The Astrophysical Journal, 696:1230–1240, 2009 May 10 doi:10.1088/0004-637X/696/2/1230

EXOPLANETARY SPIN–ORBIT ALIGNMENT: RESULTS FROM THE ENSEMBLE OF ROSSITER–MCLAUGHLIN OBSERVATIONS

Daniel C. Fabrycky[1,3] and Joshua N. Winn[2]
[1] Harvard-Smithsonian Center for Astrophysics, 60 Garden St., MS-51, Cambridge, MA 02138, USA; daniel.fabrycky@gmail.com
[2] Department of Physics, and Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
Received 2008 December 19; accepted 2009 February 6; published 2009 April 23

ABSTRACT

One possible diagnostic of planet formation, orbital migration, and tidal evolution is the angle ψ between a planet's orbital axis and the spin axis of its parent star. In general, ψ cannot be measured, but for transiting planets one can measure the angle λ between the sky projections of the two axes via the Rossiter–McLaughlin effect. Here, we show how to combine measurements of λ in different systems to derive statistical constraints on ψ. We apply the method to 11 published measurements of λ, using two different single-parameter distributions to describe the ensemble. First, assuming a Rayleigh distribution (or more precisely, a Fisher distribution on a sphere), we find that the peak value is less than 22° with 95% confidence. Second, assuming that a fraction f of the orbits have random orientations relative to the stars, and the remaining fraction $(1-f)$ are perfectly aligned, we find $f < 0.36$ with 95% confidence. This latter model fits the data better than the Rayleigh distribution, mainly because the XO-3 system was found to be strongly misaligned while the other 10 systems are consistent with perfect alignment. If the XO-3 result proves robust, then our results may be interpreted as evidence for two distinct modes of planet migration.

Key words: celestial mechanics – methods: statistical – planetary systems – stars: rotation

1. INTRODUCTION

In a planetary system, the angle ψ between the orbital axis and the stellar rotation axis may provide clues about the processes that sculpt planetary orbits. As an example, the angle between the Sun's rotation axis and the north ecliptic pole is $\psi_\odot = 7\overset{\circ}{.}15$ (see, e.g., Beck & Giles 2005). The smallness of $\psi_\odot$, along with the small mutual inclinations between planetary orbits, is prima facie evidence for formation in a spinning disk. The smallness of $\psi_\odot$ has also been used to constrain the properties of any "Planet X" or solar companion star (Goldreich & Ward 1972), and to place upper bounds on violations of Lorentz invariance (Nordtvedt 1987). That $\psi_\odot$ is not even closer to zero has been interpreted as evidence for an early close encounter with another star (Heller 1993) or a nonaxisymmetric, "twisting" collapse of the Sun's parent molecular cloud (Tremaine 1991).

For exoplanets, it has been recognized that ψ is a possible diagnostic of theories of planet migration. Some of the mechanisms that have been proposed to produce close-in giant planets would preserve an initial spin–orbit alignment (Lin et al. 1996; Ward 1997; Murray et al. 1998), while others would produce at least occasionally large misalignments (Ford et al. 2001; Yu & Tremaine 2001; Papaloizou & Terquem 2001; Terquem & Papaloizou 2002; Marzari & Weidenschilling 2002; Thommes & Lissauer 2003; Wu et al. 2007; Fabrycky & Tremaine 2007; Chatterjee et al. 2008; Jurić & Tremaine 2008; Nagasawa et al. 2008). Tides raised on the star are not expected to play a major role in altering ψ (Winn et al. 2005), but it is possible that coplanarization is more efficient than expected (Mazeh 2008; Pont 2008). For example, if a hot Jupiter migrated inward before its host star contracted onto the main sequence, the distended stellar envelope could produce more pronounced tidal effects (Zahn & Bouchet 1989; Dobbs-Dixon et al. 2004).

Independent of the interpretation, the angle ψ is a fundamental geometric property, and for this reason alone it is worth seeking empirical constraints on ψ whenever possible. We regard ψ to be on a par with the semimajor axis and the eccentricity: all of them are basic orbital parameters for which accurate and systematic measurements can lead to revealing discoveries and statistical constraints on exoplanetary system architectures.

For a generic exoplanet discovered by the Doppler method, no information about spin–orbit alignment is available. For transiting exoplanets, one may exploit a spectroscopic phenomenon known as the Rossiter–McLaughlin (RM) effect. During a transit, the planet hides part of the rotating stellar disk and causes the stellar spectral lines to be slightly distorted. The distortion is usually manifested as an "anomalous" Doppler shift of order $\Delta V = -(R_p/R_s)^2 V_p$, where R_p/R_s is the planet-to-star radius ratio and V_p is the projected rotation rate of the hidden portion of the stellar photosphere (Ohta et al. 2005; Giménez 2006; Gaudi & Winn 2007). Because photometric observations give a precise and independent measure of $(R_p/R_s)^2$, spectroscopic monitoring of ΔV reveals $V_p(t)$, thereby allowing one to chart the planet's trajectory relative to the sky-projected stellar rotation axis.

An important limitation of the RM technique is that it is sensitive only to the angle λ between the *sky projections* of the orbital and rotational axes.[4] We refer to ψ as the spin–orbit angle and to λ as the projected spin–orbit angle. In general the line-of-sight component of the stellar rotation axis is unknown. When $|\lambda|$ is small, then $|\lambda|$ is a lower limit on ψ. (The situation is a bit more complex for large $|\lambda|$, as is shown in this paper.) While the finding of a large value of λ implies that there is a large spin–orbit misalignment, with consequent implications for the system's dynamical history, the finding of a small value of λ has a more ambiguous interpretation.

[3] Michelson Fellow.

[4] Strictly speaking, the RM signal depends more directly on the angle λ' between the transit chord and the sky-projected stellar rotation axis. For an eccentric orbit, this angle may differ from the angle λ between the sky projections of the orbital axis and the stellar rotation axis. It is straightforward to relate λ to λ' when the orbital eccentricity and argument of pericenter are known from the Doppler orbit of the star. For the systems considered in this paper, the maximum difference between λ and λ' is approximately 2° (for HAT-P-2) and is in all cases much smaller than the measurement uncertainty in λ.

Table 1
Summary of RM Measurements

Exoplanet	Projected Spin–Orbit Angle λ (deg)	References
HD 189733b	-1.4 ± 1.1	1
HD 209458b	0.1 ± 2.4	2, 3, 4, 5, 6*
HAT-P-1b	3.7 ± 2.1	7
CoRoT-Exo-2b	7.2 ± 4.5	8
HD 149026b	1.9 ± 6.1	9, 6*
HD 17156b	9.4 ± 9.3	10, 11*
TrES-2b	-9.0 ± 12.0	12
HAT-P-2b	1.2 ± 13.4	13*, 14
XO-3b	70.0 ± 15.0	15
WASP-14b	-14.0 ± 17.0	16
TrES-1b	30.0 ± 21.0	17

References: (1) Winn et al. (2006); (2) Queloz et al. (2000); (3) Bundy & Marcy (2000); (4) Wittenmyer et al. (2005); (5) Winn et al. (2005); (6) J. N. Winn & J. A. Johnson (2009, in preparation); (7) Johnson et al. (2008); (8) Bouchy et al. (2008); (9) Wolf et al. (2007); (10) Narita et al. (2008); (11) Cochran et al. (2008); (12) Winn et al. (2008); (13) Winn et al. (2007b); (14) Loeillet et al. (2008); (15) Hebrard et al. (2008); (16) Joshi et al. (2009); (17) Narita et al. (2007). Where multiple references are given, the quoted result is taken from the starred reference.

The way to overcome this limitation and draw general inferences about spin–orbit alignment is to consider the ensemble of RM results. The situation is similar to the early days of Doppler planet detection. Doppler measurements give only $M_p \sin i_o$, where M_p is the planet's mass and i_o is the orbital inclination. When there were only a few detections, it was impossible to draw firm conclusions about the mass distribution of the planets, or even to be completely certain that they were planets and not brown dwarfs in face-on orbits (Mazeh et al. 1998; Stepinski & Black 2000). However, once tens of systems were known with precise measurements of $M_p \sin i_o$, the planetary mass distribution came into focus (Jorissen et al. 2001; Zucker & Mazeh 2001; Tabachnik & Tremaine 2002), under the reasonable assumption that the orbits are randomly oriented in space.

There are now 11 exoplanetary systems for which RM measurements have been reported. The results are summarized in Table 1. The time is ripe to undertake an analogous study of the statistical constraints on spin–orbit alignment. It is worth drawing attention to the entries for HD 209458 and HD 149026, for which we are using updated determinations of λ by J. N. Winn & J. A. Johnson (2009, in preparation). For HD 209458, the revision is due to an improved analysis method taking into account correlated errors in the radial-velocity data. For HD 149026, a better transit light curve led to enhanced precision in λ.

The results in Table 1 are easily summarized: each individual system besides XO-3 is consistent with perfect spin–orbit alignment within 2σ. However, it is not obvious what exactly is ruled out by these results, or what we may conclude about the "typical" value of ψ among the transiting planets. The purpose of this paper is to provide a statistical framework for understanding statistical constraints on spin–orbit alignment that follow from RM observations, and apply it to the current data. We are primarily concerned with the empirical information about the distribution of ψ, rather than the interpretation in terms of migration theories or tidal effects, which will be the subject of future studies.

This paper is organized as follows. The geometry of this problem is defined in Section 2. The relevant probability distributions for individual systems are derived in Section 3.

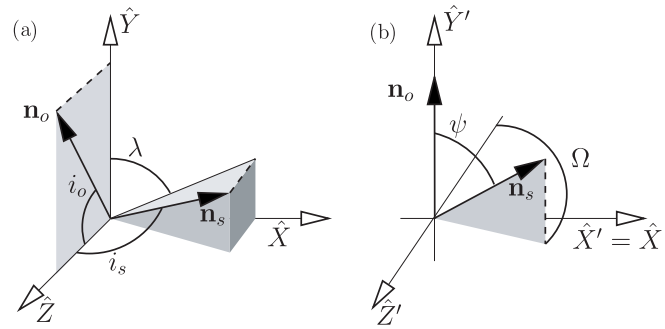

Figure 1. Two useful coordinate systems. (*a*) An "observer-oriented" coordinate system, in which $\hat{Z}$ points toward the observer and the X–Y-plane is the sky plane. (*b*) An "orbit-oriented" coordinate system in which the $\hat{Y}'$-axis is the orbital axis, and the $\hat{X}'$–$\hat{Z}'$-plane is the orbital plane. The two coordinate systems are related by a rotation of $\pi/2 - i_o$ about the $\hat{X}=\hat{X}'$-axis.

Constraints on ψ based on an ensemble of RM observations are discussed in Section 4. Section 5 gives a summary of the results, a discussion of some limitations of our analysis, and some suggestions for future work.

2. SPHERICAL GEOMETRY OF THE ROSSITER–MCLAUGHLIN EFFECT

Let $\mathbf{n}_o$ and $\mathbf{n}_s$ denote the unit vectors in the directions of the orbital angular momentum and stellar rotational angular momentum, respectively. The angle between $\mathbf{n}_o$ and $\mathbf{n}_s$ is the "spin–orbit angle," denoted by ψ. This is presumably the only angle of intrinsic physical significance in this problem, possibly bearing information about the initial condition for planet formation, the endpoint of planet migration, or the result of tidal evolution. However, ψ is not directly measurable, and we must introduce some other angles.

Figure 1 shows two useful coordinate systems. In the "observer-oriented" coordinate system shown in the left panel, $\hat{Z}$ points at the observer, $\hat{X}$ points along the line of nodes of the planetary orbit, and $\hat{Y}$ completes a right-handed triad. The ascending node of the planet (the location where the planet pierces the sky plane with $\dot{Z} > 0$) is at $X < 0$. In this coordinate system, $\mathbf{n}_o$ is in the YZ-plane and is specified by the inclination angle $i_o = \arccos(\mathbf{n}_o \cdot \hat{Z})$, which ranges from 0 to π. Specifying $\mathbf{n}_s$ requires two angles, the inclination angle $i_s = \arccos(\mathbf{n}_s \cdot \hat{Z})$ and an azimuthal angle, which by the convention of Ohta et al. (2005) we take to be λ, measured clockwise on the sky from the Y-axis to the sky projection of $\mathbf{n}_s$. In summary,

$$\mathbf{n}_o = \hat{Y} \sin i_o + \hat{Z} \cos i_o, \tag{1}$$

$$\mathbf{n}_s = \hat{X} \sin i_s \sin \lambda + \hat{Y} \sin i_s \cos \lambda + \hat{Z} \cos i_s. \tag{2}$$

For a transiting planet, i_o is measurable via transit photometry, λ (the "projected spin–orbit angle") is measurable via the RM effect. Usually there is no direct measurement of i_s, although it is possible to constrain i_s via asteroseismology (Gizon & Solanki 2003) or by combining estimates of the stellar radius, stellar rotation period, and projected rotational velocity (see, e.g., Winn et al. 2007a). By symmetry, a configuration (i_o, λ) cannot be distinguished from a different configuration $(\pi - i_o, -\lambda)$. Because of this degeneracy, we restrict i_o to the range $[0, \pi/2]$ and allow λ to range from $-\pi$ to $+\pi$. A positive (negative) value for λ means that, from the observer's perspective, the projected stellar spin axis is rotated clockwise (counterclockwise) with

respect to the projected orbit normal. Values of $|\lambda|$ greater than $\pi - i_o$ correspond to retrograde orbits.

In the "orbit-oriented" coordinate system $X'Y'Z'$ shown in the right panel of Figure 1, we define $\hat{Y}' \equiv \mathbf{n}_o$. This system is related to XYZ by a rotation of $\pi/2 - i_o$ about the X-axis. We define ψ and Ω as the polar and azimuthal angles of $\mathbf{n}_s$ in this system, namely,

$$\mathbf{n}_s = \hat{X}' \sin\psi \sin\Omega + \hat{Y}' \cos\psi - \hat{Z}' \sin\psi \cos\Omega = 0. \quad (3)$$

Equation (2) may also be rewritten using the rotation transformation equations

$$X' = X, \quad (4)$$

$$Y' = Y \sin i_o + Z \cos i_o, \quad (5)$$

$$Z' = -Y \cos i_o + Z \sin i_o, \quad (6)$$

giving

$$\begin{aligned}\mathbf{n}_s = {} & \hat{X}' \sin i_s \sin\lambda \\ & + \hat{Y}'(\sin i_s \cos\lambda \sin i_o + \cos i_s \cos i_o) \\ & + \hat{Z}'(\cos i_s \sin i_o - \sin i_s \cos\lambda \cos i_o). \end{aligned} \quad (7)$$

Setting the components of Equation (3) equal to those of Equation (8), we obtain three relations

$$\sin i_s \sin\lambda = \sin\psi \sin\Omega, \quad (8)$$

$$\cos\psi = \sin i_s \cos\lambda \sin i_o + \cos i_s \cos i_o, \quad (9)$$

$$\sin\psi \cos\Omega = \sin i_s \cos\lambda \cos i_o - \cos i_s \sin i_o, \quad (10)$$

which will be used in the following sections to derive constraints on ψ based on measurements of i_o and λ and on reasonable assumptions regarding i_s and Ω.

3. GIVEN ψ, WHAT WILL RM OBSERVATIONS SHOW?

Suppose an observer has measured the orbital inclination of a planetary system to be i_o and is about to measure the RM effect. If the spin–orbit angle of the system is ψ, then what is the probability distribution for the projected spin–orbit angle λ that the observer will measure? In this section, we calculate this function, $p(\lambda|\psi, i_o)$, which will play an important role in the calculations to follow.

We assume that for a given ψ, the probability distribution of the azimuthal angle Ω is uniformly distributed between $-\pi$ and $+\pi$. This is self-evident for a circular orbit, as there is no physical reason to distinguish any particular azimuth. For an eccentric orbit, it is conceivable that Ω is correlated with the direction of pericenter, but this possibility seems unlikely for hot Jupiters, because the torque exerted on the stellar rotational bulge by the planetary orbit will cause Ω to precess. The precession period for a hot Jupiter orbit is much shorter than the age of the system, and the secular evolution of both the argument of pericenter and the longitude of the ascending node—both defined with respect to the stellar equator—is linear in time (Roy 2005, Section 11.4.2, to lowest order in stellar shape parameters). We therefore expect an ensemble of stellar spins to have a uniform distribution in Ω, even if their orbits are eccentric. Conversely, the spin–orbit angle ψ remains constant over the precession cycle (Roy 2005, Section 11.4.2).

To derive $p(\lambda|\psi, i_o)$, we first express λ in terms of ψ, i_o, and Ω by eliminating i_s from Equations (8)–(10):

$$\lambda(\psi, i_o, \Omega) = \arctan\left(\frac{\sin\psi \sin\Omega}{\cos\psi \sin i_o + \sin\psi \cos\Omega \cos i_o}\right). \quad (11)$$

Since $\lambda(-\Omega, \psi, i_o) = -\lambda(\Omega, \psi, i_o)$, to calculate probabilities of λ and ψ we need only consider Ω and λ in the range $[0, \pi]$. The results will apply to negative values of λ as positive and negative values occur with equal probability.

Next, making use of $p(\Omega) = \pi^{-1}$, we transform variables from Ω to λ:

$$p(\lambda|\psi, i_o) = \sum_{i=1}^{N} p(\Omega_i|\psi, i_o) \left|\frac{d\Omega}{d\lambda}\right|_{\Omega=\Omega_i} = \frac{1}{\pi} \sum_{i=1}^{N} \left|\frac{d\Omega}{d\lambda}\right|_{\Omega=\Omega_i} \quad (12)$$

where the sum ranges over the N solutions Ω_i of Equation (11), for a given choice of λ, ψ, and i_o. We now find those solutions.

Letting $u \equiv \cos\Omega$, Equation (11) is equivalent to a quadratic equation in u:

$$\begin{aligned} & u^2 \left[\sin^2\psi (1 + \tan^2\lambda \cos^2 i_o)\right] \\ & + 2u \left[\tan^2\lambda \sin\psi \cos\psi \sin i_o \cos i_o\right] - \sin^2\psi \cos^2 i_o. \end{aligned} \quad (13)$$

When $\psi < i_o$ and $\sin\lambda < \sin\psi/\sin i_o$, this equation has two roots:

$$\begin{aligned} u_{1,2} &= \cos\Omega_{1,2} \\ &= \frac{-\tan^2\lambda \cos\psi \sin i_o \cos i_o \pm \sec\lambda \sqrt{\sin^2\psi - \sin^2 i_o \sin^2\lambda}}{\sin\psi (1 + \tan^2\lambda \cos^2 i_o)}. \end{aligned} \quad (14)$$

Figure 2(a) shows the geometrical interpretation of this two-root case. The ellipse on the sky that is traced out by $\mathbf{n}_s$ as Ω takes on all values (the "$\mathbf{n}_s$ ellipse") has two intersections with the ray that corresponds to the given value of λ (the "λ ray"). When $\psi < i_o$ and $\sin\lambda > \sin\psi/\sin i_o$, there are no real roots of Equation (14). Geometrically, the $\mathbf{n}_s$ ellipse is too small to intersect with the λ ray. The last case is $i_o \leqslant \psi \leqslant \pi - i_o$, for which Equation (14) has one real root. Figure 2(b) shows an example of this case. The $\mathbf{n}_s$ ellipse encloses the origin, and there is one intersection point with each λ ray. This single root corresponds to Ω_1 in Equation (14). Thus, in Equation (12), $N = 0$, 1, or 2, depending on the values of ψ, i_o, and λ.

Because of the derivatives $|d\Omega/d\lambda|$ in Equation (12), it is easier to derive the cumulative distribution $P(\lambda < \lambda_c|\psi, i_o)$ than to solve directly for $p(\lambda|\psi, i_o)$. The results are

$$P(\lambda < \lambda_c|\psi, i_o) = \left\{\begin{array}{ll} 1 + \frac{1}{\pi}(\Omega_1 - \Omega_2), & \psi < i_o \text{ and } \lambda_c < \arcsin(\sin\psi/\sin i_o) \\ 1, & \psi < i_o \text{ and } \lambda_c \geqslant \arcsin(\sin\psi/\sin i_o) \\ \frac{\Omega_1}{\pi}, & i_o \leqslant \psi \leqslant \pi - i_o \\ \frac{1}{\pi}(\Omega_1 - \Omega_2), & \psi > \pi - i_o \text{ and} \\ & \quad \lambda_c > \pi - \arcsin(\sin\psi/\sin i_o) \\ 0, & \psi > \pi - i_o \text{ and} \\ & \quad \lambda_c \leqslant \pi - \arcsin(\sin\psi/\sin i_o) \end{array}\right\}, \quad (15)$$

where $\Omega_{1,2}$ from Equation (14) are evaluated at λ_c, ψ, i_o.

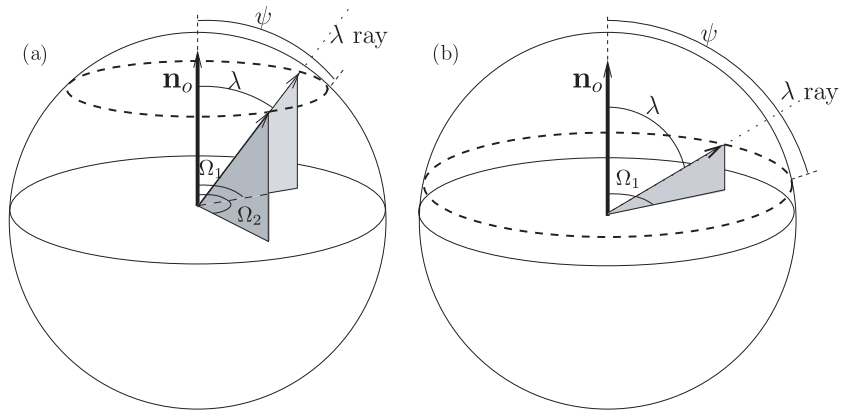

Figure 2. Geometry of the roots of Equation (14). For given values of ψ and i_o there are either 0, 1, or 2 possible values of λ. In these panels, the dashed line is the ellipse on the sky (the "$\mathbf{n}_s$ ellipse") that is traced out as $\mathbf{n}_s$ (thin vectors) sweeps around $\mathbf{n}_o$ (thick vector), with Ω taking on all values, and the dotted line is the ray corresponding to a given value of λ (the "λ ray"). (*a*) $\psi < i_o$ and $\sin\lambda < \sin\psi/\sin i_o$. There are two intersections of the $\mathbf{n}_s$ ellipse and λ ray. (*b*) $i_o < \psi < \pi - i_o$. There is only one intersection. For $\psi < i_o$ and $\sin\lambda > \sin\psi/\sin i_o$ (not shown), the $\mathbf{n}_s$ ellipse is too small to intersect the λ ray.

The probability densities are obtained by differentiation. First, we evaluate the derivatives $d\Omega_i/d\lambda$:

$$\frac{d\Omega_{1,2}}{d\lambda} = \left(\frac{2\tan\lambda\,\sec^2\lambda}{\sin\Omega_{1,2}\sin\psi(1-\tan^2\lambda\cos^2 i_o)}\right) \times \left(\cos\psi\sin i_o\cos i_o + \sin\psi\cos^2 i_o\cos\Omega_{1,2} \pm \frac{\cos\lambda(\sin^2 i_o - \sin^2\psi)}{2(\sin^2\psi - \sin^2 i_o\sin^2\lambda)^{1/2}}\right). \quad (16)$$

Finally, we calculate $p(\lambda|\psi, i_o)$ by inserting these derivatives into Equation (12). The full expressions are too large to reproduce here; instead we give the expressions into which Equation (16) may be substituted:

$$p(\lambda|\psi, i_o) = \left\{\begin{array}{ll} \frac{1}{\pi}\left(\frac{d\Omega_1}{d\lambda} - \frac{d\Omega_2}{d\lambda}\right), & \sin\psi < \sin i_o \text{ and } \sin\lambda < \sin\psi/\sin i_o \\ 0, & \sin\psi < \sin i_o \text{ and } \sin\lambda \geqslant \sin\psi/\sin i_o \\ \frac{1}{\pi}\frac{d\Omega_1}{d\lambda}, & \sin\psi \geqslant \sin i_o \end{array}\right\}. \quad (17)$$

For a transiting planet, i_o is always close to $\pi/2$. When $i_o = \pi/2$ exactly, the results are simplified as

$$P(\lambda < \lambda_c|\psi, i_o = \pi/2) = \left\{\begin{array}{ll} \frac{2}{\pi}\arccos\left[\frac{1}{\sin\psi}\left(1 - \frac{\cos^2\psi}{\cos^2\lambda_c}\right)^{1/2}\right], & |\lambda - \pi/2| \geqslant |\psi - \pi/2| \\ 1, & \psi \leqslant \lambda_c < \pi/2 \\ 0, & \psi > \lambda_c > \pi/2 \end{array}\right\} \quad (18)$$

and

$$p(\lambda|\psi, i_o = \pi/2) = \left\{\begin{array}{ll} \frac{2}{\pi}\frac{\cos\psi}{\cos\lambda(\cos^2\lambda - \cos^2\psi)^{1/2}}, & |\lambda - \pi/2| \geqslant |\psi - \pi/2| \\ 0, & |\lambda - \pi/2| < |\psi - \pi/2| \end{array}\right\}. \quad (19)$$

In the degenerate case $i_o = \pi/2$ and $\psi = \pi/2$, λ is observed to be either $-\pi/2$ or $\pi/2$ with equal probability.

Figure 3 shows the probability densities and cumulative distributions for $i_o = 90°$ and $i_o = 80°$ and some representative values of ψ. To gain an intuitive appreciation of the results, consider an edge-on orbit with $\psi < 90°$, shown in the left halves of the upper two panels. In this case, the spin–orbit angle ψ is an upper bound on its sky-projected version λ. For $\psi = 30°$, the chance of observing λ to be smaller than ψ by a factor of 2 is approximately 35%. In contrast, for $\psi = 85°$, the chance of observing λ smaller than ψ by a factor of 2 is only ≈5%. In this sense, λ is a more faithful indicator of ψ when ψ is large. For orbits that are not viewed edge-on (the lower panels), the maximum value of λ is increased, and for a nonedge-on orbit with ψ near 90°, it is possible to observe any value of λ.

3.1. Given λ from RM Observations, What May Be Inferred About ψ?

Suppose an observer has just measured i_o and λ for a particular transiting system. What may be reasonably inferred about the spin–orbit angle ψ? We appeal to Bayes's theorem:

$$p(\psi|\lambda, i_o) \propto p(\lambda|\psi, i_o)p(\psi), \quad (20)$$

where $p(\lambda|\psi, i_o)$ was calculated in the previous section, and $p(\psi)$ is the "prior" distribution, quantifying the observer's assumptions prior to the measurement. In this section, we adopt a prior distribution $p(\psi) = \sin\psi$, implying that $\mathbf{n}_s$ and $\mathbf{n}_o$ are uncorrelated and $\mathbf{n}_s$ is randomly oriented in space. This is the most uninformative or conservative assumption, in the sense that if $\mathbf{n}_o$ and $\mathbf{n}_s$ are instead highly correlated (with consequent implications for the theory of planet migration or tidal evolution), this fact should be demonstrated based on the data, rather than assumed from the outset. Hence, $p(\psi|\lambda, i_o)$ may be obtained by multiplying Equation (17) by $\sin\psi$ and renormalizing. For brevity, we give here only the analytic results for the case of an edge-on orbit and $\lambda < \pi/2$, measured with no error:

$$p(\psi|\lambda, i_o = \pi/2) = \left\{\begin{array}{ll} 0, & \psi < \lambda \\ \frac{\cos\psi\sin\psi}{\cos\lambda(\cos^2\lambda - \cos^2\psi)^{1/2}}, & \psi \geqslant \lambda \end{array}\right\} \quad (21)$$

and the corresponding cumulative probability function is

$$P(\psi < \psi_c|\lambda, i_o = \pi/2) = \left\{\begin{array}{ll} 0, & \psi_c < \lambda \\ \left(1 - \frac{\cos^2\psi_c}{\cos^2\lambda}\right)^{1/2}, & \psi_c \geqslant |\lambda| \end{array}\right\}. \quad (22)$$

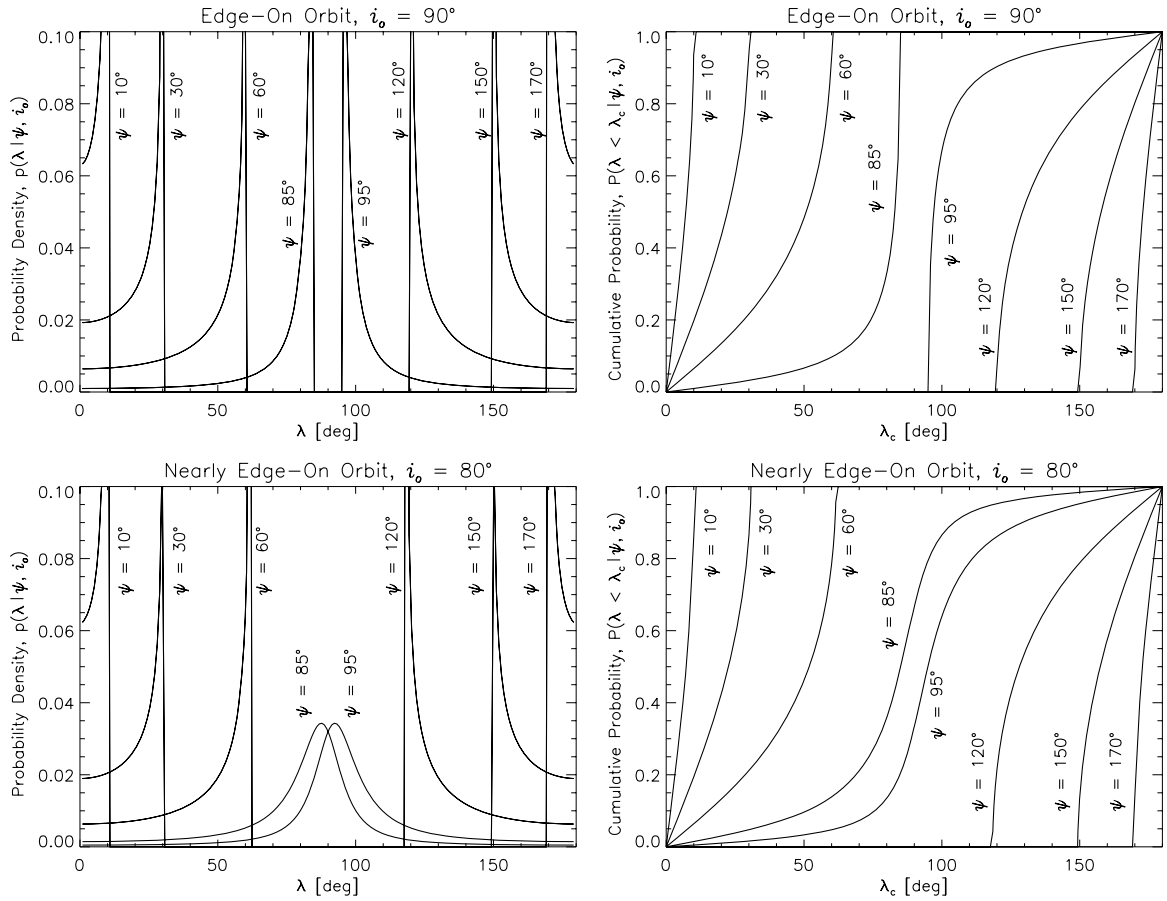

Figure 3. Probability distributions for the projected spin–orbit angle λ conditioned on the true spin–orbit angle ψ and the orbital inclination i_o. See Equations (14)–(19). The upper panels show results for an edge-on orbit ($i_o = 90°$) and various choices of ψ. The lower panels show results for a nearly edge-on orbit ($i_o = 80°$) and various choices of ψ.

These results are plotted in Figure 4 for some representative choices of λ. When λ is observed to be small, the a posteriori probability distribution of ψ has a very narrow spike near λ and extends broadly from $\psi = \lambda$ to 90°. When λ is observed to be large, it is more likely that ψ is close to λ. Just as M_p cannot be constrained strongly by $M_p \sin i_o$ for a Doppler-detected planet, we find that ψ cannot be constrained strongly by λ for an RM-detected planet, although the nature of the constraint is more complex in the latter case. In particular, for the edge-on case, it is possible to distinguish whether an orbit is prograde or retrograde without ambiguity, even though the value of ψ is quite uncertain.

As mentioned in Section 2, in principle one may obtain the missing information about i_s via asteroseismology, or via the combination of estimates of the stellar radius R_s, projected rotation velocity $v \sin i_s$, and stellar rotation period P_s. Although asteroseismology has never been undertaken for a transiting exoplanetary system, the other method has been employed for HD 189733 (Winn et al. 2007a; Henry & Winn 2008) and CoRoT-Exo-2 (Bouchy et al. 2008), and in both cases it was found that $\sin i_s$ is consistent with 1 (i.e., the equator is edge-on). However, near $\sin i_s = 1$ a small error in the measured $\sin i_s$ leads to a big error in i_s. Therefore, the data exclude only highly misaligned systems ($\psi \gtrsim 45°$). In the calculations to follow regarding the entire ensemble, we chose not to make use of these constraints specific to HD 189733 and CoRoT-Exo-2, for simplicity and because the extra information does not lead to significantly more powerful constraints.

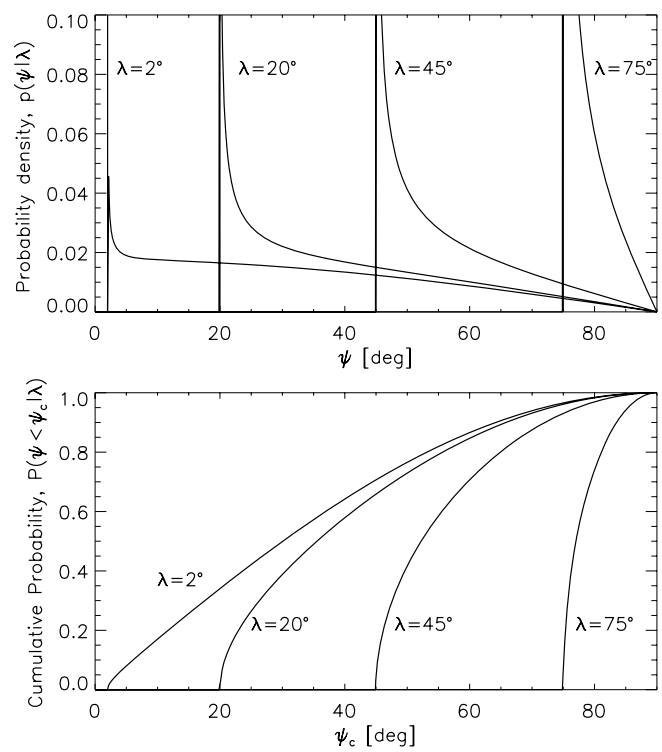

Figure 4. Probability distributions for the true spin–orbit angle ψ conditioned on the projected spin–orbit angle λ, assuming an edge-on orbit ($i_o = 90°$) and random spin–orbit alignment. See Equations (21)–(22).

Even when the rotation period is not available, one may exclude very nearly pole-on configurations of the star because they would require the star to be rotating unrealistically rapidly. Certainly, the rotation rate cannot exceed the breakup speed, although in practice we find that in realistic cases, applying this constraint does not make a perceptible difference in the distribution for ψ. One might go further by applying an a priori constraint on i_s to enforce agreement with the "typical" rotation rate for a star of the given spectral type and age. For the present study, we chose not to apply any such constraint, to avoid complications due to the uncertainties in the stellar types and ages, the intrinsic scatter in rotation rates, and the possibility that the rotation rates of stars hosting close-in giant planets may be systematically different from stars in general (due to tidal torques, earlier generations of "swallowed" planets, or other unforeseen effects).

4. INFERENCES FROM AN ENSEMBLE OF SYSTEMS

We have seen that ψ cannot be tightly constrained in an individual system, even when λ has been measured to within a few degrees, and even when $\sin i_s$ is constrained by a measured rotation period, stellar radius, and projected rotation rate. The purpose of this section is to derive stronger constraints by combining the results from different systems. The first such attempt, by Winn et al. (2006), demonstrated that the three measurements of λ available at that time were strongly inconsistent with an isotropic distribution of spin–orbit angles. This conclusion has been strengthened with the addition of many more systems with small values of $|\lambda|$, and it is now clear that $\mathbf{n}_s$ and $\mathbf{n}_o$ are correlated. The next natural question is: given the RM data, what is the distribution of spin–orbit angles? For instance, (1) is there a "typical" value which describes the mean and dispersion, and (2) is a single smooth distribution a good description of the data, or is there evidence for more than one population?

To answer these question, we suppose that the spin–orbit angles of the systems under consideration were drawn from a probability distribution $p(\psi)$ (the "model"), and we use the data to constrain the mathematical form of $p(\psi)$. Winn et al. (2006) already showed that the isotropic model, $p_I = \frac{1}{2}\sin\psi$, is untenable. A good theory of planet formation, migration, and evolution should be able to supply $p(\psi)$, or at least its general form. We will not attempt to develop such a theory here. Instead, we will use simple mathematical forms of $p(\psi|\mathbf{a})$ with a few free parameters and derive the probability distribution for those parameters, conditioned on the data.

Let the model parameters form a vector $\mathbf{a}$. The data consist of measurements of λ and i_o for $N_s = 11$ systems. These "measurements" are themselves probability distributions for λ and i_o. We neglect the error in i_o and denote by $p_{\mathrm{obs},k}(\lambda)$ the probability distribution for λ based on the observations of system k (from Table 1). We approximate the measurements as Gaussian distributions with the quoted λ and σ_λ as the mean and standard deviation.[5] The model $p(\psi|\mathbf{a})$ implies a certain probability distribution for λ, given by

$$p'(\lambda|i_o, \mathbf{a}) = \int_0^\pi p(\lambda|\psi, i_o)p(\psi|\mathbf{a})d\psi, \tag{23}$$

where $p(\lambda|\psi, i_o)$ is given by Equation (17). In practice, this integral is problematic because it integrates over the singularities visible in Figure 3, but we found that the singularity handlers in *Mathematica* are able to perform the integral numerically. For edge-on orbits ($i_o = \pi/2$), the simplified version of $p(\lambda|\psi, i_o = \pi/2)$ given by Equation (21) is applicable, and the transformation

$$y = \left(1 - \frac{\cos^2\psi}{\cos^2\lambda}\right)^{1/2} \tag{24}$$

removes the singularity. This provided a useful check on the ability of *Mathematica* to handle the singularities; for $i_o = \pi/2$ the numerical integrals were identical whether or not the transformation of Equation (24) was employed.

We may write the conditional probability as

$$p(\mathrm{data}|\mathbf{a}) = \prod_{k=1}^{N_s} \int_{-\pi}^{+\pi} p_{\mathrm{obs},k}(\lambda)p'(\lambda|i_{p,k}, \mathbf{a})d\lambda, \tag{25}$$

and then use Bayes's theorem,

$$p(\mathbf{a}|\mathrm{data}) \propto p(\mathrm{data}|\mathbf{a})p(\mathbf{a}), \tag{26}$$

where $p(\mathbf{a})$ is the prior probability density that is assigned to the parameters $\mathbf{a}$. Next, let us choose distributions to test.

4.1. A Fisher Distribution

If ψ were a Cartesian coordinate instead of a polar angle, one might model $p(\psi)$ as a Gaussian distribution with zero mean and variance σ, and derive the probability distribution for σ conditioned on the data. In the theory of directional statistics, the function that plays the same widespread and suitably generic role as the Gaussian distribution is the Fisher (1953) distribution

$$p_F(\psi|\kappa) = \frac{\kappa}{2\sinh\kappa}\exp(\kappa\cos\psi)\sin\psi, \tag{27}$$

where κ is the concentration parameter, a measure of the concentration of the probability distribution around $\psi = 0$. For $\kappa = 0$, the distribution becomes the isotropic distribution $p_I(\psi) = \frac{1}{2}\sin\psi$. For $\kappa \gg 1$ and $\psi \to 0$, the distribution becomes a Rayleigh distribution

$$p_R(\psi|\sigma) = \frac{\psi}{\sigma^2}\exp\left(-\frac{\psi^2}{2\sigma^2}\right) \tag{28}$$

with a width parameter $\sigma = \kappa^{-1/2}$. Mathematical properties of the Fisher distribution, and its relation to other distributions, can be found in Watson (1982). Some examples of Fisher distributions are plotted in Figure 5.

Let us assume that the spin–orbit angles are drawn from a Fisher distribution, and derive the probability distribution for κ conditioned on the data. We choose a prior distribution $p(\kappa) \propto (1 + \kappa^2)^{-3/4}$. This has the desirable limits $p(\kappa) \to$ constant for $\kappa \to 0$ (for broad distributions it is uninformative in κ), and $p(\sigma) \to$ constant for $\kappa \to \infty$ (for narrow, Rayleigh-like distributions it is uninformative in $\sigma = \kappa^{-1/2}$). The particular values of κ for which the distributions are illustrated in

[5] In some cases, even when the radial-velocity measurement errors are Gaussian, the posterior distribution $p_{\mathrm{obs},k}(\lambda)$ is not Gaussian. This is especially true of systems with slow stellar rotation rates or small transit impact parameters (see, e.g., the TrES-2 system (Winn et al. 2008) or the HAT-P-2 system (Winn et al. 2007b; Loeillet et al. 2008)). We investigated the sensitivity of our results on the assumption of a Gaussian distribution in λ by using the actual posterior distribution for λ whenever we had enough information to compute it. We found that the ensemble results were not significantly affected, because the most non-Gaussian cases were those with large errors, which were already downweighted in the Bayesian analysis.

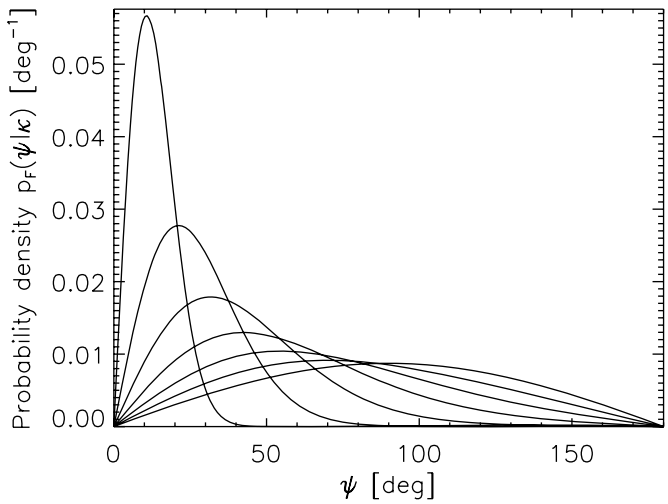

Figure 5. Fisher probability distribution (see Equation (27)), for some representative values of the concentration parameter κ. For $\kappa \to 0$, the Fisher distribution becomes an isotropic distribution ($\frac{1}{2}\sin\psi$), and for $\kappa \to \infty$ and $\psi \to 0$ the Fisher distribution becomes a Rayleigh distribution with width parameter $\sigma = \kappa^{-1/2}$. The displayed curves have $\kappa = 0.0$, 0.4, 0.9, 1.6, 3.1, 7.1, and 29. These values were chosen because our prior, $p(\kappa) \propto (1+\kappa^2)^{-3/4}$, assumes equal probability for each interval between adjacent values of κ.

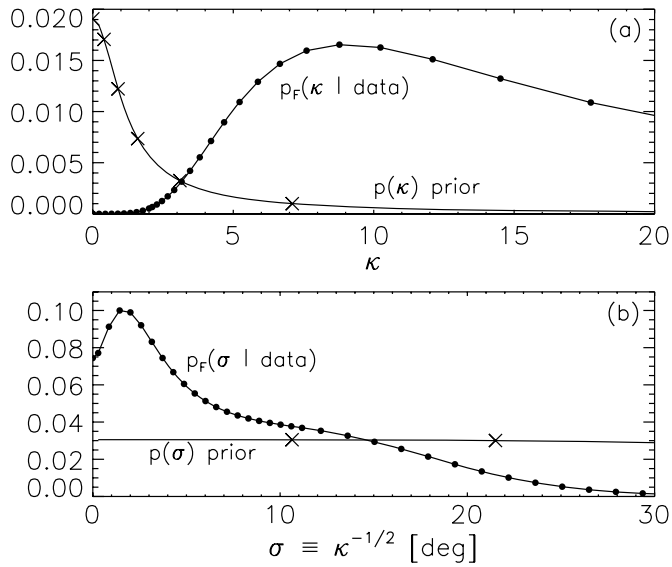

Figure 6. Modeling the RM ensemble with a Fisher distribution. (*a*) The line with dots shows the probability of the concentration parameter κ, conditioned on the data. The dots show the specific values of κ for which we computed the posterior probability. The line with *x*'s shows the assumed prior distribution of κ. The *x*'s show the values of κ for which the Fisher distributions are illustrated in Figure 5. (*b*) Correspondence with the more familiar Rayleigh distribution (Equation (28)), using the equivalence $\sigma \equiv \kappa^{-1/2}$ that is motivated by the high-κ, low-σ limit. Lines and points have the same meaning as in panel (*a*). In neither panel are the prior distributions normalized. The posterior distributions are normalized by the "evidence" for the Fisher distribution model (see Equation (31)).

Figure 5 were chosen because each interval in κ is equally likely, according to our prior.

The probability density for each value of κ, conditioned on the data, is the product of the prior and the probability density of the data given the value of κ, according to Equation (26). Figure 6(a) shows the resulting function $p(\kappa|\text{data})$, based on the 11 available RM measurements. It has been suitably normalized to unit probability. (The prior $p(\kappa)$ is also displayed in Figure 6(a), for reference.) Based on this result, we find that the characteristic concentration parameter is $\kappa > 7.6$ with 95% confidence.

The results for κ can be converted to a characteristic angular dispersion σ (in degrees) using $\sigma \equiv \kappa^{-1/2}$, bearing in mind that the Fisher distribution with concentration parameter κ is the same as the Rayleigh distribution with width parameter $\kappa^{-1/2}$ in the limit of $\kappa \to \infty$. The resulting distributions $p_F(\sigma|\text{data})$ and $p(\sigma)$ are plotted in Figure 6(b). The width parameter σ is less than 22° with 95% confidence.

We now examine the sensitivity of these results to certain aspects of the input data. First, we recompute the results without including the XO-3 data. This is because the finding of a strong misalignment in that system was considered tentative by the observers themselves. A standard RM model does not provide a statistically acceptable fit to the XO-3 data (Hebrard et al. 2008). It may be relevant that some of the data were contaminated by bright moonlight, requiring significant corrections to be applied to the spectra, and some of the data were taken at very high air masses.

Figure 7 shows the results when XO-3 is ignored and the other 10 systems are included as before. The results are very different: the distribution is tightly constrained near zero: $\sigma < 7°$ with 95% confidence. That these 10 systems are consistent with perfect alignment ($\sigma = 0$) can be seen by computing $\chi^2 = \sum(\lambda/\sigma_\lambda)^2$ from Table 1: it is 11.7, with 10 degrees of freedom. Therefore, apart from XO-3, the nonzero values of λ are consistent with observational errors alone. The $p(\sigma)$ distribution maximum likelihood is at nonzero σ because the reduced χ^2 is greater than 1, but this departure from $\sigma = 0$ is not statistically significant.

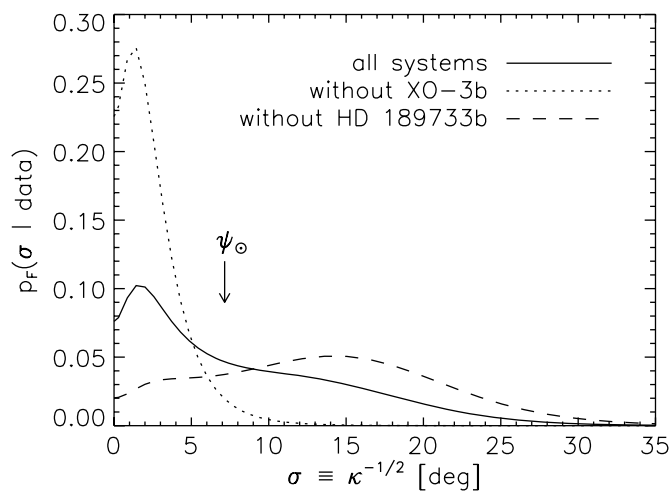

Figure 7. Same as Figure 6(b), but restricting the analysis to 10 systems instead of 11, by removing either XO-3 or HD 189733. The results are very different in those cases. Thus, the results depend strongly on the few systems for which λ is measured with the best precision, and the single system with an apparently large misalignment. The solar value of ψ is shown for reference.

Second, we investigate how sensitive are the results to the few most precise RM measurements. The dashed line in Figure 7 shows the results if we ignore HD 189733, the system with the smallest error bar σ_λ, and include all 10 of the remaining systems. In this case, the XO-3 result has enough statistical weight to pull the peak of the distribution strongly away from zero. Also plotted for reference is $\psi_\odot$, the solar spin–orbit angle. It must be remembered, though, that $\psi_\odot$ is a particular spin–orbit angle while σ describes the dispersion of ψ. The solar value of $\psi_\odot$ is typical of the "allowed" spin–orbit angles of hot Jupiters with their host stars.

With an eye toward future statistical analyses using the Fisher distribution, we note that the most computationally challenging aspect of the analysis was performing the integral (Equation (23)). A major simplification is available for nearly edge-on orbits, if the data are already known to favor a highly concentrated distribution ($\kappa \gg 1$). In this case, the hypothesized distribution for ψ is a Rayleigh distribution of parameter $\sigma \ll 1$, and the distribution of λ is Gaussian with a standard deviation σ. The problem of constraining the distribution of ψ is reduced to the problem of determining the true standard deviation of a distribution from which several noisy data points have been drawn. The dispersion σ can be estimated by finding the value of σ that gives $\chi^2 = N_s$ when it is added in quadrature with the measurement errors. The 10 systems besides XO-3 are in this simple regime. The simplified procedure described in this paragraph gives $\sigma = 1\overset{\circ}{.}1$, in agreement with the maximum likelihood value given for those 10 systems in Figure 7.

4.2. A Sum of Two Distributions: Isotropic and Perfectly Aligned

An alternative and equally simple way to describe the data is to suppose that all systems are drawn either from an isotropic distribution (with probability f) or from a very well-aligned distribution (with probability $1 - f$). We further suppose that the well-aligned distribution is sufficiently concentrated around $\psi = 0$ that none of the current measurements would be able to distinguish it from a delta function. This toy model will be a useful baseline for limiting the fraction of planets that migrate by various channels, some of which yield a nearly isotropic distribution of spin–orbit angles (specific examples are cited in Section 5).

The probability of the data $(\lambda_k, \sigma_{\lambda,k})$ given such a model is

$$p(\text{data}|f) = \prod_{k=1}^{N} \left[f\frac{1}{2\pi} + (1-f)\frac{2}{\sqrt{2\pi\sigma_k^2}} \exp\left(-\frac{\lambda_k^2}{2\sigma_{\lambda,k}^2}\right)\right]. \tag{29}$$

The factor of 2 in the numerator of the Gaussian portion above arises from the $\pm\lambda$ degeneracy mentioned in Section 2. Both of the terms in the sum are independent of i_o. Adopting a uniform prior for f, and using Bayes's theorem, we plot the result for $p(f|\text{data})$ in Figure 8. The data demand that $f < 0.36$ with 95% confidence. The favored value of f is 0.1, implying that approximately one system out of 11 is drawn from the isotropic distribution, clear indication that this result is being driven by XO-3. If we remove XO-3 from the analysis (to check the sensitivity of the analysis to this one system), $f < 0.25$ at 95% confidence, and the maximum likelihood is $f = 0$ exactly.

4.3. A Sum of Two-Fisher Distributions

Another possible model is a sum of two-Fisher distributions with different concentration parameters. This could describe, for instance, two different channels by which giant planets migrate to close-in orbits, which produce different final distributions of spin–orbit angles. Brown (2001) has found that a sum of two Fisher-like distributions is needed to fit the inclination distribution of Kuiper belt objects, and this multicomponent model has been a very useful constraint on dynamical theories. In this "two-Fisher" model, a fraction f of systems are Fisher-distributed with concentration parameter κ_1, and the remaining fraction $(1 - f)$ of systems are Fisher-distributed with concentration parameter κ_2. Thus, the two-Fisher model has three free parameters. The results of the previous section correspond to the case $f \approx 0.1$, $\kappa_1 = 0$, $\kappa_2 \to \infty$.

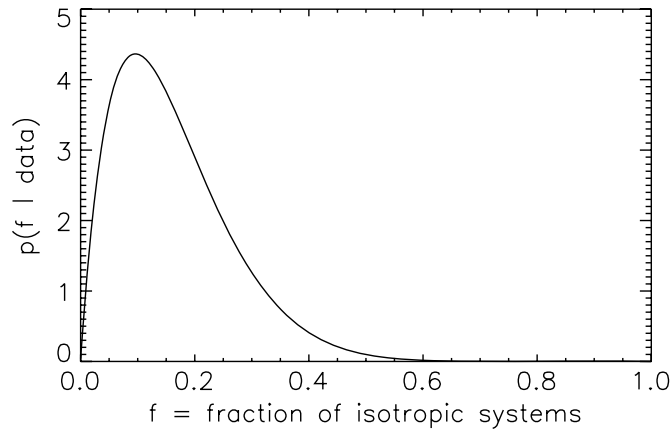

Figure 8. Modeling the RM ensemble with the sum of an isotropic distribution and a perfectly aligned (delta-function) distribution. A fraction f of the systems are drawn from the isotropic distribution, and the remaining fraction $(1 - f)$ are drawn from the delta-function distribution. Plotted is the posterior probability distribution for f, given the data. We find $f < 0.36$ with 95% confidence.

To determine the joint posterior probability distribution for f, κ_1, and κ_2, we computed

$$p_{2F}(f, \kappa_1, \kappa_2|\text{data}) = p(f)p(\kappa_1)p(\kappa_2) \times [f p_F(\text{data}|\kappa_1) + (1-f)p_F(\text{data}|\kappa_2)]. \tag{30}$$

The difficult integrals implied by this equation were already computed for the single-Fisher model. The results are $f = 0.10^{+0.13}_{-0.05}$, $\kappa_1 = 0.22^{+1.04}_{-0.22}$, and $\kappa_2 = 110^{+230}_{-76}$. The first Fisher distribution is consistent with an isotropic distribution. The second Fisher distribution is approximately equivalent to a Rayleigh distribution with width parameter $\sigma_2 = 5.5^{+4.3}_{-2.4}$ deg. The probability contours show that the correlations between the parameters are relatively small, especially within the 1σ preferred region.

4.4. Choosing Among Different Models

The results of the three-parameter, two-Fisher model are consistent with the results of the single-parameter, isotropic + perfectly aligned model given in the previous section. The greater complexity of the two-Fisher model makes it less appealing. This loss of appeal can be quantified within a Bayesian framework, which has a quantitative expression of Occam's razor.[6] Each model has an associated "evidence,"

$$E \equiv \int p(\text{data}|\mathbf{a})p(\mathbf{a})d\mathbf{a}, \tag{31}$$

where $\mathbf{a}$ is the vector of model parameters. This is the normalization factor used in Bayes's theorem, which turns proportionality (Equation (26)) into an equation. The presence of the first factor, $p(\text{data}|\mathbf{a})$, is the quantitative expression of the intuition that the data may be taken as evidence for the model only when the model predicts the data are probable. The second factor, $p(\mathbf{a})$, assigns greater evidence to models that concentrate their predictive power in the region where the data are found. This is because the integral over the prior distribution $p(\mathbf{a})$ is normalized to unity; if a prior distribution is spread too thinly over the

[6] A lucid discussion of Bayesian model choosing is given by MacKay (2003, chap. 28).

parameter space of **a**, then it cannot give much weight to models that are consistent with the data.

We computed the evidence for the models described in the three previous sections. The single-Fisher model has $E = 14.4$, the isotropic + perfectly aligned model has $E = 1927$, and the two-Fisher model has $E = 726$. The model that mixes the two extreme distributions does the best. It is favored by a factor of $\sim$130 over the single-Fisher model.

The difficulty with the single-Fisher model is that it cannot simultaneously account for the majority of systems that are well aligned while also including XO-3. A small value of κ makes XO-3 probable, but the 10 other systems are somewhat less probable, and the multiplication of these 10 lessened probabilities according to Equation (25) means there is little evidence for small κ. A large value of κ makes most of the data probable, but then the XO-3 result is very improbable, and the result is poor evidence for large κ. The distribution that mixes an isotropic distribution and a perfectly aligned distribution overcomes this difficulty by allowing *both* the majority of systems *and* XO-3 to be reasonably probable. The only free parameter is f, and the constraints on f are loose. Thus, the data provide substantial evidence for a large fraction of the parameter space.

The poor showing of the two-Fisher model relative to the isotropic + perfectly aligned model indicates that the two-Fisher model is a needless complication. The extra two parameters open up two more dimensions in the model's parameter space. As a result, much of the prior probability is "wasted" on regions of parameter space that are ruled out by the data. We hope that some day there will be enough high-quality RM data to justify a more complicated model such as the two-Fisher model. In this regard, we note that 379 Kuiper belt objects were discovered before their inclination distribution was modeled this way (Brown 2001).

We have already shown that the XO-3 measurement has an especially strong influence on the results. Unfortunately, as mentioned earlier, this measurement is considered suspect because of the possibility of systematic errors. Following the referee's suggestion, we may bring this suspicion under the umbrella of the Bayesian analysis by including a "degree of belief" parameter p_r, which gives the a priori probability that the XO-3 measurement will prove to be correct. Since the nature of the systematic errors (if any) is not known, p_r is rather subjective and open to debate. Our goal is not to determine the value of p_r, but rather to ask what is the minimum value of p_r that is required for our conclusion to hold that the isotropic + perfectly aligned model is preferred.

Each of the three models—single-Fisher, perfectly aligned plus an isotropic fraction, and two-Fisher—is fitted to the 11-member ensemble including XO-3, and also fitted to the 10-member ensemble excluding XO-3. Then the evidence for each model is computed as a weighted sum, with weight p_r applied to the 11-member set and weight $(1 - p_r)$ to the 10-member set. The evidence values of the models are computed and compared as a function of p_r. The result is that for $p_r < 0.95$, all three models fit the data equally well; they have evidence values within a factor of 3 of each other. We conclude that unless one has >95% confidence that the XO-3 result is robust, then a single smooth distribution of ψ is a perfectly viable description of the ensemble.

5. DISCUSSION

The angle between $\mathbf{n}_s$ and $\mathbf{n}_o$ is a fundamental geometric property of exoplanetary systems. A good theory of planet formation, migration, and evolution ought to predict the statistical relationship between $\mathbf{n}_s$ and $\mathbf{n}_o$ for hot Jupiters. This relationship is potentially measurable via RM observations. In this paper, we have overcome an inherent limitation of RM observations—they are sensitive only to the angle between the sky projections of the orbital axis and the stellar rotation axis—by showing how to analyze the whole ensemble in a Bayesian framework.

We modeled the 11 published measurements using a Fisher distribution, and found that the concentration parameter $\kappa > 7.6$ with 95% confidence. In this limit of a rather concentrated distribution, the Fisher distribution is equivalent to a Rayleigh distribution with width parameter $\sigma = \kappa^{-1/2}$. Based on the 11 data points, $\sigma < 22^\circ$ with 95% confidence. For comparison, the solar obliquity is 7°. If we set aside XO-3 (for instance, if the "tentative" detection of a strong misalignment is contradicted by higher-precision data), then the width parameter is $< 6.6^\circ$ with 95% confidence. In that case, the hot Jupiters are just as well aligned as the solar system.

The 11 data points also provide statistical evidence for two distinct populations within the ensemble, which might be interpreted as two different migration channels. Specifically, a model in which the systems are drawn from the sum of isotropic and perfectly aligned distributions fits the data better than a model with a single smooth distribution (Section 4.4). This is a reasonable conclusion, given that XO-3 shows the only evidence for a strong misalignment. However, due to the projection effect in converting from ψ to λ, it was not obvious prior to our analysis that the XO-3 result cannot be accomodated as part of the "tail" in a smooth distribution of spin–orbit angles. In fact, the data *do not* imply that XO-3 is the *only* system in the ensemble that is likely to be drawn from an isotropic distribution. We conclude only that fewer than 36% of the systems are drawn from an isotropic distribution, with 95% confidence. It is possible that several members of the ensemble were drawn from an isotropic distribution.

There is plenty of room for improvement in the quantity and quality of the RM data that are needed for this type of study. We have found that the present data are sufficient only to constrain single-parameter models for the ensemble. We also showed that the current results are highly sensitive to the few systems with the finest measurement precision. Measurements of λ with a precision of a few degrees are still in high demand. The "tentative" result for XO-3 needs to be followed up with more definitive data, as that single result weighs heavily on the Bayesian calculation. In addition, a uniform analysis of the data across many systems would be useful. Among the nonuniform aspects of the data analyses are whether or not $v \sin i$ was treated as a free parameter or subject to a constraint based on the observed line broadening; whether or not the uncertainties in the photometric transit parameters were taken into account; whether or not correlated noise in the radial-velocity data was assessed and taken into account if necessary; and whether or not the effects of spectral deconvolution and cross-correlation algorithms were calibrated. (These algorithms need calibration because they are generally intended to measure Doppler shifts, rather than model the actual RM spectral distortion that only superficially resembles a Doppler shift.)

There is also a potential bias regarding which systems are selected for measurement of the RM effect, as pointed out by Winn et al. (2008). Stars with low $v \sin i$ tend to be avoided, as the radial-velocity anomaly would be small and the achievable precision in λ is comparatively poor. However, stars with low $v \sin i_s$ are more likely to be viewed pole-on and therefore have

a large ψ. In the present work, we have not attempted to correct for such a selection effect.

Despite the limitations of the current data, the relative success of the two-component model (isotropic + perfectly aligned) leads us to speculate on the implications for theories that attempt to explain the presence of hot Jupiters. The chain of logic begins with the assumption that the system begins very well aligned ($\psi \approx 0°$). A natural prediction of in situ formation theories, or theories involving migration due to torques from the protoplanetary disk, is that the orbit of the planet remains very well aligned with equatorial plane of the star. (However, in the latter case the planet could conceivably misalign with the protoplanetary disk, and thus the stellar spin, depending on which resonant and secular torques dominate the planet–disk interaction; Borderies et al. 1984; Ward & Hahn 1994; Lubow & Ogilvie 2001.) In contrast, a very broad ψ distribution can be produced by mechanisms involving Kozai (1962) eccentricity cycles due to a distant companion star (Fabrycky & Tremaine 2007; Wu et al. 2007). A nearly isotropic ψ distribution can be produced by dynamical relaxation or planet–planet scattering (Narita et al. 2008). It is possible that the two components in our statistical model correspond to two different channels for migration, one that preserves the initial spin–orbit alignment and one that randomizes spin–orbit alignment to a significant degree.

A possible confounding factor is tidal damping. Based on the currently observed system parameters, it is expected that tidal coplanarization (also called "inclination damping") is not a major influence on ψ (Winn et al. 2005), but this is not a watertight argument. The timescales of long-term tidal processes are poorly known. Nevertheless, if tides raised on the star are large enough for substantial coplanarization, then the planet is in imminent danger of spiraling in and being engulfed (Levrard et al. 2009). Moreover, one would think that tidal damping of ψ by dissipation in the star should be *fastest* for systems with the most massive planets (see, e.g., Fabrycky et al. 2007; Jackson et al. 2008). Hence, it is intriguing that XO-3b, the most massive transiting planet with an RM measurement, and thus the one for which tidal dissipation should have been the most important, is the only system showing evidence for misalignment.[7] Therefore, the misalignment of XO-3b suggests that tidal damping is *not* responsible for low ψ values, and that the observed low $|\lambda|$ values should be interpreted as a relic of the planet formation and migration processes.

In this paper, we have been concerned with constraining the parameters of simple and fairly generic mathematical models for the distribution of spin–orbit angles. A priority for future work is to use the Bayesian framework developed in this paper to constrain the parameters of realistic, physically motivated models, based on the specific predictions of migration theories. In this vein, we encourage theorists to make quantitative predictions about the *distribution* of ψ. After deriving a theoretical distribution for ψ by whatever means, one may find the corresponding λ distribution for edge-on orbits using Equation (19). This requires a convolution similar to that of Equation (23). The theoretical predictions may then be directly compared with the data. With RM measurements of sufficient quantity and quality and with theories of sufficient specificity, it may be possible to rule out certain migration theories, or to derive the likely fraction of systems that migrated through different channels.

We thank Ed Turner for helpful conversations about astrostatistics and Scott Tremaine for comments on the manuscript. We thank the referee, Frederic Pont, for suggesting we quantitatively analyze non-Gaussian and catastrophic errors, for the XO-3 system in particular. D.F. gratefully acknowledges support from the Michelson Fellowship, supported by the National Aeronautics and Space Administration and administered by the Michelson Science Center. This work was partly supported by a grant from the NASA Keck PI Data Analysis Fund (JPL 1326712), and from the NASA Origins program (NNX09AD36G).

Note added in proof. Winn et al. (2009) have reobserved the Rossiter–McLaughlin effect in the XO-3 system, have found it to be indeed misaligned, and have updated the quantitative statistical inferences of this paper.

REFERENCES

Beck, J. G., & Giles, P. 2005, ApJ, 621, L153
Borderies, N., Goldreich, P., & Tremaine, S. 1984, ApJ, 284, 429
Bouchy, F., et al. 2008, A&A, 482, L25
Brown, M. E. 2001, AJ, 121, 2804
Bundy, K. A., & Marcy, G. W. 2000, PASP, 112, 1421
Chatterjee, S., Ford, E. B., Matsumura, S., & Rasio, F. A. 2008, ApJ, 686, 580
Cochran, W. D., Redfield, S., Endl, M., & Cochran, A. L. 2008, ApJ, 683, L59
Dobbs-Dixon, I., Lin, D. N. C., & Mardling, R. A. 2004, ApJ, 610, 464
Fabrycky, D. C., Johnson, E. T., & Goodman, J. 2007, ApJ, 665, 754
Fabrycky, D., & Tremaine, S. 2007, ApJ, 669, 1298
Fisher, R. 1953, Proc. R. Soc. A, 217, 295
Ford, E. B., Havlickova, M., & Rasio, F. A. 2001, Icarus, 150, 303
Gaudi, B. S., & Winn, J. N. 2007, ApJ, 655, 550
Giménez, A. 2006, ApJ, 650, 408
Gizon, L., & Solanki, S. K. 2003, ApJ, 589, 1009
Goldreich, P., & Ward, W. R. 1972, PASP, 84, 737
Hebrard, G., et al. 2008, A&A, 488, 763
Heller, C. H. 1993, ApJ, 408, 337
Henry, G. W., & Winn, J. N. 2008, AJ, 135, 68
Jackson, B., Greenberg, R., & Barnes, R. 2008, ApJ, 678, 1396
Johnson, J. A., et al. 2008, ApJ, 686, 649
Jorissen, A., Mayor, M., & Udry, S. 2001, A&A, 379, 992
Joshi, Y. C., et al. 2009, MNRAS, 392, 1532
Jurić, M., & Tremaine, S. 2008, ApJ, 686, 603
Kozai, Y. 1962, AJ, 67, 591
Levrard, B., Winisdoerffer, C., & Chabrier, G. 2009, ApJ, 692, L9
Lin, D. N. C., Bodenheimer, P., & Richardson, D. C. 1996, Nature, 380, 606
Loeillet, B., et al. 2008, A&A, 481, 529
Lubow, S. H., & Ogilvie, G. I. 2001, ApJ, 560, 997
MacKay, D. J. C. 2003, Information Theory, Inference, and Learning Algorithms (Cambridge: Cambridge Univ. Press)
Marzari, F., & Weidenschilling, S. J. 2002, Icarus, 156, 570
Mazeh, T. 2008, in EAS Publications Series, Vol. 29, Tidal Effects in Stars, Planets, and Disks, ed. M. J. Goupil & J. P. Zahn (Les Ulis, France: EDP Sciences), 1
Mazeh, T., Goldberg, D., & Latham, D. W. 1998, ApJ, 501, L199
Murray, N., Hansen, B., Holman, M., & Tremaine, S. 1998, Science, 279, 69
Nagasawa, M., Ida, S., & Bessho, T. 2008, ApJ, 678, 498
Narita, N., Sato, B., Ohshima, O., & Winn, J. N. 2008, PASJ, 60, L1
Narita, N., et al. 2007, PASJ, 59, 763
Nordtvedt, K. 1987, ApJ, 320, 871
Ohta, Y., Taruya, A., & Suto, Y. 2005, ApJ, 622, 1118
Papaloizou, J. C. B., & Terquem, C. 2001, MNRAS, 325, 221
Pont, F. 2008, MNRAS, in press (arXiv:0812.1463)
Queloz, D., Eggenberger, A., Mayor, M., Perrier, C., Beuzit, J. L., Naef, D., Sivan, J. P., & Udry, S. 2000, A&A, 359, L13
Roy, A. E. 2005, Orbital motion (4th ed.; Bristol: Institute of Physics Publishing)
Stepinski, T. F., & Black, D. C. 2000, A&A, 356, 903
Tabachnik, S., & Tremaine, S. 2002, MNRAS, 335, 151
Terquem, C., & Papaloizou, J. C. B. 2002, MNRAS, 332, L39

[7] There is also dissipation within the planet, which is less efficient for higher mass planets. However, this mode of dissipation is less relevant because the torque on the planetary bulge does not couple strongly to the stellar obliquity ψ (Fabrycky et al. 2007).

Thommes, E. W., & Lissauer, J. J. 2003, ApJ, 597, 566
Tremaine, S. 1991, Icarus, 89, 85
Ward, W. R. 1997, Icarus, 126, 261
Ward, W. R., & Hahn, J. M. 1994, Icarus, 110, 95
Watson, G. S. 1982, J. Appl. Probab., 19, 265
Winn, J. N., et al. 2005, ApJ, 631, 1215
Winn, J. N., et al. 2006, ApJ, 653, L69
Winn, J. N., et al. 2007a, AJ, 133, 1828
Winn, J. N., et al. 2007b, ApJ, 665, L167
Winn, J. N., et al. 2008, ApJ, 682, 1283
Winn, J. N., et al. 2009, ApJ, submitted (arXiv:0902.3461)
Wittenmyer, R. A., et al. 2005, ApJ, 632, 1157
Wolf, A. S., Laughlin, G., Henry, G. W., Fischer, D. A., Marcy, G., Butler, P., & Vogt, S. 2007, ApJ, 667, 549
Wu, Y., Murray, N. W., & Ramsahai, J. M. 2007, ApJ, 670, 820
Yu, Q., & Tremaine, S. 2001, AJ, 121, 1736
Zahn, J.-P., & Bouchet, L. 1989, A&A, 223, 112
Zucker, S., & Mazeh, T. 2001, ApJ, 562, 1038