Spin current in the hydrodynamic limit. The local magnetization density $\mathcal{M}$ obeys a continuity equation

$$\partial_t \mathcal{M} + \mathbf{\Omega}_0 \times \mathcal{M} = -\nabla_j \mathcal{J}_j \quad (S1)$$

where $\mathbf{\Omega}_0 = (0,0,\Omega_0)$ with $\Omega_0$ the Larmor frequency due to an external field, and $\mathcal{J}_j$ is the spin current density. In Eq. (S1) and below, the $j$ subscript indicates a spatial direction, $\partial_t$ is a time derivative, and bold quantities denote vectors in Bloch space. The magnetization density is a vector whose magnitude is $|\mathcal{M}| = (n_\uparrow - n_\downarrow)/2$. Whereas $\mathcal{M}$ is conserved in the frame rotating with the external field, the spin current is in general not conserved. In the hydrodynamic limit, the local steady-state spin current obeys

$$\mathcal{J}_j + \mathcal{J}_j \times \mu \mathcal{M} + D_0 \nabla_j \mathcal{M} = 0 \quad (S2)$$

whose solution is

$$\mathcal{J}_j = -D_{\text{eff}} (\nabla_j \mathcal{M} - \mu \mathcal{M} \times \nabla_j \mathcal{M} - \mu^2 (\mathcal{M} \cdot \nabla_j \mathcal{M}) \mathcal{M} - \mu^2 (\mathcal{M} \cdot \mathcal{N}) \mathcal{M}) \quad (S3)$$

where $D_{\text{eff}} = D_0/(1 + \mu^2 \mathcal{M}^2)$. If $\mu = 0$, (S3) reduces to “static diffusion”, described by Fick’s law, $\mathcal{J}_j = -D_0 \nabla_j \mathcal{M}$.

Transverse and longitudinal spin currents. The general expression for the current, Eq. (S3), can be broken into two components. Writing $\mathcal{M} = \mu \mathcal{M}$ with $\mathcal{M} = |\mathcal{M}|$,

$$\nabla_j \mathcal{M} = (\nabla_j \mathcal{M}) m + (\nabla_j m) \mathcal{M} \quad (S4)$$

The first term is parallel to $\mathcal{M}$, and the second term is perpendicular in Bloch space to $\mathcal{M}$, since it is the derivative of a unit vector.

For gradients $\nabla_j \mathcal{M}$ that are parallel to $\mathcal{M}$, the resultant current is called longitudinal. Here, the second term in Eq. (S3) is zero, the third term becomes proportional to $\mu^2 \mathcal{M}^2 \nabla_j \mathcal{M}$, and thus

$$\mathcal{J}_j^\parallel = -D_0^\parallel \nabla_j \mathcal{M} \quad (S5)$$

where the $\mu$ dependence has cancelled out, and only the bare longitudinal diffusivity is left.

For gradients $\nabla_j \mathcal{M}$ that are perpendicular to $\mathcal{M}$, the resultant current is called transverse. In this case, the third term in Eq. (S3) is zero, and we have

$$\mathcal{J}_j^\perp = -D_{\text{eff}}^\perp (\nabla_j \mathcal{M} + \mu \mathcal{M} \times \nabla_j \mathcal{M}) \quad (S6)$$

where $D_{\text{eff}}^\perp = D_0^\perp/(1 + \mu^2 \mathcal{M}^2)$. Since at low temperature diffusivity can differ between the longitudinal and transverse cases, they are labelled $D_0^\parallel$ and $D_0^\perp$ respectively. In Landau Fermi liquid theory [1], $D_0^\parallel$ and other damping coefficients such as viscosity and conductivity scale as $\sim T^{-2}$ at low temperature, while $D_0^\perp$ remains anomalously finite [2–4]. Also note that Eq. (S5) is independent of $\mu$, so the transverse spin current alone is sensitive to the LR effect.

Transverse polarization current. Magnetization is observed through the density-weighted trap average of dimensionless polarization $\mathcal{M} = 2\mathcal{M}/n$ with $|\mathcal{M}| \leq 1$. Paired with $\mathcal{M}$ is a polarization current $\mathcal{J}_j = 2\mathcal{J}_j/n$. Treating only the transverse component, (S6) becomes

$$\mathcal{J}_j^\perp = -D_{\text{eff}}^\parallel \nabla_j \mathcal{M} - D_{\text{eff}}^\perp \mathcal{M} \times \nabla \mathcal{M} \quad (S7)$$

where we have ignored spatial gradients of density, and $\gamma \equiv \mu n/2$, such that $\gamma \mathcal{M} = \mu \mathcal{M}$. Equation (S7) is Eq. (1) in the main text.

Now, consider the continuity equation (S1). Assuming a static density profile, $\partial_t n = 0$ and ignoring spatial gradients of density,

$$\partial_t \mathcal{M} + \mathbf{\Omega}_0 \times \mathcal{M} = -\nabla \mathcal{J}_j^\perp \quad (S8)$$

where repeated indices are summed.

Experimental methods. Fermionic, spin-polarized $^{40}\text{K}$ atoms are cooled sympathetically with bosonic $^{87}\text{Rb}$ atoms. Initially both species are trapped in a microfabricated magnetic trap, where $^{87}\text{Rb}$ is evaporated directly. A subsequent stage of evaporative cooling is performed in a crossed-beam optical dipole trap, with $^{40}\text{K}$ atoms in the $|f = 9/2, m_f = -9/2\rangle$ state and $^{87}\text{Rb}$ atoms in the $|f = 1, m_f = 1\rangle$ state. Here, $f$ and $m_f$ denote the total angular momentum and the corresponding magnetic quantum number, respectively. At the end of cooling, residual $^{87}\text{Rb}$ atoms are removed with a resonant light pulse typically leaving $N = 4 \times 10^4$ $^{40}\text{K}$ atoms. The trap has a mean trapping frequency $\omega/2\pi = 470(20)$ Hz and an aspect ratio of 4:1:1. The Feshbach field $B_3$ and the magnetic field gradient $\nabla_3 B_3$ are applied along a tight axis of the trap. The $|-z\rangle$, $|+z\rangle$, and rf-probe states in the main text refer to the high-field states adiabatically connected to the low-field $m_f = -9/2$, $-7/2$, and $-5/2$
states of the $f = 9/2$ hyperfine manifold of the electronic ground state.

The initial temperature is determined by imaging the density distribution after time-of-flight expansion, and fitting it to a Fermi-Dirac distribution. For this measurement, the gas is fully polarized in the $|f = 9/2, m_f = -9/2\rangle$ state so no interaction corrections are required. For the data presented in Fig. 4, $N = 40\times10^4$ atoms with an initial temperature of $250(40)$ nK. The initial Fermi energy is then $E_F/T = 29(4)$ kJ, such that $(T/T_F) = 0.18(4)$, where $T_F = E_F/k_B$. The global (trap-wide) Fermi energy $E_F$ is subtly different from the local Fermi energy $\epsilon_F = \hbar^2 k_F^2/2m$ with a local $k_{F,\text{pol}} \equiv (6\pi^2 n)^{1/3}$ in a polarized gas and $k_{F,\text{mix}} = k_{F,\text{pol}}^2/2^{1/3}$ in an unpolarized gas. In both cases, $n$ is the local total number density. In Fig. 4 and in theory discussion, $k_F \equiv k_{F,\text{mix}}$ as is conventional. However, we use the set of definitions for the polarized gas to define the initial reduced temperature $(T/T_F)$. In order to tune the initial temperature for data presented in Fig. 3, we vary the loading and evaporation sequence, affecting both the absolute temperature of the gas and the total atom number. The uncertainties stated for atom numbers and temperatures are a combination of statistical and calibration uncertainties.

Since the theory to which we compare our experimental results are based on uniform models, we match the local reduced temperature $k_B T/\epsilon_{F,\text{pol}}$, where $\epsilon_{F,\text{pol}} = \hbar^2 k_{F,\text{pol}}^2/2m$. At the center of the trap, this local reduced temperature is minimal, since the density is highest. The peak density is given by $n_p = \lambda_T^{-3} f_{3/2}(z_p)$, where $\lambda_T = \sqrt{2\pi\hbar^2/mk_B T}$ is the thermal de Broglie wavelength, $f_{3/2}$ is the statistical function for a uniform Fermi gas, and $z_p$ is the local fugacity at peak density. The fugacity (and chemical potential) are constrained by the temperature and total particle number per spin component, according to $6f_{3/2}(z_p) = (k_B T/E_F)^{-3}$.

During depolarization, energy and entropy increase, changing the temperature. As discussed in [5], this ‘intrinsic heating’ effect can be calculated at unitarity using an experimentally measured density equation of state. At low temperature, the effect is strongest: a $(T/T_F)_i = 0$ cloud will heat to $(T/T_F)_f \approx 0.35$, and at our lowest $(T/T_F)_i \approx 0.20$, the final reduced temperature is $(T/T_F)_f \approx 0.40$. At high temperature, the released entropy and interaction energy becomes negligible, and the effect vanishes. The rise in $T$ is also reduced for smaller $A_0$.

We measure the polarization decay for various initial magnitudes of the transverse polarization $|M_{xy}| = \sin(\theta)$ using a $\theta - \pi - \pi/2$ pulse sequence. The first and second pulse have the same phase, but the final pulse has a variable relative phase lag. Varying this phase lag reveals the magnitude $A = |M_{xy}|$ and phase $\phi = -\arg(iM_{xy})$ of the transverse polarization in the oscillation of the relative population with. The amplitude and phase of $M_{xy}$ is then determined from a sinusoidal fit to this oscillation. We are sensitive to the relative frequency between the drive $\omega$ and the atomic frequency $\omega_0$, with a precision of roughly $1/t$, where $t$ is the hold time. For $t \gtrsim 1.5$ ms, we find that our field stability (roughly 1 kHz, or a few parts in $10^4$) is insufficient to preserve a reproducible relative phase, resulting in a randomized phase for long hold times. To avoid this, we choose a magnetic field gradient such that the timescale for depolarization is smaller than this coherence time.

The analysis technique developed utilizes both the amplitude and phase data for three separate mixing angles per value of $\gamma$ and $D_{0}^{\perp}$ reported. The mixing angles used are $\theta \approx 0.32\pi$, $\theta \approx 0.50\pi$, and $\theta \approx 0.74\pi$ such that the pulse area in the rf sequence is varied simply by changing the pulse duration after optimizing a $\pi$ pulse. To extract $\gamma$ and $D_{0}^{\perp}$, we first fit the amplitude data $A(t)$ for each mixing angle using an exponential decay with a free exponent, $A_0 \exp(-t/\tau)$, to extrapolate the initial amplitude $A_0$. Typically, this amplitude is slightly different from the desired value due to imperfect pulse area. For the mixing angles where $A_0 \neq 1$, assuming $M_{0}^2 + A_0^2 = 1$, we rescale the phase data and plot $\phi(t)$ as a function of $M_z \log[A(t)/A_0]$. A linear fit to this rescaled data provides a single value of $\gamma$ for both mixing angles. Fixing $\gamma$, we then fit the amplitude of the data set where $M_z \approx 0$ using Eq. (3) to obtain an initial guess for $D_{0}^{\perp}$. Using this value as a starting point we fit the amplitude data for all three mixing angles using Eq. (3) and minimize the residuals to extract a single $D_{0}^{\perp}$. This method is used to extract the values of $\gamma$ and $D_{0}^{\perp}$ in Fig. 4 and the closed circles in Fig. 3.

The open circles in Fig. 3 are extracted from the amplitude data alone. In this case, only the $M_z \approx 0$ mixing angle is used (we have included reanalyzed data from [5]). A single fit of the amplitude data for this mixing angle to Eq. (3) provides both $|\gamma|$ and $D_{0}^{\perp}$. However, this fit is extremely sensitive to slight nonlinearities in the amplitude and often fails for data in which $\gamma \lesssim 0.5$. The phase-sensitive method greatly reduces scatter, and is sensitive to the sign of $\gamma$.

We control the magnetic field and its gradients through a combination of magnetic field coils and microfabricated wires on an atom chip located about $200$ $\mu$m from the atoms. We tune the field $|B|$ = $B_0$ near $202.10(2)$ G, at which the two states $|\pm z\rangle$ undergo a Feshbach resonance. The field is calibrated by measuring the $|\pm z\rangle$ to $|\pm z\rangle$ transition frequency and converting the frequency to magnetic field through the Breit-Rabi formula. During a measurement at a single value of $(k_F a)^{-1}$ the field drifts by as much as $0.02$ G which translates to a systematic uncertainty of $\pm 0.02$ in $(k_F a)^{-1}$ at unitarity. We control the field gradients $\nabla_x B_3$ and $\nabla_y B_3$ by adjusting the sum and difference of small currents through parallel chip wires near the atoms, setting $\nabla_x B_3 = 0$. We calibrate the gradients by repeating spectroscopy measure-
ments on a cloud translated by piezo-actuated mirrors on the trapping beams.

Our imaging scheme allows us to simultaneously count the populations of atoms in states $|\pm z\rangle$. This is achieved with a Stern-Gerlach pulse to separate the trapped spin states, if state manipulation during time of flight in a gradient, and finally, resonant absorption imaging of the $|f = 9/2, m_f = -9/2\rangle$ to $|f = 11/2, m_f = -11/2\rangle$ cycling transition. Imaging occurs after jumping the magnetic field to 209 G, the zero crossing of the $s$-wave scattering resonance, to minimize interaction effects during time of flight.

Kinetic theory. The primary theoretical calculation to which we compare our experimental data (solid lines in Fig. 3 and Fig. 4) uses the Boltzmann equation to find the non-equilibrium time evolution of the spin distribution function in response to the applied magnetic field gradient. Collisions between fermions of unlike spin use the many-body T-matrix computed in the medium of surrounding fermions [see 6 and below]. This corresponds to the Nozières-Schmidt-Rink approximation used to compute the Fermi-liquid parameters, and it is justified as the leading-order term in a systematic large-N expansion in the number of fermion flavors [7]. The calculation does not include any finite-size effects. We calculate transport parameters in the limit of large imbalance, which corresponds to initial conditions, and suppresses superfluidity in the calculation.

We have solved the kinetic theory for a homogeneous, fully polarized Fermi gas. Previous work has shown that for temperatures $0.3T_F \lesssim T \lesssim T_F$, the transverse diffusivity $D_\perp \approx 2\hbar/m_\perp$ is robust and does not differ much from the longitudinal diffusivity $D_\parallel$ [6]. From the transverse scattering time $\tau_\perp$ (see Eq. (53) in Ref. 6) we compute both

$$D_{\perp} = \frac{\tau_\perp}{2M} \int \frac{d^3k}{(2\pi)^3} \sum_i v_{ki} v_{kj} (f_{k\uparrow} - f_{k\downarrow})$$  \hspace{1cm} \text{(S9)}$$

and the spin-rotation parameter $\gamma$. The latter is a weighted momentum average of the many-body T matrix $T(q, \omega)$ [6]:

$$\gamma = -\frac{n\tau_\perp^2}{4\hbar D_{\perp} M^2} \int \frac{d^3k_1}{(2\pi)^3} \frac{d^3k_2}{(2\pi)^3} v_{1j} (v_{1j} - v_{2j}) (f_{1\uparrow} - f_{1\downarrow}) \times (f_{2\uparrow} - f_{2\downarrow}) \text{Re}T(k_1 + k_2, \xi_{\uparrow\downarrow} + \epsilon_{\downarrow})$$ \hspace{1cm} \text{(S10)}

Here 1, 2 are shorthand for $\vec{k}_1, \vec{k}_2$, $f_{k\sigma} \equiv \langle \exp(\beta \epsilon_{k\sigma}) + 1 \rangle^{-1}$ is the Fermi distribution for $\epsilon_{k\sigma} \equiv (\hbar k)^2/(2m) - \mu_\sigma$, and $v_{kj}$ is the $j$th Cartesian component of the velocity. This result is used to calculate $\gamma$ and the “momentum-averaged” $\lambda_0 \equiv -h\gamma/(2mD_{\perp})$ shown in Figs. (3) and (4) in the main text.

We relate our numerical results for the homogeneous system to the measurements in the trapping potential using the local-density approximation. The experimental response is dominated by the center of the trap with the highest local density, and in Fig. 3 we show the diffusive of a homogeneous system of that same density and reduced temperature, as described above. The spin-rotation parameter $\gamma$ thus estimated agrees well with the experiment at unitarity. Away from unitarity, at weak coupling $|k_F a| \to 0$, $T(\vec{0}, 0) \propto a$ (see below), giving $\gamma \sim -1/a$ for a homogeneous system, and $\lambda_0 \sim a$. For $\lambda_0$, the geometry dependent diffusive scattering time $\tau_\perp$ drops out, and we find qualitative agreement between the homogeneous and trapped systems.

The T-matrix on the upper and lower branches. In the simplest ladder approximation [8],

$$T = \begin{pmatrix} 1 & g \end{pmatrix}$$

where $g = 4\pi\hbar^2 a/m$ is the bare interaction vertex and the straight lines are non-interacting Fermi Green’s functions. This leads to

$$T^{-1}(q, \omega) = \frac{m}{4\pi\hbar^2} \left[ \frac{1}{\alpha} + i \hbar \sqrt{m(\hbar\omega + \mu_\uparrow + \mu_\downarrow - k^2q^2/4m)} \right]$$

$$+ \int \frac{d^3k}{(2\pi)^3} \frac{f_{k\uparrow} + f_{k-q\downarrow}}{\hbar\omega - \xi_{k\uparrow} - \xi_{k-q\downarrow}}. \hspace{1cm} \text{(S11)}$$

Physically, $T(q, \omega)$ can be thought of as the effective interaction between a pair of spin $\uparrow$ and $\downarrow$ fermions close to the Fermi surface, with centre-of-mass momentum $\vec{q}$ and energy $\hbar\omega$ with respect to the sum $\mu_\uparrow + \mu_\downarrow$ of the chemical potentials. It is renormalized from the bare $s$-wave interaction $g$ by many-body effects. $T(q, \omega)$ is strongly peaked about $q = 0, \omega = 0$. Replacing it by this value, (S10) reduces to [2]

$$\gamma = -\frac{T(\vec{0}, 0)n\tau_\perp}{\hbar}. \hspace{1cm} \text{(S12)}$$

We have confirmed that (S12) provides a good approximation to (S10), in particular the location of the zero-crossing which indicates the onset of a pairing instability (see below). (S10) remains finite at this point, however, whereas (S12) diverges. Equation (S12) thus means that that the LR parameter is sensitive to the effective interaction between spin $\uparrow$ and $\downarrow$ fermions. Moreover, the fact that $\tau_\perp > 0$ means that the sign of $\gamma$ determines whether the effective interaction is attractive $|T(\vec{0}, 0) < 0|$ or repulsive $|T(\vec{0}, 0) > 0|$. A change in the sign of $\gamma$ [$T(\vec{0}, 0)$ from negative to positive [positive to negative] as some parameter is tuned indicates the onset of a pairing instability.

The equilibrium state of the system is characterized by chemical potentials $\mu_\sigma$ obeying the thermodynamic
condition [8]

\[ n_\sigma = \sum_k f_{k,\sigma} + \frac{\partial}{\partial \mu_\sigma} \frac{1}{\beta} \sum \ln T(\mathbf{q}, i\nu_m). \quad (S13) \]

Using these values in (S11), the effective interaction \( T(\mathbf{0}, 0) \) is attractive throughout the entire BCS–BEC crossover and, at low enough temperatures, fermions are paired up. We emphasize that this is true despite the fact that the bare interaction vertex changes sign at unitarity.

At the same time, the effective interaction need not be everywhere attractive in the excited “upper branch” state [9, 10]. In this state, the fermions are unbound, in scattering states. In the limit \( k_F a \to 0^+ \), the chemical potentials are given by their ideal gas values (when \( T \ll T_F \)) \( \mu_\sigma = \hbar^2 (6\pi^2 n_\sigma)^{2/3}/2m \), and are not negative, as happens in this limit on the lower branch (for a spin balanced mixture), where the ground state is a Bose-Einstein condensate of dimer molecules. With these ideal gas values of the chemical potentials, (S11) reduces to the expected “hard-sphere” result [9] \( T(\mathbf{0}, 0) \to 4\pi \hbar^2 a/m > 0 \) in the \( k_F a \to 0^+ \) limit of the upper branch, corresponding to a negative value for the LR parameter.

To characterize this excited metastable state outside the weak-coupling limit \( k_F a \to 0^+ \), we solve (S13) self-consistently, but remove the isolated molecular pole in the T-matrix to determine the upper branch chemical potentials [10]. In this way, we exclude the possibility of the formation of bound states, which would correspond to the equilibrium state with dimer molecules when \( k_F a > 0 \). Using the resulting values for \( \mu_\sigma \) at \( T = 0.5T_F \) and \( M_z = (n_\uparrow - n_\downarrow)/n = 0.25 \), we find that \( T(\mathbf{0}, 0) \) diverges in the vicinity of \( (k_F a)^{-1} \approx 0.8 \), becoming negative for smaller values. As noted above, this sign change indicates the onset of an instability of the upper branch towards the formation of pairs and the resulting evolution to the equilibrium lower branch [11, 12]. In agreement with previous work [11], the critical value at which this happens depends very weakly on \( M_z \).

Ignoring LFL corrections, \( D_0^+ = (2\epsilon_F/3m)\tau_\perp \) and \( \lambda_0 = -\hbar\gamma/(2mD_0^+) \). Combining these with (S12) gives

\[ \lambda_0^{-1} = \frac{4\epsilon_F}{3n} T^{-1}(\mathbf{0}, 0). \quad (S14) \]

In the inset of Fig. 4c in the main text, we show the upper branch value of \( \lambda_0^{-1} \). Its zero-crossing is consistent with both the experimental values and the sign change of \( \lambda_0 \) found with the momentum-averaged T-matrix (S10).