Rapid Communications

Machine learning for achieving Bose-Einstein condensation of thulium atoms

E. T. Davletov,[1,2] V. V. Tsyganok,[1,2,5] V. A. Khlebnikov,[1] D. A. Pershin,[1,3] D. V. Shaykin,[1,2] and A. V. Akimov[1,3,4,*]

[1]*Russian Quantum Center, Business Center "Ural," 100A Novaya street, Skolkovo, Moscow 143025, Russia*

[2]*Moscow Institute of Physics and Technology, Institutskii per. 9, Dolgoprudny, Moscow Region 141701, Russia*

[3]*P. N. Lebedev Institute RAS, Leninsky Prospekt 53, Moscow 119991, Russia*

[4]*Texas A&M University, 4242 TAMU, College Station, Texas 77843, USA*

[5]*National University of Science and Technology, Leninsky Prospekt 4, Moscow 119049, Russia*

(Received 24 March 2020; accepted 17 June 2020; published 14 July 2020)

Bose-Einstein condensation (BEC) is a powerful tool for a wide range of research activities, a large fraction of which is related to quantum simulations. Various problems may benefit from different atomic species, but cooling down novel species interesting for quantum simulations to BEC temperatures requires a substantial amount of optimization and is usually considered to be a difficult experimental task. In this work, we implemented the Bayesian machine learning technique to optimize the evaporative cooling of thulium atoms and achieved BEC in an optical dipole trap operating near 532 nm. The developed approach could be used to cool down other novel atomic species to quantum degeneracy without additional studies of their properties.

DOI: 10.1103/PhysRevA.102.011302

For a number of problems in quantum simulations [1], long-range interactions are of great interest [2]. To address this demand, lanthanide atoms such as erbium and dysprosium were cooled down to quantum degeneracy and successfully used for quantum simulations [3–5]. Thulium also belongs to the lanthanide series of chemical elements, but compared to previously cooled Er and Dy, it has only one hole in the f shell. This leads to a slightly smaller dipole moment in the ground state of $4\mu_B$, but thulium has a simpler level structure and a less dense Fano-Feshbach resonances spectrum, thus simplifying the control of the interactions for this atom [6]. The lower dipole moment could be overcome by using thulium atoms in a 532-nm optical lattice, thus increasing the interaction strength via shorter interwell distance or lattice spacing.

Reaching Bose-Einstein condensation (BEC) is usually done via evaporative cooling [7]. The evaporative sequence needs to be carefully optimized to minimize the loss of atoms while maximizing the phase space density growth. Many studies have been devoted to the elaboration of a better recipe for the evaporation sequence optimization process [8–10]. Most of these techniques require accurate characterization of the trap geometry and loss mechanisms. They also use several simplifications, such as a high and fixed truncation parameter η (i.e., assuming that atoms with energy larger than $\eta k_B T$ are immediately evaporated [11] and η is constant over the evaporation process) or adiabaticity; these simplifications are not always satisfied. Furthermore, these recipes usually do not consider difficulties and/or opportunities that arise in specific cases such as dynamically shaped traps and the existence of dipolar interactions. Hence, most groups use a simple time-consuming stepwise optimization procedure, with a proposed sequential adjustment of parameters controlling the trapping potential at each time step [12–18].

The initial cooling of thulium atoms was realized in a magneto-optical trap (MOT) operating on a narrow-line transition with a wavelength of approximately 530.7 nm and Zeeman slowing at a strong 410-nm transition [19]. An optical dipole trap (ODT) is realized using a 532-nm laser [20]. The narrow transition of approximately 530.7 nm that is only 345 kHz wide could cause some minor resonance scattering and additionally complicate the evaporative cooling [21]. The following difficulties complicate the search for a good evaporative sequence. (1) Thulium has a short lifetime in the dipole trap [21] (2.8 s at the ODT at temperatures just below the MOT) compared to other lanthanide species (>7 s for Er [12], >21 s for Dy [13], and >140 s Cr [14]), but there is not a complete understanding of the loss mechanisms at the moment. (2) The maximum depth of the dipole trap ~ 110 μK is only 4.5–5.5 times higher than the temperature of the cloud (15 for Dy [13] and 13 for Er [12]). (3) There are imperfections in trap characterization especially due to scanning of the trap beam [9,20–22], which is changing during the evaporation sequence and the complex dependence of atom polarizability on the ODT beam geometry and light polarization [21]. (4) There are generally no models that are able to capture the process of transferring atoms from a single-beam ODT to a crossed-beam ODT, and this process strongly influences the evaporation efficiency. (5) In addition, the intermediate stages of the cooling transition from a single ODT to a crossed ODT (which is needed due to finite available power for the dipole trap) create additional ambiguity in the measured number of atoms due to the presence of hot atoms inside one of the traps which are not captured in the crossed ODT region.

This list makes the building of an evaporative process model almost impossible and demands a substantial amount of additional work to characterize the trap and atom

*akimov@physics.tamu.edu

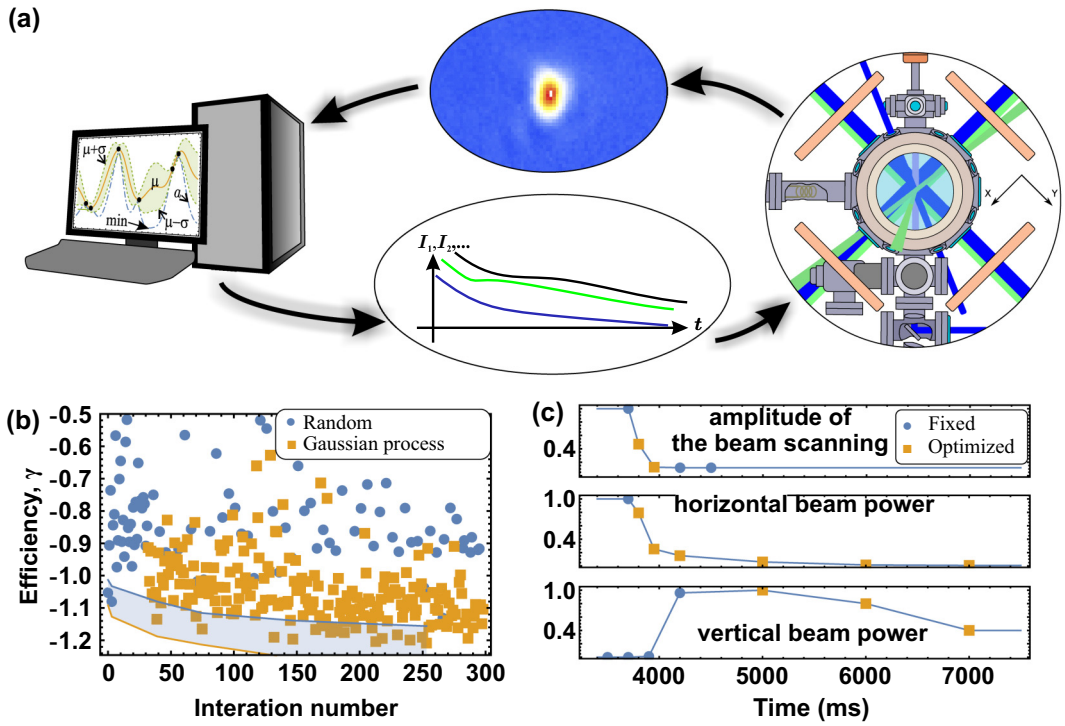

FIG. 1. (a) Idea of the optimization. The parameters to regulate are the parameters of evaporation cooling—the power of the beams and the amplitude of beam scanning. (b) Example of the optimization algorithm result. In this run, the γ parameter was optimized. The blue dots are generated randomly (the first 30 points seed the GP model, and each additional five sample unbiased data). The orange dots are iterations of the algorithm. The gray area is the confidence interval (σ_{conf} wide) for the points minimizing the GP model $\mu_n(x)$ at certain iterations. (c) Final evaporation sequence. The yellow dots are the parameters learned by the algorithm. The horizontal beam intensity and the scanning and vertical beam intensity are linearly interpolated.

properties. The experimental optimization of many mutually related parameters is quite difficult on its own.

We note that while some of these problems (1, 5, and partly 3) could be resolved through additional efforts—including experimentally determining the reason for the high loss rate and eliminating it, developing a better imaging system allowing more precise control of atomic number during the whole evaporation process, fully characterizing the trap in all possible scenarios of trap loading, etc.—there is generally no need for those efforts from the viewpoint of evaporative sequence optimization as long as there is a sequence that provides efficient BEC formation. Instead, to overcome these difficulties, we use an adaptive experimental design based on a Bayesian machine learning technique. Adaptive Bayesian optimization is a well-known statistical approach for optimization of expensive-to-evaluate functions [23]. It is widely used to tune model hyperparameters in machine learning algorithms [24] and to solve some difficult optimization problems in physics [25–29]. The idea of this approach is to avoid the direct optimization of the real system and instead build a surrogate model of the system at each step and utilize it to wisely choose the next point to sample from the real system [30].

In more detail, we use the most common class of surrogate models, the Gaussian process (GP) models [31]. The GP model outcome for experiment $f(\vec{x})$ (the evaporation efficiency γ in our case; see below) started with the vector of the controlling parameters $\vec{x}$ as a random variable. At any given vector $\vec{x}$, this model assumes that the distribution $p[f(\vec{x})]$ is a normal distribution $N[\mu(\vec{x}), \sigma(\vec{x})]$ with a mean $\mu(\vec{x})$ [see the inset in Fig. 1(a) for the case with one-dimensional (1D) $\vec{x}$] and a standard deviation $\sigma(\vec{x})$ [Fig. 1(a)]. These parameters are calculated at each step (see the Supplemental Material [32]) using already observed experimental results. With the GP model, one can choose a new vector of parameters $\vec{x}^*$, which is chosen such that it minimizes the function $a(\vec{x}^*) = \mu(\vec{x}^*) + \alpha\sigma(\vec{x}^*)$ [Fig. 1(a)]. The $a(\vec{x}^*)$ is usually called the upper confidence bound (UCB) [33] acquisition function, which has some scalar fixed parameter α (see the Supplemental Material [32]). Finally, the optimization process implemented using software packages mentioned in [27,34] can be represented as follows:

(1) The algorithm measures n outcomes of the experiment $[f(\vec{x}_1), \ldots, f(\vec{x}_n)]$ at some randomly chosen points $[\vec{x}_1, \ldots, \vec{x}_n]$ and uses them to build (seed) the initial GP model [see Fig. 1(b), blue points].

(2) The GP model algorithm calculates the acquisition function $a(\vec{x})$ and searches for the $\vec{x}^*$ that minimizes its value $\vec{x}^* = \arg\min[a(\vec{x})]$ [see Fig. 1(c)].

(3) The experiment starts at the new value of the control parameters $\vec{x}^*$, $f(\vec{x}^*)$ is measured [see Fig. 1(b), orange points], and the GP model is rebuilt (see the Supplemental Material [32]).

(4) Steps 2 and 3 repeat until the desired number of iterations is reached. Typically, the number of iterations was chosen large (one million) in our experiments, and the code

was stopped when no improvement could be seen in the optimization criteria.

To perform optimization, one needs to choose the optimization criteria $f(\vec{x})$ and define the optimization parameters $\vec{x}$. The larger the number of controlled parameters, the closer the algorithm can converge to the global minimum of the problem. The convergence time also increases with the number of parameters, thus forcing the researcher to limit possible "knobs." In our case, after the precooling stages in the MOT and the atoms are spin polarized, they are loaded into the horizontal dipole trap, the aspect ratio of which is dynamically shaped by an acousto-optical modulator (AOM) to increase the efficiency of loading atoms from the MOT to the ODT [20]. In other words the beam forming the ODT is scanned spatially in one direction so that the beam waist sweeps the distance exceeding the beam waist by several times. The scanning is done with speed, exceeding atomic motion, so that atoms effectively see a wider but somewhat shallow trap [9,20–22] (see the Appendix). After the MOT is switched off, one needs to first efficiently turn off the beam scanning to increase the atomic density; then, during the evaporative cooling, the second beam (the vertical beam in our case) must be turned on to provide a better localization of atoms and a higher density in the trap. Finally, evaporation cooling is performed until BEC is achieved. For the entire procedure, the possible degrees of freedom are the shape of the ramp for the amplitude of the beam scanning, the shape of the ramp for the intensity of the horizontal trap beam, and the shape of the ramp for the intensity of the vertical trap beam. Each initial and final time point for the ramps are the parameters and the actual shapes.

The optimization goal $f(\vec{x})$ was set to be the efficiency of evaporation cooling γ:

$$\gamma = -\ln\frac{D_{PS}}{D_{PS0}} \Big/ \ln\frac{N}{N_0}, \tag{1}$$

where D_{PS0} is the initial phase space density, D_{PS} is the final phase space density, N_0 is the initial number of particles in the ODT, and N is the final number of particles in the ODT.

To approximate the ramp shapes, the entire sequence was split into linear fragments. The algorithm was able to optimize the end points of each segment. Thus, we picked several specific times at which the parameters could be changed by the optimization algorithm. The end and start times of each segment were not varied, but the power of the beams and the horizontal beam waist were varied [see Fig. 1(c)].

The performance of the optimization algorithm for a magnetic field of 4.04 G in step 1 is demonstrated in Fig 1(b). The algorithm clearly converges, although there is a rather high level of noise, which is mainly due to the fit error in the estimation of the temperature and the number of atoms (see the Supplemental Material [32] for more details). By performing a search in the parameter space, the algorithm converges to a γ value of approximately 1.6 ± 0.1 via the

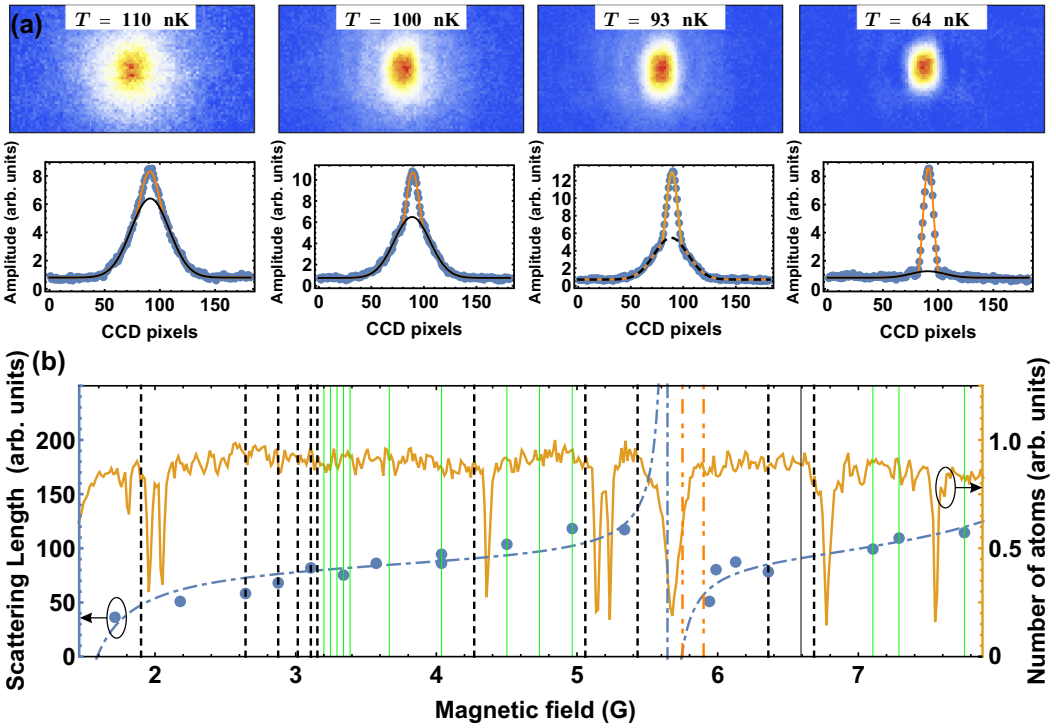

FIG. 2. (a) Typical photos of the atomic cloud with BEC inside and the corresponding bimodal density distribution at different temperatures. The photos were taken at 14.5 ms of expansion. The orange solid line represents the fit by bimodal distribution (see the Supplemental Material [32] for details); the black dashed line shows the contribution of the thermal cloud. (b) Measured scattering length in atomic units versus magnetic field (dots) and the fit of the scattering length (blue dash-dot line), taking into account only broad Feshbach resonances at 1.34, 5.68, and 9.53 G. The orange solid line represents the loss spectrum in arbitrary units [6]. The vertical black dashed line indicates the magnetic fields in which the formation of BEC failed, and the vertical green solid lines indicate the points in which the formation of BEC was successful.

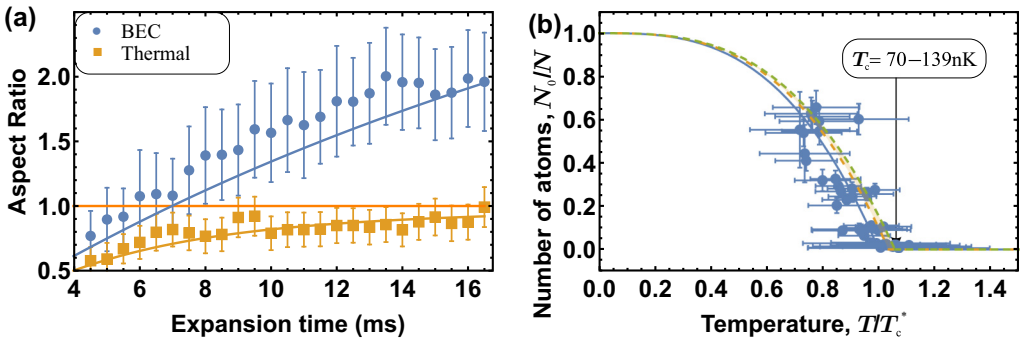

FIG. 3. (a) Aspect ratio for the BECs and the visible sizes of the thermal cloud versus the time of flight after releasing atoms from the ODT. The lines represent theoretical simulations (see the Appendix), and the dots represent experimental data. The red horizontal line indicates the asymptotic behavior of the thermal cloud. (b) Number of condensed atoms N_0 normalized by the total number of atoms N versus the ratio between the thermal part temperature T and the critical temperature T_c^* of the cloud, including interaction corrections [38]. The dots are measured values, and the solid line is the theoretical dependence, including the dipole-dipole interaction between the atoms and the finite number of atom corrections. The dashed line represents the simulations without accounting for interactions for the range of identified critical temperatures to the solid line plot (see the Supplemental Material [32] for more details).

evaporation sequence, as demonstrated in Fig. 1(c). The best stepwise optimization that was achieved manually was $\gamma \approx 0.78$, which was not enough for BEC formation. The sequence found by the algorithm finally led to a bimodal distribution of the atomic densities, which is characteristic for the Bose-Einstein condensation [Fig. 2(a)]. Although the value of γ achieved by the algorithm is still slightly smaller than that found for the Er and Dy experiments [12,13], it is enough to achieve BEC.

We note that while this optimization does not require any prior knowledge of the collisional properties of thulium, the optimization results will obviously depend on them. Figure 2(b) demonstrates the availability of BEC with the cooling sequence, as found by the algorithm at a fixed magnetic field $B = 4.04\,\mathrm{G}$ and at different magnetic fields, along with the measured scattering length (see the Appendix). One can see that outside of the sharp resonances, the scattering length continuously changes due to the presence of broad Feshbach resonances. Although the variation in the scattering length strongly affects the formation of BEC, there are rather wide ranges of positive scattering lengths in which the sequence found does lead to BEC formation. Thus, the identified sequence is quite stable with respect to changes in the strength of the interatomic interactions and potentially some associated losses.

The bimodal distribution [35] depicted in Fig 2(a) shows a striking manifestation of BEC formation. To demonstrate the coherence properties of our atomic cloud, we performed additional experiments. One of the manifestations of coherence in the atomic cloud is its asymmetric expansion, which is even more asymmetric in the presence of dipole-dipole interactions [36]. As a coherent state, BEC should "diffract"; i.e., the more constrained direction of the cloud should expand more rapidly. Indeed, this behavior could be observed in our cloud, as demonstrated in Fig. 3(a). At small expansion times, our imaging system is unfortunately not able to correctly measure the cloud size due to diffraction on the atomic cloud and resolution limitations, but as the cloud expands, the image starts to follow the predicted behavior (see the Supplemental Material [32]). In contrast, the thermal cloud tends to expand more and more symmetrically with time, with the aspect ratio of the cloud sizes asymptotically approaching 1 because wave interference does not take place in this case. Indeed, just above the critical temperature, we observe uniform expansion of the atomic cloud [see Fig. 3(a)].

Nevertheless, a similar distribution and an asymmetric expansion could appear in the hydrodynamic regime [37]. This regime may occur if the Knudsen criterion, $\mathrm{Kn} \ll 1$, is satisfied (see the Appendix). The Knudsen number Kn is basically the ratio of the time between collisions to the period of one oscillation in the trap. If this number is above 1, the atoms manage to do several oscillations in the trap before they collide. As the trap gets released the velocities of the atoms are the same as those that were in the trap, and therefore if the Knudsen number is high, the atoms are not likely to collide during the expansion. However, if this number is low, collisions dominate the dynamics in the trap and therefore during expansion collisions are very probable. We found that $\mathrm{Kn} > 1$ for each axis in all stages of the thermal cloud till the formation of BEC (see the Supplemental Material [32]). In addition, at the hydrodynamic limit, the velocity distribution of the cloud along the tight axis should have a peak over the thermal distribution while having velocities reduced beneath the thermal distribution along the weak axis. In contrast, in our experiment, velocity distributions along both the weak and tight directions have peaks (see the Supplemental Material [32]).

The bimodal distribution obtained by evaporative cooling with the slightly different initial number of atoms provides the experimental dependence of the condensate fraction on the temperature [see Fig. 3(b)]. This result is in agreement with the well-known result for the number of atoms in the condensate depending on temperature:

$$\frac{N_{\mathrm{BEC}}}{N} = 1 - \left(\frac{T}{T_C}\right)^3, \tag{2}$$

where N_{BEC} is the number of condensed particles, N is the whole number of particles, and T_C is the critical temperature (see the Supplemental Material [32] for more details).

In conclusion, we have demonstrated the Bose-Einstein condensation of thulium atoms using a statistically justified method (namely, GP-based Bayesian optimization) to choose the evaporative sequence. We found that while there is no complete characterization of the experimental setup, the GP model is able to optimize the evaporative cooling to achieve BEC without much knowledge of the atomic system. Although it is less descriptive, this approach is goal oriented and far more universal and can thus be used in a wide range of experimental tasks. More specifically, we demonstrated the formation of BEC of thulium atoms, confirmed the results with a bimodal distribution, demonstrated the dependence of the condensate fraction on the temperature, and showed the coherent behavior of the cloud. We also explored BEC formation for a wide range of background scattering lengths, demonstrating the robustness of the evaporative sequence to variations in the scattering length.

This research was supported by the Russian Science Foundation Grant No. 18-12-00266.

APPENDIX

After reloading polarized atoms from the MOT into the one-dimensional horizontal ODT in the scanning regime, we typically have approximately 6×10^6 atoms with a temperature of $20\,\mu$K, which corresponds to a phase space density of approximately 2.3×10^{-5}. The trap frequencies were measured by the standard technique of trap-frequency measurements [21] to be in the scanning trap ν_x; ν_y; $\nu_z = 5.8 \pm 0.1$; 160 ± 2; 1538 ± 20 Hz. The scanning of the ODT beam is realized using an acousto-optical modulator, which was fed by a calibrated voltage control oscillator [22]. Our evaporative cycle took place in the presence of a small magnetic field of 4.04 G, which is not very far from Feshbach resonance. We also measured the scattering length in this magnetic field, $a_0 = 90 \pm 11\, a_{\mathrm{bohr}}$, where $a_{\mathrm{bohr}} = \frac{\hbar}{m_e c \alpha}$ is the atomic unit for the scattering length. The typical frequencies in the BEC regime were measured to be $\nu_z = (197 \pm 1)$ Hz, $\nu_x = (23 \pm 1)$ Hz, and $\nu_y = (137 \pm 6)$ Hz.

To calculate the theoretical curve in Fig. 3(a), we used equations from [36] but changed the sign before the terms containing derivatives of the function f, as there appeared to be a typo in the original paper.

The main uncertainties in the measurements of the atomic numbers and temperatures are as follows: an uncertainty in the number of atoms of 10% (see the Supplemental Material [32]); errors in the temperature of approximately 6% and 10% in the z and x directions, respectively; and errors in the trap frequencies estimated [21] as 1% and 4% for the z and x directions, respectively.

The Knudsen number is defined as $\mathrm{Kn} = \omega_i \tau_c$. Here, ω_i with $(i \in \{x, y, z\})$ is the trap frequency, and $\tau_c = \sqrt{2} n_0 \overline{v_{th}} \sigma$ is the average time between atomic collisions in the trap with a peak density n_0, a thermal velocity $\overline{v_{th}}$, and an elastic-scattering cross section σ.

[1] I. Bloch, J. Dalibard, and S. Nascimbène, Nat. Phys. **8**, 267 (2012).

[2] T. Lahaye, C. Menotti, L. Santos, M. Lewenstein, and T. Pfau, Rep. Prog. Phys. **72**, 126401 (2009).

[3] S. Baier, M. J. Mark, D. Petter, K. Aikawa, L. Chomaz, Z. Cai, M. Baranov, P. Zoller, and F. Ferlaino, Science **352**, 201 (2016).

[4] L. Chomaz, D. Petter, P. Ilzhöfer, G. Natale, A. Trautmann, C. Politi, G. Durastante, R. M. W. van Bijnen, A. Patscheider, M. Sohmen, M. J. Mark, and F. Ferlaino, Phys. Rev. X **9**, 021012 (2019).

[5] L. Tanzi, E. Lucioni, F. Famà, J. Catani, A. Fioretti, C. Gabbanini, R. N. Bisset, L. Santos, and G. Modugno, Phys. Rev. Lett. **122**, 130405 (2019).

[6] V. A. Khlebnikov, D. A. Pershin, V. V. Tsyganok, E. T. Davletov, I. S. Cojocaru, E. S. Fedorova, A. A. Buchachenko, and A. V. Akimov, Phys. Rev. Lett. **123**, 213402 (2019).

[7] W. Ketterle and N. J. Van Druten, Adv. At., Mol. Opt. Phys. **37**, 181 (1996).

[8] C. A. Sackett, C. C. Bradley, and R. G. Hulet, Phys. Rev. A **55**, 3797 (1997).

[9] R. Roy, A. Green, R. Bowler, and S. Gupta, Phys. Rev. A **93**, 043403 (2016).

[10] A. J. Olson, R. J. Niffenegger, and Y. P. Chen, Phys. Rev. A **87**, 053613 (2013).

[11] J. Rudolph, W. Herr, C. Grzeschik, T. Sternke, A. Grote, M. Popp, D. Becker, H. Müntinga, H. Ahlers, A. Peters, C. Lämmerzahl, K. Sengstock, N. Gaaloul, W. Ertmer, and E. M. Rasel, New J. Phys. **17**, 079601 (2015).

[12] K. Aikawa, A. Frisch, M. Mark, S. Baier, A. Rietzler, R. Grimm, and F. Ferlaino, Phys. Rev. Lett. **108**, 210401 (2012).

[13] M. Lu, N. Q. Burdick, S. H. Youn, and B. L. Lev, Phys. Rev. Lett. **107**, 190401 (2011).

[14] A. Griesmaier, J. Werner, S. Hensler, J. Stuhler, and T. Pfau, Phys. Rev. Lett. **94**, 160401 (2005).

[15] S. Kraft, F. Vogt, O. Appel, F. Riehle, and U. Sterr, Phys. Rev. Lett. **103**, 130401 (2009).

[16] T. Weber, J. Herbig, M. Mark, H.-C. Nägerl, and R. Grimm, Science **299**, 232 (2003).

[17] W. Ketterle, D. S. Durfee, and D. M. Stamper-Kurn, in *Proceedings of the International School of Physics "Enrico Fermi"*, Bose-Einstein Condensation on Atomic Gases Vol. 140 (IOS Press Ebooks, Amsterdam, 1999), pp.67–176.

[18] K. B. Davis, M.-O. Mewes, M. R. Andrews, N. J. van Druten, D. S. Durfee, D. M. Kurn, and W. Ketterle, Phys. Rev. Lett. **75**, 3969 (1995).

[19] I. S. Cojocaru, S. V. Pyatchenkov, S. A. Snigirev, I. A. Luchnikov, E. S. Kalganova, G. A. Vishnyakova, D. N. Kublikova, V. S. Bushmakin, E. T. Davletov, V. V. Tsyganok, O. V. Belyaeva, A. Khoroshilov, V. N. Sorokin, D. D. Sukachev, and A. V. Akimov, Phys. Rev. A **95**, 012706 (2017).

[20] V. V Tsyganok, V. A. Khlebnikov, E. S. Kalganova, D. A. Pershin, E. T. Davletov, I. S. Cojocaru, I. A. Luchnikov, A. V Berezutskii, V. S. Bushmakin, V. N. Sorokin, and A. V Akimov, J. Phys. B: At., Mol. Opt. Phys. **51**, 165001 (2018).

[21] V. V. Tsyganok, D. A. Pershin, E. T. Davletov, V. A. Khlebnikov, and A. V. Akimov, Phys. Rev. A **100**, 042502 (2019).
[22] V. Tsyganok, D. Pershin, V. Khlebnikov, E. Davletov, and A. Akimov, J. Exp. Theor. Phys. **128**, 199 (2019).
[23] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas, Proc. IEEE **104**, 148 (2016).
[24] J. Snoek, H. Larochelle, and R. P. Adams, in *Proceedings of the 25th International Conference on Neural Information Processing Systems NIPS' 12*, Vol.2 (Lake Tahoe, NV, 2012), pp.2951–2959.
[25] S. S. Straupe, JETP Lett. **104**, 510 (2016).
[26] C. Kokail, C. Maier, R. van Bijnen, T. Brydges, M. K. Joshi, P. Jurcevic, C. A. Muschik, P. Silvi, R. Blatt, C. F. Roos, and P. Zoller, Nature **569**, 355 (2019).
[27] 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, C. C. N. Kuhn, I. R. Petersen, A. N. Luiten, J. J. Hope, N. P. Robins, and M. R. Hush, Sci. Rep. **6**, 25890 (2016).
[28] S. Ju, T. Shiga, L. Feng, Z. Hou, K. Tsuda, and J. Shiomi, Phys. Rev. X **7**, 021024 (2017).
[29] C. Li, D. Rubín de Celis Leal, S. Rana, S. Gupta, A. Sutti, S. Greenhill, T. Slezak, M. Height, and S. Venkatesh, Sci. Rep. **7**, 5683 (2017).
[30] D. R. Jones, M. Schonlau, and W. J. Welch, J. Global Optim. **13**, 455 (1998).
[31] D. Barber, *Bayesian Reasoning and Machine Learning* (Cambridge University Press, New York, 2012).
[32] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevA.102.011302 for details on Bimodal fit of the cloud; Optical resolution; Gaussian process; Sampling points from the GP model; Measurement of the cooling efficiency during the optimization procedure; Results of the algorithm optimization; Estimation of the temperature and the number of atoms; Comments on the critical temperature plot; Comments on hydrodynamic regime.
[33] N. Srinivas, A. Krause, S. M. Kakade, and M. W. Seeger, IEEE Trans. Inf. Theory **58**, 3250 (2012).
[34] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay, J. Mach. Learn. Res. **12**, 2825 (2011).
[35] C. J. Pethick and H. Smith, *Bose–Einstein Condensation in Dilute Gases*, 2nd ed. (Cambridge University Press, Cambridge, 2008).
[36] S. Giovanazzi, P. Pedri, L. Santos, A. Griesmaier, M. Fattori, T. Koch, J. Stuhler, and T. Pfau, Phys. Rev. A **74**, 013621 (2006).
[37] H. Wu and E. Arimondo, Europhys. Lett. **43**, 141 (1998).
[38] S. Giorgini, L. P. Pitaevskii, and S. Stringari, Phys. Rev. A **54**, R4633 (1996).