Feasibility of Exoplanet Coronagraphy with the Hubble Space Telescope


aNASA Goddard Space Flight Center, Greenbelt, MD, USA 20771;
bLockheed-Martin Corporation, Denver, CO, USA;
cSpace Telescope Science Institute, Baltimore, MD, USA;
dBall Aerospace and Technology Corp, Boulder, CO USA;
eConceptual Analytics, Greenbelt, MD USA

ABSTRACT

Herein we report on a preliminary study to assess the use of the Hubble Space Telescope (HST) for the direct detection and spectroscopic characterization of exoplanets and debris disks - an application for which HST was not originally designed. Coronagraphic advances may enable the design of a science instrument that could achieve limiting contrasts ~10^9 beyond 275 milli-arcseconds (4 \lambda/D at 800 nm) inner working angle, thereby enabling detection and characterization of several known jovian planets and imaging of debris disks. Advantages of using HST are that it already exists in orbit, it's primary mirror is thermally stable and it is the most characterized space telescope yet flown. However there is drift of the HST telescope, likely due to thermal effects crossing the terminator. The drift, however, is well characterized and consists of a larger deterministic components and a smaller stochastic component. It is the effect of this drift versus the sensing and control bandwidth of the instrument that would likely limit HST coronagraphic performance. Herein we discuss the science case, quantify the limiting factors and assess the feasibility of using HST for exoplanet discovery using a hypothetical new instrument.

Keywords: Hubble Space Telescope, coronagraphy, exoplanets, telescopes

1. INTRODUCTION

Coronagraphy is not new for the HST, the original instrument suite contained the Faint Object Camera (FOC) that had an f/288 coronagraphic mode. However few images were collected with it mode since the problem with HST’s primary mirror conic constant rendered it not useful for coronagraphy. Later instruments such as NICMOS, STIS and ACS/HRC had or have coronagraphs within them – however none of these have enough contrast to achieve 10^{-9} to 10^{-10} suppression of the central star; see Krist et al. for an example of the ACS/HRC coronagraph.

If a future servicing mission were available for the Hubble Space Telescope (HST) then two areas where new HST scientific instruments could have a high impact on science are (1) full survey field of view to explore the dark universe, and, (2) coronagraphy for the direct detection and characterization of exoplanets. Herein we address the latter and perform a comparison study for the performance of a potential HST coronagraph (HST-C) against a 1.5-meter aperture exoplanet probe mission (ExoP) dedicated to the detection and characterization of exoplanets. This study is not intended to be a complete report on scientific merits nor technological feasibility – it is intended to provide a comparison of HST-C against ExoP with regards to a set of assumptions and constraints.

These constraints and assumptions of the study are that: (1) the technology is to be frozen in 2-years (fall of 2012) in order to meet a target HST critical design review date (fall 2012) two years prior to launch of the instrument in late 2014. (2) HST-C and ExoP contrasts are both 0.8 x 10^{-5}. (3) Study is to be done in a manner that is not competition sensitive, and, (4) the comparison is not to advocate one brand of coronagraphy over another – this would likely be relegated to a later study if such an instrument concept for HST were later made available.

*Richard.G.Lyon@nasa.gov; phone 1-301-286-4302

1 Proceedings of SPIE 7731, Astronomical Telescopes and Instrumentation, San Diego CA, June 2010
Since the HST configuration would be for a near-term implementation, the study assessment needs to be based on coronagraphic technologies that have a high technological readiness for space flight, i.e. as opposed to what could be done in the future.

The assessment described here is based upon these three configurations. Exact numbers reported are not intended to be interpreted as absolute performance numbers since there is still room for discussion. The performance parameters are not intended to endorse a particular brand of coronagraphy – the choice of what the optimal coronagraphic approach is, is left for a future study. Nevertheless the results provide a basis for the comparison among these three configurations.

2. COMPARISON STUDY

2.1 Study Framework

Three distinct configurations were defined that span the configuration space and are used to frame this comparison; these are delineated in Table-1 and described in more detail below.

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Aperture (meters)</th>
<th>Obs</th>
<th>Orbit</th>
<th>Wavelength ((\lambda))</th>
<th>Passband ((\Delta\lambda/\lambda))</th>
<th>IWA ((\lambda/D))</th>
<th>IWA ((\lambda/\Omega))</th>
<th>Contrast at IWA</th>
</tr>
</thead>
<tbody>
<tr>
<td>HST</td>
<td>2.4</td>
<td>LEO</td>
<td>LEO</td>
<td>800</td>
<td>10%</td>
<td>4</td>
<td>275</td>
<td>(0.8 \times 10^{-9})</td>
</tr>
<tr>
<td>Current Exoplanet Probe</td>
<td>1.5</td>
<td>0.00</td>
<td>Heliocentric L2</td>
<td>800</td>
<td>10%</td>
<td>4</td>
<td>440</td>
<td>(0.8 \times 10^{-9})</td>
</tr>
<tr>
<td>Goal Exoplanet Probe</td>
<td>1.5</td>
<td>0.00</td>
<td>Heliocentric L2</td>
<td>500</td>
<td>10%</td>
<td>2</td>
<td>138</td>
<td>(1.8 \times 10^{-10})</td>
</tr>
</tbody>
</table>

**Configuration-1:** HST with an aperture diameter of \(D = 2.4\) meters with a lineal obscuration ratio of 0.33 in a low-Earth orbit (LEO) with an axial bay coronagraphic instrument that uses current state of the art coronagraphic instrument technology. The contrast is \(0.8 \times 10^{-9}\) at an inner working angle (IWA) of \(4 \lambda/D\) (275 milli-arcseconds) at wavelength of \(\lambda = 800\) nm and in a 10% \((\Delta\lambda/\lambda)\) spectral bandpass. We recognize that the contrast and the IWA assumptions may be conservative as multiple technology developments are likely to lead to improved implementations.

**Configuration-2:** Current Exoplanet Probe with an aperture diameter of \(D = 1.5\) meters with no obscurations in a heliocentric L2 orbit with an axial coronagraphic instrument that uses current state of the art coronagraphic instrument technology. The contrast is also \(0.8 \times 10^{-9}\) at an inner working angle (IWA) of \(4 \lambda/D\) (440 milli-arcseconds) at wavelength of \(\lambda = 800\) nm and in a 10% \((\Delta\lambda/\lambda)\) spectral bandpass.

**Configuration-3:** Goal Exoplanet Probe with an aperture diameter of \(D = 1.5\) meters with no obscurations in a heliocentric L2 orbit with an axial coronagraphic instrument that uses current coronagraphic instrument technology planned for 3-5 year time frame. The contrast is deeper at \(1.0 \times 10^{-10}\) at a closer inner working angle (IWA) of \(2 \lambda/D\) (138 milli-arcseconds) and at a shorter wavelength of \(\lambda = 500\) nm but with the same spectral bandpass of 10% \((\Delta\lambda/\lambda)\) as two prior configurations.

Configurations-1 and -2 are only different in the aperture diameter, which effects inner working angle, throughput, and point spread function (PSF) sharpness, and in the orbit. Also configuration-1 has an obscured aperture that also lowers the PSF sharpness thereby spreading the planet flux out over a larger region of the focal plane. Configuration-3 is a more aggressive version of configuration-2 in that it has a smaller IWA and operates down to a shorter wavelength, both of which make this configuration more difficult to achieve technologically.

The comparison proceeds by using a defined set of known radial velocity (RV) discovered planets – where the database of planets is of the known RV as of September 2009. A flux ratio is calculated for each planet in the database from the tabulated stellar distance, spectral class, visual magnitude, and maximum angular separation of the planet from the star as the ratio of planetary flux to stellar flux. Photometric flux collected by the apertures of configurations 1, 2 and 3 in units of photoelectrons per second in a transmissive spectral filter with short wavelength cutoff at 500 nm and long wavelength cutoff at 600 nm (V-band filter). The photometric flux can then be applied with the PSF sharpness and detector sampling to calculate the flux distribution in the focal plane. The results are then compared and discussed for the 3 configurations.
2.2 Flux Ratio

The flux ratio \( F_p \) is calculated from:

\[
\text{Flux Ratio} = \frac{F_p}{F_s} = \alpha \left( \frac{R_p}{R_s} \right)^2 \phi(\beta)
\]

where \( \alpha \) is the geometric albedo, \( R_p \) is the radius of the planet, \( R_s \) is the planet's radial distance from the star. \( \phi(\beta) \) is the orbital phase function given by

\[
\phi(\beta) = \frac{\sin \beta + (\pi - \beta) \cos \beta}{\pi}
\]

where \( \beta \in [0,\pi] \) is the phase angle and where \( \cos \beta = -\sin(i) \sin(2\pi \Phi) \) where \( i \in \left[ 0, \frac{\pi}{2} \right] \) is the inclination angle. An inclination angle of 0 implies the star system is seen face-on, and an inclination of \( \pi/2 \) means the star system is seen edge-on. The orbital phase is given by \( \Phi \in [0,1] \). For an inclination angle of 0 then \( \phi(\beta) \approx 0.32 \) independent of phase angle. For an inclination angle of 90 degrees the phase function reaches a maximum of ~0.32 and thus we used a value of 0.32 throughout this analysis. For Jupiter the geometric albedo is 0.52 and we assumed this value throughout our analysis. The phase function is 0.32 for point of longest elongation, this is not necessarily where planet is brightest but where planet is visible. From the planetary database the planets radius was not always known thus we estimated it in two ways and took the average of the two estimates. The mean over the database is 1.15 +/- 0.29 Jupiter radii, and figure-5 of Fortney et al. (2007) gives ~1.3 Jupiter radii and is a slowly varying function with respect to mass and age. Thus we used the estimated value of \( R_p \) = 1.23 Jupiter radii for those planets where the radius was unknown.

Using these values with the other parameters in the database resulted in the list of 21 candidate RV planets as shown in figure-1. The known RV planet is listed to the left followed by its mass in units of Jupiter masses, and period in days, stellar distance in parsecs, maximum angular separation in arcseconds. Columns 6 and 7 are the angular separation converted to units of \( \lambda/D \) for HST and the ExoP at the center of the V-band wavelength (550 nm). Note the angular separation used is not what is shown as the maximum angular separation but what was reported on in the database. The maximum separation is an estimated parameter, while the database parameter is measured. Column 8 is the V-band stellar magnitude and column 9 the spectral class. Column 10 is stellar radius in solar radii. Columns 11 and 12 are the stellar aperture fluxes for HST and ExoP respectively in units of photoelectrons per second collected in V-band and contains the transmission losses of the optics and detector quantum efficiency. Column 13 is the planet to star flux ratio and column 14 is the time to signal-to-noise ratio (SNR) = 10. The flux ratios and SNR incorporate the leakage flux, local zodiacal light, exozodiacal light (assumed 1 Zodi), detector dark current, but no read noise (photon counting detectors are assumed for all configurations).
2.3 Selection Criteria

For the 358 candidates in the database selection criteria were used to select that subset which for HST met the criteria that (1) the angular separation is > 275 milli-arcseconds (mas), (2) the flux ratio is greater than or equal to 0.8 x 10^-9 and that (3) the time on target is less than 1 month. This is shown graphically in figure-2 and results in the selection of 7 candidate RV sources as shown in the green region of figure-2. This same analysis was performed for the two ExoP configurations (configurations 2 and 3 in table-1).

The results are that of the known 358 (as of fall 2009) that HST (2.4 m, 4 A/D IWA, 800 nm) can reach 7 RV planets; the current ExoP (1.5 m, 4 A/D IWA, 800 nm) can reach 2 of the RV planets; and goal ExoP (15 m, 2 A/D IWA, 500 nm) can reach 16 RV planets.

2.4 Flux Ratio vs Angular Separation

Figure-3 plots the flux ratio versus angular separation for the 358 candidate sources. The majority of the RV stars are at small angular separations and high flux ratios. This is believed to be symptomatic of the ability of radial velocity measurements being biased towards close-in faster moving planets (shorter orbital periods). Most of the currently known RV planets are unlikely to reachable by any currently known coronagraphic approach.

In figure-3 HST-C stars are shown in green. The inset on left panel is expanded and becomes the right panel of figure-3. On right panel for figure-3 stars above and to the right of the blue line are reachable by HST-C; stars above and to the right of the magenta line are reachable by current ExoP, and stars above and to the right of reddish line are reachable by goal ExoP.

Based on this analysis HST-C is superior with respect to the number of reachable RV planets unless the goal ExoP can reach an IWA of less than or equal to 2.5 A/D. Note that there may be planets that have not yet been found by RV surveys and if so then the number of planets is likely to increase for the 3 configurations.
2.5 Integration Time Comparison for $10^{-9}$ Planet

If we assume that all planets are $10^{-9}$ of their parent stars brightness and plot the integration time on target versus stellar magnitude for the candidate stars. The integration time on target is approximately the same for the HST-C and current

![Figure-3: Flux Ratio versus Angular Separation for 358 Candidate RV Stars. HST-C stars are shown in green. Inset on left panel is expanded into right panel. On right panel stars above and to the right of the blue line are reachable by HST-C; stars above and to the right of the magenta line are reachable by current ExoP, and stars above and to the right of reddish line are reachable by goal ExoP.](image)

ExoP cases. Although HST has a larger aperture, 2.4 m vs 1.5 m, than ExoP, HSTs is obscured by the secondary mirror housing, and secondary mirror mounting struts and mounting pads on the telescopes primary mirror. This contains diffracting structure that diffracts light from the core of the PSF and spreads the flux in the focal plane resulting in a less sharp PSF and hence a lowering of the sharpness metric thus increasing integration time. The goal ExoP has significantly shorter integration time and thus could potentially detect more planets than either the HST-C or current ExoP cases. With the current technology assumptions the current ExoP 1.5m is comparable to HST-C.

2.6 Observing Efficiency

HST requires 'clock time' greater than the needed exposure time and LEO constraints make most targets visible for < 50% of each orbit. Additionally the South Atlantic Anomaly (SAA) limits exposures during several orbits each day. For observations taking longer than the viewing window available during an orbit the target will be lost and have to be reacquired necessitating wavefront control and this will lower the observing efficiency. This has not been quantified for HST-C or for the two ExoP configurations. However, the two ExoP configurations would be in a heliocentric L2 orbit.
and would be more stable and requiring less wavefront control likely enabling long uninterrupted observations. Optimal scheduling may allow significant use of the continuous viewing zone for HST. It is expected for equivalent mission duration and coronagraphic performance that at dedicated 1.5m ExoP would make more efficient use of time and have higher observing efficiency. The HST-C observing efficiency would need to be quantified and is still an open issue.

2.7 Summary of RV Science Return vs Technology Risk

Pushing the IWA to less than \(4 \lambda/D\) incurs increasing technology risk and hence cost and schedule risk. For example pushing HST-C to 3 \(\lambda/D\) opens up the number of reachable RV planets from 7 to 12, almost a factor of two.

With today’s coronagraphic technology it is likely that 4 \(\lambda/D\) could be reached and brought up to NASA technology readiness 6 required to transition to a true flight mission. This implies that current coronagraphic technology would need to be pushed to TRL-6 and flight qualified in the next two years to meet a launch date of late 2014. It is likely that the combination of the OTA ‘breathing’ and optimizing the coronagraph for aperture obscurations will place a hard limit on what could be achieved with any coronagraph technology on HST and this would need to be further quantified prior to preceding.

3. BREATHING AND POINTING OF HST-C

3.1 HST ‘Breathing’ of Optical Telescope Assembly

Images taken with HST show an orbit dependent drift in the image quality with morphological structure in the image, such as diffraction spikes and speckle moving on a small scale within a temporal sequence of images, a phenomena referred to as breathing. It is believed to be caused by drift of the secondary mirror in the HST telescope. Lallo et al. used phase retrieval techniques to estimate the breathing and found that it consists of both a deterministic and a stochastic components, where the deterministic component is periodic with the HST orbital period (~96 minutes). The focus variation is \(-\lambda/20\), the astigmatism variation is \(-\lambda/50\) and the coma variation is \(-\lambda/100\) rms. This implies that HST will require an approach that compensates for or senses and controls this variation and such an approach is likely to require a deformable mirror with the associated sensing and control algorithms and architecture within the coronagraphic instrument.

To estimate the level of control needed a simulation was conducted using an idealized coronagraph consisting of bandlimited occulting mask with a Lyot stop (figure-6). Zernike polynomials, one at a time were generated and scaled such that the rms of each Zernike term was 0.1 nm rms wavefront error (WFE), 1.0 nm rms WFE, 10.0 nm rms WFE. The fields from the OTA were scalar diffractively propagated through the coronagraph without any other source of errors and tabulated to conduct a sensitivity analysis. The left of figure-6 shows the HST aperture and the column to the right of the aperture shows images of the 5 Zernike polynomial terms used. The 3rd column from the left shows at the top the lineal bandlimited mask used in the focal plane and the middle image shows the occulted PSF without any Zernike
errors, and the bottom of the $3^{rd}$ column shows the Lyot in the reimaged HST pupil (exit pupil of coronagraph). No attempt was made to optimize the Lyot stop it was chosen to overlap the diffracted aperture structure after it had been spread out by the propagation through the occulting mask. The Lyot stop efficiency is low ~24% due to the diffractive structure, more optimal Lyot stop designs may exist that could increase the Lyot efficiency but likely only by small amounts. The 3 columns to the right of figure-6 form a matrix of images at the final image plane for each the amounts of Zernike terms (labeled at top of columns of images), with the Zernike terms in the same order as those in the $2^{nd}$ column of figure-6. The width of these focal plane images are 5 arcseconds on the sky and each image is shown on a log scale but with different grey scale stretches. The images are all V-band images with wavelength sampling of 2 nm (50 wavelengths across 100 nm V-band passband and these images represent the leaked flux through the coronagraph due to the Zernike terms.

The leakage PSF was tabulated at $4 \lambda/D$ and plotted vs the rms WFE of the Zernike in figure-7. It was found that to hold $10^5$ contrast that the individual terms would need to be for coma $> 0.3$ nm rms WFE, for astigmatism $< 3.0$ nm rms WFE, and for focus $< 2.0$ nm rms WFE. However this is for the terms taken one at a time. If the terms are all present and uniformly allocated into an error budget then the terms are reduced by square root of 5, to $< 0.134$ nm rms WFE for
coma, 1.342 nm rms WFE for astigmatism, and 0.894 nm rms for focus or an overall rms WFE of 1.618 nm rms. While challenging, wavefront control to levels exceeding this has been accomplished on coronagraphic testbeds.

3.2 HST Pointing Error
Pointing errors induce leakage through a coronagraph by walking the image off the occulter such that the center of the image and occulter do not line up. A linear bandlimited coronagraph will be insensitive to mispointing errors along the long axis of the occulter but will be sensitive to mispointing in the cross occulter direction. To assess this effect a cross-occulter mispointing errors were induced into the coronagraph by tilting the wavefront and the leakage PSF tabulated at 4 λ/D. The results were plotted in figure-8. The net result is that over the range of HST pointing (shown in red) of 2 – 7 mas the mispointing error does not contribute significantly to the loss of contrast and this is significant since it implies that HST would not require a fine steering mirror to achieve and hold contrasts of 10⁻⁹.

3.3 Follow up Topics
During this preliminary study a number of topics were identified that need follow on study. There are still issues such as thermally driven distortion at the critical mid-spatial frequencies, sensing and control of low order wavefront errors from the HST telescope – these topics are listed in Table-2 with a path forward for resolving them.

Table-2: Follow-on Topics.

<table>
<thead>
<tr>
<th>Topic</th>
<th>Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thermally driven mirror distortion at mid-spatial frequencies</td>
<td>HST: assess with on-orbit data and modeling</td>
</tr>
<tr>
<td>Sensing and control of low spatial frequencies</td>
<td>1.5m: modeling and TBD thermal-vac testing</td>
</tr>
<tr>
<td>Instrument control frustrated by observing constraints</td>
<td>HST: focus/astig/coma control loop modeling</td>
</tr>
<tr>
<td>Can goal 1.5m achieve IWA &lt; 2.5 λ/D with delta-mad &gt; 22.5?</td>
<td>1.5m: thermal and bus jitter integrated modeling</td>
</tr>
<tr>
<td>Are settling times reasonable</td>
<td>Model system drift with wavefront sensing and control in observing timeline</td>
</tr>
<tr>
<td>Comprehensive TRL assessment</td>
<td>Surpassing HST IWA requires technology demo of excellent low-order WF control, including TBD telescope testing</td>
</tr>
<tr>
<td>Optimal instrument architecture</td>
<td>Assess observatory stability and instrument tuning overheads vs timeline constraints</td>
</tr>
</tbody>
</table>

4. SUMMARY AND CONCLUSIONS

4.1 Summary of Results
Three configurations were compared, (1) HST with current technology, (2) 1.5m ExoP with current technology, and (3) 1.5m ExoP with future technology. The prime comparison metric is the number of reachable RV planets with respect to and the results are tabulated in Table-2.

HST is superior to with respect to the number of reachable RV targets unless the goal 1.5 m ExoP can reach an IWA of less than 2.5 λ/D. The dedicated goal ExoP mission is superior in number of RV targets reachable. HSTs coronagraphic throughput is comparable to a dedicated 1.5m mission since it requires a more aggressive Lyot stop than either of the ExoP configurations due to diffracting structure in the aperture; this effectively lowers coronagraphic throughput of
HST's larger aperture to that of 1.5 meter aperture and lowers the PSF sharpness spreading the focal plane planetary flux. Both dedicated ExoP missions, for equivalent mission duration, would make more efficient use of time due to the higher observing efficiency and less constraints on pointing to a given target. HST will require active control of the telescope breathing: all three configurations are likely to require the use of deformable mirror that has yet to be flight qualified. HST would not require a fine steering mirror, it is likely that the 1.5m current ExoP would not require a fine steering either since with current technology body pointing is achievable to a few mas (HST is a prime example of this).

<table>
<thead>
<tr>
<th>Configuration</th>
<th>D (m)</th>
<th>Orbit</th>
<th>( \lambda ) (nm)</th>
<th>IWA (mas)</th>
<th>Contrast at IWA</th>
<th>N RV Planets</th>
<th>Obs Efficiency</th>
<th>FSM</th>
<th>DM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Hubble Space Telescope</td>
<td>2.4</td>
<td>LEO</td>
<td>800</td>
<td>275</td>
<td>( 0.8 \times 10^{-4} )</td>
<td>7</td>
<td>Lower</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>2 Current Exoplanet Probe</td>
<td>1.5</td>
<td>L2</td>
<td>800</td>
<td>440</td>
<td>( 0.8 \times 10^{-9} )</td>
<td>2</td>
<td>High</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>3 Goal Exoplanet Probe</td>
<td>1.5</td>
<td>L2</td>
<td>500</td>
<td>138</td>
<td>( 1.8 \times 10^{-10} )</td>
<td>16</td>
<td>High</td>
<td>Possibly</td>
<td>Y</td>
</tr>
</tbody>
</table>

**Table-3: Summary of 3-configurations results for RV exoplanets.**

Advantages of a HST coronagraph are: it has a larger aperture than any visible planet imager in next 5-years, it has well understood PSF and telescope, demonstrated pointing accuracy and stability, competitive capabilities available with today’s coronagraph technology, telescope and spacecraft exist with known performance, ground support infrastructure exists for observing, scheduling and data processing.

Disadvantages are that the aperture obscurations require a throughput limiting mask and reduces PSF sharpness, breathing variations require sensing and control for baseline performance or contrast will be reduced at small angular separations, and need to retire risk that thermal variations in the primary will induce unacceptable contrast degradation.

If a future servicing mission were available then HST could provide compelling exoplanet science with technology available today and could be observing exoplanets in as little time as four years from now. Developing a coronagraphic instrument for HST, in the near term, would provide technology advancement of coronagraphy and its associated technologies needed for a future flagship mission such as TPF-C.

**REFERENCES**


