circuit with a rotating weak link, Phys. Rev. Lett. **110**, 025302 (2013).

[19] S. Eckel, J. G. Lee, F. Jendrzejewski, N. Murray, C. W. Clark, C. J. Lobb, W. D. Phillips, M. Edwards, and G. K. Campbell, Hysteresis in a quantized superfluid 'atomtronic' circuit, Nature (London) **506**, 200 (2014).

[20] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, Machine learning and the physical sciences, Rev. Mod. Phys. **91**, 045002 (2019).

[21] Q. Zhou, P. Tang, S. Liu, J. Pan, Q. Yan, and S.-C. Zhang, Learning atoms for materials discovery, Proc. Natl. Acad. Sci. USA **115**, E6411 (2018).

[22] A. J. Barker, H. Style, K. Luksch, S. Sunami, D. Garrick, F. Hill, C. J. Foot, and E. Bentine, Applying machine learning optimization methods to the production of a quantum gas, Mach. learn.: sci. technol. **1**, 015007 (2020).

[23] S. Mendelson, A. J. Smola, and J. van Leeuwen, *Advanced Lectures on Machine Learning*, Lecture Notes in Computer Science Vol. 2600 (Springer-Verlag, Berlin, 2003).

[24] D. Jones, M. Schonlau, and W. Welch, Efficient global optimization of expensive black-box functions, J. Global Optim. **13**, 455 (1998).

[25] G. Bonaccorso, *Mastering Machine Learning Algorithms*, 1st ed. (Packt Publishing, Birmingham, United Kingdom, 2018).

[26] I. Nakamura, A. Kanemura, T. Nakaso, R. Yamamoto, and T. Fukuhara, Non-standard trajectories found by machine learning for evaporative cooling of ^{87}Rb atoms, Opt. Express **27**, 20435 (2019).

[27] T. A. Bell, J. A. P. Glidden, L. Humbert, M. W. J. Bromley, S. A. Haine, M. J. Davis, T. W. Neely, M. A. Baker, and H. Rubinsztein-Dunlop, Bose-Einstein condensation in large time-averaged optical ring potentials, New J. Phys. **18**, 035003 (2016).

[28] G. Gauthier, Transport and turbulence in quasi-uniform and versatile Bose-Einstein condensates, Ph.D. thesis, The University of Queensland, 2019.

[29] K. C. Wright, R. B. Blakestad, C. J. Lobb, W. D. Phillips, and G. K. Campbell, Threshold for creating excitations in a stirred superfluid ring, Phys. Rev. A **88**, 063633 (2013).

[30] A. I. Yakimenko, K. O. Isaieva, S. I. Vilchinskii, and E. A. Ostrovskaya, Vortex excitation in a stirred toroidal Bose-Einstein condensate, Phys. Rev. A **91**, 023607 (2015).

[31] N. Murray, M. Krygier, M. Edwards, K. C. Wright, G. K. Campbell, and C. W. Clark, Probing the circulation of ring-shaped Bose-Einstein condensates, Phys. Rev. A **88**, 053615 (2013).

[32] T. Haug, R. Dumke, L.-C. Kwek, C. Miniatura, and L. Amico, Machine-learning engineering of quantum currents, Phys. Rev. Res. **3**, 013034 (2021).

[33] P. B. Wigley, P. J. Everitt, A. van den Hengel, J. W. Bastian, M. A. Sooriyabandara, G. D. McDonald, K. S. Hardman, C. D. Quinlivan, P. Manju, and C. C. N. Kuhn, Fast machine-learning online optimization of ultra-cold-atom experiments, Sci. Rep. **6**, 25890 (2016).

[34] G. Gauthier, I. Lenton, N. M. Parry, M. Baker, M. J. Davis, H. Rubinsztein-Dunlop, and T. W. Neely, Direct imaging of a digital-micromirror device for configurable microscopic optical potentials, Optica **3**, 1136 (2016).

[35] G. Gauthier, T. A. Bell, M. Baker, H. Rubinsztein-Dunlop, and T. W. Neely, Feedforward optimisation of optical trapping potentials for ultracold atoms, in *Proceedings of the 2020 Conference on Lasers and Electro-Optics Pacific Rim (CLEO-PR)* (IEEE, Sydney, NSW, Australia, 2020), pp. 1–2, doi:10.1364/CLEOPR.2020.C8C_4.

[36] M. R. Andrews, C. G. Townsend, H.-J. Miesner, D. S. Durfee, D. M. Kurn, and W. Ketterle, Observation of interference between two Bose condensates, Science **275**, 637 (1997).

[37] S. Eckel, F. Jendrzejewski, A. Kumar, C. J. Lobb, and G. K. Campbell, Interferometric measurement of the current-phase relationship of a superfluid weak link, Phys. Rev. X **4**, 031052 (2014).

[38] P. B. Wigley, Generating and observing soliton dynamics in Bose-Einstein condensates, Ph.D. thesis, Department of Quantum Science, Research School of Physics and Engineering, The Australian National University, 2017.

[39] P. Mehta, M. Bukov, C.-H. Wang, A. G. Day, C. Richardson, C. K. Fisher, and D. J. Schwab, A high-bias, low-variance introduction to machine learning for physicists, Phys. Rep. **810**, 1 (2019).

[40] With very few fringes present, the fringes themselves wrap around the entire interference region. The detection algorithm effectively counts periodicity of density in these regions, meaning that such wrapping obscures the measurement.

[41] H. Kiehn, V. P. Singh, and L. Mathey, Superfluidity of a laser-stirred Bose-Einstein condensate, Phys. Rev. A **105**, 043317 (2022).

[42] A. W. Baggaley and N. G. Parker, Kelvin-Helmholtz instability in a single-component atomic superfluid, Phys. Rev. A **97**, 053608 (2018).

[43] D. Hernandez-Rajkov, N. Grani, F. Scazza, G. Del Pace, W. Kwon, M. Inguscio, K. Xhani, C. Fort, M. Modugno, F. Marino *et al.*, Universality of the superfluid Kelvin-Helmholtz instability by single-vortex tracking, arXiv:2303.12631.

[44] M. T. Reeves, T. P. Billam, B. P. Anderson, and A. S. Bradley, Identifying a superfluid Reynolds number via dynamical similarity, Phys. Rev. Lett. **114**, 155302 (2015).

[45] C. Raman, M. Köhl, R. Onofrio, D. S. Durfee, C. E. Kuklewicz, Z. Hadzibabic, and W. Ketterle, Evidence for a critical velocity in a Bose-Einstein condensed gas, Phys. Rev. Lett. **83**, 2502 (1999).

[46] J. S. Stießberger and W. Zwerger, Critcal velocity of superfluid flow past large obstacles in Bose-Einstein condensates, Phys. Rev. A **62**, 061601(R) (2000).

[47] W. J. Kwon, G. Moon, S. W. Seo, and Y. Shin, Critical velocity for vortex shedding in a Bose-Einstein condensate, Phys. Rev. A **91**, 053615 (2015).

[48] We note that the lower winding numbers in O.III and O.IV result from the strong weighting on the stirring time punishment term in the cost functions. This observation highlights the tradeoffs introduced in the cost functions which are likely the limiting factors preventing more complete optimization in more complicated scenarios.

[49] R. Menchon-Enrich, A. Benseny, V. Ahufinger, A. D. Greentree, T. Busch, and J. Mompart, Spatial adiabatic passage: A review of recent progress, Rep. Prog. Phys. **79**, 074401 (2016).

[50] D. Guéry-Odelin, A. Ruschhaupt, A. Kiely, E. Torrontegui, S. Martínez-Garaot, and J. G. Muga, Shortcuts to adiabatic-

ity: Concepts, methods, and applications, Rev. Mod. Phys. **91**, 045001 (2019).

[51] S. Pandey, H. Mas, G. Drougakis, P. Thekkeppatt, V. Bolpasi, G. Vasilakis, K. Poulios, and W. von Klitzing, Hypersonic Bose-Einstein condensates in accelerator rings, Nature (London) **570**, 205 (2019).

[52] M. Edwards, Atom SQUID, Nat. Phys. **9**, 68 (2013).

[53] C. Gross, T. Zibold, E. Nicklas, J. Estève, and M. K. Oberthaler, Nonlinear atom interferometer surpasses classical precision limit, Nature (London) **464**, 1165 (2010).

[54] S. Levy, E. Lahoud, I. Shomroni, and J. Steinhauer, The a.c. and d.c. Josephson effects in a Bose-Einstein condensate, Nature (London) **449**, 579 (2007).

[55] R. E. Packard and S. Vitale, Principles of superfluid-helium gyroscopes, Phys. Rev. B **46**, 3540 (1992).

[56] C. W. Woffinden, A. J. Groszek, G. Gauthier, B. J. Mommers, M. W. J. Bromley, S. A. Haine, H. Rubinsztein-Dunlop, M. J. Davis, T. W. Neely, and M. Baker, Viability of rotation sensing using phonon interferometry in Bose-Einstein condensates, SciPost Phys. **15**, 128 (2023).

[57] M. Gajdacz, P. L. Pedersen, T. Mørch, A. J. Hilliard, J. Arlt, and J. F. Sherson, Non-destructive Faraday imaging of dynami-

cally controlled ultracold atoms, Rev. Sci. Instrum. **84**, 083105 (2013).

[58] K. E. Wilson, Developing a toolkit for experimental studies of two-dimensional quantum turbulence in Bose-Einstein condensates, Ph.D. thesis, The University of Arizona, 2015.

[59] C. Bauckhage and R. Sifa, k-maxoids clustering, in *Proceedings of the LWA 2015 Workshops: KDML, FGWM, IR, and FGDB, Trier, Germany, October 7-9, 2015, CEUR Workshop Proceedings*, edited by R. Bergmann, S. Görg, and G. Müller (CEUR-WS.org, 2015), Vol. 1458, pp. 133–144.

[60] A. Rakonjac, A. L. Marchant, T. P. Billam, J. L. Helm, M. M. H. Yu, S. A. Gardiner, and S. L. Cornish, Measuring the disorder of vortex lattices in a Bose-Einstein condensate, Phys. Rev. A **93**, 013607 (2016).

[61] G. Gauthier, Creation and dynamics of Onsager vortex clusters, in *Transport and Turbulence in Quasi-Uniform and Versatile Bose-Einstein Condensates* (Springer, New York, 2020), pp. 139–169.

[62] G. Gauthier, M. T. Reeves, X. Yu, A. S. Bradley, M. A. Baker, T. A. Bell, H. Rubinsztein-Dunlop, M. J. Davis, and T. W. Neely, Giant vortex clusters in a two-dimensional quantum fluid, Science **364**, 1264 (2019).