Rapid detector of dark matter transit with GNSS clock data

Rory Lipkis, *Stanford University, Department of Aeronautics and Astronautics*

BIOGRAPHY

Rory Lipkis is a graduate student at Stanford University.

ABSTRACT

Certain models of dark matter predict interaction with atomic clocks. The GNSS system allows the clock biases of satellites to be continually monitored and analyzed. A realtime detection tool could monitor incoming transmissions and search for possible dark matter signatures in the form of propagating wavefronts, information which could inform more in-depth radio astronomy observation. The tool is successfully validated with hardware-in-the-loop testing, and the accuracy is shown to be highly dependent on the sampling time.

INTRODUCTION

In the standard Lambda-CDM cosmological model, only five percent of the universe consists of ordinary baryonic matter. Since the late 19th century, it has been noted that the observed rotational speed of galaxies is significantly higher than expected, suggesting that they contain additional "dark" matter that gravitationally binds them against the centrifugal forces. The development of particle physics in the 20th century permitted theories of particles which do not interact with photons, and the scientific community began to propose candidate dark matter particles. However, by its very nature, any candidate particle will likely be difficult to detect.

One particular dark matter theory posits a hypothetical mechanism known as dilatonic interaction. The dilaton field couples to standard model fields in such a way as to locally alter the fine-structure constant, which determines atomic transition rates. Arvanitaki, Huang, and Tilburg (2015) suggest that this theory could be effectively tested with atomic clocks, whose accuracy depends on counting atomic transitions, and calculate that dark matter interaction could generate clock biases on the order of a nanosecond.

Absent its ground-based segment, the GNSS system represents a massive distributed network of atomic clocks with sub-nanosecond precision, which continually transmit their biases. Dark matter is believed to be weakly interacting, which explains the apparent lack of dark bodies (planetary- and stellar-scale structures would generate r^{-2} gravitational potentials that are not observed). Assuming that it only forms galactic-scale (but nonetheless distinct) structures, any transit of the GNSS system by DM would take the form of a moving wavefront. Time-separated correlations in GNSS clock biases could reveal the signature of such a wavefront.

APPROACH AND RELATED WORK

In the last few years, a multi-affiliate research group (Roberts, et al.) has unsuccessfully searched for dilaton interactions in GPS data. Although this project has involved the processing of large quantities of data, its scope is fairly limited. The researchers have searched for a particular dark matter signature, that of dilatonic interaction due to a wavefront propagating at a single velocity, approaching in the opposite direction of Earth's motion through the galaxy. This is an entirely reasonable approach, given the specific theoretic motivation, the tremendous quantities of data, and the computational complexity of the search.

The more data-driven approach of this paper involves the consideration of an arbitrary wavefront direction and velocity, as well as a generic interaction with atomic clocks. This approach is less physically motivated than previous searches, but could be viewed as a more general probe of undiscovered cosmological phenomena. Due to the dramatic increase in computational burden caused by the relaxing of the underlying model and the resulting increase in the dimensionality of the parameter space, it is much more difficult (and perhaps less scientifically useful) to analyze historical GNSS data. Rather, this paper describes the development and testing of a realtime dark matter transit detector. In the event of a candidate transit, this program can rapidly instruct radio telescope operators where to observe, as well as provide experimentally measured parameters describing the dark matter wavefront and interaction for the purpose of theoretical reconstructions. This

represents a “responsible data science” approach to physics — one whose primary goal is to inform and facilitate more principled research.

PROBLEM STATEMENT

Mathematical formulation

An infinite flat wavefront with normal vector $\mathbf{n}$ and containing the point $\mathbf{r}_0$ can be described implicitly by the equality

$$(\mathbf{r} - \mathbf{r_0}) \cdot \mathbf{n} = 0.$$

If the wavefront moves at constant velocity $\mathbf{v}$ and passes the origin at time t_0, the expression becomes

$$\left(\mathbf{r} - \mathbf{v}(t - t_0)\right) \cdot \frac{\mathbf{v}}{|\mathbf{v}|} = 0,$$

which is solved for time to yield

$$t = t_0 + \frac{\mathbf{r} \cdot \mathbf{v}}{|\mathbf{v}|^2}.$$

This expression gives the time of arrival at any point $\mathbf{r}$ for a moving wavefront fully described by the parameters $\mathbf{v}$ and t_0. For

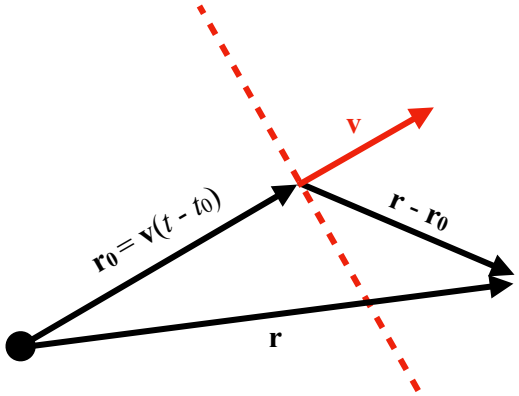

Fig 1. Geometry of wavefront propagation.

a set of *events*, each consisting of a time-position pair, the parameters can be estimated by a nonlinear least squares solution. The Levenberg–Marquardt algorithm, efficiently implemented by the scipy Python package, provides a robust method of solving such problems by computing the Jacobian

$$\frac{\partial t}{\partial(t_0, \mathbf{v})} = \left(1, \quad \frac{\mathbf{r}}{|\mathbf{v}|^2} - 2\left(\frac{\mathbf{r}}{|\mathbf{v}|^2} \cdot \hat{\mathbf{v}}\right)\hat{\mathbf{v}}\right).$$

Characterization of degeneracies

This calculation is used to inform a gradient descent over the objective function of the sum of the squares of the residuals. The nonlinear complication of considering time dependence is necessary; it is not sufficient to fit a line through a set of 3D points. Although this works for nearly collinear events (for which a reduced number of wavefront parameters can be determined), one can imagine a roughly spherical distribution of events, for which any purely spatial analysis would fail to distinguish a particular wavefront direction. It is only the information about time dependence that allows the problem to be well determined.

Due to the relatively small number of events (limited by the number of GNSS satellites), several degeneracies can still manifest in the system under certain conditions. One such issue occurs when the events are insufficiently distributed in three dimensions. In the extreme case of a regularly-spaced collinear sequence of events, the most natural wavefront fitting has a velocity vector parallel to the line, with some magnitude v that appropriate satisfies the time components of the events. However, if the wavefront is tilted by some angle $\theta < \pi/2$, the time interval between events can be preserved with the transformation $v \mapsto v \cos\theta$, since the fundamental time ordering of the events is not changed. Therefore, an entire half-sphere of propagation directions will yield valid solutions. Similarly, a planar or nearly planar set of events yields a family of valid parameters.

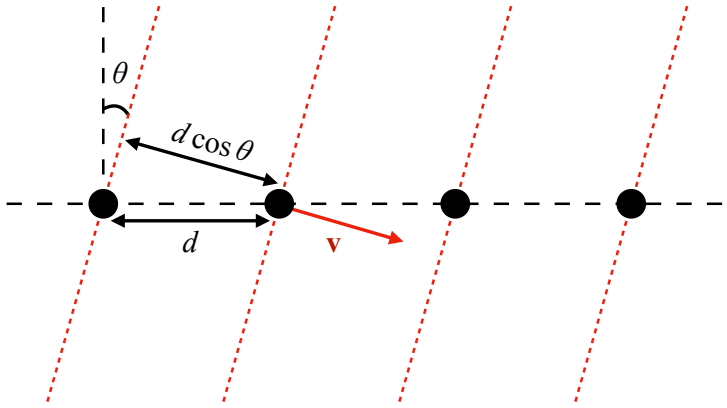

Fig 2. Example of degeneracy due to collinearity of events.

For the most part, these degeneracies manifest in pathological cases involving simple geometries and regular spacing, neither of which exists in full GNSS constellations (although in cases of limited GNSS coverage, simpler geometries could be temporarily recovered). However, limited forms of the degeneracy still occur in any configuration for which a slight rotation of the wavefront velocity can be offset by an appropriate modification of its magnitude. This is closely related to the condition that the rotation does not change the time ordering of the events (the order in which the wavefront encounters the event positions). For any two events for which $t_1 > t_2$, the corresponding velocity condition is seen to be

$$(\mathbf{r}_1 - \mathbf{r}_2) \cdot \mathbf{v} > 0.$$

If the velocity is rotated by some matrix R, the new expression can be bounded by the Cauchy-Schwarz inequality as

$$-|\mathbf{r}_1 - \mathbf{r}_2||\mathbf{v}| < (\mathbf{r}_1 - \mathbf{r}_2) \cdot R\mathbf{v} < |\mathbf{r}_1 - \mathbf{r}_2||\mathbf{v}|.$$

We can see that although R can change the expression sign, thus violating the ordering condition, $|\mathbf{v}|$ cannot. In other words, if a rotation is small, a velocity magnitude can be found to recapture the original time difference and produce and equally valid fit. If the rotation is too large, no velocity magnitude produces a valid fit. This phenomenon characterizes the limited degeneracies seen in practice, which result in small but detectable uncertainties in the wavefront fit.

APPROACH

Overview

The general approach of the program design emphasizes rapidity, concurrency, and stream processing. Since the GNSS clock biases are continually transmitted to Earth, the detector needs to be able to monitor dark matter transit signatures in realtime, and analyze the measurements fast enough to send information to radio telescopes before the transit is over. Because the wavefront velocities would likely be galactic-scale (multiple hundreds of kilometers per second), there is little margin for costly memory management: only useful data can be generated and retained. The realtime capabilities of the program are tested with a hardware-in-the-loop simulation, which sends the detector a data stream of lightly processed GPS transmissions.

Simulator

The simulator uses a GNSS ephemeris file to simulate transmissions and realistic orbits for 32 satellites. To allow for hardware-in-the-loop validation, the simulator generates output only when called by the detector. This allows for increased program flexibility and vastly improved memory characteristics. The parameters of the wavefront and its interaction with atomic clocks are specified, and its effect is recorded in the clock bias transmissions in addition to random noise: atomic clocks are assumed to experience a small level of fluctuation over a range of frequencies. The condition for a satellite's interaction with the wavefront is found from the time-of-arrive expression derived above. A satellite is behind the wavefront (and has thus interacted with in) if

$$\mathbf{r} \cdot \mathbf{v} \leq |\mathbf{v}|^2 (t - t_0).$$

Detector

The detector accepts a data stream consisting of satellite positions (in an Earth-centered inertial frame) and clock biases and returns a list of probable events. In the first step of the pipeline, the time difference between the clock bias for each satellite is computed: this difference determines the detection of an event. Special care must be taken not to process false positives from the clock noise — since there are a maximum of 32 events, a single invalid event can completely ruin the fit. The time differencing filters out any contribution from frequencies much below the detector sample rate, and higher frequency fluctuation is explicitly limited by a detector sensitivity parameter, which is set to a level that excludes low-amplitude noise.

If an event is considered probable, the true position must be estimated. Because of the finite sample rate of the detector, the precise time of the event is not known. The event position is estimated to be at the halfway point between the adjacent time steps. In practice, this places a hard lower limit on the accuracy of the time-of-transit (t_0) fit, which differs from the true value by roughly half the sampling period.

Analyzer

As events are detected, they are accumulated in the analyzer. When the detection terminates, the program performs a nonlinear least squares fit, which outputs the optimal wavefront parameters (t_0, $\mathbf{v}$) and the associated 4×4 covariance matrix. An important guard against false positives is implemented here, in the form of resampling. The residuals magnitudes are computed for the fit and the outliers (those greater than three times the average) are removed from the event list. A single false event significantly degrades the quality of the fit, and the difference manifests clearly in the magnitude of the corresponding residuals. In cases where the number false positives overwhelm the actual events, this resampling is simply repeated iteratively until it produces no change.

There is one remaining case of false positive that must be caught: instances in which the false event occurs close to the wavefront in such a way to slightly corrupt the final estimate without manifesting as a particularly high residual. Although these errors are unlikely, they can result in increased uncertainty with seemingly undetectable cause. Such cases are handled with jackknife resampling. If the final parameter fit is unacceptably poor after outlier removal, the list of events is exhaustively traversed, momentarily removing each element and recomputing the fit. If the fit cannot be improved, the resampling is repeated, removing every combination of two elements. Since the kth pass of the jackknife algorithm requires $O(n^k)$ complexity, the filter terminates after two passes. If the fit is still poor, the detection can be considered a failure. In practice, however, such resampling is rarely necessary.

Ultimately, the desired information is not the computed components of the velocity vector but the azimuth and elevation of the wavefront origin on the celestial sphere, the direction of greatest interest to dark matter astronomers. Accordingly, the velocity vector and its uncertainty is calculated in spherical coordinates. Although the vector transformation is straightforward, the standard transformation of uncertainty relies on linearization formulas that apply poorly to the singularities and discontinuities of the spherical coordinate system. Instead, we employ a Monte-Carlo simulation, drawing a large number of samples from the multivariate distribution, transforming them to spherical coordinates, and computing statistics on the resulting population.

RESULTS

The program performs efficiently and accurately, detecting and analyzing galactic-speed wavefronts with velocity accuracy on the order of meters per second and angular accuracy on the order of arcseconds. On a 2014 MacBook Pro with a 2.6 GHz Dual-Core Intel Core i5 processor, the program proceeds at an average of 355 samples per second. In other words, with non-simulated data, it could support support a minimum sampling time of 2.8 milliseconds. However, the realities of the current GNSS system necessitate strongly qualifying this claim, which will be discussed in the conclusion.

The most direct contribution to error is the detector sampling time, with which the error scales linearly, as summarized in Table 1. As discussed earlier, the uncertainty in the estimation of t_0 is approximately half the sampling time, and this uncertainty propagates to the velocity estimates through the covariance matrix. The floor of the sampling time, of course, is set much higher by the transmission of the GNSS satellites, each of which broadcasts its navigation message over the course of several minutes. A theoretical error floor is also set by the inherent degeneracies in the mathematical formulation discussed earlier, which may prevent an imperfect fit even in cases when the events precisely coincide with sampling.

Table 1 demonstrates the effect of sampling time sensitivity for simulated wavefront with a speed of v = 400 km/s, a source direction with azimuth of 120° and elevation -45°, and a time of transit of t_0 = 150 seconds.

Table 1.

Sampling time	**v** [km/s]	**azimuth** [deg]	**elevation** [deg]	$\mathbf{t_0}$ [s]
0.01	399.9958 ± 0.0054	120.0003 ± 0.0011	-45.0000 ± 0.0007	150.0060 ± 0.0005
0.05	400.0185 ± 0.0253	120.0030 ± 0.0053	-44.9989 ± 0.0034	150.0290 ± 0.0024
0.1	400.0597 ± 0.0564	120.0002 ± 0.0116	-44.9997 ± 0.0079	150.0561 ± 0.0054
0.5	400.2843 ± 0.3233	119.9977 ± 0.0669	-44.9957 ± 0.0443	150.2375 ± 0.0311
1.0	400.0609 ± 0.6343	120.2005 ± 0.1314	-45.0166 ± 0.0869	150.5160 ± 0.0610
10.0	405.8272 ± 5.1814	118.4279 ± 1.0875	-45.8324 ± 0.7879	155.4136 ± 0.5054
100.0	438.3547 ± 81.7563	119.4463 ± 13.8810	-41.2214 ± 9.3802	199.5537 ± 6.2997

Table 2 demonstrates the effect of noise and subsequent filtering when the clock bias noise standard deviation is raised 50% from its nominal value to 1.5×10^{-10} s, for a fixed sample rate of 1.0 s. Here, the higher tail of the clock noise is on the same order as the wavefront-induced signal (one nanosecond), and therefore cannot be pruned by time differencing or explicit filtering. Without resampling, the accuracy greatly suffers: the analyzer cannot determine the wavefront direction within a degree, or the velocity within tens of kilometers per second. With the application of jackknife resampling, the estimates immediately return to their nominal quality.

Table 2.

Resampling	**v** [km/s]	**az** [degrees]	**el** [degrees]	$\mathbf{t_0}$ [s]
None	396.6879 ± 20.0241	114.8896 ± 4.2976	-46.2869 ± 2.8832	149.7432 ± 2.0077
Outlier removal	400.0922 ± 0.6452	120.1852 ± 0.1360	-45.0062 ± 0.0901	150.5091 ± 0.0631
Jackknifing	400.0922 ± 0.6452	120.1852 ± 0.1360	-45.0062 ± 0.0901	150.5091 ± 0.0631

CONCLUSION AND FUTURE WORK

The current program performs excellently for a theoretical GNSS system that can rapidly report its clock biases on the order of the sample rate, but this is not consistent with how existing GNSS systems operate. A currently attainable sample rate is on the order of a few tens of seconds (the total GNSS transmission time divided by the number of reporting satellites) at the minimum, and most samples will contain data from only a few satellites. Since the wavefront estimate errors scale with the sample time, this could be problematic.

The quick sampled time was assumed for the purposes of developing a tool to monitor the GNSS system and the various algorithms it employs. Further work is required to adapt the mechanics of the tool for the sparse sampling required for realistic use. An additional modification would be to extend the data stream design principles to the entire processing

pipeline, enabling the tool to output and refine a wavefront estimate in realtime from the very first event. This would enable telescope authors to begin their observations significantly earlier. Nonetheless, in its current state, the tool serves as a demonstration that wavefronts parameters can be accurately estimated from a relatively sparse set of transit events, and shows that the data stream paradigm provides a fast and convenient method of accomplishing this computation.

ACKNOWLEDGEMENTS

The author would like to thank Prof. Grace Gao and Shubh Gupta for their teaching over the last quarter.

CODE

The full project code is available at `https://github.com/rlipkis/dark-matter`.

REFERENCES

1. Arvanitaki, Asimina, Junwu Huang, and Ken Van Tilburg. "Searching for dilaton dark matter with atomic clocks." Physical Review D 91.1 (2015): 015015.
2. Roberts, Benjamin M., et al. "Search for domain wall dark matter with atomic clocks on board global positioning system satellites." Nature communications 8.1 (2017): 1-9.