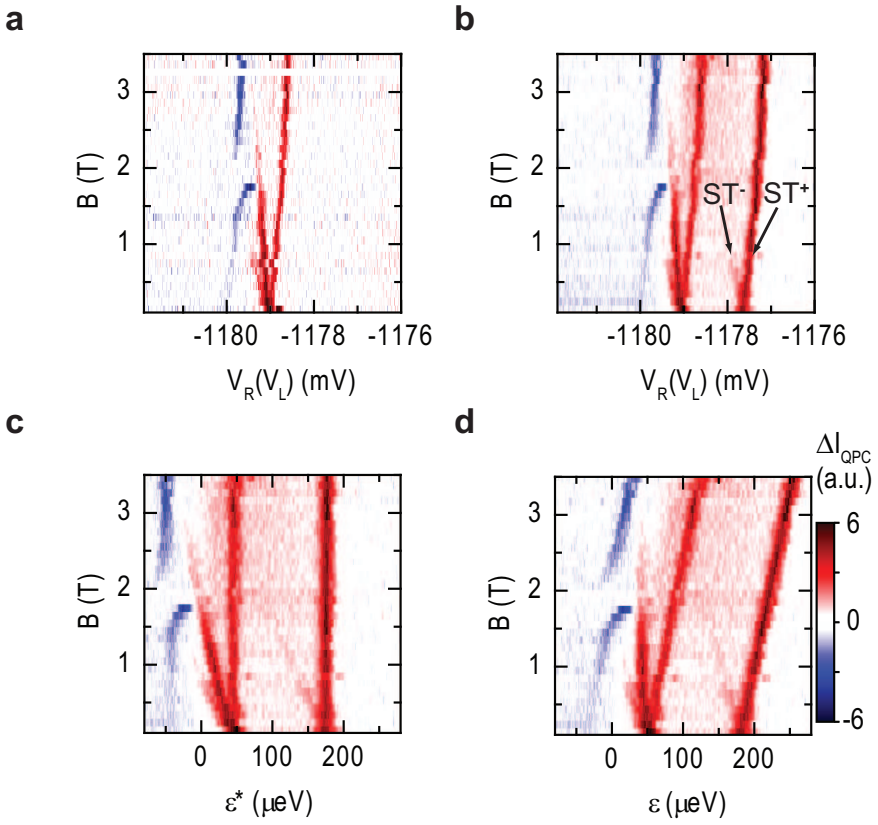
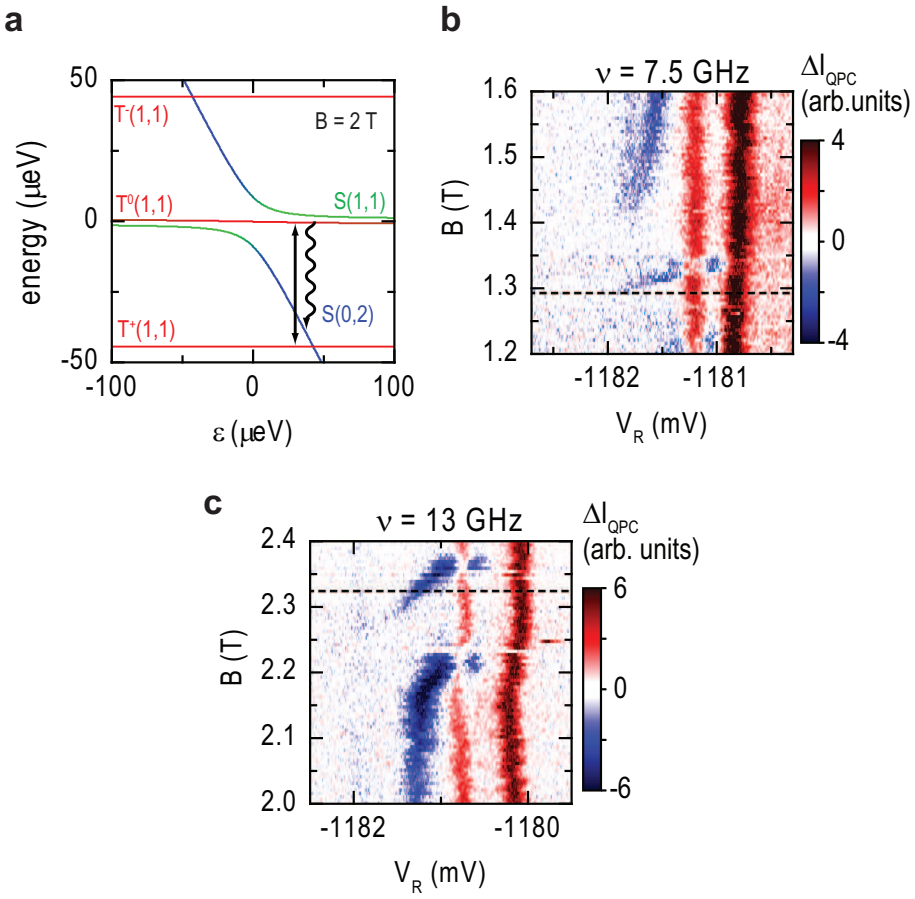
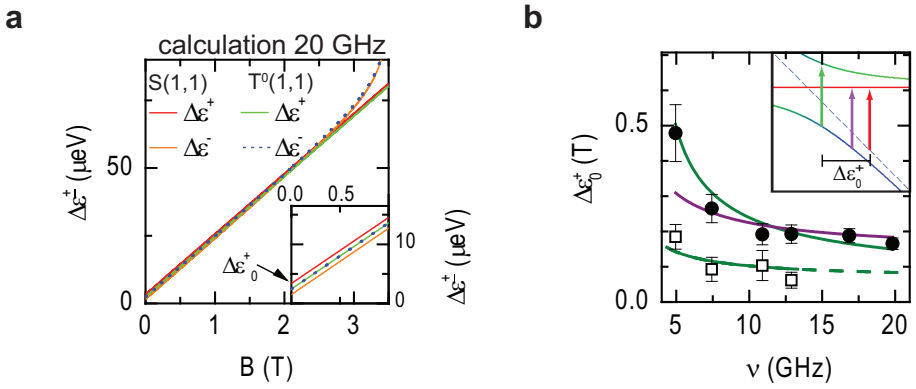


Supplementary Figure S1: **Calibration of the detuning axis for the PAT spectra. a**, PAT spectrum as measured without pulses (see inset of Fig. 1b) with a microwave frequency of 11 GHz. The voltage applied to the right side gate V_R and left side gate V_L is swept simultaneously to follow the detuning axis indicated in Fig 1b. **b**, same PAT spectrum, but with 200 ns pulses that intersperse the microwave bursts every 5 μs (see top panel of Fig. 1d). A red reference line appears at positive DD detuning, when the detuning pulse reaches exactly the ST^+ anti-crossing. Also a weak line due to mixing at the ST^- anti-crossing is observed. Both lines are marked by arrows. The more red background is due to ST^0 mixing. **c**, PAT spectrum with calibrated detuning axis. The horizontal lines of the spectrum are shifted so that the reference pulse appears at the detuning position, which equals the pulse amplitude P_ε converted into energy (see text) along the detuning axis. Here $\varepsilon^* = 0$ is equal to the ST^+ anti-crossing for every magnetic field. **d**, To convert the ε^* scale of the detuning axis into the ε scale, where $\varepsilon = 0$ equals the $S(1,1)$-$S(0,2)$ anti-crossing, the spectrum is sheared by the Zeeman-energy using the electron g-factor that is independently determined as explained in the main article. At low magnetic fields this transformation of scales is wrong and gives rise to a curvature of the spectrum, which is a function of the tunnel coupling.



Supplementary Figure S2: **Triplet resonance in the PAT spectra.** **a**, Energy eigenstates of the two-electron spin states plotted with the same color code as in Fig. 2c of the main article. Microwaves are resonant to the $T^+(1,1)/T^0(1,1)$ transition. Spontaneous relaxation from $T^0(1,1)$ to the metastable $S(0,2)$ enables the detection via ΔI_{QPC}. **b**, The PAT spectrum as measured with a microwave frequency of 7.5 GHz exhibits a nearly horizontal blue feature as visible in Fig. 2f. Also at this microwave frequency the line starts from the magnetic field, at which we expect the electron spin resonance condition to be fulfilled (black dashed line) $h\nu = g\mu(B + b_0)$. **c**, PAT spectrum recorded at 13 GHz with corresponding line indicating electron spin resonance.



Supplementary Figure S3: **Analysis of the $\Delta m = 0$ PAT resonance.** **a**, $\Delta\varepsilon^+$ and $\Delta\varepsilon^-$ calculated for the parameters $|g| = 0.382$, $t_s = 8.7$ μeV, $b_0 = 109$ mT, $\nu = 20$ GHz and two possible scenarios: The $\Delta m = 0$ line is resonant between the $S(0,2)$ ground state $S(1,1)$ (red and orange line) and $T^0(1,1)$ (green and blue line). Zoom-in for low-magnetic fields in the inset.**b**, The detuning offset $\Delta\varepsilon_0^+$ between the transition $\Delta m = +1$ for $B \to 0$ mT (red arrow in the inset) and the $\Delta m = 0$ transitions from $S(0,2)$ to $S(1,1)$ (green arrow) and to $T^0(1,1)$ (violet arrow), respectively, is plotted as a function of the microwave frequency ν. The error bars are determined from the linear extrapolation $B \to 0$ mT. Shown are offsets measured at $t_s = 8.7$ μeV (filled circles) and at $t_s \approx 4$ μeV (open rectangles). The green and violet lines are least-squares fits to the measured offsets assuming a pure $S(0,2)$ to $S(1,1)$ and to $T^0(1,1)$ transition, respectively.

Supplementary Note 1: Calibration of the detuning axis

The photon-assisted tunneling (PAT) spectra in Figs. 2a and 2b of the main article are measured both with a high energy resolution along the double dot (DD) detuning axis and over a wide external magnetic field range. In the experiment, we observe a monotonous, reproducible drift of the stable charge regions predominantly along the right side gate voltage as we change the magnetic field. Changing the voltages applied to the left and to the right side gate accordingly, we partially compensate for this drift. We then record an 11 GHz PAT spectrum as displayed in Supplementary Figure S1a. In order to precisely calibrate the detuning axis for all B, a reference signal is generated together with the PAT spectrum by interspersing the microwaves every 5 μs by a 200 ns detuning pulse with amplitude P_ε towards negative detuning (Supplementary Figure S1b). 200 ns are found to be sufficient to mix the $S(0,2)$ state entirely with the $T^+(1,1)$ state at their anti-crossing ε_{ST+}, so that a Pauli spin blocked $T^+(1,1)$ signal is observed at a detuning $\varepsilon = \varepsilon_{ST+} + P_\varepsilon$. The magnitude of P_ε is chosen such that the reference signal appears at a detuning position far away from the PAT signal. The pulses do not alter the detuning position of the PAT resonances. In addition to ST^+ mixing, mixing of the S with $T^0(1,1)$ and $T^-(1,1)$ is observed due to the pulsing. The former gives rise to a positive ΔI_{QPC} background on the left of the reference signal in Supplementary Figure S1b. The latter generates a weak second reference line that overlaps with the ST^+ reference line for $B \to 0$ T, but shifts towards negative detuning as B is increased.

In a post-processing step, we separately fit the position of the ST^+ peaks for all B and shift every row of the spectrum, such that the peaks are vertically aligned at $\varepsilon^* = P_\varepsilon$ as shown in Supplementary Figure S1c. Thus, $\varepsilon^* = 0$ is at the ST^+ anti-crossing, by definition. All data points at a given B are shifted by the same amount in detuning. This ε^*-detuning axis is well-defined but 'moves' with respect to the ε-detuning axis as a function of B, since $\varepsilon^* = \varepsilon - \varepsilon_{ST+}(B)$. The lever arm for the voltage to energy conversion is read from the voltage distance of the second and third $\Delta m = 0$ PAT line at $\nu = 11$ GHz (see Fig. 4a), which equals the photon energy $h\nu$. For the analysis of all spectra we used the same lever arm.

For better readability of the PAT spectra, we finally convert the ε^*-scale to the ε-scale found in literature, for which $\varepsilon = 0$ is defined by the $S(1,1)$ to $S(0,2)$ anti-crossing. To

do so, we additionally shift all data points at a given B by $|g|\mu_B|B|$ towards positive detuning (Supplementary Figure S1d). However, since $\varepsilon_{ST+} = |g\mu_B B|$ holds true only for $|g\mu_B B| \gg t_s$, where t_s is the interdot tunnel coupling, the detuning axis conversion fails for low magnetic fields. As a result the PAT resonance-lines bend towards positive ε for $B \to 0$ T in Supplementary Figure S1d and in the spectra shown in the main article (Fig. 2a,b).

Supplementary Note 2: Triplet spin resonance

In Fig. 2f of the main article, a PAT feature is observed that is due to a transition from the $T^+(1,1)$ ground state to the $T^0(1,1)$ excited state. The $T^0(1,1)$ state can relax via spontaneous phonon emission to the singlet bonding state, a superposition of $S(1,1)$ and $S(0,2)$. This state is metastable, since the spontaneous phonon relaxation is suppressed by the small energy difference to the $T^+(1,1)$ ground state (see Supplementary Figure S2a), which makes the transition detectable by ΔI_{QPC}. A peculiarity of the $T^+(1,1)$ to $T^0(1,1)$ resonance is its slope in the PAT spectrum, which might be a result from a gradient magnetic field along the magnetic field direction as discussed in the main article. Note that in the same detuning range, we observe also direct PAT transitions from the $T^+(1,1)$ ground state to the singlet anti-bonding state and singlet bonding state at lower and higher magnetic fields, respectively.

Here we investigate the position of the $T^+(1,1)$ to $T^0(1,1)$ resonance, as a function of the microwave frequency ν. The Supplementary Figures S2b and S2c show raw PAT spectra (without any post-processing step applied as explained above) recorded with $\nu =$ 7.5 GHz and $\nu = 13$ GHz, respectively. The dashed lines mark the magnetic field, at which the electron spin resonance condition $h\nu = g\mu_B(B + b_0)$ is fulfilled. Here we use the absolute electronic g-factor $|g| = 0.382$ and longitudinal magnetic field offset $b_0 = 109$ mT as determined by the Figs. 3a and 3b of the main article. Alternatively, we might use the $T^+(1,1)$ to $T^0(1,1)$ resonance feature to determine $|g|$. If we use the center magnetic field of this feature as the resonant field, we calculate $|g| = 0.384 \pm 0.005$ from the microwave frequency dependence in good agreement with the $|g| = 0.382$ found in Fig. 3a in the main article.

Supplementary Note 3: The $\Delta m = 0$ PAT transition

In the main article, we discuss whether the red ($\Delta m = 0, \Delta n = 1$) PAT resonance is dominantly due to a transition from the $S(0,2)$ ground state to the $S(1,1)$ state or to the $T^0(1,1)$ state. It is difficult to spectroscopically resolve these transitions, since they differ only by the exchange energy $J(\varepsilon)$. In Fig 3a of the main article, we use $\Delta\varepsilon^{\pm}(B)$, the difference in detuning between the red $\Delta m = \pm 1$ and the $\Delta m = 0$ lines, to assign the PAT resonance. Here, we support this argument by calculating the expected $\Delta\varepsilon^{\pm}(B)$ functions for both extreme scenarios: a pure $S(0,2)/S(1,1)$ and a pure $S(0,2)/T^0(1,1)$ transition. For the calculation, we use the determined $|g| = 0.382$, $t_s = 8.7$ μeV and $b_0 = 109$ mT values. The result is displayed in Supplementary Figure S3a for a microwave frequency $\nu = 20$ GHz. Obviously, the scenario of a pure singlet transition results in $\Delta\varepsilon^{+}(B) > \Delta\varepsilon^{-}(B)$, whereas $\Delta\varepsilon^{+}(B) = \Delta\varepsilon^{-}(B)$ is found for a transition to the triplet state. In the experiment, we clearly observe $\Delta\varepsilon^{+}(B) > \Delta\varepsilon^{-}(B)$ and therefore the $\Delta m = 0$ transition is dominantly a singlet transition.

In both scenarios, the $\Delta\varepsilon^{-}(B)$ increases non-linearly towards high B, since the resonance $\Delta m = -1$ becomes sensitive to the curved singlet anti-crossing at $E_z \approx h\nu - t_s$, where E_z is the Zeeman energy. This curvature is not observed in the experiment as the $\Delta m = -1$ transition fades out at $B \approx 1.5$ T. In both scenarios, the $\Delta\varepsilon^{+}(B)$ functions are linear, which holds true only if $E_z \ll t_s - h\nu$. The choice of a high ν and an appropriate B-range allows to extract $|g|$ from $\Delta\varepsilon^{+}(B)$ by a simple linear least-squares fit as demonstrated in the main article.

As a final step, we analyze the splitting of the $\Delta\varepsilon^{+}(B)$ and $\Delta\varepsilon^{-}(B)$ function quantitatively. Their offsets depend upon the exchange energy $J(\varepsilon, t_s)$, which we can experimentally vary by t_s or indirectly by ν. For $h\nu \to 2t_s$, the detuning position of the $\Delta m = 0$ resonance becomes strongly affected by J. As shown in the inset of Supplementary Figure S3b, the $S(0,2)/S(1,1)$ transition (green arrow) shifts more towards negative detuning ε than the potential $S(0,2)/T^0(1,1)$ transition (violet arrow). We measure $\Delta\varepsilon^{+}(B)$ for various ν and determine the $\Delta\varepsilon_0^{+} = \Delta\varepsilon^{+}(B \to 0)$ by a linear fit at a sufficiently high magnetic field range. This procedure turned out to be impractical with $\Delta\varepsilon_0^{-}(B)$, because the $\Delta m = -1$ line fades out at high magnetic fields. Note that the linear extrapolation of the $S(0,2)/T^{+}(1,1)$ transition to zero field (red arrow) exhibits a different detuning position than the $S(0,2)/T^0(1,1)$

transition (violet arrow), because the linear extrapolation follows the dashed blue line in the inset of Supplementary Figure S3b, i.e. the linear extrapolated detuning position is not affected by the hybridization of the singlets. The detuning position of the $S(0,2)/T^0(1,1)$ PAT transition, however, is affected by the singlet hybridization, since it lowers the energy of the initial state $S(0,2)$. In summary, the value of $\Delta\varepsilon_0^+$ is always larger than zero, but also depends upon the nature of the $\Delta m = 0$ PAT resonance.

In the inset of Supplementary Figure S3b, $\Delta\varepsilon_0^+$ for the $S(0,2)/S(1,1)$ transition is drawn. Obviously, $\Delta\varepsilon_0^+$ is considerably smaller, if the $S(0,2)/T^0(1,1)$ PAT resonance dominates over the $S(0,2)/S(1,1)$ resonance. The analysis of $\Delta\varepsilon_0^+$ is complicated by the remanence of the μmagnet, which is not exactly known, but should be smaller than the fully magnetized field of $b_0 = 109$ mT. Due to the remanence, the linear extrapolation of the $S(0,2)/T^+(1,1)$ transition towards zero external magnetic field, leaves an additional offset on $\Delta\varepsilon_0^+$. This offset, however, is independent from ν and t_s. In Supplementary Figure S3b, the extrapolated $\Delta\varepsilon_0^+$ values are plotted as a function of the microwave frequency ν for two tunnel couplings (filled circles and open squares). Fitting the filled circles with the well-known $t_s = 8.7$ μeV, we determine a reasonable fit by assuming the $\Delta m = 0$ PAT resonance to be purely singlet (green line). The fit with a potential $S(0,2)/T^0(1,1)$ transition (violet line) fails at $\nu =$ 5 GHz. The only fit parameter used here is the remanence of the μmagnet, which was found to be $\approx$ 70 mT for the fit function assuming a $S(0,2)/S(1,1)$ transition, and $\approx$ 140 mT assuming a $S(0,2)/T^0(1,1)$ transition. The latter is very unlikely, since only a maximum magnetization of $b_0 = 109$ mT was found at an external magnetic field of 2 T.

As a final check, we take $\Delta\varepsilon_0^+$ values into account, which were determined when the DD was tuned to a smaller $t_s = 4$ μeV (open squares). These data points cannot be fitted by the fit function that assumes a $S(0,2)/T^0(1,1)$ transition at all, since the $\Delta\varepsilon_0^+/g\mu_B$ values observed are already smaller than the remanence of $\approx$ 140 mT, which would stay valid for the altered tunnel coupling. Obviously, this leads to a contradiction, since all $\Delta\varepsilon_0^+$ would become negative after subtracting the remanence. Only the assumption of a purely singlet $\Delta m = 0$ PAT transition in combination with the smaller remanence of $\approx$ 70 mT, as fitted above, allows reasonable fitting. Our conclusion from the main article is therefore further supported.

Supplementary Note 4: Spin flip-tunneling mechanism

As noticed in the main text, there is a direct matrix element between the $S(0,2)$ and the $(1,1)$ triplet states. This can occur due to the nuclear spins in the barrier between the dots and due to spin-orbit (SO) interaction. We look at a toy model to examine the relative importance of these two processes. Specifically, we consider the hopping matrix element for a single electron with spin $\vec{S}$ between two wavefunctions associated with an electron on the left ($|L\rangle$) and on the right ($|R\rangle$) via the perturbation:

$$V = \frac{\hbar}{m^*\lambda_{SO}}[-\alpha(S_{\tilde{z}}p_{\tilde{y}} - S_{\tilde{y}}p_{\tilde{z}}) + \beta(p_{\tilde{y}}S_{\tilde{y}} + p_{\tilde{z}}S_{\tilde{z}})] + Av_0\sum_j \delta(r - r_j)\vec{I}^j \cdot \vec{S} \tag{S1}$$

where we have absorbed the Rashba (α) and Dresselhaus (β) terms into a single spin-orbit interaction with a characteristic spin-orbit length $\lambda_{SO} \sim 10$ μm. We recall that m^* is the effective electron mass, $A \approx 100\mu$eV, v_0 is the unit cell volume and I^j is the nuclear spin at $\vec{r}_j$.

We now wish to estimate the spin-flip tunneling for the single electron case, given by averaging over the orbital dipole:

$$\langle R|\, V\, |L\rangle \tag{S2}$$

This can be evaluated explicitly for $|L(R)\rangle = \exp(-(z \pm a/2)^2/4\sigma^2)/(2\pi\sigma^2)^{1/4}\phi(x,y)$ where z is the inter-dot axis (at an angle θ with the axis $\tilde{z}$ from the spin-orbit interaction in Eq. S1) and $\phi(x,y)$ is the transverse-longitudinal wavefunction. As tunneling occurs only along the z-axis, matrix elements with p_y are zero. We find two tunneling matrix elements:

$$t_{SO} = \frac{\hbar^2}{m^*\sigma^2}\frac{a}{4\lambda_{SO}}e^{-a^2/8\sigma^2}\vec{n}\cdot\vec{S} \tag{S3}$$

$$t_{nuc} = g\mu_B\vec{B}_{\rm nuc,f}\cdot\vec{S} \tag{S4}$$

$$\vec{n} = -\cos(\theta)[(\alpha-\beta)\cos(\theta) + (\alpha+\beta)\sin(\theta)]\hat{z} - \tag{S5}$$

$$\sin(\theta)[(\beta-\alpha)\sin(\theta) + (\alpha+\beta)\cos(\theta)]\hat{y} \tag{S6}$$

$$g\mu_B\vec{B}_{\rm nuc,f} = Av_0\sum_j |\psi_L(r_j)||\psi_R(r_j)|\vec{I}^j \tag{S7}$$

We remark that the rms value for $\vec{B}_{\rm nuc,f}$ is given by

$$g\mu_B\sqrt{|\vec{B}_{\rm nuc,f}|^2} = A\sqrt{v_0^2\sum_j |\psi_L|^2|\psi_R|^2 I(I+1)} \approx e^{-a^2/8\sigma^2}A\sqrt{v_0^2\sum_j |\psi_L|^4} \tag{S8}$$

That is, it is the rms value for a single dot, $A/\sqrt{N}$, multiplied by $\exp(-a^2/8\sigma^2)$. Also, the size of the single-particle wavefunction, σ, is related to the orbital energy scale of a single dot by $\Delta \approx \frac{\hbar^2}{m^*\sigma^2}$. Thus, the relative strength of the two tunneling terms (including spin flip) is

$$\frac{|t_{SO}|}{|t_{nuc}|} = \frac{\Delta}{A/\sqrt{N}} \frac{3|\vec{n} \times \vec{B}_{\rm ext}|}{2|B_{\rm ext}|} \frac{a}{4\lambda_{SO}} \tag{S9}$$

where N is the number of nuclear spins in a single quantum dot.

We explore briefly how this ratio varies with dot size σ and spacing a. Specifically, $\Delta\sqrt{N} \propto \sigma^{-1}$, so a larger dot reduces the strength of spin-orbit tunneling compared to hyperfine-assisted tunneling. On the other hand, increasing the distance a increases the relative strength of spin-orbit tunneling to hyperfine-assisted tunneling. Setting in $A = 100$ μeV, $N = 4 \times 10^6$, $a = 75$ nm, $\Delta = 1000$ μeV and $\lambda_{SO} = 10$ μm, we calculate the ratio of the matrix elements $\frac{|t_{SO}|}{|t_{nuc}|}$ to be ≈ 60. We remark that in external field parallel to $\vec{n}$ prevents any SO spin-charge flips. All spin charge flips occur only via t_{nuc}.

In the bases $(T^-(1,1), |\downarrow\uparrow\rangle, |\uparrow\downarrow\rangle, T^+(1,1), S^-(0,2))$ the two electron Hamiltonian that ignores the small t_{nuc} term has the form

$$\begin{pmatrix}
g\mu_B B_z & -g\mu_B \frac{\Delta B_x^\perp + i\Delta B_y^\perp}{2} & g\mu_B \frac{\Delta B_x^\perp + i\Delta B_y^\perp}{2} & 0 & -t_{SO,y} \\
-g\mu_B \frac{\Delta B_x^\perp - i\Delta B_y^\perp}{2} & -g\mu_B \Delta B^\parallel & 0 & g\mu_B \frac{\Delta B_x^\perp + i\Delta B_y^\perp}{2} & (-it_{SO,z} - t_s)/\sqrt{2} \\
g\mu_B \frac{\Delta B_x^\perp - i\Delta B_y^\perp}{2} & 0 & g\mu_B \Delta B^\parallel & -g\mu_B \frac{\Delta B_x^\perp + i\Delta B_y^\perp}{2} & (-it_{SO,z} + t_s)/\sqrt{2} \\
0 & g\mu_B \frac{\Delta B_x^\perp - i\Delta B_y^\perp}{2} & -g\mu_B \frac{\Delta B_x^\perp - i\Delta B_y^\perp}{2} & -g\mu_B B_z & -t_{SO,y} \\
-t_{SO,y} & (it_{SO,z} - t_s)/\sqrt{2} & (it_{SO,z} + t_s)/\sqrt{2} & -t_{SO,y} & -\varepsilon
\end{pmatrix} \tag{S10}$$

where, e.g., the Larmor precession frequency of an electron spin in the left dot is $g\mu_B(B_z - \Delta B^\parallel)$.