W. W. Hansen Experimental Physics Laboratory

STANFORD UNIVERSITY

STANFORD, CALIFORNIA 94305 - 4085

Gravity Probe B Relativity Mission

S0940

A MONTE CARLO STUDY OF GP-B ERRORS DUE TO A THIRD BODY IN THE IM PEG (HR 8703) SYSTEM

October 14, 2003

Prepared by: John Chandler 10/24/2003 John Chandler Date Harvard-Smithsonian Center for Astrophysics, Scientist	Prepared by: Michael Ratner 10/24/2003 Michael Ratner Date Harvard-Smithsonian Center for Astrophysics, Guide Star Lead Scientist
Approved by: Richard Whelan 10/28/03 Richard Whelan Date Gravity Probe B Systems Engineering	Approved by: G. M. Keiser 10/29/03 George "Mac" Keiser Date Stanford University Gravity Probe B Chief Scientist
Approved by: Irwin Shapiro 10/27/03 Irwin I. Shapiro Date Harvard-Smithsonian Center for Astrophysics, Guide Star Principal Investigator	Approved by: D. K. Ross 10/29/03 Dorrene Ross FOR Date Stanford University Gravity Probe B Quality Assurance

ITAR Assessment Performed FBE for Tom Langenstein ITAR Control Req'd? ☐ Yes ☒ No

A Monte Carlo Study of GP-B Errors due to a Third Body in the IM Peg (HR 8703) System

J.F. Chandler and M.I. Ratner

Abstract

We have performed a Monte Carlo study of the effect on GP-B guiding errors of an unknown third body in the IM Peg (HR 8703) system. Such a companion could change the relation between the radio and optical positions by offsetting the effective guide point from the position predicted based upon VLBI astrometry. The hypothetical companions in the study are chosen at random from a statistical distribution in mass, luminosity, rotation velocity, and orbital parameters representative of such a system. We assume that if a third body were to be detected in the system, the GP-B errors due to that body would be determined and removed, with an uncertainty that could be estimated on a case by case basis. Thus, for each hypothetical companion, we compute the probability that it would be detected by at least one of the four available searches: VLBI astrometry, HST and ground-based imaging, and spectroscopy. We then determine the probability distribution of the companion-related GP-B guiding errors under several different assumptions for the *a priori* likelihood that a third body actually exists. Under the more plausible assumptions, we find that even the "3-sigma" (99.7%) confidence level for these guiding rate errors is well below the required 0.15 mas/yr "1-sigma" accuracy of the proper motion of the mission's effective guide point. We successfully test the robustness of this conclusion by recalculating our results under various plausible alternative parameterizations of the statistics of the companions and of the sensitivities of the companion searches. We also find that even the 3-sigma confidence level for errors in the apparent annual aberration of IM Peg is well below the 0.2 mas 1-sigma level required to calibrate the GP-B gyro readout scale factor to one part in 10^5.

I. Introduction

In regard to guide star issues, the report of Meeting 8 of the GP-B SAC states:

"We recommend that additional analyses be undertaken to estimate the probability that a companion body orbiting the IM Peg system could be present within the field of view of IM Peg, with sufficient brightness that could change the relation between its radio and optical positions. Such an analysis either could help to establish better confidence limits on telescope errors, or could suggest and motivate other observations to rule out or reduce this effect."

This memo describes the numerical study performed to implement this recommendation. The study focuses specifically on the probability distribution of the errors in proper motion and apparent annual aberration that could be caused by an undetected third body. For possible future use, we have also collected statistics for the proper acceleration error, but we do not discuss these statistics here because they do not bear directly on the GP-B mission uncertainties. We compute the probability distribution of proper motion error at the midpoint of the mission (nominally assumed to be 2004 June 1, consistent with a launch date about 2003 November 1). The results are relatively insensitive to the GP-B launch date; deferring the mission by one year makes no

more than a few percent change in the value of any calculated confidence level. We take into account both the motion of hypothetical undetected companions and a secular brightness variation in IM Peg (HR 8703) characterized by a Gaussian distribution of the rate of variation. We do not account for possible bias in the estimated proper motion of the effective guide point due to the gravitational effects of an undetected companion. (The gravitational effect of any companion with a period at least comparable to the duration of the VLBI astrometry project is already included in the astrometric VLBI analysis, and that of a much shorter-period companion would likely result in systematic VLBI residuals, as yet unseen.) Throughout this memorandum, the term "companion" refers to a third body in the system, not to the unseen secondary component known to orbit the primary with a period of about 24.6 days.

For this study we assume that each observing technique has failed, or will fail, to detect any companion to IM Peg. No observations of IM Peg have detected any to date, and, with the possible exception of the VLBI observations, we regard it as quite unlikely that new observations or refined analysis will yield any detection in the foreseeable future. If a companion is detected, its contribution to the proper-motion uncertainty, etc., will need to be estimated from its particular observed characteristics.

We use a Monte Carlo approach to compute the error distributions resulting from a plausible statistical distribution of possible companions to IM Peg. We construct the latter distribution based upon published statistical distributions of the orbital elements of stellar orbits and the distributions of masses, luminosities, and colors of stars in the solar neighborhood. As necessary, we make plausible assumptions to adapt the published results to the specific case of IM Peg. In the Monte Carlo computation, we sample this distribution many thousands of times. For each Monte Carlo "throw," *i.e.*, for each hypothetical companion, we compute the probability of detection by the most relevant observations of IM Peg already completed or currently under way, based upon a model of the noise statistics and detection limits of each type of observation. These observations include spectroscopy, ground-based images, HST images, and VLBI astrometry. We then compute the probability that the hypothetical companion could escape detection by all of these techniques. Since we are concerned with the distribution of errors due to an undetected companion, this last probability provides a weighting factor for the likelihood of the errors contributed by such a companion, or to be more precise, by the corresponding point in the multi-dimensional distribution of possible companions. However, we find it more convenient to deal with the complement of this probability, *i.e.*, the probability of detection. The mean value, d, of the detection probability averaged over all the throws is also computed for the purpose of normalizing the statistical results of the Monte Carlo calculation of error probabilities.

In addition, we must estimate the probability that IM Peg is orbited by some third body. In effect, we require the *a posteriori* probability, in light of the lack of any detection, of the existence of a third body. This *a posteriori* probability can be computed from the multiplicity fraction, t, which is what we call the *a priori* probability that a system like IM Peg includes a third body. As we show in Appendix A, t and d together scale the probability, from the Monte Carlo calculation, of any nonzero error due to such a companion by the factor $t / (1 - td)$. It is important to note that this scale factor depends on both t and d and is especially sensitive to both when both are near unity. This dependence reflects the fact that the physical implication of the lack of any detection of a third body depends on the value of t: If t is small compared to d, a non-detection strongly suggests that IM Peg is truly a binary and not a triple system; but if $t \approx 1$, the same non-detection

suggests, instead, that the expected companion has characteristics that reduce the likelihood of its detection. Thus, the value of *t* has a large effect upon the estimated probability of any given error due to a companion.

Unfortunately, the value of *t* is also hard to estimate. We know of no estimate for *t* determined for binaries similar to the IM Peg system. The most relevant paper in the recent literature is that of Tokovinin (1997). He considers the multiplicity of spectroscopic binary stars grouped into three bins of period; the period of IM Peg falls within the middle bin, which spans 10 to 100 days. He notes that physically associated visual companions have been found for only 4 of the 52 systematically studied stars in the Batten *et al*. (1989) catalog of spectroscopic binaries with periods in this range and estimated distances of less than 100 pc. However, if the sample is restricted to those 7 binaries within 22.5 pc of the Sun that contain a G dwarf, as cataloged by Duquennoy and Mayor (1991), a much larger fraction, 3 out of the 7, are known to have a third body. (G dwarfs are stars similar in mass to the Sun, whereas the primary star of the IM Peg system could be twice as massive.) Tokovinin's statistical analysis of all these binary stars, independent of period, shows that the difference in the results between the Batten and the more-intensively studied Duquennoy and Mayor samples is highly significant; this difference suggests that the detection of companions in the former is significantly incomplete. Thus, we might be tempted to conclude that the best estimate of *t* lies somewhere between ~0.1 and ~0.5. The higher value would also be 25% larger than that which Tokovinin predicts under the assumption that the probability of finding a long-period orbit with any given period is the same for a short-period binary as for a single star, as long as the longer period is at least 100 times the shorter period. (However, he goes on to argue that this assumption is false, in the sense that, in at least one sample, the existence of a long-period orbit triples the probability that a short-period orbit exists.) The comparable multiplicity statistic for Tokovinin's own sample of nearby K and M dwarfs, which are stars less massive than the Sun, is none out of 5. On the other hand, with continued improvement in observational techniques, the number of companions can only increase, and there is little basis upon which to argue strongly for any upper bound less than unity. In light of this uncertainty, we compute error probabilities for four different values of *t*: 0.1, 0.5, 0.8, and 1.0. Then, for each value of *t*, we compute the size of the error corresponding to each of four confidence levels, namely, the 1-sigma, 2-sigma, 3-sigma, and 4-sigma confidence levels of a normal distribution.

II. Monte Carlo Selection of Hypothetical Companions

Using a pseudo-random number generator, we assign values to eleven variables, each according to a prescribed distribution. These are: companion absolute V magnitude, mass, orbital period, orbital eccentricity, inclination on the plane of the sky, ascending node, argument of periapse, initial mean anomaly, projected equatorial rotation velocity, rate of brightness change for IM Peg, and mass of the IM Peg primary. The values of the variables are all assigned independently, except as described below. In addition to making a nominal run of the calculation, we have also made numerous additional runs to test the sensitivity of our results to plausible changes in the values of many of the fixed parameters that are used in the calculation. To the maximum possible extent we take advantage of the deterministic nature of the pseudo-random number generator to construct the same set of Monte Carlo companions for use in each of the runs. The virtue of this procedure is that it reduces the amount of sampling noise in the difference between the results

obtained in different runs. For any given Monte Carlo sample size, this lower noise level allows us to estimate more accurately the sensitivity of our results to the various parameter choices. The sensitivity tests provide some assurance that plausible changes in the values we adopt for various fixed parameters would not lead to unacceptable errors in our final results.

1. Absolute Magnitude

We choose the masses and magnitudes of the companions based upon the local initial mass function (Reid *et al*. 2002, Reid 2003). We avoid choosing masses in proportion to the actual observed present-day mass function because the observed sample is not drawn from the same probability distribution as are the hypothetical companions of IM Peg, as shown by the following argument based on well-established aspects of stellar evolution. The present-day mass function is a result of the on-going star formation in the Galactic disk, combined with the finite lifetimes of ordinary (main-sequence) stars. These lifetimes decrease from tens of billions to merely millions of years with increasing mass. Consequently, the present-day mass function is missing many stars of at least solar mass that have already evolved into bright giants, dim white dwarfs, or (rather rarely) other compact objects. In contrast, it is reasonable to assume that any third body in the IM Peg system will have retained approximately its original mass and brightness if that mass was significantly less than that of the IM Peg primary, since the latter is currently in the relatively short-lived red-giant stage that follows the main-sequence stage of stellar evolution. On the other hand, if the mass of the companion was originally greater than that of the primary, the probability is very high that the companion is now a white dwarf. Thus, we choose the mass and luminosity of the companion by taking account of these two qualitatively different possible states of the companion, which in turn requires that we first select (at random) a value for the primary mass. We make that selection from a "composite Gaussian" distribution, in which values below the median of 1.5 M_{Sun} are distributed according to a Gaussian with a standard deviation of 0.2 M_{Sun}, and values above the median are distributed according to a Gaussian with a standard deviation of 0.4 M_{Sun}. We adopt this distribution somewhat arbitrarily, but in light of the 1.5 ± 0.2 M_{Sun} mass estimate of Berdyugina *et al*. (1999) and its possible bias. (While their estimate and uncertainty on the low-mass side appear to us to be justified, we enlarge the uncertainty on the high-mass side because we are unconvinced of the validity of a subsidiary result required for their mass estimate, namely, that the IM Peg secondary must be fainter than a G8 main-sequence star. It appears to us that, for example, the existing data would allow the secondary to be a rapidly rotating main-sequence F star, much brighter than a G8 star, and as a result allow the primary mass to be at least as high as 2.3 M_{Sun}. On the other hand, Strassmeier *et al*., 1993, listed all the known masses of stars in chromospherically active binaries, allowing us to infer that the median mass for the cool components lies closer to 1.0 M_{Sun} than 1.5 M_{Sun}. Our higher standard deviation of 0.4 M_{Sun} is chosen to be roughly consistent with the frequency of cool-component masses exceeding 2.3 M_{Sun} in this catalog: only 5 of 72.)

For convenience in selecting masses and magnitudes from various hypothetical alternative distributions, we use a procedure of interpolation from an input table of stars representing the desired distribution. In the case of the initial mass function, which Reid approximates by a pair of power laws, the input table is constructed of stars evenly spaced in cumulative probability according to the adopted mass function, with absolute magnitudes assigned according to an empirical mass-luminosity relation (Henry and McCarthy 1993). The corresponding table for the present-day mass function is simply a complete roster of observed stars within a suitable volume

of space (see below), with observed absolute magnitudes and with masses assigned according to the same mass-luminosity relation.

We choose Monte Carlo absolute magnitudes by linearly mapping the 0-to-1 range of a uniformly distributed random variable onto the range 0 to *S*+1, where *S* is the number of stars in the input list. The mapped random variable thus indexes into the list of stars (sorted by absolute magnitude), such that a value of 1.0 corresponds to the first star, and so on. For values from 1.0 to *S*, we interpolate linearly between the absolute magnitudes of the two bracketing stars. (For a table of observed stars selected from the present-day population, this procedure correctly computes the probability that an *S*+1st star chosen from the same population would be either brighter or dimmer than all *S* of the stars on the list.) At either end of the range, we extrapolate the absolute magnitude as follows. On the dim end, we simply use the dimmest star's absolute magnitude. On the bright end, we use statistics appropriate to the type of input table: for the theoretical table based on the initial mass function, the extrapolation is simply another table with finer resolution constructed from the same mass function; for the sets of observed stars representing the present-day mass function, we use the observed luminosity function for bright stars as tabulated by Gilmore and Zeilik (2000). In both types of tables, the dimmest "star" is at least 20 magnitudes dimmer than IM Peg, while the brightest star is about as bright as IM Peg itself. In fact, we modify any companion selected with a mass larger than that of the primary by "converting" the companion into a white dwarf with an absolute V magnitude of about 11. Thus, both ends of the distribution have little effect, and so the details of our extrapolation are of little concern.

For those hypothetical companions that would have evolved into white dwarfs, we use a different procedure for assigning both luminosity and mass. For such a white dwarf, we arbitrarily assign a V magnitude according to the formula $M_{Vnew} = 14 - 0.3\ M_{Vold}$. This formula was selected to have two main properties: first, the originally most massive stars become the dimmest white dwarfs at the present day (because they were the first to evolve and so their initially hot surfaces have had the longest time to cool), and, second, it yields a range of luminosities typical of the white dwarfs among the 315 nearest stars (Henry 2003). For a discussion of the white dwarf masses, see the next section.

We also consider alternative statistical distributions of companion luminosity based on actual samples of nearby stars. This simpler, more direct approach has the virtue of not forcing the luminosity distribution into any particular mathematical form, which might fail to adequately reflect the complexity of the actual luminosity distribution. One of these present-day samples of stars is a list of the nearest known 315 stars (Henry 2003). With the exception of 19 white dwarfs in the list, all the stars appear to be main-sequence stars. The absolute V magnitudes in this list range from 0.58 to 50. (Note: for convenience, we refer to all of these objects as "stars," even though the assigned mass of the dimmest is only three times that of Jupiter.) Henry (2003) suggested that a truly complete survey would contain about an additional 20% of main-sequence stars distributed perhaps uniformly in mass in the range 0.08 to 0.5 solar mass. Consequently, we constructed a second such list by taking the 315 known stars and adding 61 randomly chosen stars (for a total of 376) from a flat distribution of mass in that range and then applying the same empirical mass-luminosity relation to obtain the corresponding absolute magnitudes. We also constructed a third such list by adding, instead, 15 bright stars with M_V ranging from 0.8 to 3.2 to account for possible sampling noise at the extreme and also to allow for the possibility that the

companion is itself a double. Finally, we used the list of the 100 nearest stars (Henry 2003) as the fourth such list.

The most obvious difference between the two types of mass function is that the present-day mass function peaks in the middle of the M stars, even after adding extra M stars to the list, while the initial mass function follows a straight power law and thus rises steadily all the way to the cutoff at 0.08 solar mass. A subtler, but equally important, distinction is that the present-day mass function falls off more rapidly at high masses. Nevertheless, we find below that choosing the alternative mass functions does not alter our conclusions.

2. Mass

We use the same mapping technique for selecting the mass (more precisely, for selecting the logarithm of the mass in units of M_{Sun}). The high-mass and low-mass tails of the distribution are treated in the same way as for the distribution of magnitudes. The only complication is the special treatment needed for white dwarfs, which appear explicitly in the lists of present-day observed stars, and whose masses are not included in those lists. In effect, we partition each such list into two separate lists, one for white dwarfs and one for normal stars, simply by noting where the random pointer falls in relation to the two bracketing stars in the list. When the random pointer falls closer to a white dwarf than to the next star in the list, we say that the Monte Carlo throw has produced a white dwarf; otherwise, we say it has produced a normal star. Masses of normal stars are selected by interpolation as if the white dwarfs were omitted from the list. On the other hand, white dwarf masses are chosen independently and at random, according to the observed distribution of 50 white dwarf masses from McMahan (1989), as reported by Bergeron *et al.* (1992); each hypothetical white dwarf is assigned randomly to one of the bins in the distribution's histogram with a probability proportional to the bin's height. For stars with masses greater than that of the IM Peg primary (which have thus presumably evolved into white dwarfs), we modify the Monte Carlo throw accordingly. Since we do not have a detailed model of the evolution of white dwarfs, we modify the mass by arbitrarily substituting 1/3 of the chosen mass. Such a procedure results in masses typical of the white dwarfs observed by McMahan (1989).

3. Period

Duquennoy and Mayor (1991) analyzed the orbits of 82 nearby stellar systems with solar-type primaries and determined the distribution of their periods. We adopt Duquennoy and Mayor's parameterization of the distribution as a Gaussian distribution of the (base 10) logarithm of the period in days. This Gaussian has a mean of 4.7 and a standard deviation of 2.3. However, the period chosen in this manner is subject to a pair of stability criteria. On the recommendation of Holman (2003), we reject any selection whose periapse distance is less than 2.3 times the separation of the known binary of IM Peg. (Holman has performed a large set of numerical integrations of the orbits of triple systems with low-mass third bodies, starting from a wide variety of initial conditions. His results indicate that a third body with a smaller periapse distance than our limiting value is extremely likely to be either ejected from the system or destroyed in a collision.) From the known binary orbital period and the total mass of the pair, estimated to be 2.3 ± 0.3 solar masses (Berdyugina *et al.* 1999), we calculate the binary separation as about 0.22 AU. The hypothetical companion's periapse distance is calculated from the sum of its mass and

the pair mass, in combination with the Monte Carlo period and eccentricity. At the other extreme of orbit size, if the apoapse distance is greater than 20 pc, we replace the star, since such systems are unstable in interactions with other bodies in the Galaxy.

When a selection is rejected because of a stability criterion, the whole set of Monte Carlo parameters is dropped, and a replacement is built with all new parameters. In other words, the replacement is simply the next star in a deterministic list produced by the random number generator. This procedure thus minimizes the changes in the list of candidate stars due to any change in the rejection criterion, and so reduces the sampling noise in any comparison of two Monte Carlo runs that differ in the parameter values that govern the selection criteria.

4. Eccentricity

Following Duquennoy and Mayor (1991), we assume one of two different distributions of eccentricity, depending on orbital period. We approximate their observed eccentricity distribution for periods shorter than 1000 days with a Gaussian of mean 0.25 and standard deviation 0.14. The non-physical eccentricities outside the range of 0 to 1 are rejected in the same way as unstable orbits discussed in the previous subsection. Similarly, for periods longer than 1000 days, we approximate their distribution with a probability density function that rises linearly from zero at $e = 0$ to a maximum at $e = 0.4$, then stays flat to $e = 0.9$, and then drops sharply back to zero.

5. Inclination

Orbits are assumed to have a uniform distribution in spatial orientation. To create this distribution, we select the cosine of inclination of the orbit to the plane of the sky as a random variate uniformly distributed on the interval $[-1, +1]$.

6. Ascending node, Argument of Periapse, Initial Mean Anomaly

Each of these three angular orbital elements is chosen independently of all other parameters with a uniform distribution over the range 0 to 360 degrees.

7. Brightness Variation of IM Peg

To model the effect of the change of the primary's brightness, we make the simplifying assumption that the luminosity of the system varies linearly during the GP-B mission. This assumption is favored by the historical behavior of the orbit-averaged brightness of IM Peg (see, *e. g*., Kolodziejczak 2002). A linear change in the brightness ratio maps into an approximately linear drift in the effective guide point, contributing in turn an error of the same size to the relativistic drift ultimately determined by GP-B. The fractional rate of luminosity variation for the unresolved IM Peg system is chosen from a Gaussian distribution with zero mean and standard deviation of 0.05 per year (Kolodziejczak 2002). To provide some insight into the relative size of GP-B errors due to such brightness changes, as compared to those due to the motion of a luminous companion about the center of mass of the binary, we also repeat some of our computations assuming a standard deviation of zero, *i.e.*, assuming that the mean rate of brightening during the mission happens to be zero.

8. Rotation Velocity

Because the sensitivity of spectroscopic searches depends in part on the projected rotation velocity, we assign this parameter in the Monte Carlo selection from a uniform distribution between zero and a maximum that depends on spectral type. We take the maximum to be twice the mean equatorial rotation velocity tabulated by Lang (1992), based upon inspection of the observed distribution (Gray 1992) of these velocities.

III. Detectability

We apply four types of detection criteria to the companions chosen as described above. In each case, the detectability is not taken as an absolute yes-no switch. Rather, each criterion is assigned a statistical probability, based on the available estimates of the relevant measurement noise. We take each criterion to be independent of the others, so that the joint probability of any given hypothetical companion escaping detection is the product of all the individual probabilities of escaping detection. This approach is appropriate, since the detection probabilities mainly represent the effects of measurement noise in each observation and since the noise is presumably uncorrelated between any two observations. The four detection types are VLBI astrometry, HST imaging, ground-based imaging, and spectroscopy. We describe each in turn below.

For these detectability calculations, we convert the absolute magnitudes and physical distances to apparent magnitudes and angular separations using a distance of 95 pc, based on the VLBI-measured parallax of IM Peg. In three of the tests described below, we then convert these absolute magnitudes into relative magnitudes with respect to IM Peg, for which we assume an absolute V-magnitude 1.12, corresponding to the apparent V-magnitude 6.01 adopted by Kolodziejczak (2002). This magnitude is based on measurements sensitive to the total luminosity of the known primary and secondary, as well as any companion that may be present. We also use his adopted apparent magnitudes B = 7.15, R = 5.10, and I = 4.54. We apply the color of a K2 giant from Drilling and Landolt (2000) to get U = 8.33.

1. VLBI Astrometry

We assume that the existence of a bound companion would be the most probable explanation of a reliably detected proper acceleration of IM Peg. Thus, we would equate astrometric “detection” with the determination of a proper acceleration three or more standard deviations from zero. We consider a set of 34 VLBI position determinations for IM Peg (some completed, others projected) from 1991 through 2004. We first compute, for the epoch of each VLBI observation, the plane-of-sky offset of the center of mass of the binary due to its orbit about the hypothetical companion. We then fit an 11-parameter model to these offsets, just as we expect to do when we estimate the proper motion of IM Peg from the full set of VLBI data. In the fit, we determine a quadratic time polynomial for each sky coordinate, as well as coefficients for terms with the signatures of parallax and the binary orbit. We assume a position standard error of 0.54 mas in each direction at each epoch, since 0.54 mas is the standard error needed to make chi-squared per degree of freedom equal to unity in the current 11-parameter fit to the existing VLBI data (from 1991 through July of 2002). The resulting two acceleration components are taken as the hypothetical

proper acceleration of the centroid of the stellar radio emission. Finally, assuming a Gaussian distribution of the errors in each of the two estimated acceleration components, with a standard deviation equal to that found in our 11-parameter fit, we calculate the probability that the measured proper acceleration would lie within a circle of radius three standard deviations centered on the origin. This result is the probability that the companion escapes detection via our (long-period) VLBI astrometry.

We also have a short-period test based on analysis of our VLBI residuals to date. For any orbit with a period between 1.5 and 5.0 years, we would have detected the systematic (periodic) residuals if the amplitude were above some threshold. We estimate the one-standard-deviation detection level to be about 0.5 mas. Thus, we calculate the amplitude of the plane-of-sky motion of the center of mass of the IM Peg binary due to its orbit about the hypothetical companion and determine the probability that the measured amplitude would be less than three standard deviations.

Two aspects of the VLBI test deserve further comment. As is well known, for an isotropic, two-dimensional, zero-mean Gaussian probability distribution, the median offset from zero is 1.18 standard deviations. Therefore, we would not be surprised to find a result this large in the final VLBI fit, even without considering possible non-Gaussian systematic errors. Because of such systematic errors, any arbitrary detection limit is somewhat problematic; nonetheless the customary 3-sigma level represents a plausible detection threshold. In addition, one could implement a detection criterion based on a search for a slow variation in the spectroscopically determined radial velocities of the IM Peg primary. However, because this test is relatively insensitive compared to the VLBI test for most of the hypothetical orbits, we have not modeled it in the present study. Inclusion of this or other plausible additional tests could not increase and would almost surely decrease the error estimates produced by our study.

2. HST Imaging

The HST WFPC2 camera obtained 30 images of IM Peg in support of the GP-B mission through various spectral filters on 1997 June 28, and these have been examined for possible companions (Kolodziejczak 2003). Again we ask, what is the probability that a given Monte Carlo companion would escape detection by this method at this epoch? Using image-processing software and a manual search of I-band images, Kolodziejczak has established, as a function of color and separation on the sky from IM Peg, an upper limit on the apparent brightness of any possible undetected star in the images, as well as the rms noise level of the background. From these values, we can compute the likelihood that the companion's image would have been too faint at I band to be “noticed” above the background. At U band, we consider anything less than three standard deviations a non-detection. In order to compare brightnesses between IM Peg and each hypothetical companion, we compute the U and I magnitudes using main-sequence colors from Drilling and Landolt (2000) and white-dwarf colors from Bergeron *et al.* (1995). The noise estimates (Kolodziejczak 2003) cover angular separations from 0.04″ to 3.2″ for U band and from 0.02″ to 7″ for I band, as shown in Table 1. We assume that the monotonic detection-level decrease flattens out at the outermost distance in the table.

The situation is further complicated by the irregular shape of WFPC2 images, as well as the superposition of two such images to form a mosaic, which is done to reduce the effects of the

WFPC2 point-spread function and cosmic rays. For simplicity, we approximate the mosaic shape by a fully-covered circular zone of radius 15″ centered on IM Peg, plus another circular zone out to 92″ missing a wedge about 60 degrees wide centered on position angle ~125 deg. Also, rather than introduce the complexity entailed in computed anisotropic error distributions, we do not use the actual orientation on the sky of the HST images, but instead suppress at random one sixth of the otherwise possible detections between 15″ and 92″ from the center. Objects more than 92″ away are, more likely than not, outside the WFPC2 field of view, and hence are deemed undetectable by these observations. (In any case, they will almost inevitably be outside the GP-B telescope field of view, too.) We show below that our results are not highly sensitive to any of the values of these parameters describing the WFPC2 image shape.

Table 1. HST noise level at two wavelengths as a function of offset from the primary. Each level is expressed as the differential magnitude (companion – IM Peg) of a spot either at 3 standard deviations brightness (for 334 nm) or at the actual limiting magnitude of the exhaustive search (for 1020 nm). The differential magnitude of a spot of 1 standard deviation brightness is also given for 1020 nm. In each column, each value applies from the specified inner radius out to the radius of the next entry in the column (and indefinitely beyond the last entry).

Annulus inner radius	Differential limiting magnitude		Background as differential magnitude
(arcsec)	334 nm	1020 nm	1020 nm
0.02		2.7	3.2
0.04	2.3		
0.07		4.5	5.4
0.08	5.3		
0.11		5.0	5.7
0.16		5.5	6.1
0.20	7.9	6.8	7.8
0.25		6.9	7.5
0.34		6.9	7.8
0.40	9.5		
0.50		9.1	9.7
0.75		9.7	10.4
1.0	11.5	11.2	11.9
1.5		12.5	13.0
2.0		13.3	14.0
2.5		13.6	14.3
3.0		14.3	14.9
3.2	14.8		
3.5		15.2	15.9
4.0		15.9	16.5
5.0		15.7	16.3
6.0		16.2	16.8
7.0		16.3	17.1

3. Ground-based Imaging

On several nights from Summer 2002 to Spring 2003, Arne Henden obtained CCD images centered on IM Peg at the Flagstaff station of the USNO. Each night, he obtained a series of images through UBVRI filters. From these, he computed the night-by-night and mean magnitudes of all stars detected in the images (Henden 2003). He estimates the limiting magnitude (three-standard-deviation noise level) at V band to be 19 (13 magnitudes dimmer than the primary), valid at separations of 8″ or more. The analysis of these images thus plugs the holes in HST coverage outside of 15″, albeit with reduced sensitivity. Since the orbital period of a companion that far from the primary would be thousands of years, the motion over the course of a few months would be negligible, and all the Flagstaff images together can be considered as one test of detectability.

4. Spectroscopy

We can also use high-resolution spectroscopy to search for evidence of absorption lines indicative of a luminous third body. This search can place useful bounds on the brightness of a companion because of the Doppler separation between the lines of the primary and the corresponding lines of a companion. This separation is guaranteed by our stability criterion for the companion orbits, since the orbital Doppler shift of a companion in any stable orbit will at all times be considerably less than the amplitude of the observed periodic Doppler shift of the primary, and since the Doppler rotational broadening of the observed lines is also less than that amplitude. (The largest possible companion orbital Doppler shift is only 1/3 that of the primary. This maximum occurs only near the times of periapse for the tightest, most eccentric allowed orbit.) Spectroscopic observations of IM Peg obtained by Berdyugina *et al.* (1999) include several at epochs for which the primary's orbital Doppler shift is near maximum.

The major considerations, then, are the relative brightness and the Doppler broadening due to stellar rotation. By inspecting the sample plots of spectroscopic residuals (after subtracting the inferred spectrum of the primary) provided by Berdyugina (2003), we can estimate the “3-sigma” detection level for a hypothetical companion, in terms of the ratio of the companion's brightness to the IM Peg system brightness in the V band. The detection level depends on the companion's spectral type in two ways: first, because the intrinsic line strengths vary with spectral type and, second, because the distribution of stellar rotation velocity depends on spectral type. M stars (and white dwarfs) are too dim to be detected spectroscopically. For K and G stars, which rotate slowly, we estimate the detection level to be about 4%. For F and earlier spectral types, we use different detection levels according to the projected rotation velocity, which we assign randomly, as described in the previous section. Table 2 shows the detection levels in terms of $v \sin i$, the projected rotation velocity. We recognize, however, that a certain minimal sensitivity stems from the basic classification of IM Peg as a K giant with colors that agree with the indications of the relative strengths of absorption lines in its spectrum. Consequently, we impose a (very conservative) maximum brightness on any hypothetical companion at half of the total brightness of the IM Peg system.

Table 2. Spectroscopic “3-sigma” detection level expressed as a percentage of the V brightness of the IM Peg system for various projected rotation velocities

v sin *i* (km/s)	detection level (%)
0	5
10	6
25	10
50	15
75	50

5. Detection statistics

The probability of detection is one of the important statistics collected in this simulation. For each Monte Carlo run, we tally the sum of the detection probabilities for all the individual throws, both separately for each detection test and jointly for detection by any test. A run with 600,000 randomly constructed companions gave the probabilities shown in Table 3.

Table 3. Companion detection probabilities as percentages
(sample of 600,000 throws)

Percent	Detected by
35.59	proper acceleration (VLBI)
7.62	short-period motion (VLBI)
31.58	HST resolution at U
43.64	HST resolution at I
15.60	ground-based optical images
3.64	spectroscopy
82.50	all methods combined

It is apparent from the table that the sum of the test-by-test tallies exceeds the overall detection rate -- naturally so, since some stars are detectable by more than one technique. Also, note that, although the six tests are independent in the sense that the measurement noise is uncorrelated from one to the next, the test criteria "overlap" to varying degrees. For example, there must be some correlation among the three optical imaging techniques, as bright stars would tend to be detected by all three, while the dimmest stars would escape all three.

IV. GP-B errors due to an undetected companion

The two error signals of principal interest are the apparent proper motion of the GP-B guide point and the aberration-like component of annual motion of the guide point. It is a straightforward calculation to obtain the instantaneous offset of the guide point (from the orbit-averaged optical centroid of the IM Peg binary) as a function of time and to extract the two signals of interest. We start by noting that the roll-averaged response of the GP-B telescope has four regimes with respect to the angular separation between IM Peg and any additional source in the sky: (1) for a

source close to IM Peg, the guide point is approximately the centroid of the total light from the system; (2) for a source more than about 1″ away, the guide-point offset as a function of separation begins to level off and asymptotically approaches a value that is proportional to the brightness of the additional source; (3) for a source with separation between about 45″ and 75″, vignetting sets in, and the offset gradually drops to zero as the separation increases; and (4) for a source beyond about 75″, the offset is zero. The transition in regime 2 is known empirically (Goebel 2000, Kolodziejczak 2002), but has no detailed model. By trial and error, we find that the shape of the transition is reasonably well approximated by the function

$$g = \frac{Lra}{\sqrt{r^2 + a^2}} \qquad \text{(IV-1)}$$

where g is the guide-point offset, L is the proportionality constant, r is the angular separation of the source from IM Peg, and a is the unscaled asymptotic limit. Clearly, this function has the right limiting behaviors for regimes 1 and 2: for $r << a$, $g \approx Lr$, and for $r >> a$, $g \approx La$. We also find that, among the family of functions of the form $ra/(r^n+a^n)^{1/n}$, this function most nearly approximates the transition between the two extremes. From Kolodziejczak (2002), we have the statement that, in the asymptotic limit of regime 2, a value of 1 mas for g corresponds to a value of 0.00044 for L. From that limit and equation (IV-1), we obtain a value of 2.27″ for a. For simplicity, we model the transition in regime 3 as a linear roll-off between two sharp edges by adding a vignetting factor R to equation (IV-1). Thus, we write

$$g = \frac{RLra}{\sqrt{r^2 + a^2}} \qquad \text{(IV-2)}$$

where R is unity in regimes 1 and 2, varies linearly in regime 3 (from 1 at r = 45″ to 0 at r = 75″), and is zero in regime 4.

The factor L is not strictly the ratio of luminosities between the other source and IM Peg, but rather the ratio of detector signals due to the two sources. We follow Kolodziejczak (2002) in approximating L as the sum of contributions from the four bands (V, B, R, and I) that comprise the detector bandpass. However, we compute only one L for each Monte Carlo star, rather than four (one for each detector pair), by taking the arithmetic average of Kolodziejczak's four detector-pair weighting factors. (The error introduced into our results by this simplification should be small compared to other uncertainties in our study.) We use the same set of main-sequence star colors here as in the HST detectability calculations described earlier. (Note: although, in the limit of a dim hypothetical companion, the factor L reduces to the simple ratio of the luminosity of the companion to that of the binary, L is more properly computed as the ratio of the detector signal due to the companion to that of all three components together. We use the latter formulation.)

To quantify the effect of each hypothetical companion on the guide point, we first tabulate the effective guide-point offset at 50 evenly spaced epochs spanning the projected one-year science mission. Next, we use a least-squares fit to solve simultaneously for eight parameters: the mean offset, proper motion, and proper acceleration (all in two dimensions), plus the signatures

mimicking annual parallax and aberration. Note that the only effect of interest here is due to the light of the companion; the orbital motion of the known IM Peg binary about the overall system center of mass is already included in the VLBI astrometry, albeit with a mean epoch different from that of the GP-B mission. In practice, the GP-B relativity test will be sensitive to the total motion of the guide point, including the effects of both the companion's light and its mass. The proper motion determined by VLBI astrometry (reflecting the motion of the radio emission tied to the primary) will then be added, thus largely removing the effect of the companion's mass as a source of error.

We can consider each error signal as being composed of two parts: a kinematic effect due to the relative motion of the luminous companion and a photometric effect due to the changing brightness ratio between IM Peg and the companion during the mission. We allow for a range of mean rates of brightness change during the mission by including the brightness variation as one of the Monte Carlo parameters, as described in section II.7. However, by the time of the post-mission analysis, luminosity measurements of IM Peg over the course of the mission will be available, so that this mean rate will be known. We therefore report on the probability distributions of both the full, simulated signal and the signal produced when the brightness variation has been set to zero.

Table 4 shows the confidence levels in proper motion for the two error distributions, each for four alternative values of t. Although, in general, allowing IM Peg to vary in brightness should tend to increase the expected error signal and the 1-sigma confidence level, the table shows that the higher thresholds (the 3- and 4-sigma levels) differ very little between the two distributions. Figure 1 shows the corresponding cumulative probability distribution. For all values of t, the 1-sigma confidence limits in Table 4 are negligible, or even zero (being zero whenever the probability of an undetected companion is below 33%). However, both Table 4 and Figure 1 illustrate that the distribution is extremely non-Gaussian and that the results depend strongly on the value of the multiplicity fraction, t. In particular, for t greater than about 0.8, the 3-sigma confidence limits exceed three times the 0.05 mas/yr 1-sigma requirement (Kolodziejczak 2002) on the uncertainty of the difference between the optical and radio proper motions, while the smaller (and more plausible) values of t lead to much lower 3-sigma confidence limits.

Thus our best estimate of the 3-sigma confidence limit is that it falls in the range of values in Table 4 for $t = 0.1$ to $t = 0.5$, namely the range 0.012 mas/yr to 0.031 mas/yr. We conclude that the 3-sigma (99.7%) confidence level for the guiding rate error is well below even the required 0.05 mas/yr "1-sigma" accuracy of the proper motion of the effective guide point of the GP-B mission. Moreover, it is also well below the limit on the 3-sigma uncertainty of the relativity tests due to gyro readout noise alone. (The 1-sigma uncertainty due to noise alone is expected to be about 0.1 mas/yr from each gyro, yielding 0.05 mas/yr if the signals from all four gyros can be averaged together; Kolodziejczak 2002.)

The results above are computed allowing for the observational constraints provided by the VLBI data set that we project will be available by the end of 2004, but it is also useful to compute the confidence limits that reflect the current set of reduced data. These limits characterize the error distributions due to all companions that are unknown currently, even though some of them may be detected later using our VLBI results. We therefore repeated our nominal run using only the VLBI epochs through July 2002. For the reasonably conservative case of $t = 0.5$, the resulting

1-, 2-, and 3-sigma confidence levels for the guiding rate error are 0.0, 0.0007, and 0.027 mas/yr, respectively.

Table 4. Confidence levels for absolute errors (mas/yr) in proper motion due to undetected companions of IM Peg, with and without brightness variation in the primary

Multiplicity fraction, t	67% confidence	95.4% confidence	99.7% confidence	99.994% confidence
	(1-sigma)	(2-sigma)	(3-sigma)	(4-sigma)
	Varying Brightness Ratio (Run 1)			
1.0	0.00053	0.0063	0.32	4.7
0.8	0.00004	0.0023	0.13	2.9
0.5	0.0	0.00056	0.031	1.5
0.1	0.0	0.0	0.0017	0.27
	Constant Brightness Ratio (Run 2)			
1.0	0.00030	0.0053	0.32	4.7
0.8	0.00002	0.0016	0.13	2.8
0.5	0.0	0.00033	0.030	1.5
0.1	0.0	0.0	0.0012	0.27

Table 5 shows the corresponding confidence levels for the apparent aberration-like signal, for the same two runs as shown in Table 4. Again, we see that the errors are extremely non-Gaussian. For all values of *t*, the 3-sigma confidence levels are small enough to allow the gyro readout scale factor to be calibrated, as desired, to better than one part in 10^5 from the known ~20″ amplitude annual aberration of IM Peg's apparent position.

Table 5. Confidence levels for absolute errors (mas) in apparent aberration due to undetected companions of IM Peg, with and without brightness variation in the primary.

Multiplicity fraction, t	67% confidence	95.4% confidence	99.7% confidence	99.994% confidence
	(1-sigma)	(2-sigma)	(3-sigma)	(4-sigma)
	Varying Brightness Ratio (Run 1)			
1.0	4.2×10^{-6}	0.0025	0.17	2.4
0.8	1.1×10^{-9}	0.00056	0.065	1.5
0.5	0.0	7.6×10^{-6}	0.015	0.80
0.1	0.0	0.0	0.0003	0.14
	Constant Brightness Ratio (Run 2)			
1.0	3.5×10^{-6}	0.0025	0.17	2.4
0.8	2.2×10^{-11}	0.00056	0.065	1.5
0.5	0.0	7.3×10^{-6}	0.015	0.80
0.1	0.0	0.0	0.0003	0.14

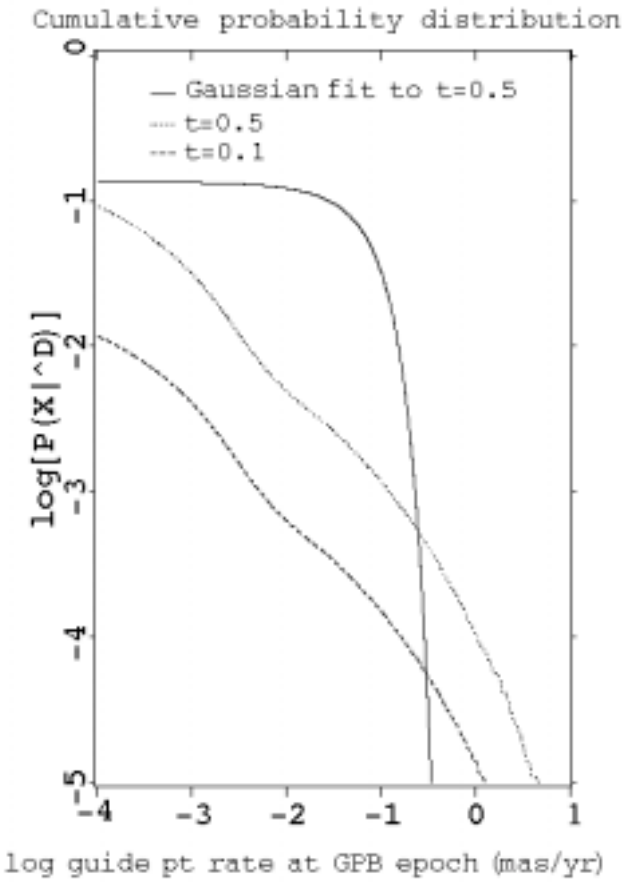

Figure 1. Conditional probability, $P(X|\hat{}D)$, of exceeding a given error in the effective guide-point proper motion due to an undetected companion, for the two more plausible assumed values of multiplicity fraction, *t*. We also show a Gaussian with the same standard deviation as the distribution of nonzero errors for $t = 0.5$, and scaled to the same cumulative probability of nonzero error. We plot $\log[P(X|\hat{}D)]$ rather than $P(X|\hat{}D)$ (or the closely related cumulative conditional probability distribution of the error) in order to show more clearly the part of the distribution where $P(X|\hat{}D)$ is near maximum. Thus the horizontal scale corresponds to 0.0001 to 10.0 mas/yr. Note that, because there is a finite probability that the error will be exactly zero, $P(X|\hat{}D)$ does not approach 100% even in the limit of zero error. (The error due to an undetected companion lying outside the GP-B field of view during the entire mission will be exactly zero. Also, by definition, if there is no companion, or if there is one but it is detected, the error is again zero.) Furthermore, because the probability of any nonzero error depends on the value of *t* only through the scale factor $t/(1 - td)$, which is the same for all nonzero errors, on the log-log plot the two curves are identical except for a vertical offset of 0.893, *i.e.*, a factor of 7.81 in the conditional probability. Interested readers, if any, can readily compute the curves for any other value of *t*, using the value $d = 0.8250$ from Table 3. The corresponding Gaussian for $t = 0.1$ (not shown) would be identical to that for $t = 0.5$, except for the same vertical offset. These curves show that the error distribution is extremely non-Gaussian. Indeed, for any value of *t*, the median nonzero absolute error is about one hundred-fold smaller than for the corresponding Gaussian. On the other hand, the tail of the error distribution falls off much less rapidly than that of a Gaussian. Indeed, while the 95.4% (2-sigma) and 99.7% (3-sigma) confidence levels are acceptably small in the context of the GP-B error budget, the 99.994% (4-sigma) confidence level is far more than a factor of 4/3 larger than the 3-sigma level, reaching 0.27 and 1.5 mas/yr for the $t = 0.1$ and $t = 0.5$ cases, respectively (see Table 4).

V. Sensitivity to the Parameter Values

To determine the sensitivity of our results to the particular choice of parameters, we have made a number of comparison runs with a variety of changes in the configuration. Table 6 shows a summary of these runs. Run 1 is the nominal run performed using the parameters described in the previous sections. Run 2 is the other run discussed above, for which we assume a constant rather than steadily changing IM Peg brightness. None of the remaining sensitivity runs change the 1-sigma, 2-sigma, or 3-sigma confidence levels of either of the two error signals by more than a factor of 2, for any value of *t*, with only a few, unimportant exceptions. The exceptions include Run 3, which displays large relative changes (up to a factor of 10) at the 1-sigma level of the apparent aberration (though the absolute errors in question are still negligible, see Table 5). Some runs, of course, actually decrease the error signals, such as Run 4, which has the effect of excluding the closest companions (which are the hardest to detect optically). Since none of these confidence levels are within a factor of 2 of the allowed limits, we conclude that our results are relatively insensitive to any of these modifications to the model. In any case, we cannot claim that this calculation, or even a perfect calculation based on the current state of astronomical knowledge, is capable of finding 2-sigma or higher confidence limits reliable to any better than a factor of 2.

Table 6. Comparison Runs

Run	Description
1	Nominal assumptions as described in text
2	Set the brightness variation of IM Peg to zero (from a Gaussian distribution with standard deviation 5% per year)
3	Changed the standard dev. of IM Peg's brightness variation to 15% per year (from 5%)
4	Raised the minimum stable periapse to 4 times the binary separation (from 2.3)
5	Changed the break between short- and long-period distributions of eccentricity to 100 days (from 1000)
6	Doubled the spectroscopic detection noise sigma (from 0.0125)
7	Raised the ground-based imaging detection sigma by 4 magnitudes to IM Peg + 8 (from IM Peg + 12)
8	Raised the maximum stable apoapse to 50 pc (from 20)
9	Raised the U-band HST 3-sigma detection level at 3.2″ separation to IM Peg + 13 (from IM Peg + 14.8)
10	Doubled the I-band HST noise levels
11	Changed the random number starting seed to 2 (from 1)
12	Changed the random number starting seed to 3 (from 1)
13	Changed the random number generator to that from Numerical Recipes (Vetterling 1985) (instead of the "linear congruence" type generator used in other runs)
14	Raised the spectroscopic detection noise sigma for F and earlier stars to 0.125 (from 0.0125)
15	Narrowed the HST full-coverage zone to 10″ (from 15″)
16	Widened the HST full-coverage zone to 90″ (from 15″)
17	Flattened the HST U-band 3-sigma detection level beyond 0.4″
18	Halved the HST I-band noise levels
19	Lowered the HST U-band 3-sigma detection level at 0.04″ to IM Peg + 5 (from IM Peg + 2.3)
20	Narrowed the HST full-coverage zone to 5″ (from 15″)
21	Doubled the I-band HST noise levels and also raised the level for U-band by 1 magnitude at all distances
22	Substituted a list of the 315 nearest stars (Henry 2003) for the power-law table
23	Added 61 fictitious low-mass stars to the list of 315 nearby stars to compensate for presumed incompleteness (Henry 2003)
24	Added 15 fictitious high-mass stars to the list, ranging from 1.4 to 2.3 solar masses, to allow for possible twinning effect in multiple systems
25	Substituted the list of 100 nearest stars (Henry 2003), excluding two that had been tabulated without magnitude or without mass
26	Deferred the mission by one year (from midpoint of 2004 June 1)
27	Doubled the mean eccentricity for short-period orbits (from 0.25)
28	Tripled the standard deviation of the eccentricity distribution for short-period orbits (from 0.14)

Appendix A. Scaling from probability conditional upon a companion to probability conditional upon a non-detected companion

In this appendix, we derive the relation between the conditional probability of an event, given the existence of a third body, such as we compute in a Monte Carlo study, and the corresponding conditional probability, given the non-detection of any third body. In terms of conditional probabilities, using the event composition method, we write

$$P(X) = P(X|D)\,P(D) + P(X|\hat{}D)\,P(\hat{}D) \tag{A-1}$$

where X is some event, and events D and $\hat{}D$ are, respectively, the mutually exclusive (and collectively exhaustive) detection and non-detection of a third body. The events of interest for this study are those of the form "error signal due to an undetected third body exceeds a specified threshold." From here on, we restrict X to events of this nature. We also assume that, if any third body in the system is detected, either its effects on the GP-B measurements can be modeled with negligible errors, or these errors will be evaluated for the particular characteristics of the detected body. Consequently, $P(X|D)$ and hence the entire first term on the right-hand side of equation (A-1) vanish, so that

$$P(X) = P(X|\hat{}D)\,P(\hat{}D) \tag{A-2}$$

or

$$P(X|\hat{}D) = P(X) / P(\hat{}D)\ . \tag{A-3}$$

Determining the dependence of $P(X|\hat{}D)$ on the error threshold defining X is the goal of this study.

Aside from the question of the detection or non-detection of a third body is the more basic question of whether a third body actually exists. If there is none, then the simulation of "its" effect is trivial. Therefore, as a practical matter, our Monte Carlo calculations need only cover the cases where a third body exists. The possibility that there is no third body can be accounted for by scaling the probability estimates obtained from the Monte Carlo calculation. As above, we write

$$P(X) = P(X|T)\,P(T) + P(X|\hat{}T)\,P(\hat{}T) \tag{A-4}$$

where events T and $\hat{}T$ are, respectively, the existence and non-existence of a third body. For any event X of the type we are considering here, namely, those defined as the error signal exceeding a given threshold, $P(X|\hat{}T)$ must be zero for any non-zero threshold. Thus,

$$P(X) = P(X|T)\,P(T)\ . \tag{A-5}$$

Combining equations (A-3) and (A-5), we get

$$P(X|\hat{}D) = P(X|T)\,P(T) / P(\hat{}D)\ . \tag{A-6}$$

$P(X|T)$ is the probability that X occurs for a "random" Monte Carlo throw. More precisely, the Monte Carlo experiment is designed to yield $P(X|T)$ in the limit of an infinite number of throws, or a sufficiently accurate estimate of $P(X|T)$ after a large, but not impractically large, number of

throws. In contrast, $P(T)$, the probability that any binary like IM Peg is a triple in the absence of any observational evidence on the question, must be inferred from the astronomical literature. Only $P(\wedge D)$ would then remain to be evaluated. Using the event composition method again, we write

$$P(D) = P(D|T)\,P(T) + P(D|\wedge T)\,P(\wedge T)\,. \tag{A-7}$$

Next, for convenience, we introduce three abbreviations:

$$\begin{aligned} t &= P(T) \\ d &= P(D|T) \\ f &= P(D|\wedge T). \end{aligned} \tag{A-8}$$

(To summarize: t is the multiplicity fraction, d is the conditional probability of detection, which we estimate directly within the Monte Carlo study, and f is the probability of false detection, which we will eventually assume to be negligible.) In the new notation, equations (A-6) and (A-7) become

$$P(X|\wedge D) = P(X|T)\,t\,/P(\wedge D) \tag{A-9}$$

and

$$P(D) = td + (1 - t)\,f\,. \tag{A-10}$$

Finally, by combining (A-9) and (A-10), we get

$$P(X|\wedge D) = P(X|T)\,t\,/\,[1 - td - (1 - t)\,f\,]\,. \tag{A-11}$$

We can interpret this result in terms of our bookkeeping in the Monte Carlo calculation of $P(X|T)$ as follows. In this calculation, the probability that each randomly selected, hypothetical star would have escaped detection, is added up for all those hypothetical stars whose error contribution exceeds the threshold defining X. By the definition of conditional probability, dividing this sum by the number of throws N gives a "close enough" approximation to $P(X|T)$ for sufficiently large N. On the other hand, the desired conditional probability $P(X|\wedge D)$ can be expressed as the same sum divided by the total number of undetected cases, including both binaries and multiples. When we make N Monte Carlo throws, Nd is the expected number of companions detected, and N/t the corresponding expected total number of cases. Also, $(N/t - N)\,f$ of the $N/t - N$ binary cases would, on average, suffer from false detections. Hence, the net number of undetected cases is $[N/t - Nd - (N/t - N)\,f]$. Dividing the number of throws N by the net number of undetected cases yields $t\,/\,[1 - td - (1 - t)\,f\,]$, exactly the scale factor in (A-11) for converting the one conditional probability into the other.

In addition to characterizing $P(X|\wedge D)$, for possible future use, we compute the leading even moments of the probability distribution. (The inherent symmetry guarantees that all the odd moments are zero.) The same arguments as above apply to the scaling of these moments, since each must be the average over all cases with no detected companion.

In practice, the value of *f*, the probability of false detection, should be negligible for our purposes. We are adopting reasonably conservative detection criteria (three standard deviations on all tests). As shown by equation (A-11), the largest error probabilities occur when $t = 1$, and the dependence on *f* vanishes for that case. More generally, the probabilities depend far less on *f* than on *t*. Consequently, we choose to neglect *f* altogether (*i.e.*, we set *f* to 0), giving

$$P(X|\wedge D) = P(X|T)\, t / (1 - td) \,. \qquad \text{(A-12)}$$

Oddly, the denominator in equation (A-12) seems to imply that the probability of exceeding any given error threshold X increases with increasing conditional detection probability, *d*. This dependence occurs because the $P(X|\wedge D)$ that we are computing is the probability of errors when guiding upon the IM Peg system given that it has no detectable companion. In that case, for any particular value of *t*, the larger the fraction of hypothetical companions that could have been detected, the more "weight" the denominator gives to those hypothetical companions that are likely to escape detection. Nevertheless, the numerator, too, in equation (A-12) decreases with increasing *d*, since the error due to any detected companion is taken to be negligible, under the assumption that its effects could be adequately modeled in the GP-B data analysis process.

One can also derive equation (A-12) by another approach that may seem more intuitive to some readers. Let us consider the *a priori* probability density function of companions to be $P(\boldsymbol{\alpha})$, where $\boldsymbol{\alpha}$ represents the set of all relevant parameters. Let us denote the probability that a companion specified by $\boldsymbol{\alpha}$ will escape detection by $P(\wedge D(\boldsymbol{\alpha}))$, and the *a priori* probability that IM Peg has no detected companion by $P(\wedge D)$. Then, with the definitions in (A-8)

$$\begin{aligned} P(\wedge D) &= P(\wedge T) + P(T)\, P(\wedge D|T) \qquad \text{(A-13)} \\ &= P(\wedge T) + P(T) \int d\boldsymbol{\alpha}\, P(\boldsymbol{\alpha})\, P(\wedge D(\boldsymbol{\alpha})) \\ &= (1 - t) + t\,(1 - d) \\ &= 1 - t\,d \end{aligned}$$

and

$$\begin{aligned} P(X|\wedge D) &= [P(\wedge D)]^{-1}\, P(T) \int d\boldsymbol{\alpha}\, P(\boldsymbol{\alpha})\, P(\wedge D(\boldsymbol{\alpha}))\, \tau(X, \boldsymbol{\alpha}) \qquad \text{(A-14)} \\ &= [P(\wedge D)]^{-1}\, P(T)\, P(X, \wedge D|T) \end{aligned}$$

where $\tau(X, \boldsymbol{\alpha})$ is a "truth" function with value 1 if the error due to companion characterized by $\boldsymbol{\alpha}$ exceeds the threshold defining *X,* and with value 0 otherwise. Furthermore, since we assume that a nonzero error can only be contributed by a companion that is not detected, the integral in (A-14) is simply the probability that a randomly selected companion leads to an error exceeding the threshold defining *X,* which is the conditional probability $P(X|T)$. Thus, rewriting the integral as $P(X|T)$, and $P(T)$ as t, and using (A-13) to substitute for $P(\wedge D)$, equation (A-14) becomes

$$\begin{aligned} P(X|\wedge D) &= [\,1 - t\,d\,]^{-1}\; t\; P(X|T) \qquad \text{(A-15)} \\ &= P(X|T)\, t / (1 - td) \end{aligned}$$

just as in (A-12). The Monte Carlo calculation essentially calculates the integrals in (A-13) and (A-14) as sums over randomly selected $\boldsymbol{\alpha}$, normalizing each sum by the total number of throws.

Appendix B. Programming Considerations

The main results of this study were achieved using Monte Carlo calculations with 600,000 throws. The entire computation, for all four values of t, took ~25 minutes on a Sun SPARC Ultra 5. To minimize this time, we computed the results for each confidence level and value of t from the same set of Monte Carlo throws. This required that we construct a table containing, for each throw, the detection probability and the magnitude of the error should it go undetected. Once this table was sorted by the error magnitude, it was straightforward to locate the error magnitude corresponding to any desired confidence level. For each value of t, equation (A-12) relates any desired confidence level given the lack of detection of a third body to a corresponding confidence level in the distribution of $P(X|T)$ embodied in the sorted table.

There are two commonly used methods for generating pseudo-Gaussian-distributed random variates. The simpler method relies on the fact that the sum of independent random variables, all with the same distribution, approaches a Gaussian distribution as the number of summed variables increases. However, this approximation is poor in the tail of the Gaussian distribution. Since a major interest of this study is the tail of the error distribution, we avoid this method in favor of the second one.

The second method uses a direct analytic transformation from the finite domain of the uniformly distributed variates to the whole set of real numbers. The first step is to select a uniformly distributed point within a unit circle centered at the origin. The simplest method of selecting such a point is to take two uniformly distributed random variates x and y and scale each onto the interval $[-1, +1]$, *i.e.*, convert them to Cartesian coordinates of a point within a two-by-two square centered at the origin. If the point does not fall within the unit circle, the algorithm discards it and tries additional pairs until a suitable pair is found. This process of discarding unusable pairs has little waste; only 1.27 pairs are needed on average to obtain a point within the circle. The next step is to convert to polar coordinates: $r = (x^2 + y^2)^{1/2}$, $\theta = \tan^{-1}(y/x)$. A simple radial transformation then maps the points within the circle onto the full plane as follows: $u = (-2 \ln r^2)^{1/2}$. The angle remains the same. This transformation has the effect of turning the circle "inside out" by mapping the entire circumference onto the origin and the origin onto infinity. It can be shown that because $(-2 \ln r^2)^{1/2}$ is the inverse of a Gaussian function, the transformed point's Cartesian coordinates now each have a Gaussian probability distribution, and jointly have a bivariate Gaussian distribution. Recall that the normalized Gaussian probability density, as a function of a single coordinate z, is $(2\pi)^{-1/2} \exp(-z^2/2)$. The joint probability density of a bivariate Gaussian, integrated over the azimuthal angle and expressed in terms of the radial coordinate v, is thus $\exp(-v^2/2)v$. Since the probability density of a transformed variable is just the probability density of the original variable divided by the absolute value of the derivative of the transformation function, the radial probability density of u can be obtained directly from the radial probability density of r by multiplying the latter by dr/du, *i.e.*, by multiplying $2r$ by $\exp(-u^2/2)u/2r$. (The same result would have come from dividing by du/dr, but the other approach is simpler.) As already noted, the result matches the radial probability density of a bivariate Gaussian. Finally, u sin θ (or $u \cos \theta$) yields the desired Gaussian-distributed variate.

The formulation of an arbitrary probability distribution by suitable mapping from a uniformly distributed random variable requires finding a transformation that produces either the desired probability density function or the corresponding cumulative probability distribution, whichever

is more convenient. Take, for example, the distribution described in Section II.4 for the eccentricity of long-period companions: rising linearly from zero at $e = 0$ to a maximum at $e = 0.4$, then staying flat to $e = 0.9$, and then dropping sharply back to zero. If we normalize the probability density of the desired distribution by requiring the integral to be unity, we find $p(e) = 25e/7$ for $e < 0.4$ and $p(e) = 10/7$ for $0.4 < e < 0.9$. If follows that the probability of hitting the zone of flat probability density is 5/7. Thus, starting from a uniformly distributed variable x, we must make a linear transformation $e = f(x)$ for any value of $x \geq 2/7$, such that $f(2/7) = 0.4$ and $f(1) = 0.9$, *i.e.*, we use $f(x) = 0.2 + 0.7\,x$. For $x < 2/7$, the transformation must satisfy the requirement that df/dx be the ratio of the two probability densities, *i.e.*, $df/dx = 7/(25\,f)$. Solving for f in terms of x gives $f(x) = (0.56\,x)^{1/2}$.

A more fundamental concern is the process of generating the uniformly distributed pseudo-random variates. The most straightforward method is the linear congruence algorithm, which operates internally by generating a sequence of integers and then scales these values to the range [0, 1]. Each integer in the sequence is obtained from the previous value by multiplying by a constant and adding a second constant, modulo the maximum value representable in the computer's word size. As long as the generating constants obey certain number-theoretic requirements, the sequence will visit every possible integer value within the word size exactly once before repeating, and the correlation between successive values will be negligible (Knuth 19??). The SPARC computer used for this study has a 32-bit word, thus providing for a sequence of approximately 4 billion different pseudo-random variates, of which only a small portion is used for any one Monte Carlo run.

References

Batten, A.H., Fletcher, J.M., & MacCarthy, D.G., 1989**, Catalogue of the Orbital Elements of Spectroscopic Binary Systems**, Dominion Astrophysical Observatory, Victoria.

Berdyugina, S.V., Ilyin, I., & Tuominen, I., 1999, "The long-period RS Canum Venaticorum binary IM Pegasi. I. Orbital and stellar parameters", A&A, 347, 932.

Berdyugina, S.V., 2003, Private communication.

Bergeron, P., Saffer, R.A., & Liebert, J., 1992, "A spectroscopic determination of the mass distribution of DA white dwarfs," ApJ, 394, 228.

Bergeron, P., Wesemael, F., & Beauchamp, A., 1995, "Photometric calibration of Hydrogen- and helium-rich white-dwarf models," PASP, 107, 1047.

Drilling, A.S., & Landolt, A.U., 2000, "Normal stars," in **Allen's Astrophysical Quantities**, A. Cox, ed. (AIP Press, New York), p. 381.

Duquennoy, A., & Mayor, M., 1991, "Multiplicity among solar-type stars in the solar neighborhood. II. Distribution of the orbital elements in an unbiased sample," A&A, 248, 485.

Gilmore, G.F., & Zeilik, M., 2000, "Star populations and the solar neighborhood," in **Allen's Astrophysical Quantities**, A. Cox, ed. (AIP Press, New York), p. 471.

Goebel, J.H., 2000, S0490, Rev. B, pp. 10-11.

Gray, D.F., 1992, The **Observation and Analysis of Stellar Photospheres (**Cambridge Univ. Press, Cambridge) p. 390.

Henden, A.A., 2003, Private communication.

Henry, T.J., & McCarthy, D.W., 1993, "The mass-luminosity relation for stars of mass 1.0 to 0.08 solar mass," AJ, 106, 773.

Henry, T.J., 2003, Private communication.

Holman, M.J., 2003, Private communication.

Kolodziejczak, J., 2002, Appendix D, in S0890, Rev 1, "Error tree addendum: Proper motion of the effective guide point", M. Ratner & J. Kolodziejczak.

Kolodziejczak, J.J., 2003, Private communication.

Lang, K.R., 1992, **Astrophysical Data** (Springer, New York) p. 135.

McMahan, R.K., 1989, "Spectroscopy of the DA white dwarfs - Automatic atmospheric parameterization and mass distribution," ApJ, 336, 409.

Reid, I.N., 2003, "Canvassing the neighborhood," STSI Newsletter, 20, 18.

Reid, I.N., Gizis, J.E., & Hawley, S.L., 2002, "The Palomar/MSU nearby star spectroscopic survey. IV. The luminosity function in the solar neighborhood and m dwarf kinematics," AJ, 124, 2721.

Strassmeier, K.G., Hall, D.S., Fekel, F.C., & Scheck, M., 1993, "A catalog of chromospherically active binary stars (second edition)," A&A Suppl. Ser. 100, 173.

Tokovinin, A.A., 1997, "On the multiplicity of spectroscopic binary stars," Astronomy Letters, 23, 834.

Vetterling, W.T., 1985, **Numerical Recipes Example Book (FORTRAN)** (Cambridge University Press, New York).