APOSTLE: LONGTERM TRANSIT MONITORING AND STABILITY ANALYSIS OF XO-2b

P. Kundurthy[1], R. Barnes[1,2], A. C. Becker[1], E. Agol[1], B. F. Williams[1], N. Gorelick[3], and A. Rose[1]
[1] Astronomy Department, University of Washington, Seattle, WA 98195, USA
[2] Virtual Planetary Laboratory, Seattle, WA 98195, USA
[3] Google Inc., Mountain View, CA 94043, USA

ABSTRACT

The Apache Point Survey of Transit Lightcurves of Exoplanets (APOSTLE) observed 10 transits of XO-2b over a period of 3 yr. We present measurements that confirm previous estimates of system parameters like the normalized semi-major axis ($a/R_\star$), stellar density ($\rho_\star$), impact parameter (b), and orbital inclination (i_{orb}). Our errors on system parameters like $a/R_\star$ and $\rho_\star$ have improved by ∼40% compared to previous best ground-based measurements. Our study of the transit times show no evidence for transit timing variations (TTVs) and we are able to rule out co-planar companions with masses ⩾0.20 $M_\oplus$ in low order mean motion resonance with XO-2b. We also explored the stability of the XO-2 system given various orbital configurations of a hypothetical planet near the 2:1 mean motion resonance. We find that a wide range of orbits (including Earth-mass perturbers) are both dynamically stable and produce observable TTVs. We find that up to 51% of our stable simulations show TTVs that are smaller than the typical transit timing errors (∼20 s) measured for XO-2b, and hence remain undetectable.

Key words: eclipses – methods: numerical – planetary systems – planets and satellites: fundamental parameters – planets and satellites: individual (XO-2b)

Online-only material: color figures, machine-readable tables

1. INTRODUCTION

Observational efforts on two fronts (ground-based and space-based) have revealed a great diversity in the properties of planets (and planet candidates), and posed several new questions about the planet formation process. The search for planets and the careful measurement of planetary properties are the observational foundations upon which the physics of planet formation may be understood. The exoplanet community has co-opted the term "architecture" to embody several properties of planetary systems such as multiplicity and orbital parameters like eccentricities (e), inclinations (i_{orb}), semi-major axes (a), and orbital periods. The distributions of individual planetary properties like masses and radii (M_p and R_p, respectively) may also be included in this term. The primary goal of the Apache Point Observatory Survey of Transit Lightcurves of Exoplanets (APOSTLE) is to catalog transit light curves at high precision in order to (1) measure system parameters and (2) look for transit timing variations (TTVs) that may indicate the presence of additional planets.

The transit technique applies to those systems where the orbital inclination of an exoplanet is close to 90° (i.e., edge-on) with respect to the observer's sky-plane (see the discovery paper Charbonneau et al. 2000). In this case, the observer sees a u-shaped dip in the starlight caused when the planet eclipses the star (Winn 2011). The objective of several ground-based and space-based efforts focused on transit observations is to catalog and improve measurements of system parameters, which in turn gives us an improved picture of the architecture of the planetary system. This process is key toward developing theories of planet formation that can adequately explain the origin and evolution of all planetary systems (including our own).

The APOSTLE target XO-2b is a Hot-Jupiter on a 2.6 day orbit around an early type K dwarf ($V = 11.2$; Burke et al. 2007). The planet is known to have a mass of 0.555 M_{Jup} and a radius of 0.992 R_{Jup} (Southworth 2010). The system is not known to have any other planet-mass objects. However, Narita et al. (2011) claim to detect long-term radial velocity variation that deserves further follow-up. Transit timing measurements from the ground are consistent with a linear ephemeris (Fernandez et al. 2009; Sing et al. 2011; Crouzet et al. 2012), although the cited studies note that there are statistically significant deviations in the measurement of the orbital periods. A search for additional eclipses using the *EPOXI* mission by Ballard et al. (2011) had inconclusive results since the eclipse was not fully sampled. *Spitzer* observations of the secondary eclipse of XO-2b show IRAC 3.6, 4.5, and 5.8 μm fluxes which are consistent with the presence of a temperature inversion (Machalek et al. 2009). There have also been detections of optical absorbers in the planetary atmosphere such as potassium from narrow band optical transmission spectrophotometry (Sing et al. 2011). Early theoretical studies indicated that stellar insolation levels directly influenced the presence or absence of a thermal inversion layer (depending on the survival of atmospheric absorbers; Hubeny et al. 2003; Burrows et al. 2007; Fortney et al. 2008). The planetary atmosphere classification system developed by Fortney et al. (2008) places XO-2b in an transition zone between planets with (pM) and without (pL) thermal inversions. XO-2b is one of a handful of Hot-Jupiters in this region. However, observational evidence suggests (Machalek et al. 2008) that more complicated models need to be considered. Irradiation from stellar activity may need to be included (Knutson et al. 2010) and atmospheric chemistry may also need to be considered to provide a more complete picture (Madhusudhan 2012). The host star of the XO-2 system is known to be a high metallicity ([Fe/H] = 0.45 ± 0.02), high proper motion star ($\mu_{Tot} = 157$ mas yr^{-1}) in a visual binary (Burke et al. 2007). Spectral activity indices show that XO-2 is fairly inactive compared to other stars of similar spectral type (Knutson et al. 2010).

A primary goal of APOSTLE and other campaigns that monitor transiting exoplanets is to search for TTVs that reveal

Table 1
APOSTLE Observing Summary for XO2

T# (1)	UTD (2)	Obs. Cond. (3)	Filter (4)	Exp. (5)	Bin (6)	Phot. Ap. (7)	RMS (ppm) (8)	%Rej. (9)	Flux Norm. (10)	Error Scaling (11)
1	2008-01-09	Clear	I	0.5	45	13	557	3%	0.9561	1.1740
2	2008-02-12	Clear	I	0.5	45	24	510	<1%	0.9561	1.3129
3	2008-03-04	Clear	I	0.5	45	15	411	<1%	0.9632	1.2510
4	2008-11-23	Poor Weather	I	10, 25, 45	45	27	939	1%	0.9756	2.7482
5	2009-02-07	Poor Weather	I	45	...	31	1085	11%	0.9721	2.9868
6	2009-03-13	Poor Weather	I	45	...	46	405	11%	0.9766	1.2356
7	2010-10-25	Clear	r'	45	...	35	553	2%	0.9676	2.2760
8	2010-12-27	Poor Weather	r'	45	...	43	775	5%	0.9643	3.0092
9	2011-01-30	Clear	r'	45	...	48	354	1%	0.9646	1.4571
10	2011-03-05	Clear	r'	45	...	43	693	<1%	0.9660	2.7990

Notes. Column 1: transit number; Column 2: universal time date; Column 3: observing conditions; Column 4: observing filter; Column 5: exposure time (seconds); Column 6: bin size in seconds; Column 7: optimal aperture radius (pixels); Column 8: scatter in the residuals; Column 9: % frames rejected due to saturation or other effects; Column 10: flux normalization between the target and comparison star; Column 11: the factor by which the photometric errors were scaled.

the presence of unseen companions (Agol et al. 2005; Holman & Murray 2005). In principle, Earth-mass planets in or near mean motion resonances could perturb the orbit of the transiting planet enough to produce a sinusoidal oscillation in the mid-points of transits. However, the full range of stable orbits that can produce a detectable TTV signal has never been explored. The detection of TTVs by *Kepler* (e.g., Holman et al. 2010) has demonstrated that stable systems are capable of producing TTVs, and other studies have explored a limited range of architectures (Haghighipour & Kirste 2011), but the systematic exploration of parameter space of an Earth-mass planet in orbit near a hot Jupiter has not been undertaken. Here we examine 3.6 million possible masses and orbits of an approximately Earth-mass planet orbiting in or near the 2:1 outer mean motion resonance through *N*-body simulations. We find that stable orbits that also produce detectable TTVs do exist in the XO-2 system, and hence can exist in similar Hot-Jupiter systems.

In this paper we report observations of 10 transits of XO-2b, taken as part of APOSTLE. In Section 2 we outline our observations. In Section 3 we briefly outline (1) the data reduction, photometry (Section 3.1), (2) the transit model (Section 3.2) and light curve fitting (Section 3.3); both processes have also been described in previous work (Kundurthy et al. 2013). In Section 4 we present our estimates of the system parameters for XO-2b and in Sections 4.1 and 4.2 we present results from our study of transit depth variations (TDVs) and TTVs. In Section 5 we present results form *N*-body simulations used to study the stability of hypothetical planetary configurations at the 2:1 mean-motions resonance. Finally, in Section 6 we summarize our findings.

2. OBSERVATIONS

XO-2b was observed by members of the APOSTLE team on 10 occasions over a timespan of 3 yr from early 2008 until the spring of 2011. All observations of XO-2 were carried out using *AGILE*, a high-speed frame-transfer CCD (Mukadam et al. 2011), on the ARC[4] 3.5 m telescope at Apache Point, New Mexico. The summary of observations is given in Table 1. XO-2 was an early APOSTLE target and was observed using a variety of instrumental settings, as the team had not converged on an optimal observing strategy prior to 2010. Early observations were made in the *I*-band ($\lambda_0 = 805$ nm; Cousins 1976; Bessell

1990) with several data sets taken with short read-out (exposure times, Column 5 in Table 1). The short read-out mode allowed for fine sampling of the light curve, but, due to the lower signal-to-noise and the unsuitability for characterizing systematics, this observing mode was abandoned. In addition, the *I*-band images also contained a strong contribution from a fringe pattern due interference from the backscattering of atmospheric lines within the CCD's pixels. The removal of this fringe pattern is discussed in Section 3.1. Early in 2010, the observing strategy changed to longer readouts (typically 45–75 s) to reduce the level of correlated noise (red-noise), and were made using the r'-band which is similar to the SDSS[5] r filter ($\lambda_0 = 626$ nm; Fukugita et al. 1996), to reduce the influence of fringes. In the long read-out mode the telescope was defocused to spread the stellar point-spread function (PSF) across multiple pixels, which minimized the systematics caused by pixel-to-pixel wandering of the PSF over the imperfect flatfield. The longer exposures also allowed for a greater count rate which maximized the signal-to-noise per image. The count rate was kept below *AGILE*'s non-linearity limit of ~52k ADU and well below its saturation level of 61k ADU by small adjustments to the telescope's secondary focus during observations.

XO-2 (TYC 3413-5-1) and its visual binary companion TYC 3413-210-1 (separated by ~30″) were the only bright stars that fit in *AGILE*'s field of view. Both stars are of identical spectral type (K0V) and nearly identical brightnesses in the filters used by APOSTLE, with their Johnson *R* and *I* magnitudes at 10.8 and 10.5, respectively (Monet et al. 2003). Our uncalibrated differential photometry also showed good agreement in their brightnesses, with the out-of-eclipse, un-normalized flux ratios being different by only 4% and 3% in the *I*-band and r'-band, respectively (Column "Flux Norm." in Table 1). The observations were made over a variety of observing conditions (Column "Obs. Conditions" in Table 1). The observing conditions are classified as "Clear" or "Poor Weather" with the former implying good data with few or no interruptions in data collection, and the latter indicating that the we experienced cloud cover or poor seeing conditions resulting in lower quality data. The tabulated transits are those for which we were able to capture the whole transit or at least a partial transit. Partial transits are those where portions of the in-eclipse light curve were lost due to bad weather or instrumental failure. Several data points were lost for the nights of UTD

---

[4] Astrophysical Research Consortium.

[5] Sloan Digital Sky-Survey.

2008 September 22 (transit 4), 2009 February 6 (transit 5), 2009 March 12 (transit 6), and 2010 December 27 (transit 8). However, we do include these nights since we did manage to obtain reasonable portions of the in-eclipse and out-of-eclipse data, which make it possible to determine transit properties (albeit at a loss of accuracy and precision).

3. DATA ANALYSIS

This section outlines various stages in the analysis of light curves, starting with (1) the image reduction and photometry, and (2) the transit model and Markov Chain Monte Carlo (MCMC) analyzer (see also Kundurthy et al. 2013).

3.1. Reduction and Photometry

APOSTLE data were reduced using a pipeline developed specifically for *AGILE* images. The pipeline (written in IDL[6]) performs pixel-by-pixel error propagation, and image processing specific to the *AGILE* CCD. In addition to the standard photometric reduction steps like dark subtraction and flat-fielding, the pipeline performs non-linearity and fringe corrections specific to *AGILE*. The details on *AGILE*'s non-linearity correction are described in Section 3 of Kundurthy et al. (2013). Some of the initial observations by APOSTLE were carried out using short exposures (see Column 5 "Exp.") in Table 1. We binned these data by averaging the flux ratios in 45 s bins.

For several of the initial XO-2 data taken using *AGILE*'s *I*-band, a fringe pattern had to be subtracted to create science images. Photons from strong atmospheric lines (in the *I*-band bandpass) backscatter within the CCD pixels, and owing to the variable thickness of the pixel array, interference between these photons creates the fringe pattern. Since the fringe pattern is convolved with the illumination pattern of the CCD, the fringe correction has to be applied after dark-subtraction and flat-fielding. During the initial characterization of the *AGILE* CCD, an empirical fringe frame was produced by median combining dithered frames on the dark sky, where the flux contributions from stars were removed by outlier rejection. The resulting combined frame served as a "Model" fringe pattern (*F*), normalized to have a median of zero, and an amplitude of one, such that it could be scaled to match the fringe patterns on science frames. The science frame affected by the fringe pattern (T') is assumed to be a linear combination of the fringe-less science image (*T*) and a fringe pattern:

$$T' = T + a_0 F, \qquad (1)$$

where *F* is the model fringe frame and a_0 is a scaling factor describing the amplitude of the fringing on a given frame. The fringe amplitude is estimated by minimizing χ^2 and fitting for a_0 in the above model. The corrected images ($T = T' - a_0 F$), were found to be sufficiently corrected of fringes after visual examination. We found the fringe amplitudes (a_0) on XO-2 science frames to always be smaller than the standard deviation (i.e., scatter) of the global sky background on each frame; a_0 ranged between 6%–13% of the scatter in the sky for XO-2 data.

We extracted photometry from an optimal circular aperture centered around the target and comparison stars. In addition we extracted the counts on image products like the master-dark and master-flat, to serve as nuisance parameters for detrending, using the same aperture and centroids from photometry. The centroids

---
[6] Interactive Data Language.

Table 2
Nuisance Parameters Used for XO-2 Nights

T#	Nuisance Parameters Used (FLDC, OLDC, and MDFLDC)
T1	airmass, $msky_1$, $msky_2$, gsky, x_1, y_1, x_2, y_2, sD_1, sD_2, sF_1, sF_2, a_0
T2	airmass, $msky_1$, $msky_2$, gsky, x_1, y_1, x_2, y_2, sD_1, sD_2, sF_1, sF_2, a_0
T3	airmass, $msky_1$, $msky_2$, gsky, x_1, y_1, x_2, y_2, sD_1, sD_2, sF_1, sF_2, a_0
T4	airmass, $msky_1$, $msky_2$, gsky, x_1, y_1, x_2, y_2, sD_1, sD_2, sF_1, sF_2, a_0
T5	airmass, x_1, y_1, sD_1, sF_1, a_0
T6	airmass, x_1, y_1, sD_1, sF_1, a_0
T7	airmass, gsky, x_1, y_1, sD_1, sF_1
T8	airmass, gsky, x_1, y_1, sD_1, sF_1
T9	airmass, gsky, x_1, y_1, sD_1, sF_1
T10	airmass, gsky, x_1, y_1, sD_1, sF_1

Notes. airmass: atmospheric column; (x_1, y_1), (x_2, y_2): centroids of Target (1) and Comparison (2); $msky_1$, $msky_2$: median sky around Target (1) and Comparison (2); gsky: global median sky; sD_1, sD_2: sum of counts in aperture on master-dark; sF_1, sF_2: sum of counts in aperture on master-flat; a_0: fringe scaling (see Equation (1)).

were derived using SExtractor (Bertin & Arnouts 1996), which allowed for the use of customized donut-shaped convolution kernels for defocused PSFs. Science frames where pixels inside a photometric aperture exceeded *AGILE*'s saturation limit of 61k were rejected. Images at the other extreme, where the stars were obscured by clouds, and resulted in low signal-to-noise measurements were also rejected (i.e., where individual photometric errors were >5000 ppm). The fraction of rejected frames per night is listed in Column 9 "%Rej." in Table 1. The optimal aperture was selected after extracting photometry on a list of circular apertures with radii between 5–50 pixels at an interval of 1 pixel. The optimal aperture was selected where the scatter in the residuals of the detrended light curve minus a trial transit model (based on values from the literature) was minimized. The correction function F_{cor} (or detrending function) is modeled as a linear sum of nuisance parameters as described by the following equation:

$$F_{cor,i} = \sum_{k=1}^{N_{nus}} c_k X_{k,i}, \qquad (2)$$

where $X_{k,i}$ are the nuisance parameters, c_k are the corresponding coefficients. The index k counts over the number of nuisance parameters N_{nus}. The detrending coefficients are chosen by minimizing the χ^2 between the observed data (*O*), a model function (*M*), and the correction function,

$$\chi^2 = \sum_j^{N_{all}} \frac{(O_j - M_j - F_{cor,j})^2}{\sigma_j^2}, \qquad (3)$$

here *j* is the index which counts over the total number of data points (N_{all}). The observed, model and correction function terms are all in normalized flux ratio units. The list of parameters used for detrending each light curve in the XO-2 dataset are presented in Table 2, with variable definitions given in the footnotes. A suitable set of detrending parameters were selected for a given night by running several manual trials of linear least squares minimization on single transit light curves and its corresponding model light curve and correction function (Equation (2)). The model parameters were fixed at reasonable values and only the coefficients (c_k) were fit. Those nuisance parameters which returned large uncertainties on the coefficients were excluded since this indicated a poor match to the noise trend in a

Figure 1. Six *I*-band and four *r'*-band light curves of XO-2b. The vertical axis is in normalized flux ratio units. The horizontal axis shows time from the mid-transit time in days, computed by subtracting the appropriate mid-transit time for each transit from the best-fit values in the fixed LDC chain.

Table 3
APOSTLE Light Curve Data for XO-2

T# (1)	T − T0 (2)	Norm. Fl. Ratio (3)	Err. Norm. Fl. Ratio (4)	Model Data (5)
1	−0.1341616	1.000004100	0.001207284	1.000000000
1	−0.1315574	0.999797978	0.001102731	1.000000000
1	−0.1305157	0.997707769	0.001047412	1.000000000
1	−0.1299949	0.999906661	0.001190794	1.000000000
1	−0.1289532	0.998693635	0.001118320	1.000000000
1	−0.1279116	0.998939727	0.001098241	1.000000000
1	−0.1273907	0.999021973	0.000987372	1.000000000
1	−0.1268699	1.001381660	0.001209313	1.000000000
1	−0.1263491	0.998899046	0.001076096	1.000000000
1	−0.1258282	1.001212353	0.001160014	1.000000000
.	.	.	.	.
.	.	.	.	.

Notes. Column 1: transit number; Column 2: time stamps − mid transit times (BJD); Column 3: normalized flux ratio; Column 4: error on normalized flux ratio; Column 5: model data.

(This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

The set of parameters used for Multi-Filter version of MTQ is $\theta_{\text{Multi-Filter}} = \{t_T, t_G, D_{j...N_F}, v_{1,j...N_F}, v_{2,j...N_F}, T_{i...N_T}\}$, where t_T is transit duration, t_G is the limb-crossing duration and T_i are the mid-transit times. The filter-dependent parameters include the transit depth D_j and the limb-darkening parameters $v_{1,j}$ and $v_{2,j}$ described in Kundurthy et al. (2013). The subscripts $i...N_T$ and $j...N_F$ are used to denote multiple transits (N_T) and multiple filters (N_F) respectively. The XO-2 data were gathered using the *I*-band and *r'*-band, where N_F was 2 and the number of transits N_T was 10.

The parameter set is only slightly different for the Multi-Depth version, with $\theta_{\text{Multi-Depth}} = \{t_T, t_G, D_{i...N_T}, v_{1,j...N_F}, v_{2,j...N_F}, T_{i...N_T}\}$; the difference being the transit depth is now fit for each transit separately instead of each filter. The filters are still tracked to ensure the use of the correct limb-darkening parameters for each light curve. A single light curve is set as the "reference" light curve and used to internally compute several orbital parameters. For the XO-2 data set we used transit 3 since it had few gaps, and had the best photometric precision among the *I*-band light curves. We did not use an *r'*-band light curve as a reference since this filter is more affected by limb-darkening when compared to the *I*-band and hence estimates of the planet-to-star size ratio ($R_p/R_\star$) (used in the computation of orbital parameters) may be susceptible to degeneracies. Using the Multi-Depth model aids in understanding transit depth variations over multiple epochs.

3.3. Transit MCMC

We developed a MCMC analyzer called Transit MCMC (TMCMC), based on the Metropolis–Hastings algorithm (Gelman et al. 2003; Tegmark et al. 2004; Ford 2005), with an adaptive step-size modifier (Collier Cameron et al. 2007). MCMC routines quantify the uncertainty distributions of model parameters (given the data) using Bayes' theorem, by sampling parameter space such that samples from high probability regions (low χ^2) are selected at a greater rate than those from low probability regions. The final ensemble of sampled points the MCMC routine typically represent the uncertainty distributions of the model parameters. One must note that adaptive MCMCs are generally not considered to be truly Markovian in nature and their results

light curve. This method of selecting nuisance parameters is admittedly ad-hoc; automated parameter selection techniques could be applied in the future. The detrending is performed using the final set of nuisance parameters in conjunction with fitting the transit parameters in order to ensure that the correction process is not biased by the trial model used in the selection of nuisance parameters.

The 10 transits of XO-2b are shown in Figure 1 in normalized flux ratios (with offsets for clarity). Shown are the detrended lightcurves, the corresponding errors, and the best fit transit model; the data shown in Figure 1 are listed in Table 3. The plotted data result from the data reduction and model fitting processes described in Sections 3.1–3.3.

3.2. MultiTransitQuick

We developed a transit model called MultiTransitQuick (MTQ) in PYTHON, which is based on the analytic light curve models presented in Mandel & Agol (2002), and the PYTHON implementation of some of its functions (from EXOFAST by Eastman et al. 2013). The description of MTQ used for this study is described in some detail in Kundurthy et al. (2013). MTQ can be used to (1) fit for transit depths for data taken with multiple filters (Multi-Filter model), or (2) fit each transit depth individually (Multi-Depth).

Table 4
TMCMC Chains for XO-2

Chain	Model Vector	N_{free}	Chain Length	Corr. Length	Eff Length	χ^2	DOF
FLDC	$\boldsymbol{\theta}_{\text{Multi-Filter}}$	14	1,900,001	100	19,000	3272.94	3312
OLDC	$\boldsymbol{\theta}_{\text{Multi-Filter}}$	18	1,900,001	4,949	383	3268.48	3308
MDFLDC	$\boldsymbol{\theta}_{\text{Multi-Depth}}$	22	1,900,001	588	3,231	3197.53	3304

are valid only if adaptation diminishes with time (Roberts & Rosenthal 2009); a property that our chains do display.

For APOSTLE data sets, we explored system parameters using three different kinds of chains. Two of these were based on the Multi-Filter parameter set $\boldsymbol{\theta}_{\text{Multi-Filter}}$ described in Section 3.2 and the third used the Multi-Depth parameter set ($\boldsymbol{\theta}_{\text{Multi-Depth}}$). The two Multi-Filter chains used fixed limb-darkening coefficients (LDC) and open limb-darkening coefficients (OLDC). For the fixed LDC chains (FLDC), the coefficients were simply fixed to values tabulated for the appropriate observing filter (Claret & Bloemen 2011). For the OLDC chains, the limb-darkening parameters v_1 and v_2, for both I-band and r'-band data, are allowed to float. The ability to constrain stellar limb-darkening requires high precision data, such as those collected using the *Hubble Space Telescope* (Brown et al. 2001; Knutson et al. 2007). Previous attempts to fit for limb-darkening on APOSTLE data have resulted in Markov chains that failed to converge (Kundurthy et al. 2011, 2013). The third type of Markov chain was run on the Multi-Depth parameter set $\boldsymbol{\theta}_{\text{Multi-Depth}}$ described in Section 3.2. APOSTLE light curves were gathered over a long time-baseline, and statistically significant depth variations seen in the data may help shed light on the various phenomena responsible for depth variations (see Section 3.2), or point to limitations in the data and model.

Several of the preliminary steps for executing a chain using TMCMC are described in (Kundurthy et al. 2013). These steps include (1) setting bounds and (2) running short single-parameter exploratory chains to determine a set of suitable starting jump-sizes for the long Markov Chains. We ran long chains of 2×10^6 steps from two different starting locations for each model scenario: FLDC, OLDC, and Multi-Depth/FLDC. After completion we (1) cropped the initial stages of these chains to remove the burn-in phase, where the chain is far from the best-fit region, and (2) we exclude the stage where the chain is far from the optimal acceptance rate of 23% ± 5% (as noted for multiparameter chains; Gelman et al. 2003). We run three types of post-processing on the chains after cropping: (1) We compute the ranked and unranked correlations in the chains of every fit parameter with respect to the others. These statistics provide an estimate of the level of degeneracy between parameters in a given model. The next post-processing steps are two commonly used diagnostics to check for chain convergence, namely (2) computing the auto-correlation lengths and (3) the Gelman–Rubin $\hat{R}$-static values (Gelman & Rubin 1992). The auto-correlation lengths determine the scale over which a chain has local trends. From the auto-correlation length one can compute the effective length as the total chain length divided by the auto-correlation length, which represents the statistical significance with which the uncertainty distribution was sampled. A large effective length (>1000) represents a well-sampled distribution. The $\hat{R}$-statistic represents the level of coverage the chain has over the parameter space. When parameter space has been properly sampled the $\hat{R}$-statistic computed using chains from different starting locations will be close to 1. We deem those

chains as converged that have an effective length >1000 and an $\hat{R}$-statistic within 10% of 1.

3.3.1. TAP

The Transit Analysis Package (TAP; Gazak et al. 2012) implements the red-noise model of Carter & Winn (2009), who find that models that do not fit for red-noise are subject to inaccuracies in transit parameters on the order of 2σ–3σ and tend to have underestimated errors by up to 30%. For transit timing studies, poor estimates such as these are cause for concern, since smaller errors and large deviations from the expected time can easily lead to false claims of TTVs. Since TMCMC does not include red-noise analysis we run fits on APOSTLE light curves using TAP as a check to the results derived from TMCMC.

The typical TAP parameter set is: $\boldsymbol{\theta}_{\text{TAP}} = \{a/R_\star, i_{\text{orb}}, (R_p/R_\star)_{i...N_F}, T_{i...N_T}, \sigma_{(\text{white},i...N_T)}, \sigma_{(\text{red},i...N_T)}\}$, where $a/R_\star$, i_{orb} and $(R_p/R_\star)$ are the commonly fit transit parameters denoting the semi-major axis (in stellar radius units), the orbital inclination and the planet-to-star radius ratio, respectively. The noise analysis parameters $\sigma_{(\text{white},i...N_T)}$ and $\sigma_{(\text{red},i...N_T)}$ are the white-noise and red-noise levels for N_T transits respectively. The TAP package does not fit for the period using the transit times, and often yields poor estimates of the period, so we fixed the period to the value derived from TMCMC. The limb-darkening was fixed to values from the literature. The orbital eccentricity (e) and argument of periastron were kept fixed at 0 for XO-2b.

4. SYSTEM PARAMETERS

This section describes results from our execution of the two chains for the parameter set $\boldsymbol{\theta}_{\text{Multi-Filter}}$, and one for the $\boldsymbol{\theta}_{\text{Multi-Depth}}$ described in Section 3.3. Post processing statistics and other data for these chains are listed in Table 4. The columns "N_{free}," "Chain Length," "Corr. Length" and "Eff. Length" list the number of free parameters, the length of the cropped chain, the correlation and effective lengths, respectively. All chains were run for approximately 2 million steps, but about 100,000 of the initial steps were removed to account for "burn-in" and selection rate stabilization. The XO-2b chain with OLDC has a low effective length indicating poor Markov chain statistics. The FLDC chains (both $\boldsymbol{\theta}_{\text{Multi-Filter}}$ and $\boldsymbol{\theta}_{\text{Multi-Depth}}$) model satisfy the condition of a well sampled posterior distribution (effective length is >1000). The final two columns list the goodness of fit (i.e., lowest χ^2 in the MCMC ensemble) and degrees-of-freedom (DOF) from the respective chain. Parameters from all chains had Gelman–Rubin $\hat{R}$-statistics close to 1 indicating that the parameter space was covered evenly (though the OLDC chain was not sampled finely enough, based on the auto-correlation data).

The resulting best-fit parameter estimates are listed in Table 5 for the Multi-Filter models, and in Table 6 for the Multi-Depth models. These tables also list the derived system parameters. The transformation between the MTQ parameters to the derived system parameters are described in Carter et al. (2008) and

Table 5
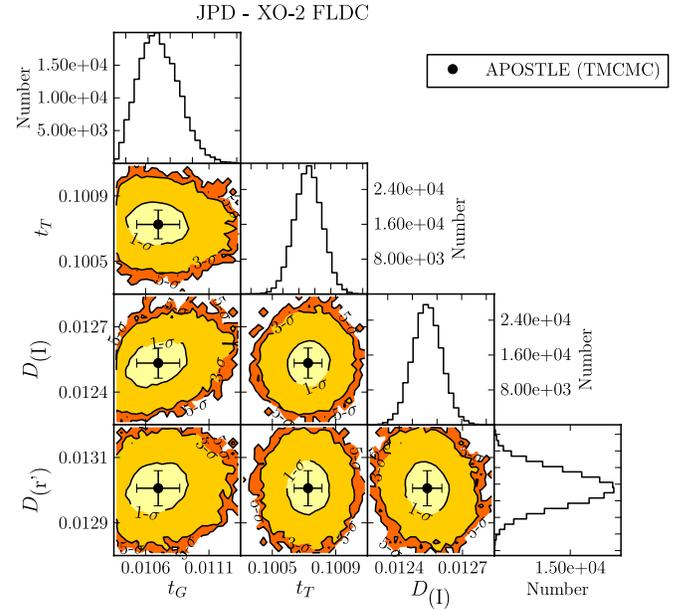XO-2 Parameters for $\theta_{\text{Multi-Filter}}$

Parameter	FLDC	OLDC	Unit
	MTQ Parameters		
t_G	0.0107 ± 0.0002	0.0108 ± 0.0002	days
t_T	0.1008 ± 0.0001	0.1004 ± 0.0002	days
$D_{(I)}$	0.0126 ± 0.0001	0.0126 ± 0.0001	...
$D_{(r')}$	0.0130 ± 0.0001	0.0131 ± 0.0001	...
$v_1(I)$	(0.5944)	0.5440 ± 0.0313	...
$v_1(r')$	(0.6980)	$0.6411^{+0.0332}_{-0.0249}$	...
$v_2(I)$	(0.1452)	0.2806 ± 0.0904	...
$v_2(r')$	(0.3524)	$0.4786^{+0.0734}_{-0.0992}$	...
	Derived Parameters		
$(R_p/R_\star)_{(I)}$	0.1030 ± 0.0003	0.1033 ± 0.0004	...
$(R_p/R_\star)_{(r')}$	0.1024 ± 0.0003	0.1029 ± 0.0004	...
b	$0.17^{+0.04}_{-0.02}$	0.21 ± 0.03	...
$a/R_\star$	8.14 ± 0.06	8.11 ± 0.07	...
i_{orb}	88.79 ± 0.15	$88.53^{+0.02}_{-0.10}$	(deg)
$v/R_\star$	19.55 ± 0.15	19.48 ± 0.17	days^{-1}
$\rho_\star$	1.49 ± 0.03	1.48 ± 0.04	g cm^{-3}
P (2.6159 days +)	-3467 ± 22	-3466 ± 22	ms

Table 6
XO-2 Parameters for $\theta_{\text{Multi-Depth}}$

| Transit Depths	Value	Units	$R_p/R_\star$	Value	Units
		I-band			
$(D)_1$	0.01338 ± 0.00013	...	$(R_p/R_\star)_1$	0.1061 ± 0.0005	...
$(D)_2$	0.01300 ± 0.00017	...	$(R_p/R_\star)_2$	0.1047 ± 0.0007	...
$(D)_3$	0.01226 ± 0.00013	...	$(R_p/R_\star)_3$	0.1019 ± 0.0005	...
$(D)_4$	0.01199 ± 0.00013	...	$(R_p/R_\star)_4$	0.1009 ± 0.0005	...
$(D)_5$	0.01209 ± 0.00039	...	$(R_p/R_\star)_5$	0.1013 ± 0.0015	...
$(D)_6$	0.01267 ± 0.00015	...	$(R_p/R_\star)_6$	0.1035 ± 0.0006	...
		r'-band			
$(D)_7$	0.01326 ± 0.00009	...	$(R_p/R_\star)_7$	0.1032 ± 0.0004	...
$(D)_8$	0.01320 ± 0.00015	...	$(R_p/R_\star)_8$	0.1030 ± 0.0006	...
$(D)_9$	0.01307 ± 0.00009	...	$(R_p/R_\star)_9$	0.1026 ± 0.0003	...
$(D)_{10}$	0.01300 ± 0.00013	...	$(R_p/R_\star)_{10}$	0.1024 ± 0.0005	...
		Other MTQ Parameters			
Parameter	Value	Units	Parameter	Value	Units
t_G	0.0108 ± 0.0001	days	t_T	0.1008 ± 0.0001	days
$v_1(I)$	(0.5944)	...	$v_1(r')$	(0.6980)	...
$v_2(I)$	(0.1452)	...	$v_2(r')$	(0.3524)	...
		Derived Parameters			
b	0.24 ± 0.02	...	$a/R_\star$	$8.02^{+0.03}_{-0.04}$	...
i_{orb}	88.29 ± 0.15	$^\circ$(deg)	$v/R_\star$	$19.27^{+0.08}_{-0.10}$	days^{-1}
$\rho_\star$	1.43 ± 0.02	g cm^{-3}	...	...	...

Figure 2. Plots of the joint probability distributions (JPDs) of parameters from the fixed LDC chains, demonstrating that parameters chosen for $\theta_{\text{Multi-Filter}}$ are generally uncorrelated. Table 5 gives the units.

(A color version of this figure is available in the online journal.)
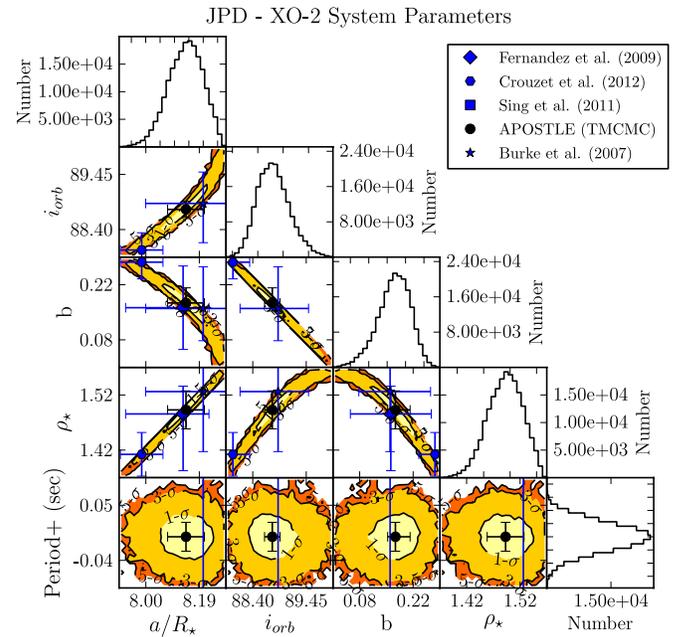
Figure 3. Plots of the joint probability distributions (JPDs) of derived system parameters from the fixed LDC chains. Parameter estimates available in the literature are overplotted. Table 5 gives the units.

(A color version of this figure is available in the online journal.)

Kundurthy et al. (2011). Contour plots showing the joint probability distributions (JPDs) for the fit and derived parameters are shown in Figures 2 and 3, respectively. There are no strong correlations between the fit parameters t_T, t_G and the transit depths, D_I and $D_{r'}$, as seen in Figure 2.

System parameters agree with previously published values in the literature, as seen by the overlap of the uncertainties in the JPD plot (Figure 3). In addition APOSTLE's measurements give tighter constraints on several of the system parameters. The errors on $a/R_\star$, i_{orb}, the impact parameter (b) and the stellar density ($\rho_\star$) are more precise than previous measurements by a factor of ~3 (Burke et al. 2007; Fernandez et al. 2009; Sing et al. 2011). However, previous studies using the com-

bination of the TMCMC Markov chain analyzer and MTQ transit model have resulted in underestimated errors since we do not include rednoise analysis (Carter & Winn 2009). Hence, the errors presented in Table 6 and Figure 3 do not truly reflect improvements in the measurements of parameters. More conservative constraints were placed on a subset of these system parameters using the TAP package. Comparisons of some parameters and their uncertainties are presented in Table 7.

It is clear that using TAP on the APOSTLE dataset and accounting for rednoise provides more conservative estimates
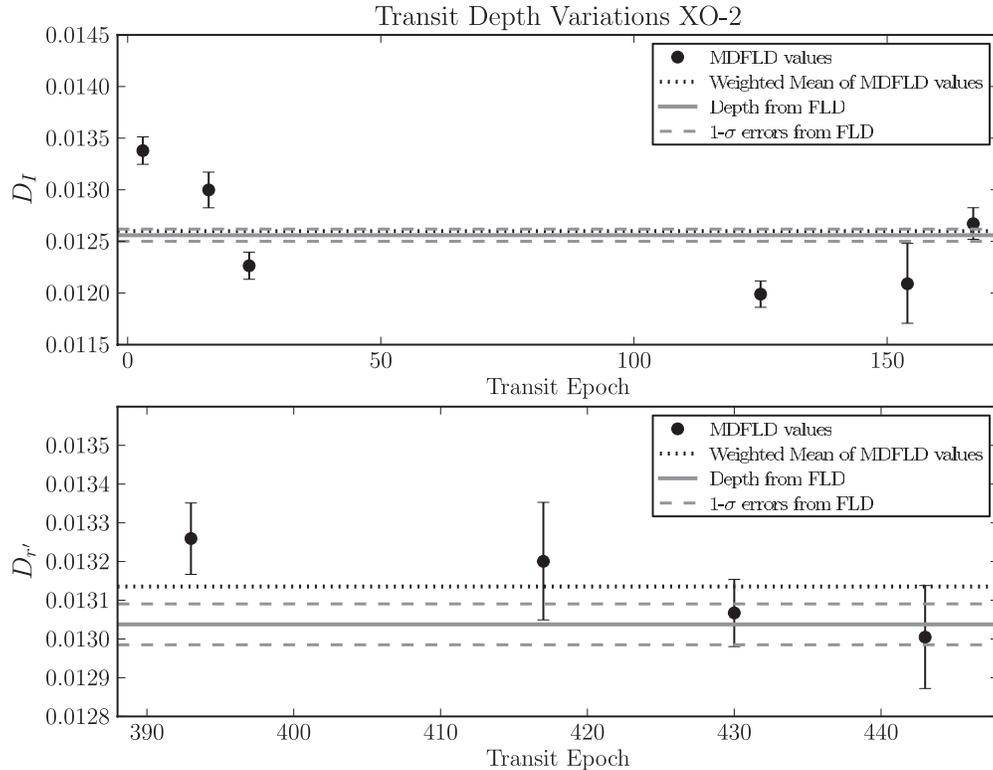
Figure 4. The transit depth D as a function of transit epoch for both I-band and r'-band observations of XO-2b. The solid horizontal and dashed lines represent the best-fit value and errors, respectively, for D from the fixed LDC TMCMC fit. The dotted line is the weighted mean of transit depth values from the Multi-Depth fixed LDC chains.

Table 7
Comparison of Estimates of System Parameters for XO-2b

Parameter	TMCMC	TAP	B07	F09	S11	C12	Units
$a/R_\star$	8.14 ± 0.06	8.13 ± 0.10	8.20 ± 0.20	8.13 ± 0.20	7.83 ± 0.17	7.99 ± 0.07	...
$i_{\rm orb}$	88.79 ± 0.15	88.80 ± 0.61	88.90 ± 0.75	...	87.62 ± 0.51	88.01 ± 0.33	(deg)
b	$0.172^{+0.040}_{-0.021}$	0.171 ± 0.085	0.158 ± 0.110	0.160 ± 0.110	0.324 ± 0.070	0.280 ± 0.044	...
$\rho_\star$	1.49 ± 0.03	1.49 ± 0.05	1.52 ± 0.11	1.48 ± 0.10	1.33 ± 0.09	1.41 ± 0.04	g cm^{-3}

Notes. TMCMC and TAP values are from independent analysis of APOSTLE light curves; B07: Burke et al. 2007; F09: Fernandez et al. 2009; S11: Sing et al. 2011; C12: Crouzet et al. 2012.

of the system parameters when compared to TMCMC values. TAP errors for $a/R_\star$ and $\rho_\star$ are better than those reported by Sing et al. (2011) by up to ~40%, whose observations were made using the 10.4 m Gran Telescopio Canarias (GTC). One must note the Sing et al. (2011) data result from narrow-band photometry and hence would have much lower photometric precision when compared to broadband observations from the same telescope. The resulting TAP errors for the impact parameter (b) and the orbital inclination are larger than those from the GTC study by ~20%. Thus we report only some improvements in the measurements of system parameters. Our system parameters agree with the current best estimates reported in Crouzet et al. (2012), but our error bars are larger by factors $\leqslant 2$.

4.1. Transit Depth Analysis

Figure 4 shows transit depth versus transit epoch for the I-band (top panel) and r'-band (bottom panel). The overall variations in the I-band depth are ~0.05% compared to the 0.01% uncertainty in $D_{(I)}$ from the joint fit to depths in the FLDC chain. Depth variations can be caused by spots. Even though the variations we present are significant we refrain from claiming the detection of spot-modulation. These deviations are likely due

to the incomplete sampling of several transits. Spots influence stellar brightness to a greater extent at shorter wavelengths, so the r'-band would be more conducive to showing depth variations. However, the overall depth variation in the r'-band light curves is of the order of 0.01% and is consistent with the errors on $D_{(r')}$ from the fit reported by the FLDC chain (Table 5). The variations seen in the I-band depths are difficult to explain. They can either be due to (1) real brightness variations caused by spot-modulations, or (2) variations arising from the transit model's inability to accurately constrain transit depths and errors given the incomplete sampling of light curves (like transits 4–6). We note that errors on the transit depth from our Multi-Depth LDC chains are more sensitive to the incompleteness of in-eclipse data rather than the photometric precision of a given light curve. For example, transit numbers 4 and 6 have 86% and 72% of the in-eclipse transit data sampled, respectively; they also have light curve residuals of 939 ppm and 405 ppm, respectively. Thus even though transit 6 has significantly better photometric precision, the depth for transit 4 is constrained slightly better due to the fact that more in-eclipse points are available. A more dramatic difference can be seen with transit 5, which has only 47% of the in-eclipse light curve sampled and the worst

Figure 5. The observed minus computed transit times for XO-2b. Values from APOSTLE's TMCMC fit, TAP, and the literature are plotted. The horizontal axis represents the transit epoch. The zero-line ephemeris is described in Section 4.2.

(A color version of this figure is available in the online journal.)

Table 8
APOSTLE Transit Times for XO2

Epoch	T0 (TMCMC) 2,400,000+ (BJD)	σ_{T0} (BJD)	T0 (TAP) 2,400,000+ (BJD)	σ_{T0} (BJD)
26	54474.73242	0.00011	54474.73252	0.00021
39	54508.73875	0.00013	54508.73877	0.00020
47	54529.66596	0.00009	54529.66604	0.00016
148	54793.86845	0.00024	54793.86836	0.00041
177	54869.72753	0.00023	54869.72735	0.00050
190	54903.73397	0.00014	54903.73403	0.00020
416	55494.91796	0.00015	55494.91794	0.00021
440	55557.69891	0.00020	55557.69892	0.00025
453	55591.70507	0.00008	55591.70511	0.00015
466	55625.71093	0.00014	55625.71112	0.00019

Fit	Period (days)	σ_P	T0 (BJD)	σ_{T0}
TMCMC	2.615860095	±0.000000209	2454474.7327333	±0.0000599
TAP	2.615860014	±0.000000346	2454474.7327964	±0.0001028

The observed minus computed mid-transit times ($O−C$) are plotted in Figure 5. There are two versions of APOSTLE transit times presented, one from the FLDC chain (TMCMC + MTQ), and the other from the TAP fit to APOSTLE light curves. These transit times are tabulated in Table 8. A linear ephemeris was fit to all the data (including literature values) using the equation,

$$T_i = T0 + \text{Epoch}_i \times P, \qquad (4)$$

resulting in a best fit ephemeris of

$$P = 2.61585988 \pm 0.00000016 \text{ days}$$
$$T0 = 2454406.720516 \pm 0.000046 \text{ BJD}$$

with a goodness of fit $\chi^2 = 105.14$ for 31 DOF. The reduced chi-squared $\chi^2/\nu = 3.39$ indicates that the linear fit is not robust, and is either due to the large timing deviations in the data or underestimated errors. For example, several of the Sing et al. (2011) data points lie far from the zero $O−C$ line, with the largest deviation being ∼106 s The scatter in the $O−C$ values as a whole is ∼39 s which makes the 106 s deviation fall within the 3σ confidence interval of the collective data set. Though the linear ephemeris does not precisely fit the timing data, the level of variation is not significant enough for us to claim unseen planets as the cause.

In order to compare the ephemerides derived with and without rednoise analysis, we fit for a linear ephemeris to the APOSTLE transit times from TMCMC and TAP, respectively (presented in Table 8). The difference between the periods derived for these subsets and the period derived from all available transit times was <7 ms. The results from the fits to the TMCMC and TAP transit times are presented at the bottom of Table 8. The reduced χ^2s for linear ephemeris fits to the TMCMC and TAP subsets were 3.96 and 1.26, respectively, confirming that TAP gives more conservative errors for the transit times due to the red-noise analysis.

We were also able to rule out sinusoidal trends in the data by running a generalized Lomb–Ccargle analysis on the $O − C$ data (Zechmeister & Kürster 2009). We fit for a period, amplitude, and phase offset on two sets of $O − C$ data. The first set included the literature dataset and all APOSTLE measurements, and the second included only the TAP measurements of APOSTLE. We found that a sinusoid of period ∼19 days and amplitude of ∼32 s improved the fit when the entire timing data set was used. The sinusoid fit yielded a $\chi^2 = 81.8$, which was an improvement

depth constraint. Other than transit 5 all transits have eclipses sampled to better than 70%, hence the depth uncertainties are all comparable. Even though the Multi-Depth LDC chain satisfied the convergence criteria described in Section 3.3, the effective length of the chain is far lesser than the value derived for the Multi-Filter LDC chain (see Table 4), indicating the lower significance of these MCMC results. In addition one must note the lack of rednoise analysis in TMCMC, which implies that the errors presented Figure 4 are probably underestimates.

4.2. Transit Timing Analysis

Several planetary systems have been observed to have TTVs (Holman et al. 2010; Lissauer et al. 2011a; Ballard et al. 2011; Nesvorný et al. 2012). In certain configurations these variations can be on the order of minutes and can be easily seen in an diagram showing observed minus computed transit times (i.e., the difference between the measured times and times expected from purely Keplerian orbital periods). Using TTVs to look for additional planets was first proposed by Agol et al. (2005) and Holman & Murray (2005), who showed that unseen planetary siblings can gravitationally influence the orbits of a known transiting planet. The TTV signals are known to be especially strong if the unseen companion lies close to mean motion resonance with the transiting planet.

Available transit times from the literature for XO-2b include those of the discovery paper by Burke et al. (2007), and the follow-up observations by Fernandez et al. (2009), Sing et al. (2011), and Crouzet et al. (2012). We excluded the 11 transit times from Burke et al. (2007) since the reported timing errors were on the order of ∼3 minutes, and are too large to provide meaningful constraints on the orbital ephemeris. We do include transit times from the other three studies. The time coordinate BJD (TDB) has become the standard system used for transit timing studies (Eastman et al. 2010) and all transit times used in this study were brought to this system. The timestamps of all APOSTLE data were converted to BJD (TDB) in the customized reduction pipeline (Kundurthy et al. 2011). The APOSTLE pipeline's time conversions have been verified by comparison to the commonly used time conversion routines made available by Eastman et al. (2010).

Table 9
Range of Parameter Space for Stability Calculations

Parameter	Unit	Min	Max	Δ	N
Mass	$M_\oplus$	1	10	1	10
a	AU	0.055	0.062	0.0016	5
e	. . .	0	0.8	0.1	9
i_{orb}	°	0.001	70	10	8
Ω	°	0	324	36	10
ω	°	0	324	36	10
M	°	0	324	36	10

of $\Delta\chi^2 = 23.3$ compared to the linear ephemeris (see above). However, the reduced χ^2 remains greater than 1, making it a non-robust fit. The periodogram analysis on only the TAP data yielded a period and amplitude of 8.5 days and 27.4 s, respectively. The $\Delta\chi^2$ showed improvement compared to the linear ephemeris model by ~ 7.6. In a manner similar to Becker et al. (2013) we tested the significance of this sinusoidal fit by repeating the analysis on 10^5 randomly cycled permutations of the amplitudes of the $O - C$ measurements, keeping the epochs fixed. We found that 76.7% of the fits showed improved $\Delta\chi^2$ at amplitudes greater than or equal to 27.4 s, indicating that the periodicity is not likely from a real TTV.

For the case when planets are in mean-motion resonance (MMR), Agol et al. (2005) showed that the analytic expression, $\delta t_{max} \sim (P/4.5j)(m_{pert}/(m_{pert} + m_{trans}))$, can roughly estimate the amplitude of the timing deviation (δt_{max}). The quantities m_{pert}, m_{trans}, P and j are the mass of the unseen perturber, the mass of the transiting planet, the orbital period of the transiting planet and the order of the resonance, respectively. For the XO-2 system, we can rule out possible system configurations given the amplitude of the weak sinusoidal fit to the TTV data (~ 27.4 s). Using the orbital period from Table 8 and the mass of XO-2b, $m_{trans} = 0.565 \, M_{Jup}$ (Fernandez et al. 2009), we compute the maximum mass perturber that could exist in the XO-2 system in the 2:1 MMR to be $\sim 0.2 \, M_\oplus$, i.e., additional planets with $M_p < 0.2 \, M_\oplus$ may exist near the 2:1 MMR given these data. At higher order resonances, this maximum mass (for a possible perturber) is larger. A more detailed analysis of TTV needs to account for a variety of orbital configurations of hypothetical companions, which is addressed in the following section.

5. SYSTEM DYNAMICS

In order to evaluate the stability of systems that produce a detectable TTV, we integrated 3.6×10^6 orbital configurations designed to mimic the XO-2 system. Each trial consisted of the known planet XO-2b and a hypothetical terrestrial-like companion.

The parameter space we cover is presented in Table 9. In this table Δ is the interval between values, all of which were varied uniformly in the range between "Min" and "Max," producing N bins. In this table, m is mass of the hypothetical terrestrial exoplanet, e is eccentricity, i_{orb} is the orbital inclination measured from the plane perpendicular to the sky, Ω is the longitude of ascending node, ω is the argument of pericenter, and M is the mean anomaly. Since our timing precisions are at the level of a signal from a terrestrial planet, we limited the mass coverage to the "super-Earth" range of 1–10 $M_\oplus$ planets. Our semi-major axis coverage spans the 2:1 resonance, with one choice directly at the commensurability. All other parameters were varied so as to fully cover all possible architectures, but still to keep the entire program tractable. Each simulation required $\gtrsim 1$ hour of CPU time, or ~ 7 million CPU hours total.
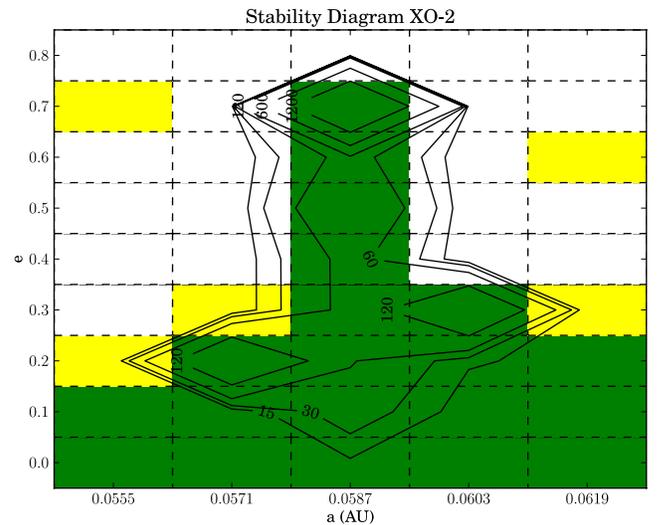
Figure 6. Contours of TTVs for orbital configurations of a 1 $M_\oplus$ companion to XO-2b, close to the 2:1 resonance. The color-coded squares note the stability of N-body simulations, with yellow (light gray) being unstable, white being failed, and green (dark gray) being stable simulations.
(A color version of this figure is available in the online journal.)

For the N-body simulations, we used the symplectic integrator code in the HNBody package (Rauch & Hamilton 2002),[7] which includes general relativity. We integrated each trial for 10^6 orbits of XO-2b (~ 7200 yr), which is enough to identify $>95\%$ of unstable orbits for other exoplanet system (Barnes & Quinn 2004). Each simulation could produce one of three outcomes: stable, unstable, or fail. Stable configurations lost no planets due to gravitational perturbations, while unstable ones did. Failed systems did not conserve energy to better than 1 part in 10^4, which is required for symplectic integrators (Barnes & Quinn 2004).

In Figure 6, a sample slice through the data set is shown. We limited the visualization to cases with $m \sim 1 \, M_\oplus$ planets with $i_{orb} = \Omega = \omega = 0$, and $M = 144°$. Green squares designate stable configurations, red unstable, and yellow failed. Most simulations from the yellow bins are probably unstable, as close approaches between planets can violate the algorithm's underlying assumption that the gravitational force from the star is much larger than that from any other planet. We note that the average fractional change in energy (dE/E) for stable cases was 7×10^{-9}.

The trials shown in Figure 6 show the that, for this value of M, the mean motion resonance at 0.0587 AU stabilizes the system. At all values of a and low e, the system is stable, as predicted by Hill stability theory (see Marchal & Bozis 1982; Gladman 1993; Barnes & Greenberg 2006; Kopparapu & Barnes 2010). However at the resonance, stability is likely for $e \leqslant 0.7$. As expected, we find that only certain values of M predict this "tongue" of stability.

The contour lines in Figure 6 show the values of the TTV signal in seconds. For the broad stable region at $e \lesssim 0.1$, the signal is at or below the detection limit. For larger e, especially in the resonance, the value can be much larger. For the bin at $a = 0.0587$ AU, and $e = 0.7$, the signal magnitude is close to 2377 s, or nearly 40 minutes. Clearly, the APOSTLE project could have detected an Earth-mass companion if it were in a favorable orbit. We do note, however, that an Earth-like planet with large eccentricity is likely to be rapidly tidally circularized

---

[7] Publicly available at http://janus.astro.umd.edu/HNBody/

Table 10
Data on the Stability Simulations of the XO-2 System

Sim. ID ... (1)	Mass ($M_\oplus$) (2)	a (AU) (3)	e ... (4)	i (°) (5)	Ω (°) (6)	ω (°) (7)	M (°) (8)	Stability ... (9)	Energy Cons. ... (10)	std(TTV) (s) (11)	std(TDV) (days) (12)	std(TBV) ... (13)
0	1.0	0.0555	0.0	0	0.0	0	0	1	2.370e-09	2.040e+00	2.656e-06	2.066e-08
1	1.0	0.0555	0.0	0	0.0	0	36	1	2.630e-09	2.062e+00	2.633e-06	2.000e-08
2	1.0	0.0555	0.0	0	0.0	0	72	1	2.610e-09	2.077e+00	2.710e-06	2.179e-08
3	1.0	0.0555	0.0	0	0.0	0	108	1	2.050e-09	2.068e+00	2.639e-06	2.165e-08
4	1.0	0.0555	0.0	0	0.0	0	144	1	1.980e-09	2.027e+00	2.684e-06	2.102e-08
5	1.0	0.0555	0.0	0	0.0	0	180	1	9.120e-10	2.029e+00	2.662e-06	1.983e-08
6	1.0	0.0555	0.0	0	0.0	0	216	1	2.210e-09	2.017e+00	2.680e-06	2.031e-08
7	1.0	0.0555	0.0	0	0.0	0	252	1	2.300e-09	2.050e+00	2.677e-06	2.140e-08
8	1.0	0.0555	0.0	0	0.0	0	288	1	2.430e-09	2.051e+00	2.673e-06	2.009e-08
9	1.0	0.0555	0.0	0	0.0	0	324	1	2.610e-09	2.031e+00	2.665e-06	1.958e-08
.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.

Notes. Column 1: simulation ID; Column 2: mass of perturber; Column 3: semi-major axis; Column 4: eccentricity; Column 5: inclination; Column 6: longitude of the ascending node; Column 7: the argument of pericenter; Column 8: the mean anomaly; Column 9: stability outcome; Column 10: energy conservation dE/E; Column 11: standard deviation of TTV—transit timing variations; Column 12: standard deviation of TDV—transit duration variations; Column 13: standard deviation of TBV—transit impact parameter (b) variations.

(This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

due to the close proximity (Rasio et al. 1996; Jackson et al. 2008; Barnes et al. 2010), although some eccentricity would be maintained by the resonance.

As noted in Table 8 (Section 4.2) the median TTV errors obtained from APOSTLE light curves of XO-2b were ~20 s. Of the 3.6 million parameter configurations, ~1.14 million were found to be stable, and approximately 51% of these stable configurations produce TTVs on the order of 20 s or less and thus would not be easily detected by APOSTLE or other surveys with similar capability. The vast amount of data generated by the stability simulations cannot be adequately discussed in this text, hence we provide the results from all 3.6 million simulations as part of an online-only table, a segment of which is displayed in Table 10. The first column "Sim. ID" denotes the simulation number and columns numbered 2–8 are the input configurations for a given simulation. The range of these simulation configurations were summarized earlier in this section and in Table 9. Columns 9–13 are the simulation statistics. Column 9 indicates the stability outcome of a given simulation, with 0 or 1 denoting unstable and stable simulations respectively. Column 10 lists the energy conservation of a given simulations as noted by the maximum fractional change in energy over a given step (dE/E). Columns 11–13 list the standard deviations in the transit timing variations, transit duration variations and the variations in the impact parameter, (b) respectively.

6. CONCLUSIONS

1. *Photometric precision.* The XO-2 system was observed over a variety of observing conditions over a period of 3 yr between 2008–2011. The best photometric precisions we obtained with our *I*-band and *r'*-band observations were 405 and 354 ppm, respectively (see Table 1).

2. *XO-2b system parameters.* Our analysis of the 10 transit light curves yielded estimates of system parameters that agree with measurements presented by other studies (see Figure 3). We were able to improve the constraints on $a/R_\star$ and $\rho_\star$ by ~40% compared to the previous best measurements from the ground (Sing et al. 2011). The

measurements are presented in Table 7; see the TAP values. We could not get the Markov chains to converge while fitting for stellar limb-darkening parameters, echoing the results from previous studies.

3. *Search for transit depth variations.* Our Multi-Depth fits show some variations in the transit depth over transit epoch (see Figure 4 and Table 6) for the six *I*-band light curves. Since three out of the six light curves were not fully sampled we cannot confidently assert real variability in the data. The seen variations could be shortcomings of the transit model's ability to fit a set of light curves with both complete and incomplete data. We do not see similar variations in the *r'*-band, where one may expect spot-modulated variability to appear, due to the greater spot-to-star contrast in the *r'*-band. There are no known reports of stellar activity on XO-2, hence the *r'*-band results are consistent with this fact.

4. *Search for transit timing variations.* The XO-2b dataset contains light curves with some of the best photometric precisions achieved with ground-based observations. Since photometric precision directly translates to transit timing precision we are able to report timing precisions as low as 12 s (after red-noise analysis, see Table 8). We were unable to detect significant timing deviations for XO-2b in our data. The linear fit was not robust, with χ^2/DOF being significantly greater than 1, indicating large scatter around the linear ephemeris fit. The transit times derived from the APOSTLE light curves using the red-noise analysis of Carter & Winn (2009) resulted in more conservative errors than those derived using TMCMC. A linear ephemeris is consistent with the transit time measurements reported from the TAP analysis. The overall variation in the $O - C$ values from our rednoise analysis was on the order of ~39 s. From a sinusoidal fit to APOSTLE's $O - C$ data, we obtained a TTV amplitude of ~27.4 s. However, checking the goodness-of-fit of sinusoids to random rearrangements of the data show that the detected periodicity unlikely to be real. The resulting amplitude rules out planetary companions more massive than 0.2 $M_\oplus$ near the 2:1 MMR, and larger companions near higher order resonances.

We conclude that the set of transit times published in the literature for XO-2b and other transiting systems in general are not suited for transit timing analysis. Lacking red-noise analysis leads to underestimated timing errors and may lead to premature reporting of timing variations. A proper analysis of transit times would need a simultaneous analysis of transit light curves using a transit model which is (1) suited for Bayesian inference (i.e., with a fairly uncorrelated parameter set; Carter et al. 2008) and (2) a transit model which can adequately account for red-noise in the data (like TAP; Gazak et al. 2012). Using large data sets of transit light curves may be inefficient due to the slowness of Markov chains with the addition of model parameters. In addition to the development of more detailed models, the utilization of fast Markov chain algorithms (for e.g., Foreman-Mackey et al. 2013) is also recommended.

Our lack of detection of TTVs in the data for the Hot-Jupiter XO-2b is also consistent with *Kepler*'s findings that (1) Hot-Jupiters tend to lack other planetary siblings (Latham et al. 2011; Steffen et al. 2012) and (2) members of multi-planet systems with short period planets (Period <10 days) are more likely to be Hot-Neptunes (Latham et al. 2011; Lissauer et al. 2011b).

5. *Dynamical study.* We ran 3.6 million *N*-body simulations of possible multi-planet configurations near the 2:1 resonance in order to test for (1) orbital stability and (2) detectability of TTV signals. We varied several properties of the hypothetical companion to XO-2b (see Table 9), to look for patterns in the resulting stability and transit times of the simulations over parameter space. Of the several stable configurations, we find that ∼51% of the simulations would display TTVs weaker than the precision limits of our survey and hence remain undetectable. The entire table of simulation statistics are presented online (see Table 10).

REFERENCES

Agol, E., Steffen, J., Sari, R., & Clarkson, W. 2005, MNRAS, 359, 567
Ballard, S., Fabrycky, D., Fressin, F., et al. 2011, ApJ, 743, 200
Barnes, R., & Greenberg, R. 2006, ApJL, 647, L163
Barnes, R., & Quinn, T. 2004, ApJ, 611, 494
Barnes, R., Raymond, S. N., Greenberg, R., Jackson, B., & Kaib, N. A. 2010, ApJL, 709, L95
Becker, A. C., Kundurthy, P., Agol, E., et al. 2013, ApJL, 764, L17
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Bessell, M. S. 1990, PASP, 102, 1181
Brown, T. M., Charbonneau, D., Gilliland, R. L., Noyes, R. W., & Burrows, A. 2001, ApJ, 552, 699
Burke, C. J., McCullough, P. R., Valenti, J. A., et al. 2007, ApJ, 671, 2115
Burrows, A., Hubeny, I., Budaj, J., Knutson, H. A., & Charbonneau, D. 2007, ApJL, 668, L171
Carter, J. A., & Winn, J. N. 2009, ApJ, 704, 51
Carter, J. A., Yee, J. C., Eastman, J., Gaudi, B. S., & Winn, J. N. 2008, ApJ, 689, 499
Charbonneau, D., Brown, T. M., Latham, D. W., & Mayor, M. 2000, ApJL, 529, L45
Claret, A., & Bloemen, S. 2011, A&A, 529, A75
Collier Cameron, A., Wilson, D. M., West, R. G., et al. 2007, MNRAS, 380, 1230
Cousins, A. W. J. 1976, MNSSA, 35, 70
Crouzet, N., McCullough, P. R., Burke, C., & Long, D. 2012, ApJ, 761, 7
Eastman, J., Gaudi, B. S., & Agol, E. 2013, PASP, 125, 83
Eastman, J., Siverd, R., & Gaudi, B. S. 2010, PASP, 122, 935
Fernandez, J. M., Holman, M. J., Winn, J. N., et al. 2009, AJ, 137, 4911
Ford, E. B. 2005, AJ, 129, 1706
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Fortney, J. J., Lodders, K., Marley, M. S., & Freedman, R. S. 2008, ApJ, 678, 1419
Fukugita, M., Ichikawa, T., Gunn, J. E., et al. 1996, AJ, 111, 1748
Gazak, J. Z., Johnson, J. A., Tonry, J., et al. 2012, AdAst, 2012, 697967
Gelman, A., & Rubin, D. B. 1992, StaSc, 7, 457
Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. 2003, Bayesian Data Analysis (2nd ed.; Boca Raton, FL: CRC Press)
Gladman, B. 1993, Icar, 106, 247
Haghighipour, N., & Kirste, S. 2011, CeMDA, 111, 267
Holman, M. J., Fabrycky, D. C., Ragozzine, D., et al. 2010, Sci, 330, 51
Holman, M. J., & Murray, N. W. 2005, Sci, 307, 1288
Hubeny, I., Burrows, A., & Sudarsky, D. 2003, ApJ, 594, 1011
Jackson, B., Greenberg, R., & Barnes, R. 2008, ApJ, 678, 1396
Knutson, H. A., Charbonneau, D., Noyes, R. W., Brown, T. M., & Gilliland, R. L. 2007, ApJ, 655, 564
Knutson, H. A., Howard, A. W., & Isaacson, H. 2010, ApJ, 720, 1569
Kopparapu, R. K., & Barnes, R. 2010, ApJ, 716, 1336
Kundurthy, P., Agol, E., Becker, A. C., et al. 2011, ApJ, 731, 123
Kundurthy, P., Becker, A. C., Agol, E., Barnes, R., & Williams, B. 2013, ApJ, 764, 8
Latham, D. W., Rowe, J. F., Quinn, S. N., et al. 2011, ApJL, 732, L24
Lissauer, J. J., Fabrycky, D. C., Ford, E. B., et al. 2011a, Natur, 470, 53
Lissauer, J. J., Ragozzine, D., Fabrycky, D. C., et al. 2011b, ApJS, 197, 8
Machalek, P., McCullough, P. R., Burke, C. J., et al. 2008, ApJ, 684, 1427
Machalek, P., McCullough, P. R., & Burrows, A., et al. 2009, ApJ, 701, 514
Madhusudhan, N. 2012, ApJ, 758, 36
Mandel, K., & Agol, E. 2002, ApJL, 580, L171
Marchal, C., & Bozis, G. 1982, CeMec, 26, 311
Monet, D. G., Levine, S. E., Canzian, B., et al. 2003, AJ, 125, 984
Mukadam, A. S., Owen, R., Mannery, E., et al. 2011, PASP, 123, 1423
Narita, N., Hirano, T., Sato, B., et al. 2011, PASJ, 63, L67
Nesvorný, D., Kipping, D. M., Buchhave, L. A., et al. 2012, Sci, 336, 1133
Rasio, F. A., Tout, C. A., Lubow, S. H., & Livio, M. 1996, ApJ, 470, 1187
Rauch, K. P., & Hamilton, D. P. 2002, BAAS, 34, 938
Roberts, G. O., & Rosenthal, J. S. 2009, J. Comput. Graph. Stat., 18, 349
Sing, D. K., Désert, J.-M., Fortney, J. J., et al. 2011, A&A, 527, A73
Southworth, J. 2010, MNRAS, 408, 1689
Steffen, J. H., Ragozzine, D., Fabrycky, D. C., et al. 2012, PNAS, 109, 7982
Tegmark, M., Strauss, M. A., Blanton, M. R., et al. 2004, PhRvD, 69, 103501
Winn, J. N. 2011, in Exoplanets, ed. S. Seager (Tucson, AZ: Univ. Arizona Press), 526
Zechmeister, M., & Kürster, M. 2009, A&A, 496, 577