Supplementary Information

Section S1. Predicting drift with a passive drag model

If the lovebirds were to use no compensation, their lateral position, y , would follow passively from the lateral drag force generated by the lateral gusts. The associated lateral displacement can be calculated as the second integral of the lateral gust forces, F_v , with respect to time:

$$
y(t) = \iint_0^t F_y(\tau) d\tau.
$$

The gust drag force can be calculated based on the drag law,

$$
F_y = C_D \frac{1}{2} \rho v_{\text{gust}}^2 S,
$$

where ρ is the density of air, S is the surface area of the bird, and C_p is a drag coefficient. The drag coefficient will be a function of wing/body position and Reynolds number, but to estimate its magnitude, we can apply a drag coefficient for a sphere with the same surface area (0.01 m^2 , based on morphometric measurements of lovebirds from the same colony) and Reynolds number ($Re = \rho v_{\text{gust}} \ell / \mu \approx (1.2 \text{ kg/m}^2)(3 \text{ m/s})(0.1 \text{ m})/$ $(1.8 * 10^{-5} \text{ kg/m} * s) \approx 20,000$, where $C_D \approx 0.5$. This is a reasonable estimate, because while the birds may be more streamlined in the frontal direction, in the lateral direction, they will experience mostly bluff body pressure drag.

We can estimate the total lateral displacement based on our gust velocity maps. We know the lateral gust speed from our hotwire measurements, which give the gust speed, v_{gust} , as a function of x and y position. We know the x and y position as functions of time from the motion-tracking data. Therefore, we can estimate the expected lateral displacement over the duration of the flight (t_{flight}) – assuming there were no compensation – to be

$$
y(t_{\text{flight}}) = \iint_0^{t_{\text{flight}}} C_D \frac{1}{2} \rho v_{\text{gust}}^2(x(\tau), y(\tau)) S \, \mathrm{d}\tau.
$$

Calculating this integral numerically for all the *crosswind* cases gives $y(t_{\text{flight}}) =$ 23 cm \pm 11 cm. We present the predicted and measured lateral drift values for the lovebirds in Fig. SF1 for both the *cross* and *shear* gust condition. Considering the birds reached the goal perch in all 366 flights, sometimes to the opposite side of the perch from what would be expected based on passive lateral drift, the birds clearly compensate for the gusts. To be sure, we tested the measured perch landing locations against the predicted drift values and found they were significantly different (Fig. SF1). The only reasonable explanation for these observations is that the birds are compensating for lateral gust forces to avoid drifting during flight.

Fig. SF1. Lovebirds compensate for the lateral gusts, showing less lateral deflection (∆**) than what is predicted by a passive drag model.** A passive drag model (Section S1) predicts nonzero lateral deflection (empty circles), whereas the measured deflection was closer to zero (filled circles). When pooled by visual condition, the predicted deflections in cross and shear were significantly different from the measured deflections (Mann-Whitney U Test, $p < 0.001$)

Section S2. Derivation and scaling of the vertical head stabilization model

To determine how head stabilization scales with body size, we use the fact that body displacement depends on the net vertical force on the body. Generalist birds and bats are known to have a mostly inactive upstroke during which they do not support much of their body weight. In contrast, hummingbirds and some insects generate appreciable weight support on the upstroke. Either way, the average net vertical force (lift minus body weight) must be 0 during straight forward flight. The corresponding body

displacements are obtained by integrating the net vertical force divided by mass twice with respect to time. The head displacement follows by multiplying the body displacement scaling law with the predicted head stabilization gain, G .

For scaling purposes, which only require order of magnitude accuracy at specific scales, we modelled the weight support (lift divided by body weight), L/mq , with a periodic function of wingbeat phase (wingbeat frequency times time), ft . In our function, the ratio of upstroke-to-downstroke weight support, μ , can range from 0 to 1 (Fig. SF2A). Generalist birds and bats have a mostly inactive upstroke ($\mu = 0$). Hummingbirds support approximately 25% of their weight on the upstroke ($\mu =$ 0.25/0.75 \approx 0.33) (1). Flies support weight equally on the downstroke and upstroke (μ = 1) (2). Our lift function can be written for any μ by incorporating the unit step function, $U(t)$:

$$
\frac{L}{mg} = \frac{1}{2}\sin(2\pi ft)\left[\pi(2-\mu)U\left(\frac{1}{2}-ft\right)-\pi\mu U\left(rt-\frac{1}{2}\right)\right].
$$
 (S1)

The coefficients in the function are tuned such that the net vertical force (lift minus body weight) is zero when averaged over a wingbeat $(0 < t < 1/f)$.

Fig. SF2. The lift and vertical body displacements can be modeled as a periodic function of wingbeat phase. (A) The lift divided by body weight, L/mq , depends on the relative contribution of the upstroke, μ . Birds and bats show a mostly inactive upstroke ($\mu = 0$), whereas flies show equal lift contribution on the upstroke ($\mu = 1$). (**B**) The maximum excursion of the body during a wingbeat, $\Delta_{z, body}$, scales with gravity, q , divided by frequency squared, f^2 . The exact value of the maximum excursion can vary up to an order of magnitude depending on the relative contribution of the upstroke.

Regardless of the lift curve's shape, the magnitude of body displacement scales with gravity divided by frequency squared. The net vertical force is equal to lift minus body weight, $L - mg$. Integrating the net vertical force divided by mass twice with respect to time gives the vertical body displacement, z_{hody} .

$$
z_{\text{body}} = \iint \frac{L - mg}{m} dt dt
$$

= $\frac{g}{f^2} \Big[\Big(\frac{ft}{2} - \frac{1}{4} \Big) \Big(ft - U \Big(ft - \frac{1}{2} \Big) \Big) + \frac{1}{8\pi} \Big(\mu - 2 + 2U \Big(ft - \frac{1}{2} \Big) \Big) \sin \left(2\pi ft \right) \Big].$ (S2)

The integration constants (0 and $ft/4$) are chosen such that z_{body} is periodic with an average height of $z_{\text{body}} = 0$. The body displacement is equal to g/f^2 times a dimensionless function (everything inside the square brackets). This solution form shows that the precise shape of the lift curve is inconsequential for scaling purposes. The peak-to-peak excursion in vertical body position over a wingbeat, $\Delta_{z, body}$, depends mostly on gravity and the wingbeat frequency. The shape of the lift curve, which is driven by the upstroke-to-downstroke weight support fraction, only affects the scaling coefficient (Fig. SF2B). The scaling coefficient can change by a factor of about 9: for generalist birds and bats ($\mu = 0$), $\Delta_{z, body} = (1/4\pi)(g/f^2) \approx 0.08g/f^2$; for hummingbirds $(\mu = 0.33), \Delta_{z, body} = (1/6\pi)(g/f^2) \approx 0.05g/f^2$; for flies $(\mu = 1), \Delta_{z, body} = ((\pi 4)/32\pi$)(g/f^2) $\approx 0.01g/f^2$. For the lovebirds, the model captures the body displacement well: the model predicts $\Delta_{z,body} = 0.08g/f^2 \approx 3$ mm, and the average peak-to-peak vertical body displacement we measured was 2.2 +/- 0.5 mm.

Using our scaling prediction, we find that wingbeat-driven eye displacements are less detrimental to the vision of hummingbirds and insects. By using published frequency and mass data for flying animals (*3*), we plotted the predicted peak-to-peak eye displacements, $\Delta_{z,\text{eye}}$, across species (Fig. 3C). For these estimates, we used the same vertical head-body gain that we measured for lovebirds: $\Delta_{z,\text{eye}} = 0.64\Delta_{z,\text{body}}$. This is a conservative estimate given that lower gains (0.25) were observed in whooper swans (*4*). To compare the displacement with eye diameter, we used an established power law for eye size as a function of body mass, $d_{\text{eve}} \sim m^{0.23}$ (5), which we passed through the measured point for our lovebirds: $m = 54$ g, $d_{\text{eve}} = 5$ mm. Changing the head-body gain would shift the prediction by an order 1 value and would not change the scaling. Similarly, using a different lift curve with nonzero upstroke weight support would decrease the predicted eye displacement by up to a factor of 9. However, a lower eye displacement would only strengthen our conclusion, which is that smaller flying animals have vertical eye displacements that are tiny fractions of eye diameter. For this reason, we predict head stabilization to be more important for larger flying animals like lovebirds and whooper swans (*4*).

Section S3. Derivation of the body yaw dynamics model

We modeled the body yaw dynamics of lovebirds to better understand how they respond to crosswind. Our second-order spring-damper-inertia model has three forcing terms: an aerodynamic restoring torque inspired by our ornithopter experiments in the wind tunnel (Fig. 5D), a damping term inspired by Flapping Counter Torque (*6*), and a proportional (P) controller term inspired by the slip angles we observed (Fig. 5A,B):

$$
\ddot{\theta}_{\text{body}} = (k/I)\theta_{\text{slip}} + (k_{\text{P}}/I)\theta_{\text{neck}} - (c/I)\dot{\theta}_{\text{body}},
$$
\n(S3)

where θ_{body} is the body yaw angle with respect to the x-direction (Fig. 2B), $\theta_{\text{slip}} =$ $\theta_{wind} - \theta_{body}$ is the slip angle (effective wind angle minus body yaw angle), $\theta_{neck} =$ $\theta_{\text{head}} - \theta_{\text{body}}$ is the neck yaw angle (head yaw angle minus body yaw angle), *I* is the body's moment of inertia about the vertical axis, and k , $k_{\rm P}$, and c are constants.

To verify that the corroborated coefficients for Eqn. S3 $(k/I, k_P/I, c/I)$ are physically reasonable, we determined expected rough magnitudes of each torque in Eqn. S3.

The first torque, $k\theta_{\text{slip}}$, is a passive torque inspired by ornithopter tests (Fig. 5D) that drives the bird to orient into the wind. The magnitude of this term in our corroborated yaw dynamics model is consistent with a classical aerodynamic analysis. The restoring torque is caused by a difference in the average thrust or drag between the two wings (Fig. SF3). Aerodynamic forces like average thrust and drag scale with $(1/2)\rho u^2C_F S$ where ρ is air density, u is the incoming flow speed, C_F is an aerodynamic force coefficient, and S is the surface area of the wing. The restoring torque on the body is proportional to the difference between these forces times $\hat{r}_3 \ell$, where \hat{r}_3 is the nondimensional third moment of wing area and ℓ is wing length (shoulder to wingtip). The moment arm of the torque also includes the distance from center of mass to shoulder, but this distance is small compared to $\hat{r}_3 \ell$ and does not affect the scaling exponents. We can estimate the magnitude of restoring torques by substituting typical physical values. The density of air at sea level is about 1.2 kg/m³. The average airspeed of the lovebirds in our experiment was 2.9 m/s. A typical non-dimensional third moment of area is 0.59 (average of 31 species in Tab. S1, Hedrick, 2011). We estimate lovebird wing length and body area to be about 120 mm and 6000 mm2 based on measurements on birds from the same colony. A typical restoring torque is therefore $\hat{r}_3\ell(1/2)\rho u^2\Delta\mathcal{C}_\text{F}S\approx$ $2\Delta\mathcal{C}_F$ N*mm, where $\Delta\mathcal{C}_F$ is the difference in force coefficient between left and right wings (Fig. SF3). Order 1 values of $\Delta C_{\rm F}$ would result in torques ≈ 2 N*mm. This estimated torque magnitude is consistent with the torque magnitudes we found in our corroborated yaw model. Based on the average corroborated k value (12 \pm 5 N*mm) and a slip angle of 10° (Fig. 5A,B), the model exhibits passive torques of $k\theta_{\text{slip}} \approx 2N^*$ mm.

Fig. SF3. A passive restoring torque due to differential wing forces acts to minimize the slip angle. The restoring torque is equal to the thrust or drag on each wing times the second moment of wing area (plus the distance from body center of mass to wing shoulder).

The second torque in Eqn. S3, $k_{P} \theta_{\text{neck}}$, is a torque that the birds actively produced using wing asymmetries to navigate towards the goal perch based on proprioceptive input. The magnitude of this term in our corroborated yaw dynamics model is consistent

with the theoretical torque that Hedrick *et al*. (2009) proposed for asymmetric wing motion: $(1/8)(\gamma-1)\rho f^2 \ell^5 \chi^{-1} \hat{r}_3^3 \Phi^2 \overline{C_F} \sin (\delta) \widehat{\omega}^2$. In this equation, γ is a coefficient of asymmetry that ranges from 0 (maximum asymmetry) to 1 (symmetric wing motion), f is wingbeat frequency, χ is wing aspect ratio (shoulder-to-wingtip length divided by mean chord), Φ is stroke amplitude, C_F is an aerodynamic resultant force coefficient of the wing, δ is the spanwise rotation angle of the wing, $\hat{\omega}$ is the non-dimensional angular velocity of the wing, and the overbar denotes a stroke-averaged value. We can estimate the expected magnitude of this torque by substituting typical physical values. We estimate a wing aspect ratio of about 2.4 based on measurements on lovebirds from the same colony. From kinematic videos, we estimate the stroke amplitude for the lovebirds to be about 140° ($\Phi \approx 2.4$ rad). The wingbeat frequency for the lovebirds was 17 + 1 Hz. The remaining term, $\overline{C_E \sin(\delta) \hat{\omega}^2}$, has been estimated to be about 31.3 for typical harmonic wing motions (6). Combined with the estimates above ($\rho = 1.2 \text{ kg/m}^3$, $\hat{r}_3 =$ 0.59), the magnitude of torque due to wing asymmetry is $(1/8)(\gamma 1)\rho f^2 \ell^5 \chi^{-1} \hat{r}_3^3 \Phi^2 \overline{C_F \sin(\delta) \hat{\omega}^2} \approx 13(\gamma - 1) N^*$ mm. Based on the average corroborated $k_{\rm P}$ value of the lovebird data (8 \pm 5 N*mm) and a neck angle of 25° (Fig. 5A,B), the yaw dynamics model exhibits active torques of $k_P \theta_{\text{neck}} \approx 3 \text{N}^* \text{mm}$. The lovebirds could therefore generate the corrective torques modeled by the P-controller by using an asymmetry coefficient of about 0.77 (1 $-$ 3/13), representing mild wing asymmetry.

The third torque in Eqn. S3, $c\dot{\theta}_{\rm body}$, represents a passive aerodynamic torque that dampens yaw motion. The magnitude of this term in our corroborated yaw dynamics model is consistent with Flapping Counter Torque theory. Passive yaw damping torque is expected to equal $\rho f \ell^5 \chi^{-1} \hat{r}_3^{-3} \Phi \overline{\mathsf{C}_F\sin(\delta) \widehat{\omega}}\, \dot{\theta}_{\rm body}$ (*6*). The term $\overline{\mathcal{C}_F\sin(\delta) \widehat{\omega}} ,$ has been estimated to be about 6.0 for typical harmonic wing motions (*6*). Combined with the estimates of other terms given above ($\rho =1.2$ kg/m³, $f = 17$ Hz, $\ell = 120$ mm, $\chi = 2.4$, \hat{r}_3 = 0.59, ϕ = 2.4 rad), the damping torque evaluates to about 0.5 $\dot{\theta}_{\rm body}$ N*mm, that is, Flapping Counter Torque theory predicts $c \approx 0.5$ N*mm*s. The average corroborated c value in our yaw dynamics model was 1 ± 1 N*mm*s. The strength of the damping torque in our model is therefore the same order of magnitude as what Flapping Counter Torque theory predicts.

Section S4: Solution to the body yaw dynamics model with unit step forcing

To gain analytical insight into how body yaw responds to crosswind, we derived a closed-form solution to Eqn. S3 when the neck or slip angle experiences a sudden change. To derive a solution, it is clearer to express all terms in Eqn. S3 in terms of the body angle:

$$
I\ddot{\theta}_{\text{body}} = k(\theta_{\text{wind}} - \theta_{\text{body}}) + k_P(\theta_{\text{head}} - \theta_{\text{body}}) - c \dot{\theta}_{\text{body}}
$$

=
$$
(k + k_P) \left(\frac{k\theta_{\text{wind}} + k_P \theta_{\text{head}}}{k + k_P} - \theta_{\text{body}} \right) - c \dot{\theta}_{\text{body}}.
$$
 (S4)

This rearrangement shows that the system responds like a damped angular spring with a stiffness of $k + k_p$ and an equilibrium angle of $(k\theta_{wind} + k_p\theta_{head})/(k + k_p)$. When this equilibrium angle experiences a sudden change in the form of a unit step function $U(t)$, we find a transient response that decays with a time constant in wingbeats of $f/(2\pi f_n \zeta)$ or $2fI/c$:

Initial conditions: $\theta_{\text{body}}(0) = \dot{\theta}_{\text{body}}(0) = 0$, Forcing function (inspired by Eqn. S4): $\frac{k\theta_{wind} + k_P \theta_{head}}{k+p} = U(t)$, $k+k$ _P Solution: $\theta_{\text{body}} = U(t) \left[1 - e^{-\frac{t}{1/(2\pi f_n \zeta)}} \left[\frac{\zeta}{\sqrt{\zeta^2 - 1}} \text{Sinh}\left(2\pi f_n t \sqrt{\zeta^2 - 1} \right) + \text{Cosh}\left(2\pi f_n t \sqrt{\zeta^2 - 1} \right) \right] \right]$, (S5)

where ζ is the damping coefficient $\left(\frac{c}{\sqrt{4(k + k_{\rm P})I}}\right)$ and $f_{\rm n}$ is the resonant frequency of the undamped system $(\sqrt{(k + k_{\rm p})/I}/2\pi)$. To understand how the lovebird's body responds to changes in neck or slip angle, we can substitute the average fitted value of ζ and f_n (Fig. SF4). The time constant of the transient response in wingbeats is $f/(2\pi f_n \zeta) = 2fI/c$, which for our data is 1.4 \pm 1.0 wingbeats. The hyperbolic sine and cosine functions cause the actual response to be slightly slower; the yaw angle settles to its new equilibrium value after closer to 4 wingbeats (Fig. SF4).

Figure SF4. The model predicts a nearly critically-damped yaw angle response. A unit step function stimulates a response in the yaw angle defined by Eqn. S5. The solution is shown with the average fitted values of ζ and f_n based on the yaw angle traces of the lovebirds: $\zeta = 2.2 \pm 2.0$ and $f_n = 2.8 \pm 0.5$ Hz.

Section S5. Scaling the body yaw response time

To determine how animals may respond to crosswind more generally, we perform a scaling analysis based on the transient time constant in wingbeats we derived for lovebirds, $2fI/c$. We first analyze the effect of body size on the time constant, after which we consider the remaining variables that influence the precise value.

The time constant depends on the moment of inertia, *, and the damping* constant, c , due to Flapping Counter Torque (δ). Moment of inertia scales roughly with body mass (m) times wing length squared (3). The damping constant c is expected (6) to equal $\rho f \ell^5 \chi^{-1} \hat r_3^{-3} \varPhi \overline{\mathsf C_F \sin{(\delta)} \widehat{\omega}}$. If all non-dimensional ratios stayed constant, the damping constant c would scale with $\rho f \ell^5$. Therefore, the time constant in wingbeats, 2*f1/c*, is expected to scale with $2fm\ell^2/(\rho f\ell^5) \sim m/\rho \ell^3$. By using published mass and wing length data for flying animals (*3*), we plotted the predicted time constant across

species (Fig. 5G). Because mass scales roughly with wing length cubed (*3*), the predicted time constant has only a weak dependence on body mass (spans 2 orders of magnitude across 6 orders of magnitude in body mass). We therefore considered how air density and non-dimensional ratios would cause deviations from isometric scaling.

We now consider variables that factor into the time constant that don't depend on body size. Many birds fly annually up to 6000m above ground during migration, some species up to 9000 m ($\bar{7}$). At these altitudes, the air density (ρ) drops to about 40% of its value at sea level. The remaining variables are order 1 non-dimensional ratios, and their average values are relatively consistent. The average wing aspect ratio (y) and nondimensional third moment of area (\hat{r}_3) do not vary much among birds, bats and insects (*8*). The average wing aspect ratio was 3.3 +/- 1.0 across 319 species of insects, hummingbirds, bats, and birds. The non-dimensional third moment of area was 0.59 +/- 0.03 across 31 species of insects and birds (Tab. S1 (9)). The stroke amplitude (Φ) can theoretically vary from 0 to π , but tends to cluster near what we observed for the lovebirds; across 42 species, the average stroke amplitude was 2.1 +/- 0.6 (Tab. S1 (*9*)). Similarly, the maximal force coefficients of wings do not vary considerably amongst taxa; across hawkmoths, bumblebees, mayflies, and quails, force coefficients differed by up to about +/- 15% over a range of incidence angles representative for flight (Fig. 8 (10)). The wing path shape factor, $\overline{\sin(\delta)}\widehat{\omega}$, depends on the flight path of the wing, but it cannot vary considerably. We can test two extremes, a sine wave and a square wave, where the stroke-averaged values of the waveform are 1 and $2/\pi$, respectively $(\int_0^1 |sgn(sin(2\pi t))|dt$ and $\int_0^1 |sin(2\pi t)|dt)$. Thus, we expect the variation in $\sin(\delta)\widehat{\omega}$ to be no more than about $\pm(\pi-2)/2\pi \approx \pm 18\%$.

Based on the preceding analysis, we expect the time constant to be minimal when the damping constant, $c = \rho f \ell^5 \chi^{-1} \hat{r}_3^3 \Phi \overline{\mathsf{C}_F \text{sin}(\delta) \widehat{\omega}}$, is maximized. Conversely, we expect the time constant to be maximal when the damping constant is minimized. The damping constant, c , is up to 4 times higher than the average value for the following combined condition extremes: sea level air density (also used to calculate the average value), high aspect ratio, high non-dimensional third moment of area, high stroke amplitude, high force coefficient, and square wave wing path. Assuming the representative maximum variation is twice the reported standard deviations (preceding paragraph), we expect each variable to increase c as follows:

The cumulative effect is that the maximum c is about 4 times more than the average c $(1 \times 1.61 \times 1.10 \times 1.57 \times 1.15 \times 1.18 \approx 3.8)$.

If we now consider the opposite extreme values for all variables, the time constant would be 0.2 times the average value. We expect the highest possible time constant when c is minimized (low air density, low aspect ratio, low non-dimensional third moment of area, low stroke amplitude, low force coefficient, sin wave wing path). High altitude flight is typically performed by animals with higher aspect ratio wings, so we will also use the high aspect ratio multiplier when calculating this extreme condition. We expect each variable to decrease c as follows:

Variable Effect on average c low air density ρ (high altitude) \times 0.40 high aspect ratio $\chi \times 1.61$ low non-dimensional third moment of area $\hat{r}_3 \times 0.90$ low stroke amplitude $\Phi \times 0.43$ low force coefficient $C_F \times 0.85$ sinusoidal $\overline{\sin(\delta)}\ \overline{\hat{\omega}}$ \times 0.82

The cumulative effect is that the minimum c is about 0.2 times the average c $(0.4 \times 1.61 \times 0.90 \times 0.43 \times 0.85 \times 0.82 \approx 0.2)$.

Section S6: Motion-Capture Uncertainties

Uncertainties in the reported x/y/z/pitch/roll/yaw data result from a combination of motion-tracking uncertainty and marker positioning uncertainty. The Qualisys motiontracking software reports an RMS error for position markers of 0.1-0.3 mm. We verified this accuracy using ground truth measurements of markers on a servo arm with a known trajectory (60° rotation in 30 ms). We also considered stretching or slippage of the harness/goggles. There are two types of uncertainty to consider: constant offset ("DC") uncertainty, caused by the fabric shifting during flight or from trial to trial, and periodic ("AC") uncertainty, caused by a systematic stretch or slip during each wingbeat.

DC (constant offset) uncertainties: The harness and goggles were made snug by custom-fitting the straps to each bird (1). However, we installed the harness and goggles by eye, so variations up to around 1 mm are possible as DC sources of uncertainty between trials. Based on the dimensions of the marker constellations, this leads to a worst-case uncertainty in angles of \pm tan⁻¹(1 mm/3 cm) $\approx \pm 2^{\circ}$. We mitigated these uncertainties by analyzing *relative* motions wherever possible. For example, the coupling between neck and body angles (Fig. 2C), which illustrates the key compensatory behavior that facilitates the gust mitigation, is mostly unaffected by DC uncertainty. In contrast, the traces of absolute angles (Fig. 5) should be considered with this uncertainty in mind.

AC (time-variant) uncertainties: Based on the dimensions of the marker constellations $(\approx 3 \text{ cm}$ across), we can estimate how much stretch or slip would be required to produce the reported angle/displacement residuals. For the residual in head pitch angle

 $(2.5^{\circ}$ +/- 0.6°), the required stretch/slip is $(3 \text{ cm})^*$ tan $(2.5^{\circ}) \approx 1.3$ mm. The required stretch/slip in other directions are comparable: x, 5.0 mm; y, 1.1 mm; z, 2.2 mm; roll, 1.0 mm; yaw, 1.0 mm. If there were stretch/slip, it would lag the body/head motions by about 1/4 of a flapping cycle (0.015 s). However, we observed no phase-lagged marker motion in our high speed visual light cameras (1MP; 500 Hz). In those videos, the 5 mm markers are 5 or more pixels wide, so we were able to resolve motions at the subpixel/sub-millimeter scale (*11*) over about 1/30 of a flapping cycle (0.002 s). Therefore, we expect the RMS error (0.1-0.3 mm) to drive the AC uncertainty, causing at worst a signal-to-noise ratio in the residuals of about 3 (1.1 mm / 0.3 mm \approx 3).

Section S7: Additional Supplemental Figures

Fig. SF5. Ground speed, wind speed, body yaw, and head yaw separated by bird (BB, NN, and RR). See Figure 1D, E in the main manuscript for additional details.

Fig. SF6. Head-neck pitch coordination separated by bird (BB, NN, and RR). See Figure 1D, E in the main manuscript for additional details.

Fig. SF7. Head-neck roll coordination separated by bird (BB, NN, and RR). See Figure 2C in the main manuscript for additional details.

Fig. SF8. Head-neck yaw coordination separated by bird (BB, NN, and RR). See Figure 2C in the main manuscript for additional details.

Fig. SF9. Yaw-roll body coordination in the *cave* **separated by bird (BB, NN, and RR).** See Figure 3A in the main manuscript for additional details.

Fig. SF10. Yaw-roll body coordination in the *lake* **separated by bird (BB, NN, and RR).** See Figure 3A in the main manuscript for additional details.

Fig. SF11. Yaw-roll body coordination in the *forest* **separated by bird (BB, NN, and RR).** See Figure 3A in the main manuscript for additional details.

Fig. SF12. Yaw-roll head coordination in the *cave* **separated by bird (BB, NN, and RR).** See Figure 3B in the main manuscript for additional details.

Fig. SF13. Yaw-roll head coordination in the *lake* **separated by bird (BB, NN, and RR).** See Figure 3B in the main manuscript for additional details.

Fig. SF14. Yaw-roll head coordination in the *forest* **separated by bird (BB, NN, and RR).** See Figure 3B in the main manuscript for additional details.

Fig. SF15. High frequency head stabilization summary separated by bird (BB, NN, and RR). See Figure 4A in the main manuscript for additional details.

Fig. SF16. Flight path summaries in the *still* **environment separated by bird (BB, NN, and RR).** See Figure 5 in the main manuscript for additional details.

Fig. SF17. Flight path summaries in the *cross* **environment separated by bird (BB, NN, and RR).** See Figure 5 in the main manuscript for additional details.

Fig. SF18. Flight path summaries in the *shear* **environment separated by bird (BB, NN, and RR).** See Figure 5 in the main manuscript for additional details.

Fig. SF19. Slip and neck angles separated by bird (BB, NN, and RR). See Figure 6A,B in the main manuscript for additional details.

Fig. SF20 Gust compensation metrics show no clear trends with increasing trial number. (A) The body yaw angle soon after takeoff (20 cm right of perch) does not show patterned changes within each gust configuration as trial number increases. The lovebirds show little to no yaw preference in the *still* environment ($\theta_{body,1st} = -5^\circ \pm 12^\circ$) but are consistently turned into the first gust in the *cross* and *shear* environments ($\theta_{\text{body,1st}} = -26^{\circ} \pm 15^{\circ}$ and $-28^{\circ} \pm 15^{\circ}$, respectively). The consistent nonzero yaw angles in the *cross* and *shear* environments suggest that any anticipatory yaw behavior in the *still* environment is small compared to yaw responses in the gusts. **(B)** The maximum body yaw speed as lovebirds enter the second half of the arena (max value for -0.4 $m < x < 0.8$ m; area between dashed lines shown on qust maps) does not show patterned changes within each gust configuration as trial number increases. The lovebirds consistently reorient into the second gust in the *shear* case ($\dot{\theta}_{\rm body,2nd} = 258^{\circ}/s \pm 91^{\circ}/s$). In contrast, they do not reorient in *still* or *cross* trials ($\dot{\theta}_{body,2nd} = -28^{\circ}/s \pm 94^{\circ}/s$ and $-31^{\circ}/s \pm 104^{\circ}/s$, respectively), even directly following *shear* trials, showing that they are not anticipating a second gust in those cases. Data separated by bird are shown in Fig. SF 23, 24.

Fig. SF21. The body yaw angle soon after takeoff (cm right of takeoff perch) does not show patterned changes with increasing flight number. We observe no decay of transient behaviors, suggesting that lovebirds are not learning to anticipate the first gust from flight to flight. The number of flights that share the same flight number in a batch (with Flight $# = 1, 2, 3...$) is shown to the right of each data set. The number of flights, n , starts at 9 (3 birds x 3 visual conditions) and then reduces, because the total number of trials varies per bird and condition. In all cases, n is 9 for the first 5 Flight #'s and thus only partially plotted to avoid clutter. Data separated by bird are shown in Fig. SF 25-27.

Fig. SF22. The maximum body yaw speed as lovebirds enter the second half of the arena (-0.4 m < < 0.8 m) does not show patterned changes with increasing flight number. We observe no decay of transient behaviors, suggesting that lovebirds are not learning to anticipate the second gust from flight to flight. The number of flights that share the same flight number in a batch (with Flight $# = 1, 2, 3...$) is shown to the right of each data set. The number of flights, n , starts at 9 (3 birds x 3 visual conditions) and then reduces, because the total number of trials varies per bird and condition. In all cases, n is 9 for the first 5 Flight #'s and thus only partially plotted to avoid clutter. Data separated by bird are shown in Fig. SF 28-30.

Fig. SF23. Figure SF20 $(\theta_{body,1st})$ separated by bird (BB, NN, and RR).

Fig. SF24. Figure SF20 ($\dot{\theta}_{\text{yaw,2nd}}$) separated by bird (BB, NN, and RR).

Fig. SF25. Figure SF21 (still) separated by bird (BB, NN, and RR). RR shows a negative yaw bias soon after takeoff $(\theta_{\text{body,1st}} = -16^{\circ} \pm 7^{\circ})$. However, his behavior doesn't decay with subsequent trials, suggesting his yaw bias is preferential rather than learned. The number of flights, n , starts at 3 (3 visual conditions) and then reduces, because the total number of trials varies per condition. In all cases, n is 3 for the first 5 Flight #'s and thus only partially plotted to avoid clutter.

Fig. SF26. Figure SF21 (cross) separated by bird (BB, NN, and RR). The number of flights, n, starts at 3 (3 visual conditions) and then reduces, because the total number of trials varies per condition. In all cases, n is 3 for the first 5 Flight $#$'s and thus only partially plotted to avoid clutter.

Fig. SF27. Figure SF21 (shear) separated by bird (BB, NN, and RR). The number of flights, n, starts at 3 (3 visual conditions) and then reduces, because the total number of trials varies per condition. In all cases, n is 3 for the first 5 Flight $#$'s and thus only partially plotted to avoid clutter.

Fig. SF28. Figure SF22 (still) separated by bird (BB, NN, and RR). The number of flights, n, starts at 3 (3 visual conditions) and then reduces, because the total number of trials varies per condition. In all cases, n is 3 for the first 5 Flight $#$'s and thus only partially plotted to avoid clutter.

Fig. SF29. Figure SF22 (cross) separated by bird (BB, NN, and RR). The number of flights, n, starts at 3 (3 visual conditions) and then reduces, because the total number of trials varies per condition. In all cases, n is 3 for the first 5 Flight $#$'s and thus only partially plotted to avoid clutter.

Fig. SF30. Figure SF22 (shear) separated by bird (BB, NN, and RR). The number of flights, n, starts at 3 (3 visual conditions) and then reduces, because the total number of trials varies per condition. In all cases, n is 3 for the first 5 Flight $#$'s and thus only partially plotted to avoid clutter.

Fig. SF31. Roll and pitch are not coupled in the way that yaw and roll are coupled (Fig. 3).

Fig. SF32. Yaw and pitch are not coupled in the way that yaw and roll are coupled (Fig. 3).

Fig. SF33. Saccade magnitudes are larger in the cave environment. (**A**) The number of saccades per flight was comparable across all conditions. Pitch saccades were rare; only one saccade was recorded in the *forest-shear* condition and none were recorded in the *cave-shear* condition. (**B**) Yaw and roll saccade magnitudes were slightly higher in the cave environment (roll, 11.0 \pm 4.2°; yaw, 14.6 \pm 5.6°) compared to the lake and forest environments (roll, $4.9 \pm 3.0^{\circ}$; yaw, $7.9 \pm 4.2^{\circ}$). No differences were observed across wind conditions. (**C**) Saccade speeds were comparable across all conditions (310°/s +/- 130°/s).

Fig. SF34. Lovebirds increased airspeed to maintain nearly constant ground speed in *gust* **and** *shear* **conditions.** Arrow lengths are mean speeds across all flights and error bars show standard deviations.

Fig. SF35. Lateral airspeed could be controlled based on neck angle. Neck angle, θ_{neck} , and lateral airspeed (lateral ground speed, \dot{y} , minus local lateral wind speed, v_{wind}) are linearly correlated for each lovebird. Colored dots, observed data; black lines, linear fit. Linear fits $(y - v_{wind} = \alpha \theta_{neck} + \beta)$: BB, $\alpha =$ 0.09 m/°s, β = 0.02 m/s, R² = 0.77; NN, α = 0.07 m/°s, β = 0.29 m/s, R² = 0.76; RR, α = 0.08 m/°s, β = -0.23 m/s, R² = 0.52.

Fig. SF36. Lovebirds can use proprioceptive cues to effectively compensate for lateral gusts. Regardless of visual condition, lovebirds effectively compensate for crosswinds ($0 =$ full drift; $1 =$ full compensation) compared to migratory species (table of avatar species in Tab. ST1). We calculated crosswind compensation by comparing lateral drift speed to crosswind speed. Lateral drift speed was calculated as the numerical derivative of y position (DerivativeFilter, Mathematica, at a Gaussian scale of one wingbeat period, 59 ms). We calculated ground speeds as the numerical derivative of x position using the same technique (Fig. 1D shows mean +/- STD). The slope of drift speed versus crosswind speed was calculated for each flight by fitting a line through zero crosswind speed and zero drift speed. We report crosswind compensation as 1 minus this slope (0 implies full drift; 1 implies full compensation).

Fig. SF37. Figures 4C (A) and 5G (B) are shown with all species data. Species data are derived from literature (*3*) for evaluating the scaling trend (blue, insects; yellow, hummingbirds; red, other birds). Table of avatar species in Tab. ST2.

Fig. SF38. CAD renderings of flight arena. (**A**) Isometric view. (**B**) Side view. The arena is shown in the *gust* wind condition (both wind generators on the same side). The black cloth used to cover the arena is not shown so that the location of the cameras and other components of the setup can be seen.

Fig. SF39. CAD rendering of wind generators. Wind generators were made from six fans arranged in a wooden housing. (**A**) A front view shows the flow straightening honeycomb. (**B**) A back view shows the six fans behind the honeycomb.

Fig. SF40. Reflective marker gear. (**A**) Front and side CAD rendering of the 3D-printed goggle sockets. (**B, C**) Reflective marker constellations were mounted on the custom-fitted goggle sockets and harness. (**D**) The custom-fit provided a snug fit while lovebirds flew in the wind arena.

Fig. SF41. Flight kinematics in one direction were mirrored. The arena was symmetrical, so the kinematics from flights in direction B were mirrored. By mirroring the flights, we made all flights occur in the positive x direction to facilitate averaging. The mirroring transformations were set up so that in the *cross* case, crosswinds came from the right of the lovebirds, and in the *shear* case, crosswinds came from the right and then the left of the lovebirds. In the *cross* environment, Direction B, crosswinds originated from the bird's left rather than right, so x and y were flipped, then y was flipped again, and the net effect was no change in y.

Section S8. Supplemental Tables

Tab. ST1. Gust compensation values taken from literature for Fig. SF35:

*Not shown in Fig. 5H

Tab. ST2. Mass and wing length values used for Fig. SF35:

***** Differs from the 54 g reported in main text, because this mass and wing length are based on two lovebirds from the same colony (masses $= 45.4$ g, 46.0 g).

Tab. ST3. Trial ordering. We first conducted experiments in the *lake* condition, then the *forest* condition, then the *cave* condition. Within each visual condition, we pseudorandomly varied the gust conditions depending on which of the three lovebirds were flying.

Tab. ST4. Detailed trial ordering

We analyzed 366 of the flights saved over 10 days of experiments.

* flights excluded from analysis due to clear misidentifications and/or blank datasets

Section S9. Supplemental Videos

Video S1: BB's first flight in the *cave-still* **environment.** Footage from an infrared camera above the arena shows bird BB flying from left to right. There were no gust generators in this trial and no visual cues except a faint point light source behind the target perch. The blue and red arrows show the orientation of the body and head based on the motion tracking data.

Video S2: BB's first flight in the *cave-cross* **environment.** Footage from an infrared camera above the arena shows bird BB flying from left to right. White arrows show the location of two gust generators producing lateral gusts in the same direction. There were no visual cues in this trial except a faint point light source behind the target perch. The blue and red arrows show the orientation of the body and head based on the motion tracking data.

Video S3: BB's first flight in the *cave-shear* **environment.** Footage from an infrared camera above the arena shows bird BB flying from left to right. White arrows show the location of two gust generators producing lateral gusts in opposing directions. There were no visual cues in this trial except a faint point light source behind the target perch. The blue and red arrows show the orientation of the body and head based on the motion tracking data.

Video S4: A bird-scale ornithopter passively reorients into the wind, minimizing slip angle. Footage from three angles (top, left, right) of the ornithopter as it is released from a slip angle of 90 degrees in a 4 m/s wind in the wind tunnel. The stroke plane angle is 25 degrees.

References

- 1. D. R. Warrick, B. W. Tobalske, D. R. Powers, Aerodynamics of the hovering hummingbird. *Nature*. **435**, 1094 (2005).
- 2. F. T. Muijres, M. J. Elzinga, J. M. Melis, M. H. Dickinson, Flies evade looming targets by executing rapid visually directed banked turns. *Science (80-.).* **344**, 172–177 (2014).
- 3. C. H. Greenewalt, Dimensional relationships for flying animals (1962).
- 4. A. E. Pete, D. Kress, M. A. Dimitrov, D. Lentink, The role of passive avian head stabilization in flapping flight. *J. R. Soc. Interface*. **12**, 20150508 (2015).
- 5. M. de L. Brooke, S. Hanley, S. B. Laughlin, The scaling of eye size with body mass in birds. *Proc. R. Soc. London B Biol. Sci.* **266**, 405–412 (1999).
- 6. T. L. Hedrick, B. Cheng, X. Deng, Wingbeat time and the scaling of passive rotational damping in flapping flight. *Science (80-.).* **324**, 252–255 (2009).
- 7. S. Åkesson, A. Hedenström, How migrants get there: migratory performance and orientation. *AIBS Bull.* **57**, 123–133 (2007).
- 8. D. Lentink, M. H. Dickinson, Rotational accelerations stabilize leading edge vortices on revolving fly wings. *J. Exp. Biol.* **212**, 2705–2719 (2009).
- 9. T. L. Hedrick, Damping in flapping flight and its implications for manoeuvring, scaling and evolution. *J. Exp. Biol.* **214**, 4073–4081 (2011).
- 10. J. R. Usherwood, C. P. Ellington, The aerodynamics of revolving wings II. Propeller force coefficients from mayfly to quail. *J. Exp. Biol.* **205**, 1565–1576 (2002).
- 11. H. Nobach, M. Honkanen, Two-dimensional Gaussian regression for sub-pixel displacement estimation in particle image velocimetry or particle position estimation in particle tracking velocimetry. *Exp. Fluids*. **38**, 511–515 (2005).
- 12. N. Sapir, N. Horvitz, M. Wikelski, R. Avissar, R. Nathan, Compensation for lateral drift due to crosswind in migrating European Bee-eaters. *J. Ornithol.* **155**, 745–753 (2014).
- 13. J. D. McLaren, J. Shamoun-Baranes, C. J. Camphuysen, W. Bouten, Directed flight and optimal airspeeds: homeward-bound gulls react flexibly to wind yet fly slower than predicted. *J. avian Biol.* **47**, 476–490 (2016).
- 14. J. W. Chapman *et al.*, Adaptive strategies in nocturnally migrating insects and songbirds:

contrasting responses to wind. *J. Anim. Ecol.* **85**, 115–124 (2016).

- 15. Y. Goto, K. Yoda, K. Sato, Asymmetry hidden in birds' tracks reveals wind, heading, and orientation ability over the ocean. *Sci. Adv.* **3**, e1700097 (2017).
- 16. R. H. G. Klaassen, M. Hake, R. Strandberg, T. Alerstam, Geographical and temporal flexibility in the response to crosswinds by migrating raptors. *Proc. R. Soc. London B Biol. Sci.*, rspb20102106 (2010).
- 17. B. M. Van Doren, K. G. Horton, P. M. Stepanian, D. S. Mizrahi, A. Farnsworth, Wind drift explains the reoriented morning flights of songbirds. *Behav. Ecol.* **27**, 1122–1131 (2016).
- 18. J. Vidal-Mateo *et al.*, Wind effects on the migration routes of trans-Saharan soaring raptors: geographical, seasonal, and interspecific variation. *Curr. Zool.* **62**, 89–97 (2016).
- 19. R. B. Srygley, R. Dudley, Optimal strategies for insects migrating in the flight boundary layer: mechanisms and consequences. *Integr. Comp. Biol.* **48**, 119–133 (2008).
- 20. C. M. Lessells, G. N. Ovenden, Heritability of wing length and weight in European bee-eaters (Merops apiaster). *Condor*. **91**, 210–214 (1989).
- 21. M. Shirai, Y. Niizuma, K. Tsuchiya, M. Yamamoto, N. Oka, Sexual size dimorphism in Streaked Shearwaters Calonectris leucomelas. *Ornithol. Sci.* **12**, 57–62 (2013).
- 22. J. Ferguson-Lees, D. A. Christie, *Raptors of the world* (Houghton Mifflin Harcourt, 2001).
- 23. R. Spaar, Flight strategies of migrating raptors; a comparative study of interspecific variation in flight characteristics. *Ibis (Lond. 1859).* **139**, 523–535 (1997).
- 24. C. Pennycuick, Wingbeat frequency of birds in steady cruising flight: new data and improved predictions. *J. Exp. Biol.* **199**, 1613–1618 (1996).