JOSEPHSON MODEL

We compare our experimental results to the Josephson model (JM) and its plasma frequency, \( \omega_p \) (Fig. 3). The JM employed here is based on the nonlinear two-mode ansatz used in [1],

\[
\Psi(r, t) = \psi_R \phi_R(r; N_R(t)) + \psi_L \phi_L(r; N_L(t)) \quad (S1)
\]

where \( \psi_{R,L} = \sqrt{N_{R,L}(t)} \exp(i \theta_{R,L}(t)) \) and \( \phi_{R,L} = (\phi_+ \pm \phi_-)/\sqrt{2} \) is a real function localized in the left (right) well, with \( \phi_+ (\phi_-) \) being the ground (first antisymmetric) state of the GPE along the splitting direction. The linearized equation gives \( Z(t) = Z(0) \cos(\omega_p t + \Delta \theta(0)) \), where \( \Delta \theta = \theta_R - \theta_L \) and the plasma frequency

\[
\omega_p^2 = \frac{1}{\hbar^2} \Delta E \left( \Delta E + N \frac{\partial \mu_{loc}}{\partial N_L} \right), \quad (S2)
\]

where

\[
\mu_{loc} = \int dr \left[ \frac{\hbar^2}{2m} (\nabla \phi_{R,L})^2 + U(r) \phi_{R,L}^2 + g N_{R,L} \phi_{R,L}^4 \right], \quad (S3)
\]

and \( \Delta E = E_- - E_+ = 2(\mathcal{K} + N \chi) \) with

\[
E_\pm = \int dr \left[ \frac{\hbar^2}{2m} (\nabla \phi_\pm)^2 + U(r) \phi_\pm^2 + \frac{1}{2} g N \phi_\pm^2 \right], \quad (S4)
\]

\[
\mathcal{K} = - \int dr \left[ \frac{\hbar^2}{2m} \nabla \phi_R \nabla \phi_L + \phi_R U(r) \phi_L \right], \quad (S5)
\]

\[
\chi = - \frac{g}{4} \int d r \phi_R^* \phi_L. \quad (S6)
\]

The plasma frequency \( \omega_p \) depends on the derivative of the single-well chemical potential \( \mu_{loc} \), and therefore takes into account the effect of transverse degrees of freedom on the effective nonlinearity determining the interaction energy. This provides an important correction to the plasma frequency, typically around 20%. Here \( \Delta E \) is the energy splitting between the ground state and the lowest antisymmetric state along the splitting direction. In our experiments, for example, \( \mu_{loc}/\hbar = 2\pi \times 1.8 \text{ kHz} \) and \( \Delta E/\hbar = 2\pi \times 3.7 \text{ Hz} \) at \( \delta_0 = 6.9 \text{ kHz} \).

HYDRODYNAMIC MODEL

To determine the behaviour of the condensate in a double well in the hydrodynamic regime, we use the continuity equation and the equation of motion for the condensate in the hydrodynamic description, ignoring the quantum pressure term:

\[
\frac{\partial \rho(r, t)}{\partial t} + \nabla \cdot [\mathbf{v}_s(r, t) \rho(r, t)] = 0 \quad (S7)
\]

\[
m \frac{\partial \mathbf{v}_s(r, t)}{\partial t} + \nabla [U(r) + g \rho(r, t) + \frac{1}{2} m \mathbf{v}_s^2(r, t)] = 0 \quad (S8)
\]

where \( \rho(r, t) \) is the local density and \( \mathbf{v}_s \) is the superfluid velocity. We assume harmonic motion of the population balance between the wells such that

\[
\ddot{Z} = -\omega^2_{HD} Z \quad (S9)
\]

where \( \omega_{HD} \) is the hydrodynamic frequency that characterizes the system.

The first time derivative of \( Z \equiv 2N_R/N \) is

\[
\dot{Z} = \frac{2}{N} \int_{V_R} \dot{\rho} \, d^3r = - \int_{V_R} \nabla \cdot (\rho \mathbf{v}_s) \, d^3r = - \int_S \hat{n} \cdot (\rho \mathbf{v}_s) \, dS \quad (S10)
\]

where \( V_R \) is the volume of the right well, \( S \) is the area of the plane separating the two wells, and \( \hat{n} \) is the unit normal vector for this plane.

The second derivative of \( Z \) is then

\[
\ddot{Z} = - \frac{2}{N} \int_S \hat{n} \cdot (\dot{\rho} \mathbf{v}_s + \rho \dot{\mathbf{v}}_s) \, dS. \quad (S11)
\]

To evaluate the frequency, \( \omega_{HD} \), we assume that the system begins at rest, such that \( \mathbf{v}_s(t = 0) = 0 \). The time derivative of \( \mathbf{v}_s \) is given by the hydrodynamic equation of motion, Eq. (S8), and

\[
\frac{\partial \rho(r, t)}{\partial t} + \nabla \cdot (U(r) + g \rho) \, dS = 0 \quad (S12)
\]

The geometry of this double well system is such that the normal vector \( \hat{n} = \hat{x} \), and the only component of the gradient which contributes is the \( x \)-component. Assuming some initial imbalance, \( Z_0 \), the frequency with which the populations oscillate is given by

\[
\omega^2_{HD} = - \frac{\dot{Z}}{Z} = - \frac{2}{mN Z_0} \int_S \rho \frac{\partial}{\partial x} (U(r) + g \rho) \, dy \, dz. \quad (S13)
\]
We calculate this initial density profile in the trap, tilted by a linear bias \( Gx \), using the Thomas-Fermi approach:

\[
\rho_{\text{TF}}(r) = \left( \mu - (U(r) + Gx) \right) / g, \tag{S14}
\]

The gradient term in the integrand of Eq. (S13) is then simply \(-G\).

The characteristic frequency is thus

\[
\omega_{\text{HD}}^2 \approx \frac{2G}{mN\Sigma_0} \int_S \rho_{\text{TF}} \, dy \, dz, \tag{S15}
\]

which indicates that the frequency can be found by simply evaluating the density at the surface between the two wells and integrating over the region by which the two halves are connected. From this expression, we see that the \( \omega_{\text{HD}}^2 \) decreases as the area connecting the wells decreases, and falls to zero when the barrier surpasses the chemical potential and the Thomas-Fermi density is strictly zero on the plane \( S \).

The equation (S13), upon substituting \( \rho_{\text{TF}} \) with the Gross-Pitaevskii ground state density, is also valid when we include a quantum pressure term in Eq. (S8). However, even with the quantum pressure, this model is not in exact agreement with the Gross-Pitaevskii equation, due to the non-harmonic component in the oscillation. At high barriers, though anharmonicity is small, Eq. (S13) is less accurate than the JM.

We have checked that the frequencies predicted by Eq. (S15) are consistent with the dynamical simulations of Eqs. (S7), (S8), done using a test particle method [5].

**GROSS-PITAЕVSKII EQUATION**

We solve numerically the time-dependent equation

\[
i\hbar \partial_t \Psi(r) = -\frac{\hbar^2}{2m} \nabla^2 \Psi(r) + \left( U(r) + g|\Psi(r)|^2 \right) \Psi(r) \tag{S16}
\]

where \( \Psi(r) \) is the complex condensate order parameter, \( U(r) \) is the double-well external trapping potential, and \( g = 4\pi\hbar^2\alpha_s/m \) with \( \alpha_s \) the \(^{87}\text{Rb} \) s-wave scattering length. While all calculations were done using Eq. (S26), an intuitive understanding of the potential emerges from the separable approximate form:

\[
U_{\text{sep}} \approx \frac{1}{2} m \omega_x^2 y^2 + \frac{1}{2} m \omega_z^2 x^2 + \alpha_2 x^2 + \alpha_4 x^4, \tag{S17}
\]

where \( \alpha_2 < 0 \). The trap exhibits an axial anisotropy, with elongation along the \( y \)-direction such that \( \omega_z \approx 4\omega_y \).

**DECAY OF POPULATION IMBALANCE**

In comparing the time series measured in the experiment with those found from GPE calculations, we see similar multiple-frequency behaviour. One striking difference is the presence of “decay” in the experimental data – the fall off of the amplitude of the populations oscillations with time. The characteristic time scale of the decay, \( \tau \), is approximately equal to two oscillation periods over all values of \( V_0/\mu \). We model this as an exponentially decaying envelope in our analysis, and include it in our fitting equation (Eq. (S27)).

In the GPE results, no such decay is observed. Figure S1 shows a comparison between one experimental run and a GPE calculation for very similar parameters (\( \delta = −0.1 \text{ kHz} \)). Indeed, GPE calculations to 64 ms show no sign of damping. Besides the possibility of the damping arising from technical sources, it may be due to thermal or other stochastic effects not included in the \( T = 0 \) mean field calculation.

**ROLE OF TRAP ANISOTROPY**

We studied the role of the trap anisotropy (i.e., \( \omega_z \neq \omega_y \)) by observing the transformation of the \( m = 0 \) and \( m = 2 \) modes as the trap is deformed from axially symmetric to strongly axially anisotropic, in presence of a purely anharmonic potential along \( z \), using the simplified potential, Eq. (S17). The dipole perturbation excites both modes as soon as the axial symmetry is broken, and the spectrum shows a second frequency growing in strength as the axial anisotropy is increased. The values of the mode frequencies as a function of \( \omega_z/\omega_y \) are shown in Fig. S2. Close to axial symmetry, the lower frequency depends only slightly on the transverse confinement, indicating that the \( m = 0 \) mode is a dominant component of the Bogoliubov excitation. Motion is primarily along the splitting direction without oscillations in the transverse directions.
FIG. S2: Mode frequencies for \( m = 0 \) (open) and \( m = 2 \) (closed) as a function of trap anisotropy. Grey arrow indicates the anisotropy used in this experiment. These calculations use the approximate potential Eq. (S17). These are calculated with fixed \( \omega_z \) and decreasing \( \omega_y \).

Sufficiently far from axial symmetry, both frequencies start to decrease with increasing anisotropy and show a similar behavior. In particular, the experimental trapping conditions correspond to the point \( \omega_z \approx 4 \omega_y \), as indicated in Fig. S2, where the two modes begin to show a similar dependence on transverse confinement. This strongly suggests that for such high axial anisotropy, each Bogoliubov mode is mainly a combination of the two original \( m = 0 \) and \( m = 2 \) modes at axial symmetry.

FULL DESCRIPTION OF RWA POTENTIAL

The double-well potential is created through a coupling between static and rf magnetic fields. In the dressed state picture, these combine to form the effective potential [2]

\[
U_{\text{RWA}}(\mathbf{r}) = m_F' \sqrt{[\hbar \omega_{\text{rf}} - g_F \mu_B B_S(\mathbf{r})]^2 + \frac{[g_F \mu_B B_{\text{rf},\perp}(\mathbf{r})]}{2}^2} \tag{S18}
\]

where \( m_F' \) is the adiabatic magnetic quantum number, \( g_F \) is the Landé g-factor, \( \mu_B \) is the Bohr magneton, \( B_S(\mathbf{r}) \) is the static magnetic field, described by an Ioffe-Pritchard potential, and \( B_{\text{rf},\perp}(\mathbf{r}) = |\mathbf{B}_S(\mathbf{r}) \times \mathbf{B}_{\text{rf}}(\mathbf{r})|/|\mathbf{B}_S(\mathbf{r})| \) is the component of the oscillating magnetic field locally perpendicular to the static field at each point, \( \mathbf{r} \).

The static magnetic trap arises a result of the combination of current flowing through the ‘Z’-wire on the chip, an external bias field, and an external Ioffe field. In combination, these create an Ioffe-Pritchard style trap, a static magnetic field \( \mathbf{B}_S = B_z \hat{x} + B_y \hat{y} + B_z \hat{z} \), whose components are described by

\[
B_z(x, z) = B' x - \frac{B''}{2} xy, \tag{S19}
\]

\[
B_y(x, y, z) = B_S(\mathbf{0}) + \frac{B''}{2} \left( y^2 - \frac{1}{2}(x^2 + z^2) \right), \tag{S20}
\]

and

\[
B_z(y, z) = -B' z - \frac{B''}{2} yz. \tag{S21}
\]

In the limit of a small cloud, the static potential is well-approximated by a harmonic trap, characterized by radial and axial trapping frequencies \( \omega_x, z \) and \( \omega_y \). In terms of these measurable values, the static trap-bottom term, the gradient term, and the curvature term are given by

\[
B_S(\mathbf{0}) = \frac{2 \hbar \omega_{\text{TB}}}{m_F' g_F \mu_B} \tag{S23}
\]

\[
B' = \sqrt{\frac{m B_S(\mathbf{0})}{m_F' g_F \mu_B} \left( \frac{\omega_x^2}{2} + \frac{\omega_y^2}{2} \right)} \tag{S24}
\]

\[
B'' = \frac{m g_F^2 \mu_B}{m_F' g_F \mu_B}, \tag{S25}
\]

respectively, where we define \( \omega_{\text{TB}} = \mu_B g_F B_S(\mathbf{0})/\hbar \) as the “trap bottom” frequency.

Typical values for the parameters in Eqs. (S25) and (S26) are: \( (\omega_x, z, \omega_y) = 2 \pi \times (1310, 10) \) Hz, \( \omega_{\text{TB}} = 2 \pi \times 787 \) kHz, \( B_{\text{rf},\perp} = 240 \) mG, \( \omega_y, 0 = 2 \pi \times 95 \) Hz, and in the \( |F = 2, m_F = 2 \) state of \(^{87}\text{Rb}\) we use, \( m_F' g_F = 1 \).

CORRECTIONS TO THE ROTATING-WAVE APPROXIMATION

To calculate our trapping potential, Eq. (S26) assumes the rotating-wave approximation (RWA), but as discussed in [3], the RWA fails for large Rabi frequencies. We study the effect of the beyond-RWA effects for our trap and find that we can account for the difference between the approximate and full potentials by simply shifting the detuning by a fixed amount.

We calculate the full potential in a two-dimensional plane at \( y = 0 \) for our trap at a particular detuning, \( \delta_0 \), as described in [3]. This 2D contour is fit to

\[
U_{\text{RWA}} = m_F' \text{sgn}(g_F) \hbar \sqrt{(\delta(\mathbf{r}) - \delta_{\text{shift}})^2 + \Omega^2}, \tag{S26}
\]

at \( y = 0 \), which is just Eq. (S26) without the compression term, where \( \delta_{\text{shift}} \) is the only fit parameter and describes a shift of the detuning. We compare the shape of the full potential to the RWA potential with the shift and find that they are very similar. The shift is calculated for all detunings used in this work and is roughly uniform for the range explored (Fig. S3). We apply a shift \( \delta_0 \rightarrow \delta_0 + 2 \pi \times 1.9 \) kHz to each of the detunings used with the RWA in this work.
The shift we find is of the opposite sign to that found in Ref. [3]. Compared to the potential used in that work, our Rabi frequency is much smaller and our detuning much closer to zero, and we have confirmed that the shift changes sign for larger Ω and large negative detunings.

**SYSTEMATIC TRAP-BOTTOM SHIFTS**

As noted in the main text, we shift the data by a fixed detuning for display in Fig. 3. Despite the discrepancy between the calculated and measured values of the detuning, the data matches the GPE in terms of the shape of the curves and the slope of the population oscillation frequency as a function of \( V_b/\mu \). For this reason, we are satisfied that the shift we are applying is acting to account for an unknown systematic uncertainty in the determination of the trap bottom value \( B_S(0) \).

After taking into account all known systematic shifts, which include the beyond-RWA effect described above and calibrations in measuring the trap bottom \( B_S(0) \), we fit the experimental data to the GPE simulation data using a single-parameter least-squares fit, where the fitting parameter acts to slide the data along the detuning axis. We find that a shift of \( 2\pi \times (5.1 \pm 0.1) \) kHz accounts for the difference between the experiment and the GPE. This shift is in the opposite direction to the beyond-RWA corrections. Possible sources of this discrepancy include systematic errors in determining the static trap bottom \( B_S(0) \), or imperfections in the polarization of \( B_{RF} \) due to the proximity of the fields to the chip and its copper support.

**ATOM NUMBER**

To calibrate the atom number, we use standard absorption imaging to measure the thermal fraction of clouds above and below the condensate temperature. We determine the total atom number by measuring the total absorption of the cloud, and the thermal number by fitting the wings to a Bose-Einstein distribution and integrating under the entire curve to extract thermal atom number. The temperature of each condensate is determined by fitting the wings to a Gaussian.

To find \( T_C \), the condensation temperature, we plot the condensate fraction as a function of temperature. We determine the temperature at which the condensate fraction is first non-zero, and find the number of atoms to which this corresponds. Using the relationship between condensation temperature and atom number, including finite size and interaction effects [4], we can determine the condensation temperature to ±9% \( (T_C = 640 \pm 40 \text{ nK}) \). Propagating this error through to atom number, we arrive at a calibration factor \( N_{\text{actual}} = N_{\text{measured}} \times (1.3 \pm 0.3) \), which accounts for the systematic uncertainty in our atom number, \( N = 6600 \pm 1700 \).

The number, \( N = 8000 \) was chosen for the calculations because this is the number within the systematic uncertainty for which the best agreement is found for mode amplitudes (Fig. 4). The same \( N = 8000 \) is used in the Josephson model and hydrodynamic approximations.

**DATA ANALYSIS**

To analyze the time series data, as in Fig. 2(a), we use a Fourier transform (FT). To prepare the data, we eliminate the offset from \( \mathcal{Z} = 0 \) components by subtracting from each point the mean, where the mean might be non-zero due to a small equilibrium imbalance in the system. We smooth the transformed data by padding the time series with zeros to a total of \( 1024 \) points.

When extracting the peak locations from the FT, we eliminate the points below the frequency given by \( 1/t_{\text{tot}} \), where \( t_{\text{tot}} \) is the longest hold time. We then find the two peaks with the maximum height and use these as our data points in Fig. 3. We plot the amplitudes-squared of the FTs in the color-map in Fig. 3 behind the data. The data are linearly interpolated numerically between the values of \( V_b/\mu (\delta_0) \) at which the data were measured.

The uncertainty in the frequency measurement is found by simulating data with the same level of noise as the original time series. The quantity of noise is determined by fitting the time series to a 2-frequency decaying exponential function

\[
\mathcal{Z}(t) = e^{-t/\tau} \left[ a_1 \sin(2\pi\nu_1(t - t_01)) + a_2 \sin(2\pi\nu_2(t - t_02)) \right],
\]

(S27)
where \( \tau \) is a time constant for decay, \( a_{1(2)} \) is the amplitude of the first (second) frequency component, and \( \nu_{1(2)} \) is the first (second) frequency component, and \( t_{01(02)} \) is the constant accounting for the phase shift of the first (second) component. The standard deviation of the residuals from this fit gives the noise level. We simulate data 100 times with the same parameters as those given by the fit, with the same total time and density of points, but with different randomized instances of Gaussian noise whose standard deviation is the same as that measured. Taking the frequency measurements from each of these trials, we determine the smallest range inside of which 68\% of the measurements lie. This confidence interval is used as the uncertainty in the frequency measurement.

The noise floor in the FT is established in a similar fashion. Using the result for the noise level from the time series, we simulate pure Gaussian noise and take the FT of this. The noise floor we show is the mean plus one standard deviation of the maximum peak amplitudes found in 100 such simulations.

The amplitudes used in determining the ratio \( R_1 \), shown in Fig. 4, are given by the values determined by the fit (Eq. (S27)). The uncertainties in these values are determined in a similar way to those in the frequencies; we use the noise level in the residuals of the fit, simulate and fit 100 sets of data with similar parameters, and use the 68\% confidence interval of these results to represent our uncertainty.

One significant difference between the calculated and measured quantities is that the calculated amplitudes display no decay. The measured values, which come from the fits to the Eq. (S27), rely upon the fitting routine to extrapolate backwards in time to determine the \( t = 0 \) amplitudes. The uncertainty associated with this process results in the scatter in the measurements, and may be a cause of some of the discrepancy between the calculated and measured values.

[1] A. Smerzi and A. Trombettoni, Phys. Rev. A 68, 023613 (2003).
[2] I. Lesanovsky et al., Phys. Rev. A 73, 033619 (2006).
[3] S. Hofferberth et al., Phys. Rev. A 76, 013401 (2007).
[4] S. Giorgini, L. P. Pitaevskii, and S. Stringari, Phys. Rev. A 54, R4633 (1996).
[5] A. Smerzi and S. Fantoni, Phys. Rev. Lett 78, 3589 (1997).