

PAPER • OPEN ACCESS

## Gravitational redshift test of EEP with RadioAstron from near Earth to the distance of the Moon

To cite this article: N V Nunes *et al* 2023 *Class. Quantum Grav.* **40** 175005

View the [article online](#) for updates and enhancements.

You may also like

- [PROBING THE INNERMOST REGIONS OF AGN JETS AND THEIR MAGNETIC FIELDS WITH RADIOASTRON. I. IMAGING BL LACERTAE AT 21  \$\mu\$ m RESOLUTION](#)

José L. Gómez, Andrei P. Lobanov, Gabriele Bruni *et al.*

- [Observing the Time Evolution of the Multicomponent Nucleus of 3C 84](#)

Brian Punsly, Hiroshi Nagai, Tuomas Savolainen *et al.*

- [THEORY AND SIMULATIONS OF REFRACTIVE SUBSTRUCTURE IN RESOLVED SCATTER-BROADENED IMAGES](#)

Michael D. Johnson and Carl R. Gwinn

# Gravitational redshift test of EEP with RadioAstron from near Earth to the distance of the Moon

N V Nunes<sup>1,\*</sup> , N Bartel<sup>1</sup>, A Belonenko<sup>2</sup>,  
G D Manucharyan<sup>2,3</sup> , S M Popov<sup>2</sup> , V N Rudenko<sup>2</sup> ,  
L I Gurvits<sup>4,5</sup> , G Cimò<sup>4</sup> , G Molera Calvés<sup>6</sup> ,  
M V Zakhvatkin<sup>7</sup>  and M F Bietenholz<sup>1</sup> 

<sup>1</sup> York University, 4700 Keele St, Toronto, ON M3J 1P3, Canada

<sup>2</sup> Sternberg Astronomical Institute, Lomonosov Moscow State University, Universitetsky pr. 13, 119234 Moscow, Russia

<sup>3</sup> Bauman Moscow State Technical University, ul. Baumanskaya 2-ya, 105005 Moscow, Russia

<sup>4</sup> Joint Institute for VLBI ERIC, Oude Hoogeveensedijk 4, 7991 PD Dwingeloo, The Netherlands

<sup>5</sup> Aerospace Faculty, Delft University of Technology, Delft, The Netherlands

<sup>6</sup> University of Tasmania, Private Bag 37, Hobart 7001, Tasmania, Australia

<sup>7</sup> Keldysh Institute for Applied Mathematics, Russian Academy of Sciences, Miusskaya sq. 4, 125047 Moscow, Russia

E-mail: [nvnunes@yorku.ca](mailto:nvnunes@yorku.ca)

Received 5 April 2023; revised 3 July 2023

Accepted for publication 10 July 2023

Published 27 July 2023



CrossMark

## Abstract

The Einstein Equivalence Principle (EEP) is a cornerstone of general relativity and predicts the existence of gravitational redshift. We report on new results of measuring this shift with RadioAstron (RA), a space very long baseline interferometry (VLBI) spacecraft launched into an evolving high eccentricity orbit around Earth with geocentric distances reaching 353 000 km. The spacecraft and ground tracking stations at Pushchino, Russia, and Green Bank, USA, were each equipped with a hydrogen maser frequency standard allowing a possible violation of the predicted gravitational redshift, in the form of a violation parameter  $\varepsilon$ , to be measured. By alternating between RA's frequency referencing

\* Author to whom any correspondence should be addressed.



Original Content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

modes during dedicated sessions between 2015 and 2017, the recorded downlink frequencies can essentially be corrected for the non-relativistic Doppler shift. We report on an analysis using the Doppler-tracking frequency measurements made during these sessions and find  $\varepsilon = (2.1 \pm 3.3) \times 10^{-4}$ . We also discuss prospects for measuring  $\varepsilon$  with a significantly smaller uncertainty using instead the time-domain recordings of the spacecraft signals and envision how  $10^{-7}$  might be possible for a future space VLBI mission.

Keywords: RadioAstron, gravitational redshift, test of Einstein Equivalence Principle, general relativity

(Some figures may appear in colour only in the online journal)

## 1. Introduction

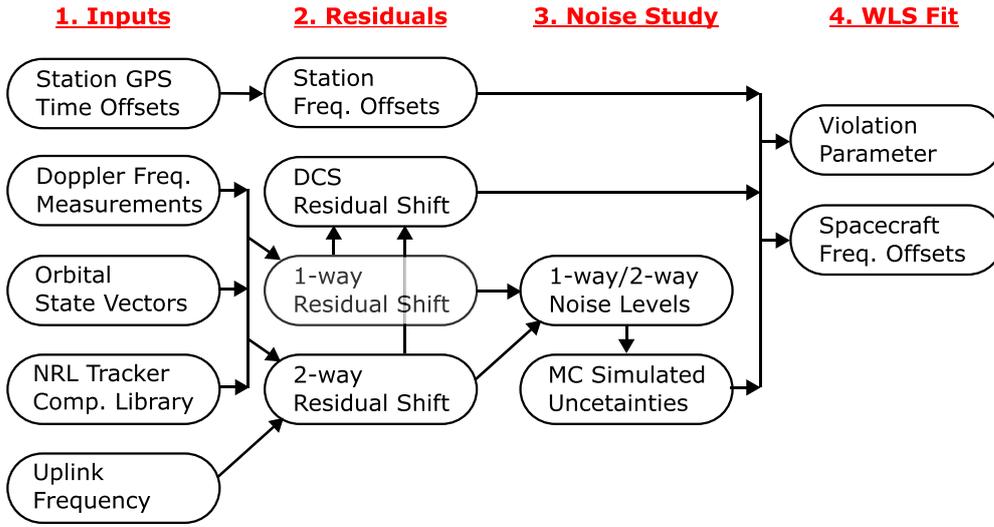
The symmetries that embody the Einstein Equivalence Principle (EEP) underly metric theories of gravitation like general relativity. However, attempts at a quantum description of gravity seem to inevitably lead to violations of the EEP [1]. A consequence of such a violation might be a departure from the predicted gravitational redshift:

$$y_{\text{grav}} \equiv \frac{\nu_e - \nu_o}{\nu_e} = (1 + \varepsilon) \frac{\Delta U}{c^2} \quad (1)$$

where  $\nu$  is the frequency of an electromagnetic signal measured at different points within a gravitational field,  $\Delta U$  is the difference in gravitational potential at the points, and  $\varepsilon$  the violation parameter in the case where identical atomic frequency standards are used by the emitter (e) and observer (o) [2]. Measuring  $\varepsilon$  is the subject of this paper.

The first high-precision laboratory experiments of this type, reaching a relative accuracy of 1%, were done in the 1960s by Pound and Rebka [3] and later improved by Pound and Snider [4]. In 1976, Gravity Probe A (GP-A) [5] was launched on a non-orbital trajectory, with an apogee altitude of 10 000 km, allowing the gravitational redshift to be measured with an accuracy of  $\sigma_\varepsilon = 1.4 \times 10^{-4}$  [6]. More recently, teams utilizing a pair of Galileo global navigation system satellites, which are in elliptical orbits with an eccentricity of 0.16, were able to refine the measurement of the violation parameter to  $(0.19 \pm 2.48) \times 10^{-5}$  [7] and  $(4.5 \pm 3.1) \times 10^{-5}$  [8] by taking advantage of an  $\sim 8500$  km variation in geocentric distance over the satellites' orbits. Optical lattice clocks have also become sufficiently accurate to allow a similar measurement of  $(1.4 \pm 9.1) \times 10^{-5}$  on Earth [9]. Proposed future experiments may allow these measurements to be further refined by several orders of magnitude [10, 11].

In this paper, we present the latest results from a gravity experiment using RadioAstron (RA), the spacecraft element of the Russian-led international space very long baseline interferometry (VLBI) mission, launched in 2011 into a highly eccentric elliptical orbit with geocentric distances under 7000 km and as large as 353 000 km [12], comparable to the Earth-Moon distance. The spacecraft carried a VCH-1010 space-qualified hydrogen maser (SHM) frequency standard [13]. The mission's two ground tracking stations in Pushchino, Russia (PU) and Green Bank, WV, USA (GB) are also equipped with hydrogen masers (H-masers). First results were published in 2020 [14] based on measurements of the spacecraft's downlink signals made with the Doppler tracking equipment (Doppler frequency measurements) also used for orbit determination but were limited to  $\sigma_\varepsilon = 3\%$  by systematics most likely due to the error in compensating for the non-relativistic Doppler shift. In this follow-up paper, we report on an analysis of Doppler frequency measurements where the non-relativistic Doppler



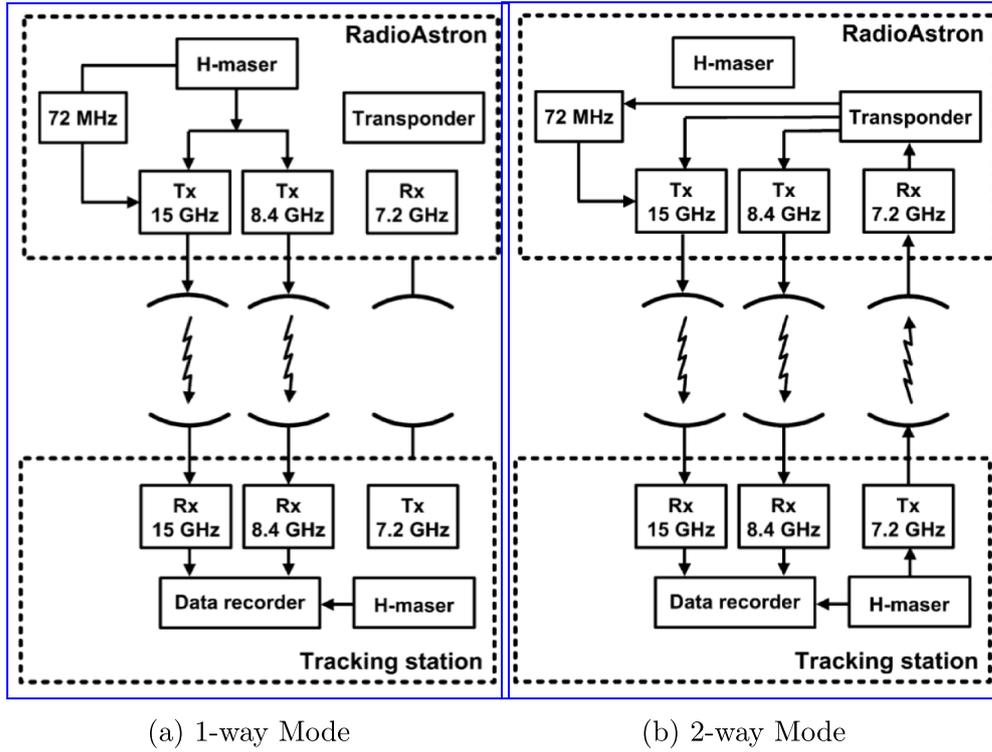
**Figure 1.** Major steps in the data analysis: 1. inputs gathered including Doppler frequency measurements, RA orbital state vectors, and Earth rotation and atmospheric models from the NRL's Tracker Component Library; 2. residuals computed using the expected frequency shifts; 3. noise in residuals studied to estimate uncertainties; and 4. weighted least squares (WLS) fit to estimate the violation parameter, spacecraft frequency offsets and their uncertainties.

shift could be suppressed resulting in a 100 times reduction in  $\sigma_\varepsilon$ . This higher accuracy was achieved through a Doppler compensation scheme (DCS) during dedicated downlink sessions in which a combination of 1-way and 2-way links similar to GP-A could be used [2, 15].

The main steps in the analysis are indicated in figure 1. In the first column are the main inputs: i. station time offsets relative to GPS time, ii. Doppler frequency measurements at the ground stations, iii. orbital state vectors for the spacecraft, iv. Earth rotation and atmospheric models from the Naval Research Laboratory's (NRL) Tracker Component Library, and v. log of the instantaneous uplink frequency transmitted to the spacecraft. The next column indicates how the inputs are initially used to compute the frequency offset of the ground station clocks and the residuals when comparing the Doppler frequency measurements to those expected from the state vectors and various models. It is at this stage that the DCS is implemented. The third column indicates that residuals are used to measure noise levels which are used to estimate uncertainties using Monte Carlo techniques. Finally, the last column indicates the use of weighted least squares to simultaneously fit the violation parameter,  $\varepsilon$ , and the frequency offset of the SHM over time. These steps will be elaborated on in detail in the remainder of this paper which is organized as follows: section 2 describes the DCS and how it was implemented with RA; section 3 describes the measurement procedures and the data acquired for the experiment; section 4 describes the data analysis; section 5 describes how  $\varepsilon$  and its uncertainty were estimated; section 6 discusses prospects for further reducing the uncertainty in  $\varepsilon$ ; and section 7 provides a brief summary and our conclusions.

## 2. Doppler Compensation Scheme

RA was equipped with two modes of onboard frequency referencing as shown in figure 2. In the 1-way mode, the observed frequency of the reference tone at 8.4 GHz and the carrier signal



**Figure 2.** RA's frequency referencing modes. (a) In the 1-way mode, the SHM provided the necessary reference to generate the 8.4 GHz reference tone and the 15 GHz carrier signal. (b) In the 2-way mode, the 7.2 GHz reference tone transmitted by the ground tracking station was used in a phase-synchronization loop to provide the onboard frequency reference. Adopted from [15].

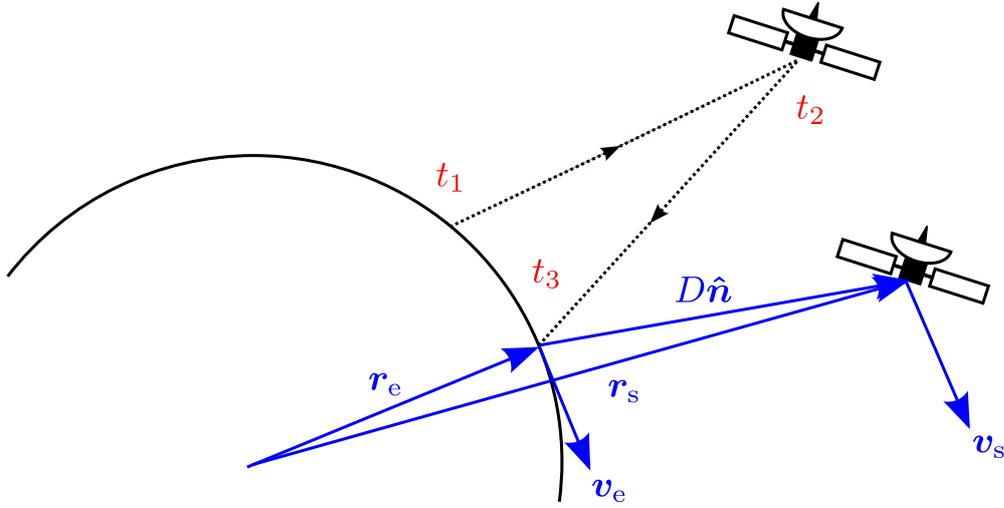
at 15 GHz ( $\nu_{1w}$ ) experienced a relative frequency shift compared to the nominal or unshifted frequency at the ground tracking station ( $\nu_0$ ):

$$y_{1w} \equiv \frac{\nu_{1w} - \nu_0}{\nu_0} \quad (2)$$

where  $y$  (lowercase) here and hereafter signifies an observed relative frequency shift,  $Y$  (uppercase) will be the corresponding expected value and  $r$ , the residual relative frequency shift or simply 'residual shift', will be the difference between them (except in figure 3 where  $r$  is a position vector). For 1-way, the expected value is given by:

$$Y_{1w} = -\frac{\dot{D}}{c} + \frac{\Delta U}{c^2} + \frac{|\mathbf{v}_e - \mathbf{v}_s|^2}{2c^2} + \frac{\mathbf{D} \cdot \mathbf{a}_s}{c^2} + Y_{\text{fine},1w} \quad (3)$$

where  $\dot{D}$  is the rate of change of the magnitude of the range vector  $\mathbf{D}$ , also called the range rate,  $\mathbf{v}_e$  and  $\mathbf{v}_s$  are respectively the velocities of the ground station and spacecraft, and  $\mathbf{a}_s$  is the acceleration of the spacecraft.  $Y_{\text{fine},1w}$  includes relativistic Doppler terms at third order or higher in  $v/c$  as well as other small effects particularly those due to the troposphere, ionosphere and phase-center motion (PCM) that are considered but omitted here for brevity. Key times and state vectors are as defined in figure 3 at  $t_3$  (observation time). Vectors at times  $t_1$  and  $t_2$ , when needed, are related to  $t_3$  using higher order corrections derived by expansion around  $t_3$ .



**Figure 3.** The key times indicated in red are: i. emission of uplink phase synchronization tone by ground station at  $t_1$ , ii. reception of uplink tone by spacecraft and emission of downlink signal at  $t_2$  and iii. reception of downlink signal by ground station at  $t_3$ . Note that ground station position ( $\mathbf{r}_e$ ) and velocity ( $\mathbf{v}_e$ ) and spacecraft position ( $\mathbf{r}_s$ ) and velocity ( $\mathbf{v}_s$ ) are all defined for simplicity at  $t_3$ , the observation time, as is the range vector ( $\mathbf{D} \equiv \mathbf{r}_s - \mathbf{r}_e = D\hat{\mathbf{n}}$ ).

The first order term, the non-relativistic Doppler shift, presents a significant challenge as near perigee it can exceed the gravitational redshift by a factor of  $10^4$ . Orbital state vectors are typically not accurate enough to estimate  $\dot{D}/c$  with sufficient precision and thus dominate the 1-way residual shift:

$$r_{1w} \equiv y_{1w} - Y_{1w}. \quad (4)$$

For example,  $\delta\dot{D} \sim 2 \text{ mm s}^{-1}$  for RA [16] which limits the measurement of  $y_{\text{grav}}$  to  $\sim 1\%$  when using only  $r_{1w}$  [14].

In the GP-A experiment, a novel technique was used to suppress the non-relativistic Doppler by taking advantage of a phase-synchronization loop locked to a reference tone uplinked from the ground tracking station to provide a second onboard frequency reference. RA's 2-way mode worked in a similar way. The observed frequency of the downlink signals in this mode ( $\nu_{2w}$ ) experienced a relative frequency shift compared to the frequency of the uplink tone from the ground tracking station ( $\nu_{\text{up}}$ ) defined as:

$$y_{2w} \equiv \frac{\nu_{2w} - F_0\nu_{\text{up}}}{F_0\nu_{\text{up}}} \quad (5)$$

where  $F_0$  is a multiplier applied by the spacecraft's electronics when generating the downlink frequencies. To second order this shift is given by:

$$Y_{2w} = -2\frac{\dot{D}}{c} + 2\frac{|\mathbf{v}_e - \mathbf{v}_s|^2}{c^2} + 2\frac{\mathbf{D} \cdot \mathbf{a}_s}{c^2} - 2\frac{\mathbf{D} \cdot \mathbf{a}_e}{c^2} + Y_{\text{fine},2w} \quad (6)$$

where  $\mathbf{a}_e$  is the acceleration of the ground station and  $Y_{\text{fine},2w}$  includes smaller effects, again omitted for brevity as in equation (3). Notice that  $y_{2w}$  does not contain  $\frac{\Delta U}{c^2}$ . We can again define a residual shift as:

$$r_{2w} \equiv y_{2w} - Y_{2w} \quad (7)$$

which is also dominated by errors in estimating  $\dot{D}/c$ . However, with a DCS, both observed relative frequency shifts are combined to form a DCS relative frequency shift or simply ‘DCS frequency shift’ defined as:

$$\Delta y \equiv y_{1w} - \frac{1}{2}y_{2w} \quad (8)$$

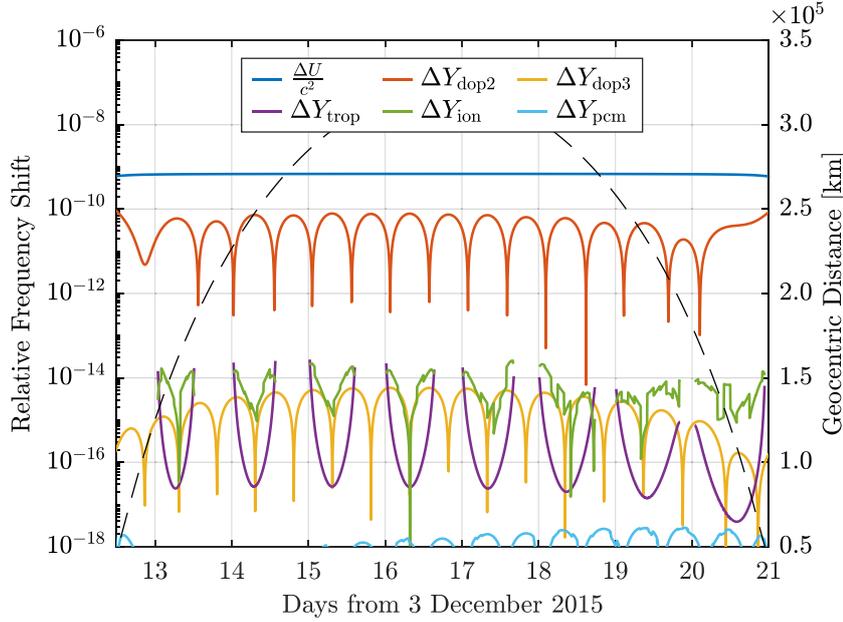
where  $\dot{D}/c$  has canceled and  $\frac{\Delta U}{c^2}$  remains as the leading effect in the expected DCS frequency shift:

$$\begin{aligned} \Delta Y &\equiv Y_{1w} - \frac{1}{2}Y_{2w} \quad (9) \\ &= \frac{\Delta U}{c^2} - \frac{|\mathbf{v}_e - \mathbf{v}_s|^2}{2c^2} + \frac{\mathbf{D} \cdot \mathbf{a}_e}{c^2} + \Delta Y_{\text{fine}}. \end{aligned}$$

The DCS also significantly diminishes the effects of the troposphere, ionosphere and PCM with only differential effects remaining in  $\Delta Y_{\text{fine}}$ . In order to achieve the desired accuracy, relative frequency shifts as small as  $10^{-15}$  must be considered, including terms at third order in  $v/c$  which are included in  $\Delta Y_{\text{fine}}$ :

$$\begin{aligned} \Delta Y_{\text{fine}} &= \Delta Y_{\text{trop}} + \Delta Y_{\text{ion}} + \Delta Y_{\text{pcm}} \quad (10) \\ &- \frac{D}{c^3} [\mathbf{D} \cdot \mathbf{j}_e + (\mathbf{v}_e - \mathbf{v}_s) \cdot \mathbf{a}_s + \nabla U_s \cdot \mathbf{v}_s] \\ &- \frac{\dot{D}}{c^3} \mathbf{D} \cdot \mathbf{a}_e + \frac{D}{c^3} [2(\mathbf{v}_e - \mathbf{v}_s) \cdot \mathbf{a}_e - \nabla U_e \cdot \mathbf{v}_e] \\ &- \frac{\dot{D}}{c^3} \left[ \Delta U - \frac{1}{2} |\mathbf{v}_e - \mathbf{v}_s|^2 \right] + O(v/c)^4 \end{aligned}$$

where  $\Delta Y_{\text{trop}}$ ,  $\Delta Y_{\text{ion}}$  and  $\Delta Y_{\text{pcm}}$  are the differential effects of the troposphere, ionosphere and PCM while  $\mathbf{j}_e$  is the jerk of the ground station due to Earth’s rotation. For RA’s orbit, the terms on the last line of equation (10) are negligible as they do not exceed  $6 \times 10^{-16}$ . Relativistic Doppler terms are estimated using orbital state vectors for the spacecraft provided by the mission (see [16]) along with state vectors for the ground stations computed using the IAU Earth rotation model as implemented in the Tracker Component Library (TCL) by the Naval Research Laboratory (see [17]). The Earth’s gravitational field is modeled using tide-free coefficients from EGM2008 (see [18]) with post-glacial rebound, polar motion, solid earth tides, oceanic tides and pole tides added following IERS conventions (see [19]). State vectors for the Moon and Sun are computed using JPL’s DE421 ephemeris (see [20]) and used to estimate their tidal effects included in  $\Delta U$ .  $\Delta Y_{\text{trop}}$  is the residual relative frequency shift due to the troposphere estimated using the VMF3 and GRAD models (see [21, 22]).  $\Delta Y_{\text{ion}}$  is the residual effect of the ionosphere, mostly due to the difference between uplink and downlink frequencies, estimated using electron densities from CDDIS (see [23]).  $\Delta Y_{\text{pcm}}$  includes the PCM effect due to the offset of the spacecraft antenna’s phase center from its center of mass (see [24]). Also added is the effect due to the significant displacement between the phase center and reference point of the GB ground tracking station. Examples of these relative frequency shifts are plotted over an orbit in figure 4.



**Figure 4.** Relative frequency shifts after implementing the DCS at GB over a particular orbit in December 2015. Dashed line is the spacecraft’s geocentric distance and is plotted against the right axes. Note  $\Delta Y_{\text{dop2}}$  are the second order relativistic Doppler effects from equation (9) and  $\Delta Y_{\text{dop3}}$  the third order Doppler terms from equation (10).

Now we introduce our main observable, the DCS residual relative frequency shift or simply ‘DCS residual shift’:

$$\begin{aligned}\Delta r &\equiv \Delta y - \Delta Y \\ &= r_{1w} - \frac{1}{2}r_{2w}.\end{aligned}\quad (11)$$

This observed relative frequency shift only contains unmodeled effects. While H-maser frequency standards may reach or even exceed a relative stability of  $10^{-15}$  over thousands of seconds, they are susceptible to a number of systematic effects that cause their frequency to drift over longer times [25]. An offset between each H-maser and Geocentric Coordinate Time (TCG) must be accounted for when comparing observed frequency shifts to prediction. Thus, we define the difference between the SHM and ground tracking station H-maser (GHM) frequency offsets, respectively  $h_s$  and  $h_e$ , as:

$$\Delta h \equiv h_s - h_e. \quad (12)$$

This difference will appear in the observed value of  $\Delta r$  along with a possible violation of  $\frac{\Delta U}{c^2}$  as follows:

$$\Delta r = \varepsilon \frac{\Delta U}{c^2} + \Delta h. \quad (13)$$

The basic approach of this experiment is thus to measure  $\Delta r$  using equation (11) and fit the resulting observations using a model function based on equation (13) to estimate  $\varepsilon$  and its uncertainty.

**Table 1.** Single mode sessions at 8.4 GHz used to measure noise levels between March 2015 and December 2017. These  $\sim 1$  hr long sessions in which RA performed space VLBI observations employed only one of the two frequency referencing modes from figure 2. Sessions up to mid-2017 were 1-way only while the remainder were 2-way only.

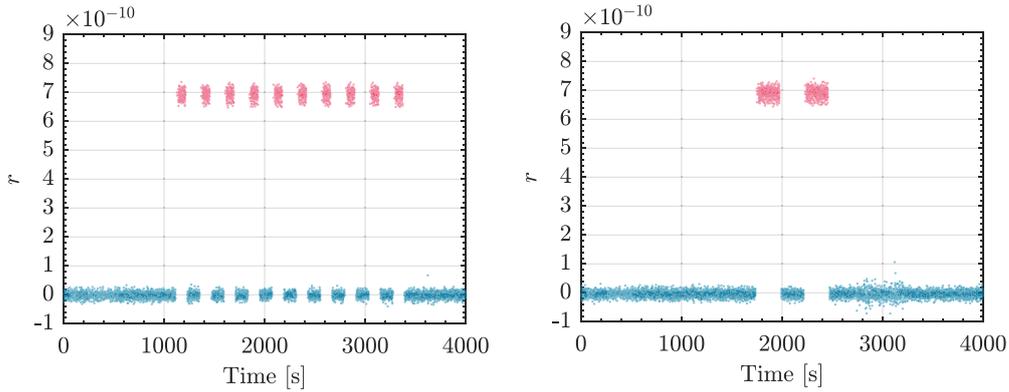
Type	Station	Time period (days from 1 January 2012)	Sessions
1-way Mode	PU	1150–1948	790
	GB	1156–1948	375
2-way Mode	PU	2107–2189	45
	GB	2108–2189	33

**Table 2.** Subset of the 199 interleaved sessions that were used in the experiment. Sessions are  $\sim 1$  hr long and have a few or many referencing mode switches. A total of 538 segments of data could be used to implement the DCS. Sessions with few segments employed a slower switching cycle resulting in segments of longer duration.

Type	Station	Time period (days from 1 January 2012)	Sessions	Usable segments	Segment duration
Few segments	PU	1206–1943	53	85	213 s
	GB	1195–1935	31	56	216 s
Many segments	PU	1392–1939	40	368	77 s
	GB	1426–1452	4	29	77 s

### 3. Measurement procedures

Although RA began scientific observations in 2012, measurements for this experiment only started in 2015 and ran until the SHM failed in mid-2017. Most RA sessions involved the real-time downlinking of astronomical space VLBI observations. During these ‘single mode’ sessions, the frequency referencing mode was held fixed and Doppler tracking equipment was used to determine the peak frequency of the spacecraft’s signals. This was done for the 8.4 GHz reference tone and the 15 GHz carrier signal by computing Fourier spectra using 80 ms of digitally sampled data with 50% overlap resulting in a 25 Hz measurement rate. A subset of these single mode sessions was used for noise analysis, as will be described later on, and are summarized in table 1. In addition, 199 ‘interleaved’ sessions, each about 1 hr long, were dedicated solely to gravitational redshift measurements. During these sessions, RA’s frequency referencing mode was switched between 1-way and 2-way modes. Of these sessions, 128 could be used for this experiment and are summarized in table 2. The time series of measurements from an interleaved session can be divided into a series of up to several dozen ‘segments’ over which a particular mode was in use. The interleaving of modes stands out clearly in figure 5 where the 1-way residual shift with the gravitational redshift added back,  $r_{1w} + \frac{\Delta U}{c^2}$ , and the 2-way residual shift,  $r_{2w}$ , have been plotted for two sample sessions. Interleaved sessions with many short segments were recorded at a variety of distances while sessions with fewer but longer segments mostly took place close to apogee where the gravitational redshift



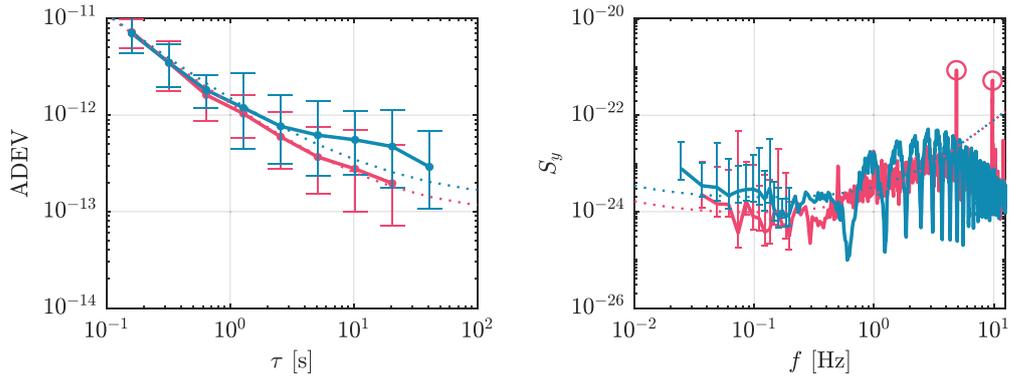
**Figure 5.** Residual shifts,  $r_{1w} + \frac{\Delta U}{c^2}$  (red) and  $r_{2w}$  (blue), at 8.4 GHz for a session on day 1451 with many mode switches (left) and another session on day 1366 near apogee with only two switches (right). Days are counted from 1 January 2012.

between the GHM and SHM is at its maximum value. The former provide sensitivity to a possible violation of the gravitational redshift while the latter largely provide sensitivity to the evolution of the frequency offset between the H-masers introduced in the next section.

## 4. Data analysis

### 4.1. Measuring the DCS Residual Shift

Values of the DCS residual shift,  $\Delta r$ , were measured by fitting polynomials to the 1-way and 2-way residual shifts,  $r_{1w}$  and  $r_{2w}$ , to interpolate simultaneous relative frequency shifts and apply the DCS. This was done using linear least squares (LLS). In the presence of only white noise, LLS would correctly estimate the interpolation error and correspondingly the uncertainty of  $\Delta r$ . However, in the presence of non-white noise, such estimates can be significantly biased. To determine the nature of the noise present, the Allan deviation (ADEV as a function of averaging time  $\tau$ ) [26] and power spectral density ( $S_y$  as a function of frequency  $f$ ) were computed using  $r_{1w}$  and  $r_{2w}$  and are shown for a typical session in figure 6. Colored noise is evident in the spectrum and can be characterized by spectral index,  $\alpha$ , where  $S_y \propto f^\alpha$ . At high frequencies, phase noises such as white phase modulation noise (WPM), with  $\alpha = 2$ , and flicker phase modulation noise (FPM), with  $\alpha = 1$ , dominate, while at low frequencies, flicker frequency modulation (FFM) noise, with  $\alpha = -1$ , begins to dominate above a floor of white frequency modulation (WFM) noise where  $S_y$  is constant. In the presence of these noises, LLS error estimates will be biased by the phase noises while the non-stationary nature of FFM will introduce a minimum error (the ‘flicker noise floor’) which cannot be overcome by fitting more data. In the presence of WFM alone, a LLS fit using all the data in a session would result in a single measurement of the DCS residual shift with the smallest possible error. However, due to the flicker noise floor, it is instead advantageous to partition the data and separately fit the frequencies in each part. The error introduced by FFM will be random from fit to fit and will thus allow the effect of the flicker noise floor to be diminished by having multiple independent values of  $\Delta r$  per session. In practice, we chose to measure  $\Delta r$  at a point in each 1-way segment where the error is expected to be smallest and for which a portion of neighboring 2-way segments could be fit to obtain a simultaneous 2-way mode frequency. A total of 538



**Figure 6.** Example of ADEV (left) and power spectra ( $S_y$ , right) at 8.4 GHz computed using residual shifts  $r_{1w}$  (red) and  $r_{2w}$  (blue) from a typical session. Error bars are 68% confidence intervals and are omitted at higher frequencies for clarity. Dotted lines correspond to the noise model, fit to the mean  $S_y$  from many sessions. Circles indicate interference spikes.

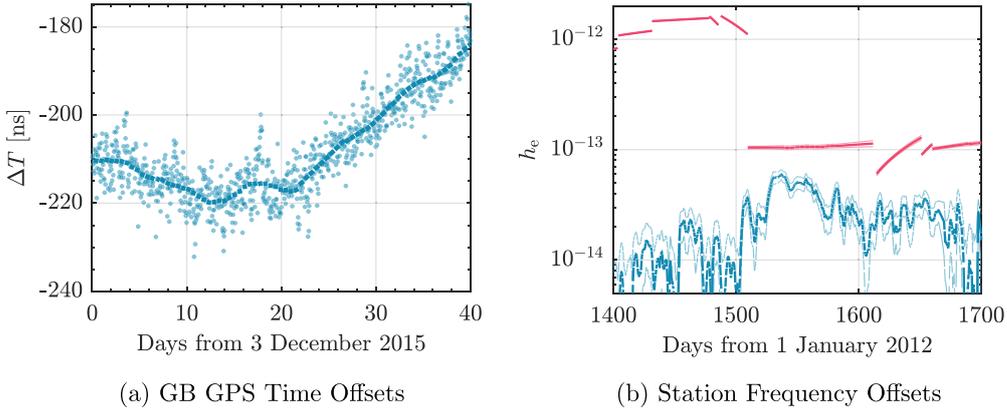
measurements of  $\Delta r$  could be made using these segments at 8.4 GHz (see table 2) with only 519 of these also being usable at 15 GHz.

#### 4.2. Measuring ground tracking station offsets

PU and GB ground tracking stations are equipped with GPS receivers that allow their local clocks to be steered such that they remain within a maximum time offset relative to GPS time. As GPS time is itself steered to follow Terrestrial Time (TT), a time series of offsets between the local clock and GPS time ( $\Delta T$ ) allow the relative frequency offset of the local clock relative to TT,  $h_e$  or station frequency offset, to be estimated using:

$$h_e(t) = \frac{d}{dt} \Delta T(t) - \frac{\mu_{\oplus}}{c^2} \frac{H_e}{r_{g,e}^2} \quad (14)$$

where the derivative is taken with respect to TT,  $\mu_{\oplus}$  is the standard gravitational parameter of Earth while  $H_e$  and  $r_{g,e}$  are the orthometric height and geoid radius at the station's location. The second term accounts for the station not being on the geoid where TT is defined. Prior to fitting, the time series is resampled into uniform hourly measurements. A portion of the resampled time series for GB is shown in figure 7(a). Note the two dominant types of noise present: (1) white noise from the GPS receiver and (2) random run phase noise due to the random walk frequency noise of the GHM. These are superimposed on the systematic drift of the maser's offset which, to first-order, is linear in frequency and therefore quadratic in the time offset. However, when an H-maser is disturbed, sudden changes in drift are possible, as seen near day 20 in figure 7(a). The first step in estimating the station frequency offset,  $h_e$ , is to divide the time series into intervals over which the H-maser was undisturbed. This was done using operator logs from GB and manually by inspecting the PU time-series looking for discontinuities. Within each interval, a single quadratic fit of the time series would be appropriate if only white noise were present. However, due to the random run, each interval must be divided into sub-intervals which are as long as possible to minimize the effect of white noise, but over which the random run does not dominate. The optimal length of these sub-intervals was determined by measuring the white noise level and using GHM specifications for



**Figure 7.** (a) Time offset measurements ( $\Delta T$ ) between GB local time and GPS time with the superimposed trend, given by non-overlapping regions from numerous sub-interval fits, showing both the effect of a systematic H-maser drift and random run noise. (b) Sample range of station frequency offsets,  $h_e$ , for PU (red) and GB (blue) with confidence intervals (higher lines) only visible when  $\sigma_{h_e}$  is sufficiently large relative to  $h_e$ .

the frequency random walk noise level. Overlapping quadratic fits of the optimal length were done and the non-overlapping regions with the lowest uncertainty from each fit were used to produce a time-series of frequency offsets, a range of which are plotted in figure 7(b). Finally, the uncertainties of the frequency offsets measured using this approach were determined using Monte Carlo simulation with randomly generated noise (see [27] for a general approach to generating clock noise) according to the determined noise levels. These uncertainties,  $\sigma_{h_e}$ , are also plotted in figure 7(b).

#### 4.3. Measuring spacecraft frequency offset and $\varepsilon$

As the station frequency offsets,  $h_e$ , can be measured independently using the method described in section 4.2, we consider it an observable along with the DCS residual shift,  $\Delta r$ , and so modify equation (13) to the following:

$$\Delta r + h_e = \varepsilon \frac{\Delta U}{c^2} + h_s. \quad (15)$$

Values of  $\Delta r + h_e$  plotted versus time appear in figure 8. The nearly linear drift of the spacecraft relative frequency offset,  $h_s$ , is apparent. Discriminating  $h_s$  from the effect of  $\varepsilon$  requires long intervals over which  $h_s$  evolves linearly and over which the range of gravitational redshifts is as large as possible. While the data cover an impressive range of redshifts,  $\Delta y_{\text{grav}} = 1.6 \times 10^{-10}$  corresponding to a distance range of 320 000 km from 26 000 km to 346 000 km, a full 2/3 of points are near apogee within 20% of the maximum value of the gravitational redshift. This has the effect of strongly correlating  $\varepsilon$  with the initial value of the spacecraft frequency offset,  $h_s$ , and requires that both be fit simultaneously.

In figure 8, we see the slope of  $h_s$  changing suddenly on two occasions. We thus divide the data into three intervals (T1, T2, T3 arranged chronologically) with boundaries at days 1464.0 and 1875.5 where days are counted starting on 1 January 2012. Sensitivity to this choice

is discussed in section 5.4. Within each interval, we assume a linear drift and so define the following model function:

$$E\left(\frac{\Delta U}{c^2}, t, \beta\right) = \varepsilon \frac{\Delta U}{c^2} + \left[ \sum_{i=1}^3 \Pi(t, i) (a_i + b_i t) \right] \quad (16)$$

where the factor  $\Pi(t, i)$  is 1 if  $t$  lies within interval  $T_i$  and 0 otherwise. The parameter vector,  $\beta$ , includes  $\varepsilon$  and the ‘ $h_s$  parameters’,  $a_i$  and  $b_i$ , that account for the drift of the SHM frequency offset. Having three  $a_i$  parameters instead of just one overall constant, allows for discontinuities at the boundary between intervals, which reduces the sensitivity of  $E\left(\frac{\Delta U}{c^2}, t, \beta\right)$  to the precise choice of boundary times. In total, this model function has 7 parameters including  $\varepsilon$  for fitting to the measurements of  $\Delta r + h_e$ .

## 5. Estimating the violation parameter, $\varepsilon$

### 5.1. Unweighted fit

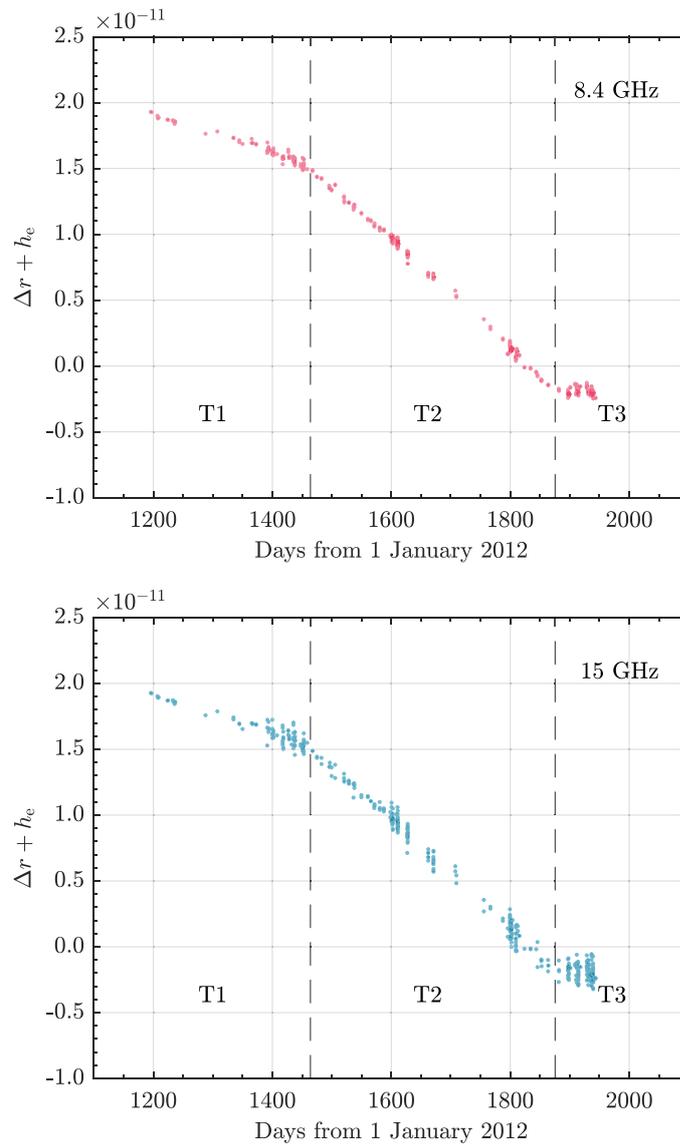
First, the model function,  $E\left(\frac{\Delta U}{c^2}, t, \beta\right)$ , was fit to the data in figure 8 using unweighted LLS. Following a preliminary fit, 7 points (1%) at 8.4 GHz and 46 points (9%) at 15 GHz were excluded using a three scaled median-absolute-deviations from the median criterion leaving  $N_{8.4} = 531$  and  $N_{15} = 473$  points. Fit residuals are shown in figure 9 and have root-mean-square (RMS) values of  $\text{RMS}_{8.4} = 2.4 \times 10^{-13}$  and  $\text{RMS}_{15} = 5.1 \times 10^{-13}$ . The values of  $\varepsilon$  from the unweighted fit are:  $\varepsilon_{\text{unw}, 8.4} = (2.2 \pm 3.4_{\text{stat}}) \times 10^{-4}$  and  $\varepsilon_{\text{unw}, 15} = (1.5 \pm 7.6_{\text{stat}}) \times 10^{-4}$ . However, given that segments are not all of the same size and noise levels varied throughout the experiment, a weighted fit incorporating the expected uncertainty in each measurement of  $\Delta r + h_e$  is more appropriate.

### 5.2. Weighted fit

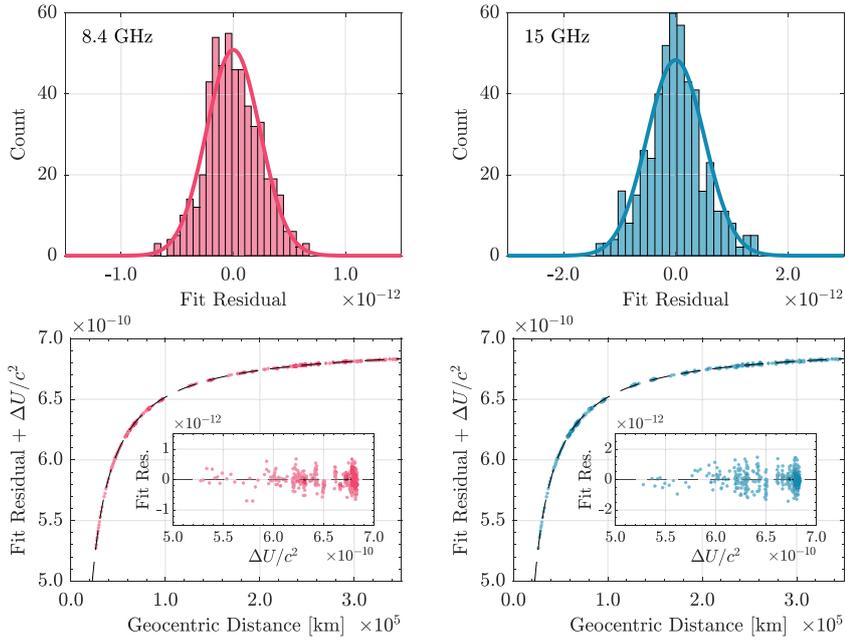
The uncertainty of the measurements,  $\sigma$ , was used as a weighting factor and, based on equations (11) and (15), includes the statistical uncertainties  $\sigma_{r_{1w}}$ ,  $\sigma_{r_{2w}}$  and  $\sigma_{h_e}$  from the fits described in sections 4.1 and 4.2. The uncertainty in the expected DCS frequency shift,  $\Delta Y$ , was ignored as it is either too small or not statistical in nature. This will be further discussed along with estimating the systematic error in section 5.4. Combining the uncertainties in quadrature gives the total uncertainty:

$$\sigma^2 = \sigma_{r_{1w}}^2 + \frac{1}{4} \sigma_{r_{2w}}^2 + \sigma_{h_e}^2. \quad (17)$$

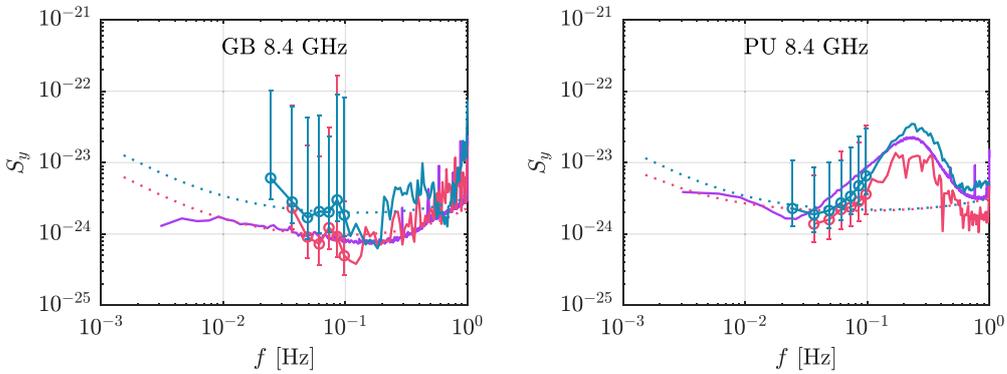
Due to the presence of non-white noise, the statistical uncertainties could not be estimated from their corresponding LLS fits. Instead, they were estimated as confidence intervals using Monte Carlo simulations of colored noise generated using models matched to observed noise power (see  $S_y$  in section 4.1). These models include four power-law noise components (WPM, FPM, WFM and FFM) and were fit to the mean  $S_y$  from many sessions (see figure 10). Only single mode sessions, all 1-way or 2-way mode, were used to compute mean  $S_y$  as their longer stretches of data are better suited to measuring non-stationary noises. The presence of additional noise at intermediate frequencies above 0.02 Hz in PU spectra after 7 May 2015 obscures the noise floor relevant at longer averaging times. Therefore, we used the FFM noise level fit to sessions between 24 February 2015 and 7 May 2015 and adjusted the WFM noise level to match that observed in the sessions coming after. As 2-way only sessions were not performed before additional noise appears in PU spectra, the 2-way noise level for PU could



**Figure 8.** The main observables in the experiment,  $\Delta r + h_e$  from equation (15), to which the model function,  $E\left(\frac{\Delta U}{c^2}, t, \beta\right)$ , was fit to obtain  $\varepsilon$  and the spacecraft frequency offset,  $h_s$ . Each of the 531 points at 8.4 GHz and 473 points at 15 GHz correspond to a 1-way segment for which the DCS could be implemented at either GB or PU. The trend is dominated by the linear drift of  $h_s$  over three intervals (T1, T2, T3) with boundaries near days 1464 and 1875 (dashed lines). The larger scatter at 15 GHz starting near day 1400 is due to PU points which have more noise. Day 1200 corresponds to 15 April 2015.



**Figure 9.** Top are distributions of residuals from fitting the model function,  $E\left(\frac{\Delta U}{c^2}, t, \beta\right)$ , to  $\Delta r + h_e$  with overlaid normal distributions (solid line) centered at zero with standard deviations corresponding to the RMS of the residuals. Bottom are the same fit residuals but with gravitational redshift added and plotted versus distance. Insets are the fit residuals versus gravitational redshift.



**Figure 10.** Mean power spectra of the residual shifts  $r_{1w}$  (red) and  $r_{2w}$  (blue) at 8.4 GHz computed using segments from interleaved sessions. Left is the mean of GB sessions and right is that of PU sessions between 7 May 2015 and 3 June 2016. Dotted lines are corresponding noise models. The mean spectra from 1-way only sessions are plotted in magenta. Some error bars are omitted for legibility. Note, the presence of additional noise above 0.02 Hz in PU spectra which is discussed in the text.

not be fit. Instead, FFM noise power was assumed to be twice that in 1-way, which is approximately the same ratio seen between 2-way and 1-way power at GB. Monte Carlo simulation showed that FFM and, to a lesser extent, WFM dominate the error. Care was taken to generate FFM noise with the appropriate characteristics using an ARFIMA (1, 0.5, 0) stochastic model following the approach of [28]. Using 1000 simulations per session, the uncertainties in the 1-way and 2-way residual shifts,  $\sigma_{r_{1w}}$  and  $\sigma_{r_{2w}}$ , were found to respectively contribute 74% and 26%, on average, to the total uncertainty while the contribution of the uncertainty in the station frequency offset,  $\sigma_{h_e}$ , is negligible. The mean total uncertainties across all segments are  $\bar{\sigma}_{8.4} = 2.1 \times 10^{-13}$  and  $\bar{\sigma}_{15} = 2.5 \times 10^{-13}$ .

Using the total uncertainties as weighting factors when fitting equation (16), the following values of  $\varepsilon$  were obtained:  $\varepsilon_{8.4} = (2.1 \pm 3.3_{\text{stat}}) \times 10^{-4}$  and  $\varepsilon_{15} = (0.7 \pm 7.6_{\text{stat}}) \times 10^{-4}$ . These results are very similar to those from the unweighted fits. The weighted fits have chi-square per degree of freedom of  $\chi_{\nu,8.4}^2 = 1.1$  and  $\chi_{\nu,15}^2 = 4.4$ . The former suggests that the weighting factors determined using Monte Carlo methods account for nearly all the scatter in  $\Delta r + h_e$  at 8.4 GHz. In contrast, the larger  $\chi_{\nu}^2$  at 15 GHz, which is evident from  $\text{RMS}_{15}$  being significantly larger than  $\bar{\sigma}_{15}$ , implies that the estimated uncertainty at 15 GHz is too low. This is not surprising given that we could not directly fit noise levels in this band and instead resorted to using the lower noise levels from 8.4 GHz. The statistical uncertainties of  $\varepsilon$  have been adjusted so they correspond to a  $\chi_{\nu}^2$  of unity.

### 5.3. Check on statistical error

We can compare  $\sigma_{\varepsilon}^{\text{stat}}$  to what is expected given the mean variation in gravitational redshift ( $\overline{\Delta y}_{\text{grav}} \sim 2.6 \times 10^{-11}$ ), the RMS of the fit residuals, the number of points given in section 5.1 and the mean correlation between  $\varepsilon$  and the  $h_s$  parameters ( $\bar{\rho}_{8.4} = -21\%$ ) as follows:

$$\hat{\sigma}_{\varepsilon,8.4} \sim \frac{1}{\overline{\Delta y}_{\text{grav}}} \sqrt{\frac{\text{RMS}_{8.4}^2}{N_{8.4}} (1 + 2\bar{\rho}_{8.4})} = 3.1 \times 10^{-4}. \quad (18)$$

Using this value, we find  $\sigma_{\varepsilon}^{\text{stat}}/\hat{\sigma}_{\varepsilon} = 1.07$  at 8.4 GHz which is fairly close to unity. For 15 GHz, this check is not useful since noise levels could not be directly measured. Thus we conclude, at least at 8.4 GHz, that the statistical uncertainty is a reliable estimate.

### 5.4. Systematic error

We tested our analysis technique for a bias when measuring  $\varepsilon$ , particularly towards  $\varepsilon = 0$ . By assuming a non-zero violation in the presence of simulated noise we confirmed that our estimate of  $\varepsilon$  is unbiased. Further, as mentioned in section 4.3, these tests confirmed that  $h_s$  cannot be measured independently from  $\varepsilon$  using the same data set without a possible violation being suppressed. This shows that our overall approach of fitting  $\varepsilon$  and  $h_s$  simultaneously is necessary.

To determine the contributions to systematic uncertainty, we considered three effects. First, we studied the effect of the interference spikes that appear in 1-way mode spectra (see figure 6). A shift,  $\Delta\varepsilon_{\text{filt}}$ , resulted when passing the DCS residual shift through a 3 Hz Butterworth low-pass filter of order 8 to remove the spikes prior to fitting. As we could not ascertain which result is more likely to be correct, we conservatively include these shifts,  $\Delta\varepsilon_{\text{filt},8.4} = 3 \times 10^{-5}$  and  $\Delta\varepsilon_{\text{filt},15} = 4.2 \times 10^{-4}$ , in the error.

Second, the uncertainty of the specific boundaries between intervals T1, T2, and T3 (see figure 8) were studied. Alternate boundaries between T1 and T2 as well as between T2 and T3

**Table 3.** Sources of systematic error: (1) variation in  $\varepsilon$  due to filtering interference, (2) variation in  $\varepsilon$  due to varying SHM frequency offset interval boundaries, and (3) error due to the uncertainty in the expected DCS frequency shift,  $\sigma_{\Delta Y}$ , estimated by dividing that uncertainty by  $\overline{\Delta y}_{\text{grav}} = 2.6 \times 10^{-11}$ .

Source	Contributions $\times 10^{-4}$	
	8.4 GHz	15 GHz
1. $\Delta\varepsilon_{\text{filt}}$	0.3	4.2
2. $\Delta\varepsilon_{\text{bound}}$	0.2	0.5
3. $\sigma_{\Delta Y} < 10^{-15}$	<0.4	<0.4
$\sigma_{\varepsilon}^{\text{sys}}$	0.5	4.3

were tried corresponding to where the  $1\sigma$  confidence intervals of the fits on either side of the boundary meet, namely at days 1470 and 1890 respectively. The differences in the fit values,  $\Delta\varepsilon_{\text{bound},8.4} = 2 \times 10^{-5}$  and  $\Delta\varepsilon_{\text{bound},15} = 5 \times 10^{-5}$ , are also added to the error.

Third, the terms in equation (9) larger than the uncertainty in the DCS frequency shift are  $\frac{\Delta U}{c^2}$ , the second order relativistic Doppler terms and the station frequency offset. As described in section 4.2, the uncertainty in the station frequency offset was estimated and is included in the weights and, therefore, the statistical error. For the other terms, the main sources of error are the uncertainty in the spacecraft state vectors ( $\delta r \sim 200$  m and  $\delta v \sim 2$  mm s $^{-1}$  [16]), and the position of PU's reference point ( $\delta r < 10$  m). The error introduced in the expected DCS frequency shift by these uncertainties does not exceed  $1 \times 10^{-15}$ , even for the closest perigee session for which the error would be the largest.

Systematic thermal and magnetic effects on the GHMs contribute to the station frequency offset and thus are taken into account. Ground testing of the SHM showed a thermal sensitivity of  $\Delta f/f = \pm 5 \times 10^{-15}/^{\circ}\text{C}$  and magnetic field sensitivity of  $\Delta f/f = \pm 2 \times 10^{-14}/\text{G}$ . During observing sessions, the thermal management system on board RA maintained the SHM temperature with an accuracy of  $1^{\circ}\text{C}$ . The resulting random frequency shift is therefore expected to be much smaller than the estimated uncertainty in the residual shift,  $\bar{\sigma}_{8.4}$ , and can be neglected. Similarly, at the distances of RA's orbit, the Earth's magnetic field is sufficiently weak ( $\ll 0.1$  G) that effects due to its variation are also negligible.

The systematic error contributions, also listed in table 3, are added in quadrature to yield total systematic uncertainties of  $\sigma_{\varepsilon,8.4}^{\text{sys}} = 0.5 \times 10^{-4}$  and  $\sigma_{\varepsilon,15}^{\text{sys}} = 4.3 \times 10^{-4}$ .

### 5.5. Sensitivity study

In table 4, we summarize our results using different subsets of the data with and without weighting. Measuring  $\varepsilon$  using only T2, the longest interval, results in a correlation of 72% between  $\varepsilon$  and the constant parameter in the frequency offset. By combining all three frequency offset intervals in a single fit, not only is the uncertainty in  $\varepsilon$  reduced, but so is the correlation falling by 6% in the case of T2 but by about 24% in the case of the other two intervals. The majority of sessions with numerous switches, performed nearer Earth, were recorded at PU. The addition of GB data makes essentially no difference at 8.4 GHz, indicating that the GB data are consistent with PU, but also that the lower noise levels at PU drive the result. The reverse is true at 15 GHz, where the additional noise at PU is partially offset by the inclusion of GB data. All the results, including those from unweighted fits, are broadly consistent with each other within the uncertainties. Furthermore, we find in all cases  $\varepsilon$  is consistent with zero within  $1.1\sigma_{\varepsilon}^{\text{stat}}$ .

**Table 4.** The result of fitting  $\varepsilon$  using different portions of the data set with  $N$  points both using and not using weights derived from simulation.

Station	Interval	Weighted	8.4 GHz		15 GHz	
			$N$	$\varepsilon \pm \sigma_{\varepsilon}^{\text{stat}} \times 10^{-4}$	$N$	$\varepsilon \pm \sigma_{\varepsilon}^{\text{stat}} \times 10^{-4}$
PU, GB	T1, T2, T3	Yes	531	$2.1 \pm 3.3$	474	$0.7 \pm 7.6$
PU, GB	T2	Yes	268	$-3.9 \pm 3.6$	232	$-1.3 \pm 8.6$
PU	T1, T2, T3	Yes	447	$0.9 \pm 3.3$	405	$-1.0 \pm 9.7$
PU, GB	T1, T2, T3	No	531	$2.2 \pm 3.4$	473	$1.5 \pm 7.6$
PU, GB	T2	No	268	$-2.9 \pm 3.8$	233	$-1.5 \pm 8.9$
PU	T1, T2, T3	No	447	$1.6 \pm 3.4$	404	$3.0 \pm 9.7$

### 5.6. Final results

Combining our estimates for  $\varepsilon$  from the weighted fit and its uncertainty we arrive at the following results:  $\varepsilon_{8.4} = (2.1 \pm 3.3_{\text{stat}} \pm 0.5_{\text{sys}}) \times 10^{-4}$  and  $\varepsilon_{15} = (0.7 \pm 7.6_{\text{stat}} \pm 4.3_{\text{sys}}) \times 10^{-4}$ . For a final result, we considered combining the measurements from the two frequency bands. However, the noise in the two bands appears strongly correlated during GB sessions and at least partially correlated during many of the PU sessions. This implies that the measurements at 8.4 GHz and 15 GHz cannot be considered statistically independent. The result at 8.4 GHz is favored since its  $\chi_{\nu}^2$  being close to unity confirms that our weighting scheme derived from colored noise simulations is reliable. Using the 8.4 GHz result and combining its statistical and systematic uncertainties in quadrature, we arrive at a final estimate for the violation parameter:

$$\varepsilon = (2.1 \pm 3.3) \times 10^{-4}.$$

## 6. Discussion

Tests of the EEP are considered an important, if not essential, probe of metric theories of gravity [29], with measuring the gravitational redshift being one of the classical tests of general relativity. Our measurements were made with the space VLBI RA mission which was not primarily designed for a gravitational redshift test. In particular, we were limited by the lack of simultaneous downlink signals in the 1-way and 2-way referencing modes and the limited observation time allocated to the experiment. The flicker noise floor of the online frequency measurements by the Doppler tracking equipment is an order of magnitude higher than that of GP-A and almost 35 times what was determined in the laboratory for the SHM prior to launch. Nevertheless, the mission allowed an accurate measurement of the gravitational redshift from near Earth to almost the distance of the Moon where it asymptotically approaches its maximum relative to Earth's surface (see bottom plots in figure 9).

In addition to the Doppler tracking measurements, time-domain recordings of the spacecraft's signal at 8.4 GHz were also made at PU and GB. These permit measuring the frequency evolution of the spacecraft's reference tone with improved offline processing techniques, such as those developed for spacecraft tracking by the Joint Institute for VLBI ERIC [30], which may allow the observed flicker noise floor to be overcome. Once frequency measurements have been made and their uncertainties estimated, the model and analysis described herein may be applied to determine  $\varepsilon$  with improved accuracy. Preliminary work on applying these offline techniques are discussed in [31] wherein it is estimated that  $\sigma_{\varepsilon} \sim 10^{-5}$  may be attainable.

Recordings of RA's signal were also made at other ground radio telescopes for which a partial DCS is possible. Including these in the final solution may allow statistical uncertainties to be further reduced.

For a future space VLBI mission in a highly eccentric orbit, we envision a setup allowing simultaneous recordings of 1-way and 2-way referenced signals in parallel to all downlinks of VLBI recordings. Over a three year period, a mission similar to RA would have  $\sim 2500$  sessions, a factor of 20 increase over the number used in this experiment. Simultaneous recordings would allow a session to be divided into 40 or more segments, a  $10 \times$  increase over our average number of segments per session. Further, an orbit with a lower inclination or a tracking station in the southern hemisphere, would allow sessions much closer to perigee increasing the variation in the gravitational redshift by a factor of 10 or more. Taken together, such a mission could improve the sensitivity of measuring  $\varepsilon$  to  $\sim 10^{-7}$ .

## 7. Summary and conclusions

In this paper we have described a test of the EEP and measurement of the gravitational redshift including:

- (i) details on Doppler-tracking frequency measurements at the PU and GB stations with RA at distances ranging from 25 000 km to the distance of the Moon,
- (ii) the implementation of a DCS, similar to GP-A, achieved by alternating RA's communication system between different frequency referencing modes,
- (iii) the model required to predict relative frequency shifts as small as  $10^{-15}$ ,
- (iv) measurements of the frequency offset of GB and PU H-masers relative to coordinate time with an accuracy exceeding  $10^{-14}$  throughout most of the experiment,
- (v) a method for measuring  $\varepsilon$  and the SHM frequency offset relative to TCG,
- (vi) using Monte Carlo simulation to determine the correct weighting of the data using GB and PU noise levels, where in both cases FFM noise was found to dominate,
- (vii) fit of 8.4 GHz data with  $\chi^2_\nu = 1.1$  validating the weights derived from simulation,
- (viii) a final result of  $\varepsilon = (2.1 \pm 3.3) \times 10^{-4}$  using the spacecraft tone at 8.4 GHz,
- (ix) the possibility to significantly improve the measurement sensitivity with existing time-domain data, and
- (x) the prospect of increasing the sensitivity further with future space VLBI missions to  $\sim 10^{-7}$ .

## Data availability statement

The data cannot be made publicly available upon publication due to legal restrictions preventing unrestricted public distribution. The data that support the findings of this study are available upon reasonable request from the authors.

## Acknowledgments

We thank D A Litvinov for his contribution to this experiment. The authors are also grateful to the anonymous reviewers for their useful comments and suggestions. The RadioAstron project is led by the Astro Space Center of the Lebedev Physical Institute of the Russian Academy of Sciences and the Lavochkin Scientific and Production Association under a contract with

the Russian Federal Space Agency, in collaboration with partner organizations in Russia and other countries. This paper was supported in part by the Russian Academy of Science Program KP19-270, ‘The study of the Universe origin and evolution using the methods of earth-based observations and space research.’ N B, M F B and N V N were supported by the National Sciences and Engineering Research Council of Canada.

## ORCID iDs

N V Nunes  <https://orcid.org/0000-0001-9691-1463>  
 G D Manucharyan  <https://orcid.org/0000-0003-0464-0040>  
 S M Popov  <https://orcid.org/0000-0001-5697-3045>  
 V N Rudenko  <https://orcid.org/0000-0002-7178-416X>  
 L I Gurvits  <https://orcid.org/0000-0002-0694-2459>  
 G Cimò  <https://orcid.org/0000-0002-1167-7565>  
 G Molera Calvés  <https://orcid.org/0000-0001-8819-0651>  
 M V Zakhvatkin  <https://orcid.org/0000-0002-0900-5786>  
 M F Bietenholz  <https://orcid.org/0000-0002-0592-4152>

## References

- [1] Will C M 2014 *Living Rev. Relativ.* **17** 1–117
- [2] Biriukov A V, Kauts V L, Kulagin V V, Litvinov D A and Rudenko V N 2014 *Astron. Rep.* **58** 783–95
- [3] Pound R V and Rebka G A 1960 *Phys. Rev. Lett.* **4** 337–41
- [4] Pound R V and Snider J L 1964 *Phys. Rev. Lett.* **13** 539–40
- [5] Vessot R F C *et al* 1980 *Phys. Rev. Lett.* **45** 2081–4
- [6] Vessot R F C 1989 *Adv. Space Res.* **9** 21–28
- [7] Delva P *et al* 2018 *Phys. Rev. Lett.* **121** 231101
- [8] Herrmann S *et al* 2018 *Phys. Rev. Lett.* **121** 231102
- [9] Takamoto M, Ushijima I, Ohmae N, Yahagi T, Kokado K, Shinkai H and Katori H 2020 *Nat. Photon.* **14** 411–5
- [10] Savalle E, Guerlin C, Delva P, Meynadier F, le Poncin-Lafitte C and Wolf P 2019 *Class. Quantum Grav.* **36** 245004
- [11] Litvinov D and Pilipenko S 2021 *Class. Quantum Grav.* **38** 135010
- [12] Kardashev N S *et al* 2013 *Astron. Rep.* **57** 153–94
- [13] Alexandrov Y A *et al* 2012 *Sol. Syst. Res.* **46** 458–65
- [14] Nunes N V, Bartel N, Bietenholz M F, Zakhvatkin M V, Litvinov D A, Rudenko V N, Gurvits L I, Granato G and Dirx D 2020 *Adv. Space Res.* **65** 790–7
- [15] Litvinov D A *et al* 2018 *Phys. Lett. A* **382** 2192–8
- [16] Zakhvatkin M V *et al* 2020 *Adv. Space Res.* **65** 798–812
- [17] Crouse D F 2017 *IEEE Aerosp. Electron. Syst. Mag.* **32** 18–27
- [18] Pavlis N K, Holmes S A, Kenyon S C and Factor J K 2012 *J. Geophys. Res. Solid Earth* **117** B04406
- [19] Petit G and Luzum B 2010 IERS Technical Note **36**
- [20] Folkner W M, Williams J G and Boggs D H 2009 *IPN Prog. Rep.* **42** 1–34
- [21] Landskron D and Böhm J 2018 *J. Geod.* **92** 349–60
- [22] Landskron D and Böhm J 2018 *J. Geod.* **92** 1387–99
- [23] Noll C E 2010 *Adv. Space Res.* **45** 1421–40
- [24] Litvinov D A, Nunes N V, Filetkin A I, Bartel N, Gurvits L I, Molera Calves G, Rudenko V N and Zakhvatkin M V 2021 *Adv. Space Res.* **68** 4274–91
- [25] Vessot R F C 2005 *Metrologia* **42** S80
- [26] Barnes J A *et al* 1971 *IEEE Trans. Instrum. Meas.* **IM-20** 105–20

- [27] Zucca C and Tavella P 2005 *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **52** 289–96
- [28] Xu C 2019 *Astron. J.* **157** 127
- [29] Damour T 2012 *Class. Quantum Grav.* **29** 184001
- [30] Molera Calves G, Pogrebenko S V, Wagner J F, Cimo G, Gurvits L I, Bocanegra-Bahamon T M, Duev D A and Nunes N V 2021 *Publ. Astron. Soc. Aust.* **38** e065
- [31] Belonenko A V, Gusev A V and Rudenko V N 2021 *Gravit. Cosmol.* **27** 383–91