Molecules are increasingly important for testing fundamental physics. They are used to probe parity violation in nuclei [1] and chiral molecules [2,3], search for changing fundamental constants [4–8], test quantum electrodynamics [9], and measure the electric dipole moments of electrons [10–12] and protons [13]. Measurements of the electron’s electric dipole moment (eEDM) using molecules tightly constrain the parameters of theories that extend the Standard Model [14]. For an atom or molecule with unpaired electrons, the eEDM induces a linear Stark shift so that a coherence time exceeding 150 ms is feasible in a molecule for measuring eEDM. It is more difficult to cool than those cooled previously because (i) it is heavier, so more photons must be scattered, (ii) its vibrational branching ratios are less favourable, so it is harder to scatter many photons, and (iii) the excited states are strongly perturbed, requiring careful design of the cooling scheme [43].

We investigate two sub-Doppler mechanisms, exploring how their effectiveness at cooling molecules depends on laser polarization and applied magnetic field. We also demonstrate a successful approach to modeling sub-Doppler cooling in such complex molecular systems. With the help of these advances, we cool a beam of YbF below 100 μK, so that a coherence time exceeding 150 ms is feasible in a beam, a fountain [40,44], or a trap [45].

Figure 1(a) illustrates the experiment. At $z = 0$, pulses of YbF are emitted from a cryogenic buffer gas source similar in design to Ref. [46]. The pulses have duration 250 μs and mean forward speed $v_z \approx 160$ m/s. The flux of $^{174}$YbF in the first rotationally excited state is $5 \times 10^9$ per steradian per pulse. After a 4 cm diameter aperture at $z = 20$ cm, the molecules pass through the 20-cm-long laser cooling region, the 5.5-cm-long clean-up region, and the detectors which are $l_1 = 41$ cm and $l_2 = 86$ cm from the end of the cooling region. A magnetic field $B$, applied along $z$, is...
uniform to within 0.1 G throughout the cooling region. Shim coils cancel the background magnetic field.

Figure 1(b) shows the relevant energy levels of $^{174}$YbF and the branching ratios between them [43,48]. The main laser-cooling transition is the rotationally closed $X^2\Sigma^+(v = 0, N = 1)$--$A^2\Pi_{1/2}(v' = 0, J' = 1/2)$ transition at 552 nm, with linewidth $\Gamma = 2\pi \times 5.7$ MHz. Additional lasers repump molecules that decay to the $v = 1, 2, 3$ vibrational states of $X^2\Sigma^+$, implementing the scheme proposed in Ref. [43]. Each ground (excited) state has 4 (2) hyperfine components, all with different intervals [49]. Those of the 552 nm transition are shown in Fig. 1(c). Acousto-optic and electro-optic modulators add the radio-frequency sidebands required to excite all ground hyperfine levels. All the light is coupled into a single-mode polarization-maintaining optical fiber that delivers to the molecular beam at right angles and is retro-reflected.

The beams are linearly polarized, one at $\pi/4$ to the $y$ axis and the other at $\pi/4 + \phi$. Each laser, with sidebands added, is separately tuned to produce the maximum fluorescence in the laser cooling region. The 552 nm laser is then detuned by the angular frequency $\Delta$, with sub-Doppler cooling expected for $\Delta > 0$ [50].

The light in the clean-up region has all the repump frequencies and none of the 552 nm frequencies, so any population remaining in $X^2\Sigma^+(v = 1, 2, 3)$ is driven to $X^2\Sigma^+(v = 0, N = 1)$. The spatial distribution of molecules in this state is measured by recording laser-induced fluorescence on one of two CCD cameras. The 552 nm probe light used in these detectors is the same as the cooling light, but independently tuned to $\Delta = 0$. This light crosses the molecular beam at right angles and is retro-reflected.

Figures 2(a)–2(c) show images obtained from CCD2 when $B = 0.8$ G and $\phi = 0$. In (a) the cooling light is absent and we observe a uniform fluorescence intensity in the $x$ direction, reflecting the uniform molecule density and detection efficiency across the field of view. The fluorescence along $z$ reflects the intensity distribution of the probe laser. In (b) the cooling is applied with $\Delta = +1.5\Gamma$ and we see a bright spot at the center while in (c), $\Delta = -1.5\Gamma$, and we observe a hole. We integrate these images along $z$, over the 8 mm range shown, then divide the distribution with cooling applied (b), (c) by the one with no cooling (a), giving the normalized fluorescence distributions shown in Fig. 2(d). The narrow peak obtained when $\Delta = +1.5\Gamma$ corresponds to a beam with a highly collimated center and is due to magnetically induced sub-Doppler cooling (see below). Further out, there is a dip where molecules that form the peak would otherwise have been. When $\Delta = -1.5\Gamma$, there is a dip at the center with broad wings on either side because the same mechanism now drives slow molecules to higher speeds. Doppler cooling also contributes to these broad wings. The distance between the minima (at $z_{\text{min}}$) for $\Delta = +1.5\Gamma$, or between maxima for $\Delta = -1.5\Gamma$, increases from CCD1 to CCD2. From this change, we infer a capture velocity for sub-Doppler cooling of $v_c \approx 0.9$ m/s. Fewer molecules are detected when the cooling light is applied. This depletion is largest (55%) when $\Delta = 0$. We have investigated and ruled out several possible causes for this loss, including deflection or heating of the molecular beam in the $y$ direction, incomplete optical pumping in the clean-up region, or decay to other rotational levels. We are investigating unexpected losses to higher-lying vibrational states.

We fit the density distributions to a sum of four Gaussians, $G_i$, with a common center, but different
amplitudes and widths. For $\Delta > 0$, $G_1$ represents the narrow central peak, $G_2$ the broad dip, and $G_3$ the even broader curvature of the baseline. $G_4$, always of low amplitude, helps to reproduce the shoulders of the narrow peak. Fits to the data in Fig. 2(d) are shown by the lines. We define the peak height and width as the amplitude and width of $G_1$, and use these parameters to quantify the effectiveness of the cooling.

Figure 2(e) shows the normalized density distributions at the two CCDs when $\Delta = 2\Gamma$, $B = 1.2$ G, $\phi = 0$ and $P_0 = 90$ mW. The peak is higher at CCD2 because the density of the uncooled beam, which provides the normalization, decreases faster than that of the cooled beam. Our simulations (see below) show that molecules in the central peak have a Boltzmann velocity distribution with no correlation between position and velocity. In this case, the temperature is given by

$$T = \frac{m v_{\text{rms}}^2}{k_B} \frac{w_2^2 - w_1^2}{l_2^2 - l_1^2},$$

(1)

where $w_{1,2}$ are the rms widths of the peaks at the two detectors. Fitting the four-Gaussian model to the data in Fig. 2(e) gives $w_1 = 1.009 \pm 0.045$ mm and $w_2 = 0.905 \pm 0.021$ mm, where the errors are the statistical uncertainties from the fit. These uncertainties translate into a temperature resolution of 100 $\mu$K. The systematic error due to uncertainty in the imaging magnification and various misalignments of the cameras are below 0.6%, so contribute negligibly to the uncertainty. Although $w_2 < w_1$, implying a negative temperature, the widths differ by only 2$\sigma$. We conclude that the temperature is below the resolution of the measurement, giving us an upper temperature limit of $T_{\text{upper}} = 100$ $\mu$K. Given the measured widths, the probability of $T$ being higher than this is only 0.13%. Fits to unnormalized data also give $T$ consistent with zero and an upper limit reduced to 80 $\mu$K. Fitting single Gaussians to the data lying within $x < x_{\text{cut}}$ gives $T$ consistent with zero, with an uncertainty below 100 $\mu$K, for all choices of $x_{\text{cut}} < x_{\text{min}}$. We have modeled the possibility that the central peak contains two separate distributions at temperatures $T_{\text{hot}}$ and $T_{\text{cold}}$, with $T_{\text{cold}} \ll T_{\text{hot}}$. This can lead to $w_2 < w_1$ because the hot, rapidly expanding component broadens the peak at CCD1 more than at CCD2. The model can reproduce the observed widths provided $T_{\text{hot}} > 0.5$ mK and $T_{\text{cold}} < 35$ $\mu$K. Though this method lowers the temperature limit, it depends on the model being the correct one, so we prefer to use $T_{\text{upper}}$ as a more conservative upper limit. Our estimate of $T_{\text{upper}}$ is below the Doppler temperature, which for these parameters is $T_D \approx 450$ $\mu$K. It is also below the minimum Doppler temperature of $T_{\text{D,min}} = h\Gamma/(2k_B) = 137$ $\mu$K.

Figure 2(f) shows the results of simulating these experiments. The cooling force and momentum diffusion coefficient are calculated as functions of velocity by solving the optical Bloch equations, following the approach of Ref. [50] extended to account for the hyperfine structure of the ground and excited states and the three frequency components of the 552 nm light. Using these results, the distribution of molecules is calculated by solving the Fokker-Planck equation [51]. For $\Delta = 1.5\Gamma$, the simulations reproduce the cooling data well, showing a central peak of width 1.17 mm, similar to the measured width of 0.87 $\pm$ 0.04 mm. The simulations predict a capture
velocity of $v_c = 0.6$ m/s, close to the value estimated above. For $\Delta = -1.5 \Gamma$, the simulations reproduce the central dip seen in the experiment, but show additional structure not seen experimentally. For the parameters corresponding to Fig. 2(e), the simulations predict a temperature of 1 $\mu$K, only six times higher than the recoil temperature and far below the temperature resolution of the experiment. Reducing the cooling force or increasing the diffusion coefficient by a factor 10 increases the predicted temperature to 10 $\mu$K.

With $\Delta > 0$, the peak width varies little for the range of parameters explored, and the temperature is too low to measure. However, the peak height varies strongly with the parameters and is a good measure of the number of ultracold molecules. Figure 3(a) plots the peak height vs $\Delta$ showing the dispersive shape characteristic of laser cooling. The height is antisymmetric about $\Delta = 0$ and largest when $\Delta \approx +2 \Gamma$. Figure 3(b) shows that the peak height increases with the length of the cooling region, $L$, but starts to level off once $L \approx 20$ cm.

The sub-Doppler cooling demonstrated here relies on the presence of dark states, and the mechanism depends on the polarization configuration, which we control through $\phi$ [50]. When $\phi = 0$, the polarization is uniform but there are standing waves of intensity. A molecule in a bright state, moving towards high intensity, climbs the potential hill arising from the ac Stark shift and is optically pumped into a dark state near the top of the hill. As it moves on towards low intensity, the magnetic field rotates the dark state back into a bright state. Thus, molecules continually climb potential hills. This mechanism, known as magnetically induced laser cooling [52–55], has been used for transverse cooling of SrF and SrOH beams [25,39]. The filled points in Fig. 4(a) show the peak height vs $B$ for this $\phi = 0$ case. The cooling is ineffective if $B$ is too small, and the peak height increases with $B$ up to $B \approx 1.2$ G. Cooling should be optimized when the Larmor precession time is about equal to the time taken for a molecule to move from node to antinode of the standing wave. Taking $v_c/2$ as a typical speed, we expect an optimum $B$ of $B_c \approx 2h v_c/(\mu_B \lambda)$.

Using $g = 1/3$, the degeneracy-weighted average $g$ factor of the ground-state hyperfine components, and $v_c \approx 0.9$ m/s as found above, we obtain $B_c \approx 1.1$ G, matching the optimum found experimentally. For higher $B$, the peak height is surprisingly insensitive to $B$, perhaps because of the wide range of $g$ factors and molecule speeds involved. We find the cooling to be effective up to $B \approx 15$ G.

When $\phi = \pi/2$, the intensity is uniform but the polarization is not, and the sub-Doppler mechanism involves nonadiabatic transitions between dark and bright states induced by motion through this changing polarization [56,57]. The open points in Fig. 4(a) show the peak height vs $B$ when $\phi = \pi/2$. The cooling is effective at $B = 0$, as expected since this mechanism does not require a magnetic field, unlike the $\phi = 0$ case. On the contrary, the data show that magnetic fields above 2 G are detrimental to this cooling mechanism.

Figure 4(b) shows how the peak height depends on $\phi$ when $B = 0$. The data are roughly symmetric around $\phi = \pi/2$ as expected, with maxima near $\phi = \pi/4$ and $3\pi/4$, minima near $\phi = 0$ and $\pi$, and a local minimum near $\pi/2$. Simulations for this polarization configuration [50] show that the sub-Doppler force is maximized for $\phi$ between $\pi/8$ and $3\pi/16$ for the $1 \rightarrow 1$ system and between $3\pi/16$ and $\pi/4$ for the $2 \rightarrow 1$ system. Our measurements are consistent with those results.

In summary, we have cooled YbF molecules to sub-Doppler temperatures by realizing the laser cooling scheme proposed in [43]. We have explored and explained how the cooling efficiency depends on the magnetic field and the polarizations of the cooling beams, and have demonstrated a successful approach to modeling the cooling of such complex molecular systems. These are key steps towards using ultracold molecules for an eEDM measurement [40], and other tests of fundamental physics. Our temperature limit of $T < 100$ $\mu$K extends the feasible coherence time to $\tau > 150$ ms. To make use of this with the current $v_c$ would require a 24-m-long experiment. A slower beam could be obtained by radiation pressure slowing [26,34,35], Zeeman-Sisyphus slowing [58], or bichromatic force.
slowing [59]. For the data shown in Fig. 2(e), there are about $1.3 \times 10^4$ molecules per shot in the ultracold part of the distribution. We are currently extending the method into 2D, which should yield more molecules, especially since the capture velocity for sub-Doppler cooling is larger in 2D [50]. The combination of a Doppler cooling period followed by sub-Doppler cooling would increase the capture velocity further while providing the same low final temperature.

Underlying data may be accessed from Zenodo [60] and used under the Creative Commons CCZero license.

This work was supported by the STFC under Grant No. ST/N000242/1, and by EPSRC under Grants No. EP/M027716/1 and EP/P01058X/1, and has received funding from the European Research Council under the European Union’s Seventh Framework Programme (FP7/2007–2013)/ERC Grant Agreement No. 320789.

* m.tarbutt@imperial.ac.uk


