Zodiac II: Debris Disk Science from a Balloon

Geoffrey Bryden\textsuperscript{a}, Wesley Traub\textsuperscript{a}, Lewis C. Roberts Jr.\textsuperscript{a}, Robin Bruno\textsuperscript{a}, Stephen Unwin\textsuperscript{a}, Stan Backovsky\textsuperscript{b}, Paul Brugarolas\textsuperscript{a}, Supriya Chakrabarti\textsuperscript{c}, Pin Chen\textsuperscript{a}, Lynne Hillenbrand\textsuperscript{d}, John Krist\textsuperscript{a}, Charles Lillie\textsuperscript{e}, Bruce Macintosh\textsuperscript{f}, Dimitri Mawet\textsuperscript{a}, Bertrand Mennesson\textsuperscript{a}, Dwight Moody\textsuperscript{a}, Zahidul Rahman\textsuperscript{a}, Justin Rey\textsuperscript{g}, Karl Stapelfeldt\textsuperscript{h}, David Stuchlik\textsuperscript{i}, John Trauger\textsuperscript{a}, and Gautam Vasisht\textsuperscript{a}

\textsuperscript{a}Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109
\textsuperscript{b}ATK Space Systems, San Diego, CA 92121
\textsuperscript{c}Boston University, Boston, MA 02215
\textsuperscript{d}California Institute of Technology, Pasadena, CA 91125
\textsuperscript{e}Northrop Grumman Corporation, Redondo Beach, CA 90278
\textsuperscript{f}Lawrence Livermore National Laboratory, Livermore, CA 94550
\textsuperscript{g}Northrop Grumman Xealit, Devens, MA 01434
\textsuperscript{h}NASA Goddard Space Flight Center, Greenbelt, MD 20771
\textsuperscript{i}NASA Wallops Flight Facility, Wallops Island, VA 23337

ABSTRACT

Zodiac II is a proposed balloon-borne science investigation of debris disks around nearby stars. Debris disks are analogs of the Asteroid Belt (mainly rocky) and Kuiper Belt (mainly icy) in our Solar System. Zodiac II will measure the size, shape, brightness, and color of a statistically significant sample of disks. These measurements will enable us to probe these fundamental questions: what do debris disks tell us about the evolution of planetary systems; how are debris disks produced; how are debris disks shaped by planets; what materials are debris disks made of; how much dust do debris disks make as they grind down; and how long do debris disks live? In addition, Zodiac II will observe hot, young exoplanets as targets of opportunity.

The Zodiac II instrument is a 1.1-m diameter SiC telescope and an imaging coronagraph on a gondola carried by a stratospheric balloon. Its data product is a set of images of each targeted debris disk in four broad visible-wavelength bands. Zodiac II will address its science questions by taking high-resolution, multi-wavelength images of the debris disks around tens of nearby stars. Mid-latitude flights are considered: overnight test flights within the United States followed by half-global flights in the Southern Hemisphere. These longer flights are required to fully explore the set of known debris disks accessible only to Zodiac II. On these targets, it will be 100 times more sensitive than the Hubble Space Telescope’s Advanced Camera for Surveys (HST/ACS); no existing telescope can match the Zodiac II contrast and resolution performance. A second objective of Zodiac II is to use the near-space environment to raise the Technology Readiness Level (TRL) of SiC mirrors, internal coronagraphs, deformable mirrors, and wavefront sensing and control, all potentially needed for a future space-based telescope for high-contrast exoplanet imaging.

Keywords: Suborbital, Coronagraph, Disks, Exoplanets

1. INTRODUCTION

Debris disks are remnants of the process of planet formation. While gas and solid material in some parts of a young protoplanetary disk may grow into gas giant and terrestrial planets, the process is never completely efficient. The Solar System, for example, has two bands of leftover material - the asteroid belt lying between Mars and Jupiter and the Kuiper Belt lying just outside Neptune’s orbit. These belts contain large Pluto-sized planetesimals, as well as smaller asteroids, boulders, and, most importantly for observations, micron-sized dust. Individual grains are removed quickly (lifetimes are generally much less than a million years), but the dust disk...
is continually replenished by the collisions of asteroids and Kuiper Belt Objects, and also the sublimation of comets. Though small in mass, the dust’s large surface area scatters enough sunlight to be seen by the naked eye - the Zodiacal light. If we could image it from a great distance, the dust surrounding the Earth’s orbit would be brighter than the Earth itself. Such an image of the Solar System’s debris disk would clearly reflect the influence of its neighboring planets. As Jupiter formed, for example, it exerted strong gravitational stirring just interior to its orbit. While there was plenty of solid material to build another terrestrial planet in that region, we instead ended up with a remnant asteroid belt, which even now is strongly shaped by resonances with Jupiter. In the outer disk, Neptune similarly sculpted the Kuiper Belt. As the young Neptune migrated outward, it pushed along the outlying material, capturing many planetesimals onto mean-motion resonances, for example as with Pluto and other plutinos (in 3:2 resonance). One of the goals of Zodiac II is to observe disk sculpting similar to that found in the Solar System, and to thereby infer the presence of underlying planets.

Beyond the Solar System, the IRAS satellite made the first detection of a debris disk orbiting around another star.\(^1\) Since then, many disks have been discovered via the dust’s thermal emission. Surveys by Spitzer have identified dust around many nearby Sun-like stars.\(^2,3\) Although these surveys are only sensitive to emission \(\sim 100\) times brighter than the Solar System’s, they find such bright disks around \(\sim 15\%\) of old solar-type stars like the Sun. To continually replenish their dust, the observed debris disks undoubtedly have planetesimals at least \(\sim \text{km}\) in size. In some cases, these systems are also known to have large planets that have been detected via their radial-velocity perturbation on the central star.\(^4\) In other cases, planets can be inferred from the dust distribution. Resolved images of debris disks often reveal asymmetries and warps that may be attributed to the gravitational influence of unseen planets.\(^5-8\)

The most dramatic example of planetary influence is the dust ring around the A star Fomalhaut, which is confined to a tight, eccentric orbit that would readily disperse in the absence of some shepherding force.\(^9,10\) As a confirmation of this shepherding, the predicted planet has now been directly imaged (Figure 1 left panel).\(^11\) As further evidence of linkage between debris and planets, we note that all of the famous systems with recently imaged planets also host prominent debris disks - in addition to Fomalhaut, HR 8799 has a four-planet imaged system (Figure 1 right panel)\(^12\) orbiting in the midst of interior and exterior belts of dust\(^13\) and Beta Pictoris has an imaged planet that may be responsible for its disk warp.\(^14\)

The morphology of a debris disk can reveal the presence of hidden planets in many ways - shepherded rings, inclined warps, eccentric offsets, cleared gaps, and resonant clumps. Observation of these features requires high-resolution imaging. While thermal images at far-IR wavelengths can provide hints of disk structure (e.g. Spitzer images of Fomalhaut\(^9\)), their spatial resolution is orders of magnitude worse than what can be achieved in scattered light images at optical wavelengths and is generally insufficient for resolving planet-induced structures. Short-wavelength observations of disks’ scattered light are required. The disadvantage of these short wavelengths, however, is the dimness of the disk relative to its parent star. Because of this limitation, very few of the debris disks suspected to encircle nearby stars (from measurements of excess infrared emission) have actually been seen, even by the Hubble Space Telescope (HST). The solution requires a high-performance coronagraph such as Zodiac II.

Five basic properties of debris disks will be revealed by Zodiac II: size, shape, brightness, color, and number: 

- **Size** - the average angular radius of the debris disk, here envisioned as a ring-like structure similar to the Asteroid and Kuiper Belts. Size provides information about the possible location of shepherding planets. The current disk size constrains the size of its original proto-planetary disk.
- **Shape** - three characteristics can be detected by imaging: decentering, waviness, and warps. The disk can be off center with respect to the star if a massive planet is influencing it, particularly one on an eccentric orbit. The disk can have a wave-like edge structure also due to planets. The disk can be warped, with the outer disk misaligned relative to the inner disk, again potentially as a result of planet influence.
- **Brightness** - the surface brightness of a disk, which is directly related to the surface density of material, and also the type of material through its albedo. A combination of infrared and visible flux data can inform us about the temperature and surface density of material, using plausible models of grain size distributions and materials.
- **Color** - the brightness variation as a function of wavelength can tell us about the type of material, the size of the individual grains, and their surface reflectivity.
- **Number** - each debris disk is likely to be individual, depending on the particular history of its planetary system.
To draw broad conclusions from the observations, statistically significant number of examples is need. Having a number of examples is also important in understanding the life cycle of debris disks as a class.\textsuperscript{16}

Taken as a whole, the goals for these five quantities determine many of the characteristics of Zodiac II. In particular, given the typical distances of the target stars, the expected brightness of the disks, and the expected physical dimensions, we can directly infer what size telescope is needed, what integration times are needed, and how many targets should be observed.

2. TARGET CHARACTERISTICS

In this section we list the observational characteristics of potential targets. We consider all known debris disks within 40 pc of the Sun, based on Spitzer detections of their thermal emission at infrared wavelengths.\textsuperscript{2,17–21} Numerous additional targets are also being identified through ongoing investigations by the WISE and Herschel space telescopes (e.g., Ref. 22). Although the disks are usually not resolved at infrared wavelengths, we can use their thermal emission to estimate the disk sizes. For each disk, we derive characteristic dust temperatures from the infrared photometry at 24 and 70 $\mu$m or from Spitzer/IRS 5-35 $\mu$m spectra when available.\textsuperscript{3,7,23,24} The thermal emission also provides for a baseline estimate for the brightness of the disk’s scattered light emission, although in practice this will depend on the dust properties (e.g. albedo) of each individual system. We assume that the dust is in thermal equilibrium with 10% dust albedo and dust emissivity from Mie theory using grain sizes just above the blowout limit of each star. These assumptions are anchored in modeling work to fit Spitzer SEDs to debris disks already imaged by HST.\textsuperscript{25} The radial distribution of the disk material is unknown; a primary science goal of Zodiac II is to measure this property. For estimating the expected signal levels, we consider uniform surface density rings with width that is 20% of their radius. This is suggested by the fact that simple single-temperature blackbodies provide a good description of the Spitzer spectral energy distribution in the majority of debris disks studied spectroscopically.\textsuperscript{23,24} A median inclination of 60\degree from face-on is assumed.

The resulting predictions for ring contrasts (with respect to star, per resolution element) are shown in Figure 2. Zodiac II’s limits in field-of-view and contrast are shown as a solid U-shaped line, where all points inside the U curve are accessible to Zodiac II in at least one wavelength band. Data points are marked according to their history of being observed by the best space system - HST/ACS before its failure (data from Ref. 26); or by their potential for future observations in the infrared by the Gemini Planet Imager (GPI), a ground-based coronagraph (data for GPI in H band from Ref. 27); or the James Webb Space Telescope’s upcoming coronagraph (JWST/NIRCAM) at K band (data from Ref. 28). Zodiac II has a high-performance coronagraph, which HST did not, such that it can out-perform HST in this area. JWST is strongly limited in its coronagraphic ability by the diffraction pattern from its segmented mirror and the lack of a sufficiently high-precision wavefront control system. GPI is limited by atmospheric effects even with its adaptive optics system.

Overall, a large region of the contrast/separation phase space is unique to Zodiac II. Based on the above observational estimates of disk size and contrast ratio, we identify 89 known debris disks as candidate targets.
Figure 2. Predicted radii and image contrasts are shown for potential target debris disks, derived from Spitzer data (left panel). The inner angular limit is for V1 band (520 nm); the outer is for I band (900 nm). The 89 filled circles within the U-shaped curve are targets unique to Zodiac II. The right panel shows signal-to-noise ratio (SNR) achieved by Zodiac II for these targets as a function of stellar magnitude and disk contrast, per resolution element and per hour, in the shortest wavelength (V1) band.

accessible only to Zodiac II (filled circles in Figure 2, left panel). For each of these targets, we calculate the expected debris disk signal-to-noise ratio (SNR) based on a 1-hour observation (Figure 2, right panel). The SNR values plotted are for a single, narrow band at V and for a single resolution element on the sky and are applicable to science questions regarding color and spatial variations in the debris disks. For simple detection of the existence of a ring, post-processing will increase these values by combining data from all points along the disk and over all 4 wavebands for a net increase in SNR by a factor of about 20.

The SNR calculation (in V band) uses a realistic net throughput of 9.6% (electrons/photon) at the science focal plane. The noise is the square root of the sum of 4 contributions: shot noise from the signal, dark counts, read noise, and the number of reads per hour, which we take to be 60. The dark and read-noise contributions are about equal with these numbers. Improved SNR from longer snapshots is expected if instrument stability permits integrations longer than 60 s. The SNR for other observing times scales roughly as the square root of the number of hours.

While debris disks are the primary focus of Zodiac II, the brightest known exoplanets can also be targeted. Four warm exoplanets, for example, have recently been detected in the HR 8799 system through near-infrared coronagraphic imaging (Figure 1 right panel). These objects have near-IR brightnesses of 16-17 magnitudes in K band, > 4 orders of magnitude fainter than the central star. Atmosphere models predict these objects should be ~5 mag fainter in the visible wavelengths observed by Zodiac II (I band) due to their low effective temperatures and strong pressure-broadened absorption of neutral potassium, but these models have never been observationally tested. The predicted I band stellar contrast of the outer two planets (b and c) is ~16 mag with separations of 0.9 and 1.7 arcsec (red stars in Figure 2), easily within reach of the Zodiac II instrument. Note that the inner two planets (d and e) fall inside the I band inner working angle (0.5 arcsec) and are too faint for observations in V band. For HR 8799 b and c, the expected SNR’s are calculated similarly to the disk SNR except that at the longer wavelength the image FWHM is larger and the number of photons is a factor of 2.25 larger. The SNR’s for a 1 hour integration are about 50 and 90 for for HR 8799 b and c, respectively; Zodiac II should have very good detections of both planets in I band. With several coronagraphs with advanced adaptive optics systems now coming online (e.g. GPI, P1640, and VLT/SPHERE), additional exoplanet targets that are suitable for optical photometry are expected to emerge in the next few years.

3. BENEFITS OF A BALLOON ENVIRONMENT
Zodiac II will operate at a balloon altitude of about 35 km (115,000 ft) to avoid the scattered light speckles that are created by the turbulent atmosphere at ground-based telescopes. This operating environment is close to space-like in the sense of very low pressure (so that electronics, for example, cannot cool by convection),
intermediate temperature (in that radiative equilibrium is controlling, with the warm Earth below and cold space above), and near vacuum path overhead (with less than 1% of the atmosphere, and an almost perfect absence of turbulence-induced seeing). While ground-based telescopes normally use adaptive optics (AO) to reduce speckles, ground-based AO is not capable of reducing raw speckles to the level needed for debris disk imaging (∼10⁷ times fainter than the star). Meanwhile the niche for ground-based telescopes, even large ones equipped with aggressive adaptive optics, will always be biased toward longer wavelengths because atmospheric spatial and temporal fluctuations are dramatically reduced at those wavelengths. The fundamental advantage of being above the atmosphere makes Zodiac II competitive with, and in many ways superior to, the best that will be done on the ground in this field.

The Earth’s stratosphere transmits a stellar wavefront with much less distortion than the troposphere due to its (100-fold) lower density and its convective stability in the vertical direction. In order to determine whether this lower level of turbulence in the stratosphere still generates significant visible wavelength speckles (vis-a-vis Zodiac II’s contrast requirements), we here quantify the seeing at balloon altitudes. We start with the measured turbulence parameters at 30-70 km altitude based on observations from the Mir space station of stellar scintillation. From these, we derive the inner scale of turbulence, outer scale of turbulence, and Fried parameter (l₀, L₀, and r₀), whose values at 35 km altitude are l₀ = 2.4 m, L₀ = 44 m, and r₀ = 41 m. The outer scale is the upper limit on the size of turbulent eddies, while the inner scale is the lower limit below which air motion is laminar such that turbulence ceases to exist. Notably, l₀ (also r₀) is ∼400 times larger than for a ground-based telescope with good seeing. This large l₀ indicates that the turbulence is dominated by large convective cells, i.e. air flow becomes laminar on scales smaller than 2 m and turbulence diminishes drastically. More recently, an independent balloon-borne stellar scintillation experiment confirmed the stratospheric value of this critical parameter, measuring l₀ = 2-4 m. Consequently, at balloon altitudes the already tenuous natural atmosphere scatters vanishingly small amounts of speckle light into angles larger than λ/l₀. Following Guyon’s formalism combined with the Hill-Andrews theory of power spectra for refractive index variations in the atmosphere, we find that when the inner scale is properly included, speckle intensities start to roll off rapidly at angles >0.05 arcsec; for angles larger than 0.2 arcsec speckle intensities are far below 10⁻⁹. Figure 3 presents the results of these calculations for the nominal case of 550 nm wavelength, 30° elevation, and 35 km float altitude. Intensities relative to the central star of speckles generated by natural index changes in the atmosphere as a function of angular position are plotted for each of the three cases (Kolmogorov, von Karman, and Hill-Andrews). Again, when the inner scale is accounted for (by the von Karman or Hill-Andrews spectra), speckle intensities start to roll off rapidly at 0.05 arcsec. Using the Hill-Andrews spectrum, we have calculated the speckle intensity for a range of wavelengths, elevation angles, and balloon altitudes. Over the range of these parameters relevant to Zodiac II, the speckle intensity is well below 10⁻⁹ in the dark hole.

Locally generated turbulence can also degrade stellar wavefronts. The magnitude of this locally generated “dome seeing” depends on radiative and conductive heat flows. To help estimate the magnitude of dome seeing effects, we flew a dedicated interferometric instrument in September 2007 to measure the amount of turbulence
in a one-meter path immediately outside the gondola of the Solar Bolometric Imager telescope. We found no turbulent contribution at the limit of measurement, with an implied speckle intensity of less than $10^{-9} - 10^{-8}$. Details of this experiment and analyses are described in Ref. 33.

4. OBSERVATORY OVERVIEW

4.1 Gondola

The Zodiac II telescope/instrument flies on a gondola (Figure 4) hanging from a zero-pressure balloon. The gondola is a rigid structure made of bolted aluminum, able to withstand a 9-g parachute opening. Its framework contains the other major subsystems and interfaces to the various balloon subsystems. The science instrument is a visible-wavelength coronagraph (§4.4) mounted to the rear of the optical bench that also supports the telescope primary mirror. Inside the gondola, the Wallops ArcSecond Pointing (WASP) system controls the telescope orientation (see pointing discussion in §4.3). At the top of the gondola a coarse-pointing rotator connects to the flight train, and thence to the balloon itself. This rotator uses azimuth data from a differential GPS Attitude Determination Unit to achieve $\sim 1^\circ$ accuracy. Solar panels attached to the back of the gondola are used during multi-night flights to recharge the onboard batteries during each day.

The gondola and telescope body pointing system are provided by the NASA Wallops Flight Facility (WFF) based on their experience with balloon payloads and in particular their previous development of a lab demo version of the WASP pointing system. As with a spacecraft, a balloon telescope must torque against an inertial load in order to point to a star. For azimuth we torque the gondola against the cables leading to the balloon (via the parachute) for slow response. The telescope is azimuth-torqued against the gondola, and the fine-pointing mirror against the telescope. Gondola elevation remains nearly vertical owing to gravity, with small oscillations pushing against the suspension cables, which in turn damp out oscillations through cable bending losses and air friction. The telescope is elevation torqued against the gondola, and the fine-pointing mirror against the telescope. Typical gondola oscillation frequencies and amplitudes have been measured and are used as baseline inputs in the design of the control system.

The WASP bearing design eliminates static friction (stiction) by keeping the bearing shaft constantly rotating at a slow speed, using a small torque motor and planetary gear reduction box. This innovation is at the heart of the fine-pointing system. This approach is utilized on both of the two telescope axes (pitch and yaw).
gimbals are driven by direct-drive brushless torque motors, which have high torque with low power inputs. A prototype system at WFF uses a 680 kg (1500 lb) dummy telescope. In lab tests, the system achieves a pointing jitter of 0.25 arcseconds RMS in both pitch and yaw over continuous periods of time.\textsuperscript{36}

A star tracker is boresighted with the telescope body, providing a reference frame for the gondola pointing. The star tracker is an ST5000 from the University of Wisconsin, which has been previously used on a number of sounding rocket launches.\textsuperscript{38} Its typical accuracy during rocket flights - 0.54 arcsec RMS error in pitch & yaw and 17 arcsec in roll - is sufficient to locate the target star within the science camera field-of-view.

Lastly, the gondola also contains telecommunications equipment provided by NASA’s Columbia Scientific Balloon Facility. On overnight flights, the standard package communicates via line-of-sight radio, providing 128 kbps downlink capability. For the long-duration flights, the telecom package supports satellite communications via the Tracking and Data Relay Satellite System (TDRSS). Nominal TDRSS communication through an omnidirectional antenna provides a 6 kbps download. There is also a high-gain TDRSS antenna that can download data at 92 kbps; this will be used during the day to download all the science data.

4.2 Telescope

The telescope has four main features that are important for Zodiac II - light weight, clear aperture, large diameter, and excellent optical quality.

Light Weight: Zodiac II requires a lightweight mirror in order to reach 35 km altitude with reasonable-size balloons. Silicon carbide (SiC) composition offers light weight, compatibility with the stress of launch and recovery, high thermal conductivity, excellent polishability, and shape retention with changing temperatures. With SiC mirrors, the telescope mass is less than 200 kg.

Clear aperture: Zodiac II has an unobscured primary mirror, with an off-axis secondary mirror, to minimize diffraction in the focal plane. The clear-aperture solution is by far the most desirable situation for an internal coronagraph. The manufacturing advantage is that a primary with no central hole is easier to polish.

Large diameter: Basic physics demands a large diameter for angular resolution and collecting power. The inner working angle and spatial resolution all improve with larger telescope diameter. Balancing performance with cost limitations and production capabilities, the optimal primary for Zodiac II is a 12-sided mirror with 1.21 m face-to-face distance and 1.1 m clear diameter. The secondary mirror has a clear diameter of 0.15 m.

Excellent optical quality: To meet the desired contrast performance, low-order (focus/tip/tilt/astig) and mid-range (2-28 cycle/m) surface accuracies are required to be within 18 and 3 nm RMS for the primary and secondary respectively. An enhanced Al reflective coating will be applied to an Si overcoat.

The support structure for the telescope is designed to be rigid and lightweight. The current design is a symmetrical metering structure that has first modes of 79 and 82 Hz. The primary mirror is mounted to the front of a stiff support plate via bipods. The coronagraph is mounted to the rear of the support plate, also via bipods.

4.3 Pointing System

Coronagraphs are very sensitive to pointing errors. Pointing offsets adversely affect image contrast in two main ways: 1) jitter and stability errors in the telescope line-of-sight pointing lead to contrast degradation due to wavefront errors generated by beam-walk, and 2) errors in centering the occulting mask accurately on the target star produce contrast degradation due to light spillage outside the mask. Ref. 39 calculated the contrast degradation sensitivities due to errors in telescope pointing and mask centering (see also Refs. 40–42). The pointing stability requirements for Zodiac II are guided by those calculations, which translate to a 0.4 arcsec RMS line-of-sight stability for the telescope body-pointing and a 0.04 arcsec RMS coronagraph mask centering error. Telescope roll around the line-of-sight is not actively controlled. The roll motions due to field rotation (<0.3 deg/min) and non-controlled gondola roll (<0.5 deg)\textsuperscript{37} are small compared to the roll pointing requirement of 1.2 deg over 60 s and do not affect the performance of the instrument.

To meet its pointing requirements Zodiac II uses a 3-stage pointing system. Figure 5 shows a control system diagram for this nested 3-stage pointing architecture. First, the gondola azimuth pointing stage provides coarse pointing of the gondola and telescope in azimuth. A GPS attitude determination unit and a single axis gyro are used to estimate the azimuth direction of the gondola; the gondola is then pointed in the desired azimuth
direction. This first stage has low bandwidth (<0.1 Hz). Second, the telescope gimbal system provides telescope pitch and yaw pointing. The telescope pointing system estimates the attitude and attitude rate of the telescope by filtering the star tracker and gyro measurements. Then it servos the pitch and yaw gimbals to point the telescope to the target star. The gimbal servo system runs at medium bandwidths (<10 Hz). For the third pointing stage, the finest scale adjustments are made by a tip-tilt system inside the coronagraph instrument which centers and stabilizes the occulting mask on the target star. The integral tip/tilt system has ±60 arcsec of tip/tilt motion, which translates to a line-of-sight range of motion projected onto the sky of ±5.2 arcsec. The sensor driving the tip-tilt is a 16x16 Shack-Hartmann wavefront sensor operating at a 500 Hz frame rate. Each subaperture is 4x4 pixels plus a 1 pixel guard band. The wavefront sensor estimates the position of the central star to 0.01 arcsec or better. This level of performance is due to the fact that the selected central stars are bright and the telescope has a large aperture. The combination provides ample signal for the wavefront sensor to support the high gain-bandwidth tip/tilt loop (<100 Hz).

The main sources of disturbances that need to be compensated for are balloon pendulations and gondola bearings disturbances. Balloon pendulations generate low amplitude (<1°) low frequency (<1 Hz) pointing errors. The fine steering mechanism could act as a disturbance by exciting the structure and inducing pointing errors. However, since the range of motion is small, and the mass participation of the moving parts is also small, the induced disturbance is negligible.

The telescope/gondola pointing system has been demonstrated in the laboratory, with the pitch and yaw gimbal system meeting the 0.4 arcsec pointing requirement (Figure 6, left panel). Meanwhile the Sunrise mission, with pointing requirements comparable to Zodiac II, successfully demonstrated that this fine level of pointing can be achieved on a balloon platform. However, they encountered unexpected pointing jitter that led to loss of efficiency and reduced science return during some observational periods. If jitter is found to be a problem for Zodiac II, mitigation approaches will be to add tuned mass magnetic dampers to damp jitter at specific locations and to add passive isolation to isolate the telescope from the gondola.

To assess the coronagraph fine pointing capability needed for centering a star on the coronagraph mask, and to evaluate the pointing stability of the telescope pointing gimbals in conjunction with the coronagraph's tip-tilt fine steering, we developed a preliminary design for the fine steering control loop and analyzed its performance against data from a realistic control closed-loop experiment of a prototype of the Wallops gimbals system. The right panel of Figure 6 shows the uncompensated pointing jitter residual from a telescope gimbal (dashed) and
the tip/tilt compensated pointing jitter (solid). As a result of this third pointing stage, the pointing error is reduced from 0.26 arcsec RMS to 0.009 arcsec RMS.

4.4 Coronagraph

To observe exoplanets and debris disks we must be able to detect very faint material very close to a bright star. This is precisely what coronagraphs are designed to do. Diffracted light is blocked by putting a precisely tapered dark spot in the focal plane that is $\sim 3$ times larger than the core size of the star ($\lambda/D_{tel}$ in angle on the sky) and surrounding the spot with a pattern of faintly darkened lines or circles. The central dark spot eliminates the main part of the starlight, and the faint pattern eliminates all the diffracted Airy rings.

The coronagraph must not only block the direct and diffracted light from the target star, as collected by a perfect telescope, but it must also block the scattered light from optical imperfections in the telescope. The scattered light forms a field of spots, called speckles, in the otherwise ideally dark focal plane. These speckles create a noise floor for any observations, and must be removed where possible. This is done by inserting deformable mirrors in the beam of light, between the telescope and the focus. The deformable mirrors add a pattern of wavy corrugations opposite to those of the telescope, effectively canceling them out. The resulting newly-dark area is called the dark hole. It is here where one can look for faint but real planets and debris disks. For an extended discussion of direct imaging by coronagraphs, see the book chapter in Ref. 45.

While all coronagraphs remove central light to some degree, the coronagraph used for Zodiac II is based on an extremely high-performance design - using a similar coronagraph, we have achieved contrasts in a stable lab environment at a level of sensitivity a factor of 1000 better than needed for Zodiac II. The heart of the Zodiac II coronagraph is a focal-plane band-limited mask that suppresses the telescope’s diffraction pattern. The incident wavefront must be nearly perfect for the mask to operate correctly, so two deformable mirrors (DM) precede it: a low-order deformable mirror (LODM) mounted in a tip-tilt mechanism. The stabilized collimated beam continues to the 48-mm square (and 48x48 element) HODM. Next, the beam is focused by OAP2 onto the coronagraph mask. The Zodiac II coronagraph is a flight version of the High-Contrast Imaging Testbed (HCIT) which has been in operation in the lab at JPL for over 6 years and has demonstrated contrast ratios of $6 \times 10^{-10}$ with a stability of $0.1 \times 10^{-10}$ RMS over 5 hours. Following the light through the system, the optical elements are as follows. Starlight is focused by the off-axis primary mirror. The light passes to an off-axis secondary mirror (SM), forming a Gregorian telescope. The star is focused inside the coronagraph. An off-axis parabola (OAP1) collimates the beam and forms an image of the primary on the LODM, which is mounted in a tip-tilt mechanism. The stabilized collimated beam continues to the 48-mm square (and 48x48 element) HODM. Next, the beam is focused by OAP2 onto the coronagraph mask. The Zodiac II
The low-order wavefront sensor (LOWFS) measures the low-order aberrations (first 15 Zernikes) using a Shack-Hartmann sensor. LOWFS data are used to drive the LODM and its associated tip-tilt stage. It receives the light reflected from the front of the coronagraph mask. The LOWFS detector is an 80x80 frame-transfer CCD that is widely used as the wavefront sensor for ground-based adaptive optics, such as the Palomar Adaptive Optics system. The LOWFS includes a field stop, collimating lens, and lenslet array.

The science camera records the images from all four wavebands in parallel, with each band falling in a separate quadrant of the array (see Figure 9). The camera also provides the information for real-time speckle suppression by the deformable mirrors. The camera detector is an e2v CCD57-10, a frame-transfer device enabling a low-speed read out, minimizing read noise, and eliminating the need for a shutter. It is thermoelectrically cooled to -35°C to reduce dark current. We baseline 60-s integrations with this detector, to balance read noise and dark count noise and to minimize snapshot time compared to the time scale of potential wavefront changes.

4.5 Thermal Environment

There are several critical thermal control areas.

Optics. The Zodiac II telescope is nominally athermal (uniform/balanced CTE’s throughout), but local temperature variation will introduce variation in optical paths and shapes. The time constant of heat transfer between the primary mirror and the rarefied environment is about 33 hours. This value would decrease to
17 hours for a design without MLI enclosing its sides and back, but we prefer the former to minimize gradients. Rather than allow the primary mirror to slowly change shape as it cools, our baseline approach maintains the primary and instrument package at near room temperature. This also minimizes dimensional and alignment changes from the laboratory conditions. The heater power required to maintain the primary’s temperature is less than 30 W, including all losses. The front-to-back thermal gradient of the primary mirror is expected to be about 0.035 K, giving a surface figure distortion of 0.7 nm, within budget.

While heating the primary mirror helps keep its shape stable, we must consider whether this heat source might produce distorted seeing above the mirror. First we note that the air flow around a warm primary will be laminar; the Grashof number is less than $10^8$ and the Reynolds number is less than $5 \times 10^5$. We performed a numerical flow analysis, which showed that the air immediately above the mirror is warmed slightly, and if a wind is present, as it always will be owing to wind shear in the stratosphere, a wedge of warm air will form. Turbulence will be totally lacking, such that the net effect of the wedge will be to deflect the starlight by about 0.02 arcsec. This offset will be taken up by the star-tracking function of the coronagraph. Residual non-flatness of this tilted wavefront will cause some speckle activity in the range 0-3 diffraction widths, so the central star will leak through slightly, and may be detectable at the $10^{-7}$ level, but the leak is expected to be weak.

Support Structure. Unlike the primary mirror, the optical support structure rapidly thermalizes to the surrounding ambient temperature, changing the location of the secondary. This is accommodated by a 3-axis active bipod support and by the LODM tip-tilt, both guided by LOWFS.

Electronics. The electronic equipment is kept above ambient temperature with thermal insulation and heaters maintaining a stable temperature of 290 K (the ground air temperature). Overheating (due to reduced effectiveness of convective cooling at low ambient pressure) is prevented by conductively coupling high-power components to heat sinks. Active temperature control is implemented using bimetallic thermostatic switches and electrical heaters for survival in the lowest-temperature conditions.

5. OPERATIONS

5.1 Mission Profile
Single-night test flights from Ft. Sumner, NM will be scheduled during the turnaround times when stratospheric winds are weakest (late spring and early fall), minimizing the amount of drift during the 24-hour flights. Each overnight flight will have about 10 hr observing time, which will be spread over ~4 targets. In order to survey most of the discovery space unique to this observatory (Figure 2), long-duration flights are required. The optimal path is to cross the Pacific from Antofagasta, Chile to Australia. For this trajectory, strong stratospheric winds toward the west are desired. Mid-December to mid-January is the best time of year, with both strong wind speeds toward the west and essentially zero winds in the north-south direction. During two such flights, each with length 5-10 days, all the objects on our current target list will be observed that are available from that latitude and time of year.

A typical overnight flight begins with a post-midnight pre-launch check-out of the coronagraph system in the hanger, a launch-pad checkout of the command and data-handling electronics, and a full-scale articulation of the complete pointing system. The balloon will be launched around 10 a.m., when the low-level winds at ~300 m altitude have calmed sufficiently. The balloon then ascends through the troposphere at about 5 m/s, slowing somewhat through the tropopause, and reaches an altitude of about 35 km at about noon. While the electronics will be powered continuously from launch, to keep them in their recommended operating range of temperatures, the telescope will be stowed and latched in a vertical direction during ascent. Solar tracking starts at ~25 km to keep the solar cells pointed toward the Sun and the telescope in the shade. Once the Sun is known to be outside of the field-of-view, the science cameras will be exercised. At sunset the telescope is pointed at a reference star to calibrate the system. Data is taken throughout the night. For each target, the star is fine-centered on the occulter mask using the signal from the science camera. Zero-point offsets are sent to the low-order deformable mirror and tip-tilt systems. The light that is not blocked by the occulter, i.e., the light from the debris ring around the star, continues on to a set of dichroics which split the light into 4 wavelength bands. The images from each bandpass are all focused onto the same science camera. Finally, when the flight draws to a close the telescope is stowed and operations shut down. The cut-down procedure is initiated from the ground when a favorable landing site is available.
5.2 Wavefront Correction

The key element for Zodiac II's overall performance will be its ability to reduce the wavefront errors (WFE) in the system to an acceptable level. WFE is introduced from optical polishing errors, optical misalignment errors, thermal gradients, and gravity changes. The errors are measured in two ways - with the Shack-Hartmann LOWFS and by the residual speckles in the science focal plane - and then corrected by a combination of 1) 3-D translation of the secondary mirror, 2) tip-tilt corrections of the LODM tip-tilt mount, and 3) surface adjustments of the two deformable mirrors (LODM and HODM). The acceptable level of residual WFE in the science focal plane after all sensing and correction is set by the required contrast level. Using the relation between contrast and wavefront errors \( \text{WFE} = \frac{N \lambda}{4 \pi^{1/2}} \) with \( N = 48 \) elements across the HODM, and \( \lambda = 550 \) nm for the average V-band wavelength, results in a limit on wavefront errors of 1.2 nm RMS to obtain an image contrast \( C \) of \( 10^{-7} \) or less. This requirement applies to the mid-spatial frequencies that contribute to the dark hole at sky angles from 3 to \( N/2=24 \) times the telescope resolution \( (\lambda/D_{\text{tel}}) \), corresponding to 3 to 24 cycles of ripple across the primary mirror. Translation of the secondary mirror can correct the wavefront over a large range of focus and displacement error, down to its resolution limit, leaving about 500 nm residual error in the wavefront. This residual is easily handled by the LOMD, which can correct about 2000 nm of WFE, down to a residual of about 40 nm RMS. The last stage, the HODM, has a correction range of about 200 nm WFE, and about 0.3 nm RMS residual WFE. This is well below the requirement of 1.2 nm, giving ample control authority over the WFE.

The preliminary budgeting of the WFE and its residual among the areas of optics, integration, environment, and operations partitions the WFE about equally to each area. The details of WFE budgeting drive the details of how Zodiac II is designed, for example whether we use a whiffletree support system for the primary mirror, or what type of thermal shrouding is needed. These allocations and decisions are all part of the early design phase.

Two wavefront sensing and control methods are used to adjust the deformable mirrors in Zodiac II. Each avoids non-common-path errors by using only imagery of a star at the science focal plane. The Gerchberg-Saxton method is used for initial telescope alignment and DM settings, based on analysis of defocused images of a target star with the coronagraph elements temporarily removed from the beam. The Electric Field Conjugation (EFC) method, which probes the wavefront amplitude and phase with four discrete settings of the DM's, is applied iteratively to suppress scattered light within the coronagraph dark field. Typical convergence rates for the EFC method are illustrated in Figure 8 which shows calibration times to reach a range of contrast levels for laboratory data (red) and corresponding simulations (blue). The speed of convergence of WFE control actions is determined by the count rate of electrons in each detector. For example, in the science focal plane, it will take 13 s to achieve a SNR of 5 on a speckle at the \( 10^{-7} \) level, for a nearby (within \( \sim 9^\circ \)) bright (\( V=4 \)) calibration star. This result indicates that the calibration is tractable in a time that is short compared to the observing time. The figure shows a family of convergence curves based on our theoretical SNR for stars of magnitude \( V=0 \).
5.3 Data Analysis

The science and engineering data are continually stored on board, with a subset sent to ground via telemetry. Post-flight, the ground data is processed to remove speckle artifacts.

All four wavelength channels are imaged onto a single detector. A realistic simulation of the resulting 1-hour images of a model debris disk with two concentric dust belts is shown in Figure 9. Enlarged images for each of the four wavelength bands are shown in the corners, while their positions on the 512x512 pixel detector are marked in the central square, to scale. The four dark holes together contain about 34000 pixels (68 kB at 2 bytes per pixel). During the observation of a target, consecutive exposures of approximately one minute each will be taken. The full CCD frame will be stored onboard (~30 MB per hour) for retrieval after payload recovery. On the multi-night flights, only one channel (V1) of the dark hole pixels will be transmitted to the ground (~1 MB per hour), such that the science data transmission rate is quite low and manageable.

All of the control subsystems such as the various pointing mechanisms (telescope pointing, tip-tilt mirror), the derived wavefront measurements, and the DM state, will generate data that can be used a posteriori to retrieve the state of the system at any given time. This information can be exploited to enhance the data reduction process. The engineering data rate is roughly a quarter that of the science rate. All data will be stored onboard and retrieved after recovery of the payload.

Once all data is recovered, advanced image processing is required to remove speckles and thereby reveal faint circumstellar structure. As seen in the simulated images in Figure 9, which include wavefront propagation through the coronagraph and correction cycles, residual scattered starlight forms a field of speckles with a surface brightness that can be comparable to a disk's. Given the relative stability of the wavefront due to the thin atmosphere, especially compared to ground-based telescopes, the Zodiac II speckle field can be largely subtracted out with various calibration techniques. The non-equatorial Zodiac II mount causes field rotation, so the target will be observed over a span of an hour or two with multiple short (~60 s) exposures to minimize sky smearing during each integration. The speckles, which appear static on the detector because they are created by the optics, can be subtracted using an iterative algorithm that solves for the instrumental effects (static features) and the sky (features that rotate with a known angle). This algorithm has been successfully used on HST data to extract images of debris disks (beta Pic;^{54} HD 207129^{25}). The same algorithm was used to extract simulated planets from actual HCIT coronagraphic images.^{46} This technique will be modified for use on symmetrical face-on disks, by introducing data from reference star images.
Zodiac II’s wavelength coverage provides additional information that can be used to differentiate between instrumental speckles and real sky sources. HCIT lab results imply that at the level of contrast probed by Zodiac II (∼10⁻⁷), the wavefront errors will be mostly phase perturbations, with amplitude perturbations playing a minor role. The resulting speckle positions will then be highly correlated with wavelength, allowing us to take advantage of the wide wavelength range covered by the 4 wavebands.

All removal methods rely on the time stability of the speckles, and hence the wavefront, as any changes would introduce residuals in the disk image. A factor of 10 reduction in background should be easily achieved. For reference, HST speckle field reductions of a factor of 25 have been demonstrated using these same techniques. There is much ongoing work in this area. One particular example is the LOCI algorithm of Ref. 55 which achieved a factor of 10 in improvement over conventional 2-roll subtraction on HST/NICMOS observations of HR 8799. Both LOCI and the interactive roll subtraction algorithms have been implemented successfully on real coronagraphic data.

6. SUMMARY

The detection and characterization of habitable planets around nearby stars is a fundamental goal of astrophysics that will drive science and technology development for many years to come. In the near term, preparation work is necessary before developing a full flagship mission. Target reconnaissance surveys must be carried out to help set the requirements for the next decade’s mission designs. Both planet and debris disk observations fall into this category - planets as the eventual targets and disks as a source of noise for those observations. Given their potential limitation on planet observations, it is critical to further explore the dusty disks known to circle nearby stars. Simply inventorying stars with disks is not sufficient; we need to know the distribution of the deleterious dust.

The Zodiac II observatory will resolve ∼30 disks in 4 optical colors. These images will contain enough information to measure the size, shape, brightness, and color of a statistically significant number of debris disks. By supplying direct visible-wavelength images of debris disks around nearby stars, Zodiac II will inform us about the general evolution of planetary disks and the specific locations where individual planets are found today.

In addition to its science goals, Zodiac II will complete the technology development necessary for a future flagship mission. While much of the needed development will take place in labs, full system testing can only take place in a space-like environment. The suborbital flights considered here are the most cost-effective way to provide a space-like testbed to advance Earth-imaging technology, while simultaneous advancing our understanding of exoplanetary systems.

ACKNOWLEDGMENTS

Some of the research described in this publication was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

REFERENCES


