Optimal cold atom thermometry using adaptive Bayesian strategies

Jonas Glatthard [1] Jesús Rubio [1] Rahul Sawant [2]
Thomas Hewitt [2] Giovanni Barontini [2] Luis A. Correa [1]

[1]University of Exeter [2]University of Birmingham

1. Introduction

Precise temperature measurements on systems of few ultracold atoms is of paramount importance in quantum technology, but can be very resource-intensive. Here [1], we put forward an adaptive Bayesian framework that substantially boosts the performance of cold atom temperature estimation.

We work with release-recapture thermometry [2], that is we release a trapped gas of a few atoms at temperature T and let it expand for a time t, after which a fraction $f(T,t)$ of the atoms is recaptured. This is given by

$$f(T,t) = \frac{1}{g(\eta)}\, g\left[\frac{\eta\, W(\tilde{t}^2)}{\tilde{t}^2}\right], \tag{1}$$

where $g(s) = 1-\mathrm{e}^{-s}$, $\eta = U_0/(k_B T)$, $\tilde{t}^2 = 4U_0 t^2/(m w_0^2)$ and $W(\cdot)$ is the Lambert function, defined implicitly as $W(s)\,\mathrm{e}^{W(s)} := s$.

Conventionally the above curve is fit to the recapture fraction at different release times. Yet, there is potentially more information in those measurements. For that we go beyond a simple fraction and build a fully probabilistic account. Starting with a single atom in the trap gives us a simple probability model for the number of recaptured atoms

$$p(n_i = 1|T,t_i) = f(T,t_i), \quad p(n_i = 0|T,t_i) = 1 - f(T,t_i). \tag{2}$$

2. Adaptive Bayesian thermometry

In Bayesian estimation the state of information for a parameter is given by a probability distribution. For scale parameters [3], complete ignorance is described by $p(\theta) \propto \theta^{-1}$. Measurements update the state of information according to Bayes' theorem

$$p(\theta|\boldsymbol{n},\boldsymbol{t}) \propto p(\theta) \prod_{i=1}^{\mu} p(n_i|\theta,t_i). \tag{3}$$

The probability distribution gives rise to a point estimator and error bar

$$\tilde{\vartheta}(\boldsymbol{n},\boldsymbol{t}) = \theta_u \exp\left[\int d\theta\, p(\theta|\boldsymbol{n},\boldsymbol{t}) \log\left(\frac{\theta}{\theta_u}\right)\right], \tag{4}$$

$$\Delta\tilde{\vartheta}(\boldsymbol{n},\boldsymbol{t}) = \tilde{\vartheta}(\boldsymbol{n},\boldsymbol{t})\sqrt{\bar{\epsilon}_{\mathrm{mle}}(\boldsymbol{n},\boldsymbol{t})}, \quad \bar{\epsilon}_{\mathrm{mle}}(\boldsymbol{n},\boldsymbol{t}) = \int d\theta\, p(\theta|\boldsymbol{n},\boldsymbol{t}) \log^2\left[\frac{\tilde{\vartheta}(\boldsymbol{n},\boldsymbol{t})}{\theta}\right]. \tag{5}$$

The mean information content of a measurement is quantified by

$$\mathcal{K}(t) = \sum\nolimits_n p(n|t) \log^2\left[\frac{\tilde{\vartheta}(n,t)}{\tilde{\theta}_p}\right], \tag{6}$$

with $\tilde{\vartheta}_p = \theta_u \exp\left[\int d\theta\, p(\theta) \log\left(\frac{\theta}{\theta_u}\right)\right], p(n|t) = \int d\theta\, p(\theta)\, p(n|\theta,t)$. We can optimise in t after each measurement, using the posterior as new prior, i.e. adaptively (cf. [4]).

3. Experiment

^{41}K atoms are loaded into an optical tweezer from a dark magneto-optical trap. The number of atoms transferred into the tweezer can be controlled by varying the magneto-optical trap loading time. The release-recapture protocol is realised by switching off the tweezer's light for a variable amount of time and then switching it on again. To measure the number of atoms left in the tweezer, we utilise fluorescence imaging switching on a series of dedicated resonant beams. The light emitted by the recaptured atoms is collected with the same microscope objective and used to determine the number or remaining atoms.

To account for the Poisson distributed variable number of initial atoms, assuming distinguishability and negligible interactions, we work with the probability model

$$\begin{aligned} p(n_i|T,t_i) &= \sum\nolimits_{N_0=0}^{\infty} P(N_0|\lambda)\, B[n_i|N_0, f(T,t_i)] \\ &= P[n_i|\lambda f(T,t_i)], \end{aligned} \tag{7}$$

where P is the Poisson and B the Binomial distribution. For the Bayesian estimation strategies, we get λ by a set of calibration measurements without release.

4. Processing measured data

To illustrate the methods we apply them to measurements on ^{41}K atoms in a trap with depth $U_0/k_B = 290\,\mu\mathrm{K}$.

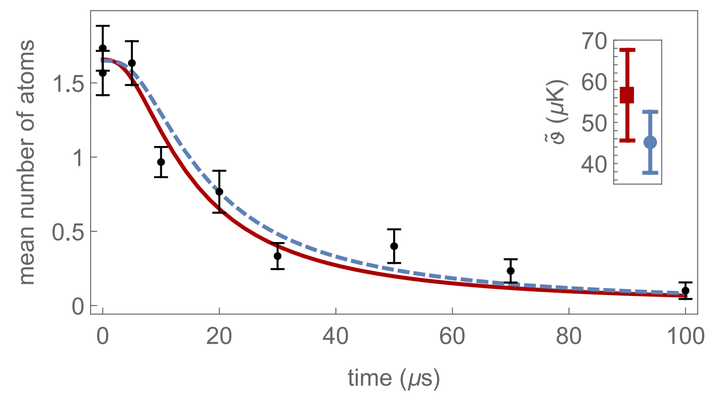

Figure 1:Empirical mean and standard error of the number of recaptured ^{41}K atoms (black dots) at each free-expansion time. The solid red curves are two-parameter non-linear fits to $\lambda f(T,t)$ in λ and T, where $f(T,t)$ is the survival probability. In turn, the Bayesian strategy applied to the full measurement record gives the parameters for the dashed blue curves.

5. Comparing strategies

To establish the relative quality of the different estimation strategies we simulate multiple runs with the same amount of data. The strategies are:

1. conventional non-linear fit (red),
2. Bayesian (with same release times as for fit, blue),
3. *a priori* optimised (optimisation done just once, orange),
4. adaptive (green).

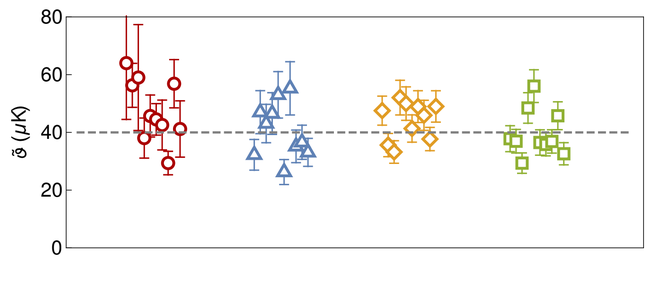

Figure 2:Simulated estimates from 270 data points (from which 60 are calibration measurements without release). The estimates are calculated via the conventional least-squares method (red circles), the Bayesian approach (blue triangles), *a priori* optimised (orange rhombs) and fully adaptive protocols (green squares). The optimised and adaptive protocol outperform the others.

6. Convergence of estimates

Here we compare the convergence speed of two of the estimates, the conventional least squares fit and the *a priori* optimised protocol.

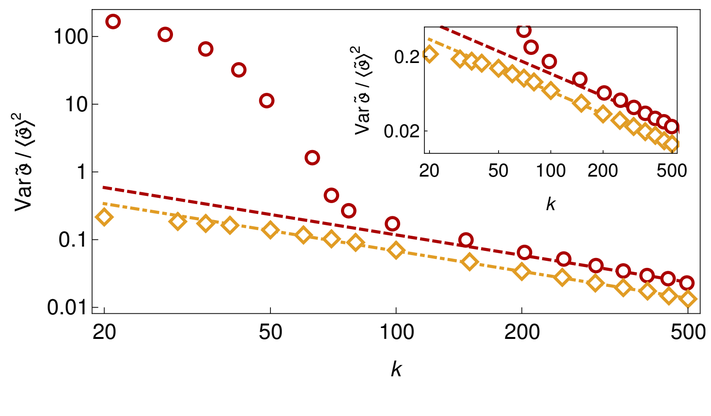

Figure 3:Variability of estimates. The conventional non-linear-fit (red circles) is compared to our *a priori* optimised protocol (orange rhombs). Note the vast superiority of the optimised method for low k, i.e., scarce data. Asymptotically, the variability takes the form $\sim 1/(kF)$. The optimised protocol enters its asymptotic regime faster. The offset between these lines indicates that the asymptotic convergence speed to the true temperature, which is roughly twice as fast for the *a priori* optimised protocol.

7. Release-time optimisation

To illustrate the impact of information maximisation, we show the convergence of a single estimate by ordering measured data from a trap with depth $U_0/k_B = 110\,\mu\mathrm{K}$ from most informative to least and *vice versa*.

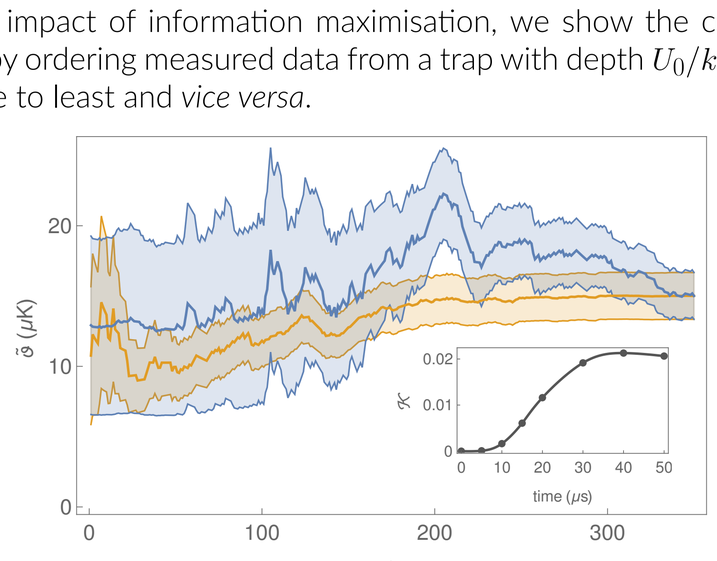

Figure 4:Estimate and error bar from the first k entries of the measurement record reordered to mimic the *a priori* optimised protocol (orange). The optimal recapture time is $t_s = 42\,\mu\mathrm{s}$. The *a priori* information gain $\mathcal{K}$ as a function of the recapture time is shown in the inset for this model (grey dots). The datapoints were chosen from most informative to least. The estimate and error bars of that exact same list, but processed in the reverse order are shown in blue, for comparison. As we can see, the orange curve already converges to the final estimate by processing the 200 most informative data. The remaining 150 data carry practically no information.

8. Adaptive strategy

Finally, we investigate when the adaptive protocol gives an advantage over the *a priori* optimised protocol.

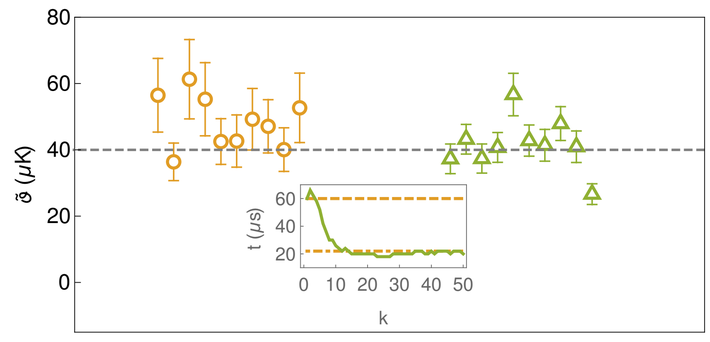

Figure 5:For a wider prior range the fully adaptive (green squares) protocol outperforms the *a priori* optimised (orange rhombs). The inset shows the optimal expansion time as a function of k for one of the simulated measurement records (solid green). The *a priori* optimal recapture time for this wider hypothesis support is indicated by the orange dashed line. The *a priori* optimal time for the narrower support is the orange dot-dashed line, falling onto the adaptive optimal time. This is why, in case of narrower support, the *a priori* optimised method performed equally well as the fully adaptive scheme.

References

[1] J. Glatthard *et al.* arXiv:2204.11816, 2022; [2] M. Mudrich *et al.* Phys. Rev. Lett. 88, 253001, 2002;
[3] J. Rubio *et al.* Phys. Rev. Lett. 127, 190402, 2021; [4] M. Mehboudi *et al.* Phys. Rev. Lett. 128, 130502, 2022.