

SIM Lite narrow-angle modeling and processing

David W. Murphy^a, Mark H. Milman^a, David L. Meier^a, Mehrdad Moshir^a

^aJet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA, USA 91109-8099

ABSTRACT

This paper examines how narrow-angle (NA) processing of data from the SIM Lite optical interferometry mission can be undertaken when realistic spacecraft and mission operational constraints are taken into account. Using end-to-end mission simulations we show that the goal of 1 μ s single measurement accuracy (SMA) is obtainable, and hence the detection of earth-like planets is achievable with the SIM Lite mission.

Keywords: Space interferometry, optical interferometry, SIM Lite, Narrow-Angle processing

1. INTRODUCTION

An overview SIM Lite mission is described elsewhere in these proceedings. The SIM Lite spacecraft at its most basic level is a rectangular box inside of which there is:

- One science interferometer, with baseline vector \mathbf{h} of length 6 m, on which the ‘science’ star with unit vector \mathbf{s} is observed. This star of course varies from observation to observation. Science stars have to lie within 7.5° of the center of the science field of regard (FOR), whose unit vector we call \mathbf{c} . Furthermore, $\mathbf{h} \cdot \mathbf{c} = 0$ to within attitude control errors.
- One guide interferometer, with baseline length 4.2 m, on which the ‘guide1’ star with unit vector \mathbf{g}_1 is observed. The guide1 star is the same as science star when the science star is a NA target. Also, $\mathbf{g}_1 = \mathbf{c}$ and the guide interferometer baseline is parallel to the science interferometer baseline to within manufacturing errors.
- One guide CCD camera on which the ‘guide2’ star with unit vector \mathbf{g}_2 is observed in two sky dimensions.
- An external metrology truss composed of 6 nodes (corner cubes), which has only 10 laser links among them (compared to a maximum potential number of 15 laser links).
- An internal metrology system to keeps track of the internal path delay between both science interferometer arms.

The primary solar constraint is that the sun must not shine on the top of the spacecraft box, whose normal direction (\mathbf{n}) points 51° away from \mathbf{c} . During the course of the SIM Lite mission a series of many successive \mathbf{c} vectors is used, and for each given \mathbf{c} vector there are observations of multiple objects. All observations undertaken with the same \mathbf{c} vector (at a given epoch) is called a ‘tile’, and typically there are about 50,000 tiles in a 5-year SIM Lite mission. The same (or very similar) \mathbf{c} vector can be, and is, re-used to obtain multiple epoch observations of the same science FOR.

Furthermore, to a good approximation, \mathbf{c} , \mathbf{n} and \mathbf{g}_2 all lie in a plane, and the angle between \mathbf{c} and \mathbf{g}_2 lies between 85° and 87° depending on guide2 star availability. Also, the SIM Lite spacecraft’s orbit plane is, to a good approximation, the ecliptic plane. (For our simulations we assume the SIM Lite spacecraft to have a circular orbit in the ecliptic plane with a 1-year period.)

The primary SIM Lite observable is the science star delay d_s , ideally (if all errors were random noise only) given by:

$$d_s = \mathbf{h} \cdot \mathbf{s} + \eta \quad (1)$$

where \mathbf{h} = baseline vector, \mathbf{s} = source unit position vector and η = noise.

In the very successful NA double blind studies¹ that were undertaken it was assumed that every NA target:

1. Is effectively observed with the same fixed known orthogonal baselines. In fact, for these studies of stellar reflex motion actual interferometer delays were never used; instead, the primary data were angular offsets along two fixed known orthogonal directions in the NA target tangent plane.
2. The noise added to these offsets was Gaussian distributed with zero mean and a fixed standard deviation.

Both of these are simple approximations. In an actual mission, the solar and guide2 constraints produce two observations of a given NA target star that are not simultaneous and do not have orthogonal baselines. Also Gaussian noise is a reasonable approximation only after a considerable amount of data processing, which we will describe. All the simulations described in this paper use our end-to-end SIM mission simulator SIM-sim², which previously has been used to simulate older SIM mission architectures (such a SIM PlanetQuest), but has now been upgraded to support the SIM Lite architecture.

For observations in the k-th tile, using a spacecraft-based XYZ-coordinate science system in which +Z lies along \mathbf{c}_k and +X lies along a nominal baseline direction \mathbf{h}_k both of which remain fixed for a tile, a better approximation for the science star delay is given by:

$$d_{s(x,y,t)} = \mathbf{h}_k \cdot \mathbf{s}(t) + (\mathbf{h}(t) - \mathbf{h}_k) \cdot \mathbf{s}(t) + C_k + \eta_{\text{BDE}} + \eta_{\text{FDE_sci}}(x,y,t) \quad (2)$$

where C_k is the so-called constant term but in fact varies from tile-to-tile, η_{BDE} is the brightness dependent error and the $\eta_{\text{FDE_sci}}(x,y,t)$ are time-variable field-dependent errors (FDEs) on the science delay.

For non-binary stars and stars with no exoplanets the temporal variations of $\mathbf{s}(t)$ can be determined by using the classic 5 astrometric parameters (APs) and the radial velocity (V_r) (assuming that accelerations are negligible). $\mathbf{h}(t)$ varies in inertial space, as the attitude control system (ACS) isn't perfect; typically in a given tile the $\mathbf{h}(t)$ pointing direction varies on the tenth of an arc-second level. If exoplanets are present then $\mathbf{s}(t) = \mathbf{s}_b(t) + \Delta\mathbf{s}(t)$ where $\mathbf{s}_b(t)$ is the barycenter unit vector (which can still be represented by 5 APs and V_r), and $\Delta\mathbf{s}(t)$ is the stellar reflex motion about the barycenter, caused by the presence of exoplanets. For an Earth-like exoplanet orbiting a star like our Sun at 1 AU at a distance of 3 pc the maximum value of $\Delta\mathbf{s}(t)$ is 1 μs . The goal of NA observations is to extract the exoplanet astrometric signature $\mathbf{h}_k \cdot \Delta\mathbf{s}(t)$, which is given by:

$$\mathbf{h}_k \cdot \Delta\mathbf{s}(t) = d_{s(x,y,t)} - \mathbf{h}_k \cdot \mathbf{s}_b(t) - (\mathbf{h}(t) - \mathbf{h}_k) \cdot \mathbf{s}(t) - C_k - \eta_{\text{BDE}} - \eta_{\text{FDE_sci}}(x,y,t) \quad (3)$$

If a NA target is observed only by itself, it is extremely difficult to determine the exoplanet signature $\mathbf{h}_k \cdot \Delta\mathbf{s}(t)$ and hence determine the reflex motion caused by exoplanets. However, by careful mission scheduling, spacecraft design, and data processing all the negative sign terms on the RHS of (3) except for η_{BDE} can be estimated with sufficiently high precision that Equation (1) can, in practice, be realized not for any arbitrary baselines but only for those which meet the solar and guide2 constraints. In fact, a least six different types of objects will be observed on the science baseline, to which will refer with the following single letter codes:

G = 'Grid' stars used for instrument self-calibration (see Section 2.2 for more details).

U = Guide stars (used to determine the orientation of the spacecraft in inertial space).

N = NA target stars.

R = Reference stars (which are used in conjunction with N stars).

W = Wide-Angle (WA) target stars.

Q = Quasars (used to removed parallax and proper motion zonal errors).

2. FOUR-STEP NA PROCESSING METHOD

To achieve the NA 1 μs SMA four steps are needs to estimate the various negative terms in equation (3) so that they can removed from the science delay.

2.1 Delay Regularization

The first term that needs to be removed from the measured science delay $d_s(x,y,t)$ is the impact of the baseline changing with time $\mathbf{h}(t)$ over the time-scale of a single science star observation, which typically lasts 10s of seconds, depending on the object being observed. By using observations of the guide stars (Hipparcos stars brighter than magnitude 7), the external and internal metrology systems, and *a priori* source position $\mathbf{s}_a(t)$ the science delay is ‘regularized’ to $d_{\text{reg}}(x,y,t)$, which is given by:

$$d_{\text{reg}}(x,y,t) = \mathbf{b}_k \cdot \mathbf{s}(t) + \underline{\Delta\mathbf{b}}_k(x,y) \cdot \mathbf{s}(t) + (\underline{\Delta\mathbf{b}}_k(x,y) + \mathbf{b}_k - \mathbf{h}(t)) \cdot (\mathbf{s}_a(t) - \mathbf{s}(t)) + C_k + \eta_{\text{FDE_sci}}(x,y,t) + \eta_{\text{BDE}} \quad (4)$$

This delay regularization process² greatly reduces the rapid baseline variations during the science star observation period, which then allows us to average the regularized science delay over that period. In equation (4) t now refers to the mid-epoch time of each star observation. However, the guide and science star positions are only known to ≈ 100 mas. Hence, the regularization process introduces the delay terms $\underline{\Delta\mathbf{b}}_k(x,y) \cdot \mathbf{s}(t) + (\underline{\Delta\mathbf{b}}_k + \mathbf{b}_k - \mathbf{h}(t)) \cdot (\mathbf{s}_a(t) - \mathbf{s}(t))$ where $\underline{\Delta\mathbf{b}}_k(x,y)$ is a tile-dependent field-dependent baseline error. In fact, since the external metrology system has FDEs, $\underline{\Delta\mathbf{b}}_k(x,y)$ is more than just a simple baseline error and the $\underline{\Delta\mathbf{b}}_k(x,y) \cdot \mathbf{s}(t)$ delay term can be represented as:

$$\underline{\Delta\mathbf{b}}_k(x,y) \cdot \mathbf{s}(t) = \underline{\Delta\mathbf{b}}_k(x,y) \cdot \mathbf{s}(x(t),y(t)) = \underline{\Delta\mathbf{b}}_k \cdot \mathbf{s}(t) + \eta_{\text{FDE_ext_met}}(x,y,t) \quad (5)$$

where $\underline{\Delta\mathbf{b}}_k$ is now a constant field-independent per-tile baseline error and $\eta_{\text{FDE_ext_met}}(x,y,t)$ are the FDEs induced by the external metrology truss system which produces field-dependent laser length measurements.

In a typical SIM Lite total mission simulation, there are 50,000 tiles with 1.3 million regularized delays that need to be further processed to extract astrometric information as well as instrument calibration parameters.

2.2 Instrument Self-Calibration/Grid Solution

In each tile there are observations of typically 6 – 7 or so G stars. Using all the typically 350,000 regularized G delays from the entire SIM Lite mission $\underline{\Delta\mathbf{b}}_k$, C_k , and the FDEs along with the 5 APs for the G stars can be determined³. Thus, the 1302 G stars used form a global grid hence their name but in subsequent WA and NA processing only the instrument self-calibration parameters $\underline{\Delta\mathbf{b}}_k$, C_k , and FDE parameters are used. In fact, the FDEs are fitted by:

$$\eta_{\text{FDE_sci}}(x,y,t) + \eta_{\text{FDE_ext_met}}(x,y,t) = \sum c_n(t) Z_n(x,y) \quad (6)$$

Where Z_n is n-th Zernike polynomial and in a typical grid solution c_4 to c_{36} are solved for every 100 tiles. c_1 , c_2 , and c_3 don’t need to be solved for as the Z_1 , Z_2 , and Z_3 FDEs are already captured by C_k and the y and x components of $\underline{\Delta\mathbf{b}}_k$. The x -component of $\underline{\Delta\mathbf{b}}_k$ is equivalent to a baseline length error and like the other FDEs in equation (6) is solved for every 100 tiles.

2.3 WA processing

Following the grid solution, the self-calibration parameters ($\underline{\Delta\mathbf{b}}_k$, C_k , c_n) are used when processing the regularized W, N, and R star delays so that their 5 APs can be determined to WA accuracy, which typically is a $4 \mu\text{as}$ 1D position error at the reference epoch. After this processing, the 5 APs WA positions as a function of time $\mathbf{s}_{\text{WA}}(t)$ for the science target can be determined and are needed for the NA processing described in the following sub-section.

2.4 NA processing

To achieve $1 \mu\text{as}$ SMA careful scheduling of N stars is required. They are typically observed with 5 reference stars (R_i) surrounding them in a N-tile, which contains the following sequence of observations:

$$\text{N-G}_1\text{-G}_2\text{-G}_2\text{-G}_4\text{-G}_5\text{-G}_6\text{-G}_7\text{-N-R}_1\text{-N-R}_2\text{-N-R}_3\text{-N-R}_4\text{-N-R}_5\text{-N-R}_5\text{-N-R}_4\text{-N-R}_3\text{-N-R}_2\text{-N-R}_1\text{-G}_1\text{-N} \quad (7)$$

By contrast observations of W stars occur in W-tiles, which typically have a much simpler observing sequence:

$$\text{U-G}_1\text{-G}_2\text{-G}_2\text{-G}_4\text{-G}_5\text{-G}_6\text{-G}_7\text{-W-G}_1\text{-U} \quad (8)$$

Some W-tiles may actually contain no W observation at all and are used solely for the purpose of observing the G stars.

The $N(t-\Delta t)$ - $R_i(t)$ - $N(t+\Delta t)$ chopping sequence removes short time-scale variations and allows the use of a chopped residual delay Δd_{chop} as the prime NA observable:

$$\Delta d_{\text{chop}} = 0.5(\Delta d_{\text{sreg}}(x_N, y_N, t - \Delta t) + \Delta d_{\text{sreg}}(x_N, y_N, t + \Delta t)) - \Delta d_{\text{sreg}}(x_R, y_R, t) \quad (9)$$

where for any object the delay residual $\Delta d_{\text{sreg}}(x, y, t)$ is given by:

$$\Delta d_{\text{sreg}}(x, y, t) = d_{\text{sreg}}(x, y, t) - (\mathbf{b}_k + \Delta \mathbf{b}_k) \cdot \mathbf{s}_{\text{WA}}(t) - C_k - \sum c_n(t) Z_n(x, y) \quad (10)$$

However, even after doing all this processing, Δd_{chop} contains a residual field dependence:

$$\Delta d_{\text{chop}}(x_N, y_N, x_R, y_R) = \alpha_k + \beta_k(x_R - x_N) + \gamma_k(y_R - y_N) \quad (11)$$

Using the 10 chops in every N-tile the per-tile coefficients α_k , β_k , and γ_k can be determined and it can be shown that:

$$\alpha_k = \mathbf{b}_k \cdot \Delta \mathbf{s}(t) + |\mathbf{b}_k| \eta_k \quad (12)$$

where η_k is the per-tile NA noise. In the absence of injecting any reflex motion $\Delta \mathbf{s}(t)$:

$$\alpha_k = |\mathbf{b}_k| \eta_k \quad (13)$$

Over many NA targets and mission realizations the statistical properties of η_k can be determined and its $\text{RMS}/\sqrt{2}$ is known as the SMA which is the standard mission metric for determining NA performance.

3. SCHEDULING

In order to allow the 4 processing steps described above to be undertaken in end-to-end mission simulations a schedule must be generated which has the appropriate number of N-tiles and W-tiles which all meet the time-variable solar constraint. Furthermore it is desirable that N-tiles for a given N star be observed in closely-spaced pairs (a few days apart) with baselines \mathbf{b}_1 and \mathbf{b}_2 that are quasi-orthogonal. These baseline pairs can't be made perfectly orthogonal due to the finite number of guide2 stars available for use. As a straw-man approach, we decided to use a total of 1304 W-tiles and 65 N-tiles and to divide the 5-year 50,000 tile schedule into W-loops which contain only W-tiles and N-loops which only contain N-tiles. At any epoch in order to meet the solar constraint each W-loop consists of about 1000 W-tiles and

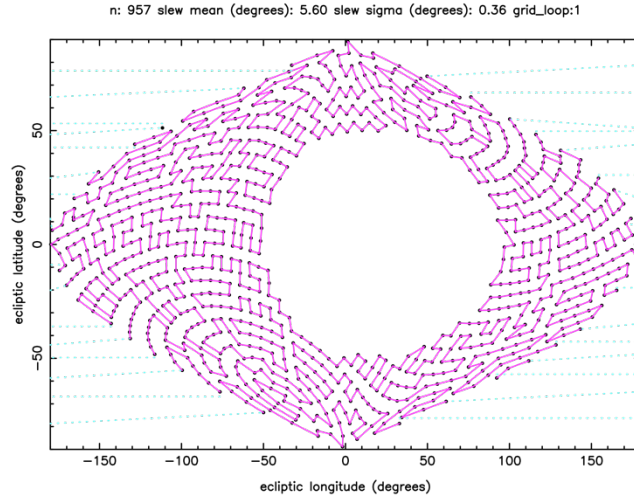


Figure 1. Example of a W-loop in the LKH/TSP schedule plotted in ecliptic co-ordinates with an equal area sinusoidal projection. Each point plotted is a tile center (\mathbf{c}_k) which is joined by straight lines to the previous and next tile center in the schedule. The solar constraint excludes W-tiles in the central quasi-circular region from being scheduled at the epoch of this W-loop.

each N-loops consists of about 45 N-tiles. The schedule is divided into 25 sections: each section starts with a W-loop followed by 20 N-loops. The baseline position angles (BPAs) in adjacent N-loops are chosen to provide pairs of quasi-orthogonal baselines \mathbf{b}_1 and \mathbf{b}_2 for each NA target separated by only a few days. For a given loop of M tiles all of which meet the solar constraints, the LKH algorithm⁴ and software are used to solve the Traveling Salesman Problem (TSP) and determine the order in which to observe the M tiles so that the total tile center to tile center slew angle is minimized. For the k-th tile with tile center \mathbf{c}_k the initial choice of baseline vector lies along direction $\mathbf{c}_k \times \mathbf{sun}$ rotated by $\pm 45^\circ$ about \mathbf{c}_k with the 45° sign changing in alternate W- and N-loops and where \mathbf{sun} is the time-variable solar unit vector. After this initial choice, there is an additional minimum spacecraft rotation about \mathbf{c}_k to select the closest guide2 star which then produces the baseline vector \mathbf{b}_k used for the schedule. For each N-tile or W-tile observations of the form shown in equations (7) and (8) are added to produce the final schedule used to simulate the SIM Lite mission. Figure 1 shows the tile centers from a W-loop in what we call the LKH/TSP SIM Lite schedule.

4. SMA STATISTICS

Figure 2 shows an example of the SMA statistic when error sources from different sub-systems of the SIM Lite spacecraft are used individually and when they combined together for the number in the top box. These SMA statistics were obtained using 2.7 million individual η_k values of the type shown in equation (13) that have been derived after

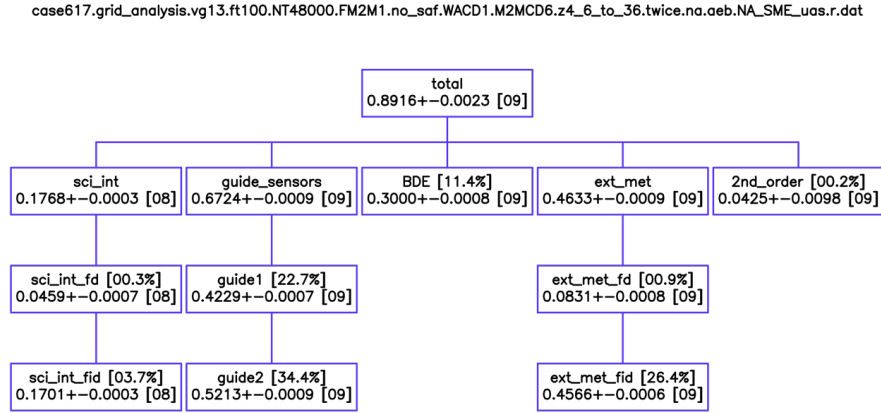


Figure 2. SMA statistic when the LKH/TSP SIM Lite schedule is used. The number in each box is the mean SMA \pm the standard deviation of the mean SMA for a particular combination of SIM Lite instrument parameters. The top box is the result obtained with the full set of SIM Lite instrument parameters.

going thru the 4 processing steps described in Section 2. These individual η_k values were obtained using 9 mission realizations for 65 NA targets with on average 190 quasi-orthogonal pairs of observations for each NA target. As can be seen, when all the instrument sub-systems are used the mean SMA is 0.89 μs . This number has a lot of caveats attached to it as the SIM Lite instrument is still being developed but it certainly indicates that using the most realistic instrument model parameters, a realizable schedule that meets the solar constraints, and the 4 steps of data processing described above, there is currently nothing to prevent a SMA of 1 μs being achieved. For individual NA targets and mission realizations, we used the Jarque-Bera (JB) statistic, which combines Skewness and Kurtosis to determine if the η_k distribution is Gaussian or not. Using 65 NA targets, 150 mission realizations and all the SIM Lite instrument errors we found that the η_k distribution is Gaussian in about 98% of 9750 datasets we studied, which shows that in the vast majority of cases the η_k Gaussian distributed assumption is correct. Further work is needed to investigate the 2% of cases where the Gaussian assumption is not correct to see if things can be improved by better scheduling or data processing.

5. ADDING AND EXTRACTING REFLEX MOTION

While simulations that have no reflex motion ($\Delta \underline{s}(t)=0$) enabled us to determine the NA SMA statistic using a realistic data processing scheme, there always is the concern that our 4-step data reduction method somehow suppresses true reflex motion, as much parameter fitting has taken place before equation (12). There is the further concern that, since the delay data no longer are taken on fixed orthogonal baselines, it will be harder to extract the reflex motion and quantities like the exoplanet orbital period and semi-major axis. Equation (12) can be recast as:

$$\alpha_k = \underline{b}_k \cdot \Delta \underline{s}(t) + b \eta_k = b \Delta \theta_k \quad (14)$$

where $b = |\underline{b}_k|$ is the baseline length, which is the same for all tiles. $\Delta \theta_k$ is the best estimate of the reflex offset in the presence of noise along the baseline direction, which is completely defined by the baseline position angle (BPA_k) for the k -th tile which occurs at time t_k . Thus, for a given NA target, there is a series of t_k , $\Delta \theta_k$ and BPA_k values that need to be processed to determine exoplanet orbital properties. Since the LKH/TSP schedule produces quasi-orthogonal pairs of $\Delta \theta_k$ reflex offset data, it is possible to convert these pairs to fixed orthogonal pairs of $\Delta \theta$ values and process this fixed orthogonal data in exactly the same way that was done for the blind tests. Doing this is inefficient, however, as it ignores any temporal changes that may take place on the timescale of the difference in time between the pairs of observations (typically a few) days, and it leads to increased noise as well. Thus, a new data processing method (the complex periodogram method) was developed by us, which treats the reflex motion as the sum of Fourier components with periods T having amplitudes $a(T)$. Both position and negative periods are allowed corresponding to anti-clockwise and clockwise reflex motion as seen by the observer. The amplitude of each Fourier component can be estimated as follows:

$$a(T) = 2\sqrt{\langle C_k \Delta \theta_k \rangle^2 + \langle S_k \Delta \theta_k \rangle^2} \quad (15)$$

where $\langle \rangle$ indicates a mean value and C_k and S_k are given by

$$C_k = \cos(2\pi t_k/T - BPA_k) \text{ and } S_k = \sin(2\pi t_k/T - BPA_k) \quad (16)$$

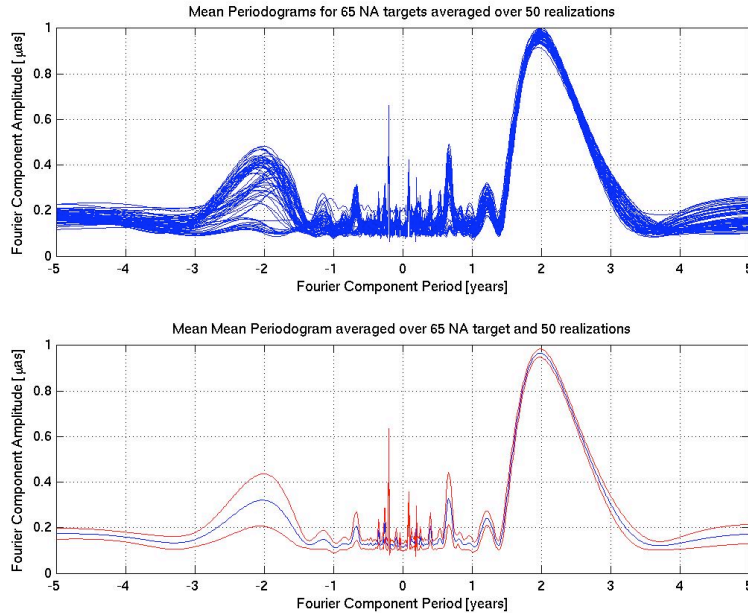
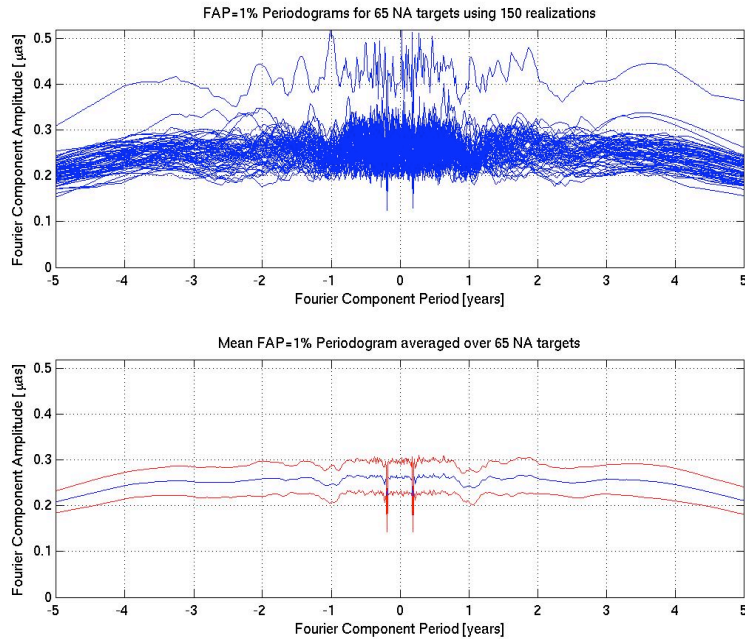


Figure 3. Top: mean complex periodogram amplitude for each of 65 NA targets averaged over 50 mission realizations. Bottom: middle curve is mean of the mean complex periodogram amplitudes. Upper and lower curves shows the \pm standard deviation of the mean amplitude over the 65 NA targets and illustrate the variation from NA target to NA target. Injected reflex motion is for a face-on circular orbit with semi-major axis = 1 μas and period = 2 years.

The blind tests used pairs of NA offset values $\Delta\theta_u, \Delta\theta_v$ taken on fixed orthogonal baselines and determined the reflex motion by examining the temporal properties of the combined power spectral density $\text{PSD} = \text{PSD}(\Delta\theta_u) + \text{PSD}(\Delta\theta_v)$. Our method for fixed orthogonal baseline is equivalent to examining the PSD of the complex signal $\Delta\theta_u + i\Delta\theta_v$ but can easily be extended to reflex offset data taken on arbitrary BPAs, in which each BPA_k can be thought of as a known phase shift whose effect is removed by the $-\text{BPA}_k$ terms in equation (16). This complex PSD method is exactly what is needed for the reflex offset data produced by SIM Lite, which has to meet the spacecraft solar and guide2 constraints. Figure 3 shows an example of the complex periodogram mean amplitudes when using equation (12) for full mission simulations with 65 NA targets and 50 mission realizations when the reflex motion is a face-on circular orbit with semi-major axis = $1 \mu\text{s}$ and period = 2 years. As can be seen from the Figure 3, there is a strong peak at a period close to 2 years with an amplitude close to $1 \mu\text{s}$. Having found this peak, more detailed work similar to what was done on the blind tests is needed to more precisely fit the data. By contrast Figure 4 shows the complex periodogram mean amplitudes when there is no reflex motion ($\Delta\mathbf{s}(t) = 0$), and as can be seen, the typical amplitude $a(T)$ produced in such a case is only $0.3 \mu\text{s}$.



11/16/09:dwm:dwm_test_fit_circular_orbit_stats.c727.aeb_default_1.0.ap_na2.n150.periodogram.crit_value

Figure 4. Top: mean complex periodogram amplitude for each of 65 NA targets averaged over 50 mission realizations. Bottom: middle curve is mean of the mean complex periodogram amplitudes. Upper and lower curves shows the \pm standard deviation of the mean amplitude over the 65 NA targets and illustrate the variation from NA target to NA target. No reflex motion has been added.

6. LINEAR MODEL JUSTIFICATION

Because the search for Earth analogues is about nearby stars, these objects may have large proper motions and parallaxes, and the basic linearized model we use needs further justification. The first order approximation of the deviation of the position of a star from its nominal is of the form:

$$\underline{\mathbf{s}} = \underline{\mathbf{s}}_0 + \delta\underline{\mathbf{s}}_T \quad (17)$$

where $\underline{\mathbf{s}}$ is the “true” position, $\underline{\mathbf{s}}_0$ is the assumed position, and $\delta\underline{\mathbf{s}}_T$ is the first order correction in the tangent plane to $\underline{\mathbf{s}}_0$.

Although this is a linearization, the *true* true general relation between \underline{s} and \underline{s}_0 involves a correction in the both the tangential and radial directions, i.e.

$$\underline{s} = \underline{s}_0 + \delta\underline{s}_T + \delta\underline{s}_R \quad (18)$$

where $\delta\underline{s}_R = \gamma\underline{s}_0$ for some γ . Using $|\underline{s}|=1$ together with $\delta\underline{s}_T \cdot \delta\underline{s}_R = 0$ enables us to solve for γ , obtaining

$$\gamma = -1 + \sqrt{(1-|\delta\underline{s}_T|^2)} \approx -|\delta\underline{s}_T|^2/2 - |\delta\underline{s}_T|^2/8 \quad (19)$$

Thus knowing the tangential component $\delta\underline{s}_T$ of the correction is sufficient for determining \underline{s} , and moreover this result is not restricted to small corrections. Note also that the angle between \underline{s} and \underline{s}_0 is totally determined by $\delta\underline{s}_T$ (independently of magnitude), since as easily shown:

$$|\underline{s}_0 \times \underline{s}| = |\delta\underline{s}_T| \quad (20)$$

so that $|\delta\underline{s}_T|$ is recognized as the modulus of the sine of the angle between \underline{s} and \underline{s}_0 . This discussion is very relevant to narrow angle astrometry applications because the observable created by the instrument is essentially the projection of the the tangent vector correction onto the instrument baseline vector. This can be seen by writing the basic delay equation as

$$\underline{b} \cdot \underline{s} = \underline{b} \cdot \underline{s}_0 + \underline{b} \cdot \delta\underline{s}_T + \underline{b} \cdot \delta\underline{s}_R \quad (21)$$

Since $|\delta\underline{s}_R| \approx |\delta\underline{s}_T|^2/2$, while $|\underline{b} \cdot \underline{s}_0| \approx 5e-5$ m (using the \approx arcsec pointing requirement of the instrument to the center of the field corresponding to the target star position), the radial term above generates a delay error on the order of less than 1e-15 m. Thus there is no loss in generality in disregarding this term, and the observable is indeed the projection of the tangent vector correction onto the baseline vector. In the course of identifying periodic motion of the parent star due to the presence of planets, the proper motion of the star is generally removed as a linear term, which amounts to detrending the data before the DFT is performed to determine harmonics in the star motion. For nearby objects with large proper motion the linear model (in time) of motion in the tangent space is in error. We will show how this is introduced into the data.

So let \underline{v} denote the proper motion vector (arcsec/yr), and assume that

$$\underline{s}_0 = \underline{r}/|\underline{r}| \quad (22)$$

where \underline{r} is the Cartesian position of the star relative to the barycenter of the solar system (normalized to unit length). Consider the motion of the target on the unit sphere given by

$$\underline{s}(t) = (\underline{r} + t\underline{v})/|\underline{r} + t\underline{v}| \quad (23)$$

and introduce the projection $\mathbf{\Pi}$ onto the tangent space to \underline{s}_0 :

$$\mathbf{\Pi} = \mathbf{I} - \underline{s}_0 \underline{s}_0^T \quad (24)$$

We are interested in the evolution of $\mathbf{\Pi} \underline{s}(t)$, since

$$\underline{s}(t) = \underline{s}_0 + \delta\underline{s}_T(t) + \delta\underline{s}_R(t) \quad (25)$$

implies

$$\delta\underline{s}_T(t) = \mathbf{\Pi} \underline{s}(t) \quad (26)$$

Since $\mathbf{\Pi} \underline{r} = 0$ we have

$$\Pi \underline{s}(t) = (t/|\underline{r} + t\underline{v}|)\Pi \underline{v} \quad (27)$$

By expanding the term $1/|\underline{r} + t\underline{v}|$, a brief bit of analysis leads to:

$$\Pi \underline{s}(t) = (t/|\underline{r}|) \Pi \underline{v} - (t/|\underline{r}|)^2 (\underline{v} \cdot \underline{s}_0) \Pi \underline{v} - \frac{1}{2}(t/|\underline{r}|)^3 |\Pi \underline{v}|^2 \Pi \underline{v} \quad (28)$$

As an example of representative magnitudes for the nonlinear terms above, take the tangential and radial velocity magnitudes to be:

$$|\underline{v}_T| = |\underline{v}_R| \approx 5e-6 \text{ rad/yr} \quad (29)$$

Then after a best linear fit the quadratic peak to peak residual is on the order of 30 μs . Note that a radial velocity component is necessary for this term to appear ($\underline{v} \cdot \underline{s}_0 \neq 0$). And the cubic peak to peak residual is on the order of $2e-4 \mu\text{s}$. Although the quadratic term is a concern for the target star it is probably safe to ignore for the reference and grid stars which presumably have much smaller proper motions as the error will decrease quadratically with the distance of the target for the same inertial velocity. The worst case scenario would require removing a quadratic term as well as the linear term to capture the proper motion in the orbital fit strategy. However, a priori estimates of the proper motion may be adequate for removing the quadratic without using parameters to fit the term. A similar analysis shows how parallax for nearby targets can be treated in much the same way, and ultimately does not affect the ability to detect Earth analogues.

7. CONCLUSIONS

This paper extends the work of the blind tests and shows that with the current SIM Lite design there is no inherent limitation on achieving a SMA of 1 μs . The work presented in this paper used full end-to-end SIM Lite mission simulations that included both solar and guide2 scheduling constraints as well a complete SIM Lite instrument model. Even though four stages of data processing are required to determine the NA reflex motion, the noise on this measurement (η_k) is Gaussian distributed for the vast majority of cases that we studied. Furthermore, the inability to observe NA targets on fixed orthogonal baselines simultaneously prevents no fundamental obstacle to reflex motion parameter estimation, provided that the complex periodogram method described in Section 5 is used. Finally, Section 6 provides a justification for the linear model we use.

ACKNOWLEDGMENTS

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

REFERENCES

- [1] Traub, W. *et al.*, "Extrasolar Planets in Multi-Body Systems: Theory and Observations," EAS Publications Series 42, 191 (2010).
- [2] Meier, D. L. and Folkner, W. M., "Simsim: An end-to-end simulator of the space interferometer mission," *Proc. SPIE* 4852, 131-142 (2003).
- [3] Milman, M. and Makarov, V., "Accuracy and covariance analysis of global astrometry with the Space Interferometry Mission," Publications of the Astronomical Society of the Pacific, 117(833), 757-771 (2005).
- [4] Helsgaun, K., "An Effective Implementation of the Lin-Kernighan Traveling Salesman Heuristic," *European Journal of Operational Research* 126(1), 106-130 (2000).
- [5] Milman, M. and Murphy, D., "High-precision narrow angle astrometry with a space-borne interferometer," *Proc. SPIE* 7013, 541-549 (2008).