# Lawrence Berkeley National Laboratory

**Title**
Observability of dark matter substructure with pulsar timing correlations

**Permalink**
https://escholarship.org/uc/item/4n61v2sc

**Journal**
Journal of Cosmology and Astroparticle Physics, 2020(12)

**ISSN**
1475-7516

**Authors**
Ramani, Harikrishnan
Trickle, Tanner
Zurek, Kathryn M

**Publication Date**
2020-12-01

**DOI**
10.1088/1475-7516/2020/12/033

Peer reviewed

# Observability of Dark Matter Substructure with Pulsar Timing Correlations

Harikrishnan Ramani,[1,2] Tanner Trickle,[1,2,3] and Kathryn M. Zurek[3]

[1]*Department of Physics, University of California, Berkeley, CA 94720*

[2]*Theoretical Physics Group, Lawrence Berkeley National Laboratory, Berkeley, CA 94720*

[3]*Walter Burke Institute for Theoretical Physics,*

*California Institute of Technology, Pasadena, CA 91125, USA*

Dark matter substructure on small scales is currently weakly constrained, and its study may shed light on the nature of the dark matter. In this work we study the gravitational effects of dark matter substructure on measured pulsar phases in pulsar timing arrays (PTAs). Due to the stability of pulse phases observed over several years, dark matter substructure around the Earth-pulsar system can imprint discernible signatures in gravitational Doppler and Shapiro delays. We compute pulsar phase correlations induced by general dark matter substructure, and project constraints for a few models such as monochromatic primordial black holes (PBHs), and Cold Dark Matter (CDM)-like NFW subhalos. This work extends our previous analysis, which focused on static or single transiting events, to a stochastic analysis of multiple transiting events. We find that stochastic correlations, in a PTA similar to the Square Kilometer Array (SKA), are uniquely powerful to constrain subhalos as light as $\sim 10^{-13}\ M_\odot$, with concentrations as low as that predicted by standard CDM.

**CONTENTS**

# I.  INTRODUCTION

The nature of the dark matter and its associated forces, known as the dark sector, remains unknown. To uncover its identity, interactions with the Standard Model have been probed through production at colliders, direct detection in laboratory experiments, indirect detection of dark matter annihilation products in the galaxy, and the impact of the dark sector on stellar and cosmological evolution. However, dark matter may interact with the Standard Model only via gravity. If this is the case, gravitational probes of dark matter substructure will be the only avenue to learn more about the underlying theory of dark matter.

The dark matter halo structure observed on cosmological and galactic scales is observed to be consistent with adiabatic density perturbations generated by inflation, and (at least at leading order) is independent of the particle nature of dark matter. On smaller mass scales, however, many theories of dark matter leave unique fingerprints on primordial density perturbations that grow into characteristic Halo Mass Functions (HMFs). The Weakly Interacting Massive Particle (WIMP), for example, features a scale-invariant adiabatic primordial power spectrum set by inflation, with a characteristic damping below $10^{-6}$ $M_\odot$ in the HMF, due to WIMP free-streaming [1]. On the other hand the QCD axion has large isocurvature fluctuations that can collapse to form very dense halos called miniclusters [2–7], enhancing the HMF on small scales. Many other theories predict enhanced matter power on small scales, including vector bosons produced during inflation [8] and theories with early matter domination [9–12]. At present, these theories are poorly constrained by observations.

The challenge of observing structure on scales much smaller than galaxies arises because, once the virial temperature of halos halos drops below the baryon temperature, baryons no longer effectively trace the dark matter halos. At masses below $\sim 10^9$ $M_\odot$, star formation is suppressed so that stars cannot be used to trace dark matter. Thus neither galaxy surveys nor observations of the Lyman-$\alpha$ absorption of the spectra of distant quasars can give information on dark matter halos on comoving scales below $\sim 0.01 - 0.1$ Mpc. Smaller structures have instead been observed with strong lensing of quasars [13] and with fluctuations in stellar streams [14], both confirming subhalos down to about $10^7$ $M_\odot$. 21 cm cosmology [15] for masses in the range $10^6 - 10^8 M_\odot$, strong gravitational lensing [16] and stellar wakes [17] for $M > 10^5$ $M_\odot$, astrometric lensing [18–20] for $M > 1$ $M_\odot$, and disruption of compact stellar systems [21] for $M > 5$ $M_\odot$, have all been proposed to extend constraints on the HMF to lower masses.

For sub-solar mass halos, microlensing of stars towards the Large Magellanic Cloud (*e.g.*, MACHO [22], EROS [23], OGLE [24]), Andromeda (*e.g.* SUBARU [25, 26]) or stars in the local neighborhood (from Gaia [19] and KEPLER [27]) constrains sufficiently dense halos to be a sub-dominant component of the dark matter. However, microlensing becomes ineffective

in detecting subhalos below $\sim 10^{-10} - 10^{-11}$ $M_\odot$ [26]. In the future, lensing of gamma ray bursts [28] and fast radio bursts [29] may be able to reach these small masses, but these searches are typically only sensitive to very compact objects rather than halos. Astrometric lensing cannot constrain halos even a thousand times more dense than the local dark matter density, while micro-lensing loses reach even for halos $10^{17}$ times more dense (see for example Ref. [18, 30, 31]); Cold Dark Matter (CDM) halos, and even axion-like or scalar miniclusters (as discussed in Ref. [5]), are often too 'fluffy' to be observed, particularly with microlensing. Recently, photometric monitoring of caustic transiting stars has been proposed as a probe of subhalos down to $10^{-15} M_\odot$ [32] and with a lower central density, although this requires dedicated monitoring by telescopes such as the Hubble Space Telescope or the James Web Space Telescope.

For low mass and low concentration subhalos, Pulsar Timing Arrays (PTAs) are a unique and powerful probe of dark matter substructure, as considered in Refs. [30, 33–39]. We previously demonstrated that individual transiting subhalos and PBHs can be detected in the future by the Square Kilometer Array (SKA) [40] in the range $10^{-11} - 10^3$ $M_\odot$ [30]. Owing to the sensitivity of an individual pulsar far exceeding that of a traditional gravitational lens, we showed in Ref. [30] that PTA constraints on dark matter substructure remain in force (over certain mass ranges) even for halo concentration typical of ordinary CDM subhalos. This implies that they will have great sensitivity to a wide range of models with even a moderate amount of additional matter power on small scales.

To cover this wide mass range, we considered four different signal types: *static* and *dynamic* signals from *Doppler* or *Shapiro* effects induced by a single transiting subhalo. The Doppler delay is an acceleration effect from the subhalos gravitationally pulling the Earth or pulsars; the Shapiro delay is a gravitational redshift effect on the travel time of photons due to metric perturbations along the photon trajectory. Static and dynamic signals are differentiated by the time scale of their events. A static signal persists over the observing time and leaves its imprint on the, usually small, second derivative of the pulsar frequency; a dynamic signal is shorter than the observing time, and gives rise to a characteristic signal shape. In Ref. [30] constraints in the dynamic regime were set using only the single strongest event (statistically drawn from a spatial distribution of halos). Because the characteristic signal shape is predictive, one can filter the data for the signal shape on an event-by-event basis; we will refer to single dynamic signals as *deterministic* throughout this work.

The natural extension of this deterministic analysis is to study the effect from an ensemble of events, where the observable is a correlation of signal shapes and the signal is stochastic in nature. The purpose of this paper is to compute the reach on dark matter substructure using a stochastic signal for the Doppler and Shapiro delays. Such a statistical observable was

considered previously in Ref. [37]; the present study improves over the previous analysis in important ways, by taking into account finite volume effects and the impact of the pulsar fit parameters on the signal-to-noise ratio (SNR). Additionally we provide the correct subtraction procedure to capture the effect of the pulsar model fit on a general signal. We perform this subtraction for the stochastic signal, as well as the deterministic signal of Ref. [30]; for the latter we find the pulsar model fit gives rise to a substantial correction on the reach. Note that the actual impact of the pulsar model fit will be dependent on the precise timing model and the pulsar sample, and we leave an analysis utilizing existing data for future work.

The outline of this paper is as follows. We begin by determining how a stochastic dark matter signal affects the PTA observable *i.e.* the residual phase, in Sec. II. This includes a general discussion of the pulsar model fit and dark matter signal in Sec. II A, a concrete calculation of the signal correlator (for both Doppler and Shapiro delays) in Sec. II B, and a derivation of the optimal signal-to-noise ratio (SNR) in Sec. II C. We then turn to dark matter model reach in Sec. III. We compute constraints for monochromatic mass distribution of PBHs and more diffuse halos in Sec. III A, and generalize these results in Sec. III B to a slightly broadened HMF, along with a CDM-like HMF. Lastly we conclude with future directions for applying our results to existing PTA data, and to a broader class of dark matter models.

## II. DARK MATTER SIGNATURES IN PULSAR PHASE CORRELATIONS

The goal of this section is to compute the signal-to-noise ratio (SNR) in PTAs generated by an ensemble of transiting dark matter subhalos. We begin with a discussion of the PTA observable, pulsar phases, and how dark matter subhalos can produce correlations in them. These correlations are the signal which we compare with PTA timing noise to construct an SNR. This SNR, for a general model of dark matter substructure, will then be our basis for projecting constraints in the next section.

### A. Pulsar Phase Correlator

Pulsars with millisecond periods, observed over decades, are known to be good clocks. This is because, while the pulsar period may fluctuate on short time scales, these fluctuations do not accumulate, such that the arrival time of light pulses can be predicted with a simple model of the pulsar phase evolution,

$$\phi(t) = \phi^0 + \nu t + \frac{1}{2}\dot{\nu}t^2, \tag{1}$$

where $\phi^0, \nu, \dot{\nu}$, are the phase offset, pulsar frequency, and its first time derivative. The success of this model implies that any deviations due to dark matter substructure can be observed or constrained. These deviations are characterized by the residual phase,

$$s(t) \equiv \phi(t) - \phi_{\text{fit}}(t), \tag{2}$$

where $\phi_{\text{fit}} = \phi_{\text{fit}}^0 + \nu_{\text{fit}}t + \dot{\nu}_{\text{fit}}t^2/2$, and $\phi_{\text{fit}}^0, \nu_{\text{fit}}, \dot{\nu}_{\text{fit}}$ are obtained by fitting the measured pulsar phase with the timing model. In the absence of dark matter substructure, this residual is well-fit by stationary white noise, $s(t) = n(t)$, where $n(t)$ is defined by its statistical properties, $\langle n(t)n(t')\rangle = \nu^2 t_{\text{rms}}^2 \Delta t\, \delta(t-t')$, with $\Delta t$ the measurement cadence and $t_{\text{rms}}$ the root-mean-square post-fit timing residual, discussed further in Appendix A.

If dark matter substructure is present, the residual phase will have additional contributions, which we quantify as

$$s(t) = h(t) + n(t), \tag{3}$$

with $h(t)$ the *subtracted* dark matter signal,

$$h(t) \equiv \delta\phi(t) - \delta\phi_{\text{fit}}(t). \tag{4}$$

Here $\delta\phi$ is the phase modification induced by the dark matter substructure, and $\delta\phi_{\text{fit}}$ is the part of the signal absorbed by the pulsar timing model fit, as detailed in Appendix B. $\delta\phi$ can be written in terms of a frequency shift,

$$\delta\phi(t) = \int_0^t \delta\nu(t')\, dt'. \tag{5}$$

We consider two gravitational effects from transiting subhalos that induce a frequency shift. The Doppler effect arises when transiting subhalos induce an acceleration in the Earth or pulsar, while the Shapiro effect is due to the change in the gravitational potential along the photon's trajectory; see Refs. [30, 37, 38, 41, 42] for more details. These shifts, for a single transiting subhalo, are given by

$$\left(\frac{\delta\nu}{\nu}\right)_D = \hat{\mathbf{d}} \cdot \int \nabla\Phi(\mathbf{r}, M)\, dt \tag{6}$$

$$\left(\frac{\delta\nu}{\nu}\right)_S = -2 \int \mathbf{v} \cdot \nabla\Phi(\mathbf{r}, M)\, dz, \tag{7}$$

for the Doppler and Shapiro delay respectively, where $\Phi$ is the gravitational potential from a single subhalo, $\mathbf{v}$ is the subhalos velocity, and $\hat{\mathbf{d}}$ is the direction from the Earth to the pulsar. In Ref. [30] these expressions were utilized to constrain the abundance of PBHs and compact halos via *single* transiting subhalos, though only the signal from the closest subhalo was considered.

The advantage of this approach is the ability to predict the specific signal shape in order to filter the data accordingly. The disadvantage is that, at small masses, even the closest subhalo does not produce a measurable signal. However, in this small mass regime there is an abundance of subhalos which could cumulatively leave a discernible signal. This signal from a *statistical ensemble* of transiting subhalos,

$$\delta\phi(t) = \sum_{i=1}^{N} \delta\phi_i(t) \,, \tag{8}$$

where $\delta\phi_i$ is the phase modification from the $i^{\text{th}}$ event, is our starting point. Correlations can then be written,

$$\langle\delta\phi(t)\delta\phi(t')\rangle = \sum_{i=1}^{N}\langle\delta\phi_i(t)\delta\phi_i(t')\rangle + \sum_{i\neq j}^{N(N-1)}\langle\delta\phi_i(t)\delta\phi_j(t')\rangle \equiv R_1(t,t') + R_2(t,t') \,, \tag{9}$$

where $R_1$ $(R_2)$ contains contributions from averaging over one (two) subhalo(s).

## B.   Dark Matter Signal Correlator

We now compute the dark matter induced phase correlation, $R$ in Eq. (9), for the Doppler and Shapiro delays. The expectation, $\langle\rangle$, averages over the random variables that determine the phase shifts from all $N$ subhalos. Similar to the treatment in Ref. [16], we take these random variables to be the subhalo masses, $M_i$, initial positions $\mathbf{r}_i^0$, and velocities $\mathbf{v}_i$. The 1-subhalo term, $R_1$, in $R$ does not include subhalo correlations, while $R_2$ depends on the subhalo correlation power spectrum, $P_\xi$, defined from the subhalo number density, $n_{\text{sub}}$ as,

$$n_{\text{sub}}(\mathbf{x}) \equiv \overline{n}\left(1 + \delta_n(\mathbf{x})\right) \tag{10}$$

$$\langle\delta_n\rangle_{\text{en}} = 0 \tag{11}$$

$$\langle\delta_n(\mathbf{x})\delta_n(\mathbf{y})\rangle_{\text{en}} \equiv \xi(\mathbf{x} - \mathbf{y}) \tag{12}$$

$$P_\xi(\mathbf{k}) = \int d^3\mathbf{x}\, e^{i\mathbf{k}\cdot\mathbf{x}}\xi(\mathbf{x}) \,. \tag{13}$$

where $\langle\rangle_{\text{en}}$ denotes an ensemble average.[1] The statistical nature of the signal is similar to Ref. [37], with important differences accounting for finite observation volume. The comparisons between the formalisms is discussed in Appendix C. We leave a discussion of $R_2$, which is non-zero when $P_\xi \neq 0$, for future work, as we expect $R_2$ to be subdominant to the leading effect

---

[1] For a monochromatic mass distribution, one can show that the matter power spectrum, $P_m$, and the correlation power spectrum, $P_\xi$, are related by $P_m = |W(k,M)|^2/\overline{n} + P_\xi\,|W(k,M)|^2$

from uniformly distributed subhalos. For notational simplicity, we will refer to $R_1$ as $R$ from here on.

Assuming $M_i, \mathbf{r}_i^0, \mathbf{v}_i$ are independent and have identical probability distribution functions, we can write $R$ in terms of the frequency shift, averaged over a single subhalo of mass $M$ with position $\mathbf{r}_0$ and velocity $\mathbf{v}$,

$$R(t, t') = \int_0^t \int_0^{t'} dt_1 dt_2 \int d^3\mathbf{v} f_{\mathbf{v}}(\mathbf{v}) \int \frac{dM}{M} F(M) \int d^3\mathbf{r}^0 \, \delta\nu(t_1; M, \mathbf{r}^0, \mathbf{v}) \delta\nu(t_2; , M, \mathbf{r}^0, \mathbf{v}) , \tag{14}$$

where $F(M) = dn/d\log M$ is the HMF, $f_{\mathbf{v}}$ is a boosted Maxwell-Boltzmann velocity distribution, with $v_0 = 230$ km/s, $v_E = 240$ km/s, and $v_{\text{esc}} = 600$ km/s. Assuming uniformly distributed subhalos means the probability distribution function in $\mathbf{r}^0$ is simply $1/V$, where $V$ is the observing volume. We expect a relatively weak dependence on the velocity distribution, and from here on will take the velocity to be an average value. The Doppler delay signal will depend on the average velocity, $\langle v \rangle_{\mathbf{v}} \equiv \bar{v} \approx 340$ km/s, whereas the Shapiro delay will depend on the velocity component perpendicular to the Earth pulsar direction, $\langle v_\perp \rangle_{\mathbf{v}} \equiv \bar{v}_\perp \approx 270$ km/s.[2]

We now compute $R$ for the Doppler and Shapiro delays.

### 1. Shapiro delay

We begin by simplifying the building block of a statistical signal by writing the Shapiro delay from a single subhalo, $\delta\nu_S$ in Eq. (7), as

$$\delta\nu_S = -8\pi i\nu GM \int_0^{z_0} d\ell \int \frac{d^3k}{(2\pi)^3} \frac{\mathbf{v} \cdot \mathbf{k}}{k^2} W(k, M) e^{ik_z(\ell-z)} e^{-i\mathbf{k}_\perp \cdot \mathbf{r}_\perp} , \tag{15}$$

where we have used the relation between the gravitational potential $\tilde{\Phi}$, density profile $\tilde{\rho}(k, M)$, and window function $W(k, M)$: $\widetilde{\Phi}(k) = -(4\pi G/k^2)\tilde{\rho}(k, M) \equiv -(4\pi GM/k^2)W(k, M)$. It will be useful to define a coordinate system with $\hat{\mathbf{z}}$ along the Earth-pulsar direction, and a plane perpendicular to $\hat{\mathbf{z}}$ such that $\mathbf{r}_\perp = \mathbf{v}_\perp t + \mathbf{r}_\perp^0 = \mathbf{v}_\perp(t - t^0) + \mathbf{b}$, with $\mathbf{b}$ the impact parameter. The position of a subhalo is thus $\mathbf{r} = \mathbf{r}_\perp + \hat{\mathbf{z}}(\mathbf{r} \cdot \hat{\mathbf{z}})$.

Evaluating the $k_z$, $z$ integrals, in the limit $k_z \lesssim z_0^{-1} \ll k_\perp$ and $0 < z < z_0$, gives (similar to Ref. [37])

$$\delta\nu_S \approx 8\pi i\nu GM \int \frac{d^2k_\perp}{(2\pi)^2} (\mathbf{v}_\perp \cdot \mathbf{k}_\perp) \frac{W(k_\perp, M)}{k_\perp^2} e^{i\mathbf{k}_\perp \cdot \mathbf{r}_\perp} . \tag{16}$$

---

[2] To compute $\bar{v}_\perp$ we calculate the expectation value of $\sqrt{\mathbf{v} - (\mathbf{v} \cdot \hat{\mathbf{d}})\hat{\mathbf{d}}}$ (the magnitude of the components perpendicular to the Earth-pulsar direction, $\hat{\mathbf{d}}$) and average over all directions of the Earth velocity, $\hat{\mathbf{v}}_E$

This expression can be written in terms of Bessel functions

$$\delta\nu_S \approx -4GM\nu\frac{\mathbf{v}_\perp \cdot \mathbf{r}_\perp}{r_\perp^2}r_\perp \int_0^\infty dk\, W(k,M)J_1(k\,r_\perp) \equiv -4GM\nu\frac{\mathbf{v}_\perp \cdot \mathbf{r}_\perp}{r_\perp^2}\mathcal{F}(M,r_\perp)\,. \tag{17}$$

We have defined a form factor $\mathcal{F}$,

$$\mathcal{F}(M,x) \equiv x\int_0^\infty W(k,M)J_1(kx)\,dk\,, \tag{18}$$

with $J_1(x)$ the first order Bessel function. In the PBH limit, $\mathcal{F} = 1$, Eq. (17) reduces to the corresponding expression in Ref. [30]. We further take $W(y/r_\perp, M) \approx W(y/b, M)$, which we expect to be reasonable as the signal is peaked near $t = t^0$. We then obtain

$$\delta\nu_S \approx -4GM\nu\,\overline{v}_\perp^2\,\frac{t - t^0}{b^2 + \overline{v}_\perp^2\,(t - t^0)^2}\,\mathcal{F}(M,b)\,. \tag{19}$$

We can now compute the signal correlator in Eq. (14). In order to account for a finite observing volume we decompose the integral over the initial position as $d^3\mathbf{r}^0 = \overline{v}_\perp dz^0\,db\,dt^0$, where $\hat{\mathbf{z}}, \hat{\mathbf{b}}, \hat{\mathbf{v}}_\perp$ are orthogonal directions. We note that in the large $t^0$ limit, $(\delta\nu/\nu)^2 \propto t_0^{-2}$, and for simplicity will extend the bounds on the $t^0$ integral to infinity. Finally,

$$R_S(t,t') = 32\pi G^2\nu^2 z_0 \int dM\, MF(M) \int db\, b\, A\left(t,t',\frac{b}{\overline{v}_\perp}\right)\mathcal{F}(M,b)^2\,, \tag{20}$$

where

$$A(t,t',\tau) \equiv \int_0^t \int_0^{t'} dt_1 dt_2\,\frac{1}{4\tau^2 + (t_1 - t_2)^2}\,. \tag{21}$$

### 2. Doppler delay

The derivation of the signal correlator for the Doppler delay begins analogously to the Shapiro delay. We write the frequency shift of an individual subhalo from the $J^{\text{th}}$ pulsar, $\delta\nu_{D,J}$,

$$\delta\nu_D^J = 4\pi i\nu^J GM \int_{-\infty}^t dt' \int \frac{d^3k}{(2\pi)^3}\frac{\hat{\mathbf{d}}^J \cdot \mathbf{k}}{k^2}W(k,M)e^{-i\mathbf{k}\cdot\mathbf{r}(t')}\,, \tag{22}$$

where we have again used the relation between the gravitational potential $\tilde{\Phi}$, density profile $\tilde{\rho}(k,M)$, and the window function $W(k,M)$: $\widetilde{\Phi}(k) = -(4\pi G/k^2)\tilde{\rho}(k,M) \equiv -(4\pi GM/k^2)W(k,M)$. $\hat{\mathbf{d}}^J$ is the the direction pointing from the Earth to the $J^{\text{th}}$ pulsar, and $\mathbf{r}(t) = \mathbf{v}(t - t^0) + \mathbf{b}$, where $\mathbf{b}$ is the impact parameter and orthogonal to $\mathbf{v}$. We evaluate the $t'$ integral with the identity $\int_{-\infty}^t dt'e^{-i\mathbf{k}\cdot\mathbf{v}t'} = \frac{i}{\mathbf{k}\cdot\mathbf{v}}e^{i\mathbf{k}\cdot\mathbf{v}t} + \pi\delta(\mathbf{k}\cdot\mathbf{v})$ and decompose $\mathbf{k}, \hat{\mathbf{d}}^J$ into the coordinate system

spanned by $\mathbf{b}, \mathbf{v}$, $i.e.$ $\mathbf{k} = \mathbf{k}_b + k_v \hat{\mathbf{v}}$, where $\mathbf{k}_b$ lies in the plane of the impact parameter (and therefore $\mathbf{k}_b \cdot \mathbf{v} = 0$). We obtain

$$
\begin{aligned}
\delta\nu_{D,J} = & -\frac{4\pi GM\nu^J}{\overline{v}} \int \frac{d^2 k_b dk_v}{(2\pi)^3} \frac{\hat{\mathbf{d}}^J \cdot \mathbf{k}_b}{k^2} \frac{W(k,M)}{k_v} e^{-i\mathbf{k}_b \cdot \mathbf{b}} e^{-ik_v \overline{v}(t-t^0)} \\
& + \frac{4\pi^2 i GM\nu^J}{\overline{v}} \int \frac{d^2 k_b}{(2\pi)^3} \frac{\hat{\mathbf{d}}^J \cdot \mathbf{k}_b}{k_b^2} W(k_b, M) e^{-i\mathbf{k}_b \cdot \mathbf{b}} \\
& - \frac{4\pi GM\nu^J \hat{\mathbf{d}}^J \cdot \hat{\mathbf{v}}}{\overline{v}} \int \frac{d^2 k_b dk_v}{(2\pi)^3} \frac{W(k,M)}{k^2} e^{-i\mathbf{k}_b \cdot \mathbf{b}} e^{-ik_v \overline{v}(t-t^0)} .
\end{aligned}
\tag{23}
$$

The first two terms can be combined and simplified when $k_v \ll k_b$ ($\mathbf{k} \approx \mathbf{k}_b$) using the identity $\int \frac{dk_v}{2\pi} \frac{e^{-ik_v v(t-t_0)}}{k_v} = -\frac{i}{2} \frac{t-t^0}{|t-t^0|}$, and written as,

$$
\delta\nu_D^J = 2\,\Theta(t-t^0)\left(\hat{\mathbf{d}}^J \cdot \hat{\mathbf{b}}\right) \frac{GM\nu^J}{b\overline{v}} \mathcal{F}(M,b) ,
\tag{24}
$$

where $\mathcal{F}$ is defined in Eq. (18). Each event causes a jump in the pulsar frequency, signaling that a series of such step functions causes the pulsar frequency to undergo a random walk.[3] The third term in Eq. (23) corresponds to a transient – the signal does not accumulate in the pulsar frequency with each passing event – as in Ref. [30]; we expect such terms to have a subdominant effect and hence drop them.

For the Shapiro correlator, the signal accumulates along an Earth-pulsar path, implying that different Earth-pulsar contributions are uncorrelated. By contrast, for the Doppler correlator, a subhalo may give rise to an acceleration of the Earth alone, indicating that one event will leave a signal across the entire array of pulsars. This allows for correlation across pulsars, $\langle \delta\phi^I \delta\phi^J \rangle \equiv R^{IJ}$, which will result in a larger SNR than contributions without this correlation.

We substitute Eq. (24) in Eq. (14), identifying $d^3 \mathbf{r}^0 = \overline{v}\, d^2 \mathbf{b}\, dt^0$. After simplification,

$$
R_D^{IJ}(t,t') = \frac{4\pi G^2 \nu^2}{\overline{v}} \left(\hat{\mathbf{d}}^I \cdot \hat{\mathbf{d}}^J\right) B(t,t') \int dM\, MF(M) \int \frac{db}{b} \left(\mathcal{F}(M,b)\right)^2 ,
\tag{25}
$$

where the correlator for a random walk process is proportional to the minimum of the time of two events:

$$
B(t,t') \equiv \int_0^t \int_0^{t'} \min(t_1, t_2)\, dt_1 dt_2 = \frac{\min(t,t')^2\,(3\max(t,t') - \min(t,t'))}{6} .
\tag{26}
$$

### C. Constructing the Signal to Noise Ratio

We now have all the ingredients to compute the SNR, which gives the significance of the measured dark matter signal over the pulsar timing noise. The dark matter signal is subtracted

---

[3] Ref. [37] dropped the second term in Eq. (23), which is necessary for obtaining this random walk behavior.

to take into account the effect of the pulsar fit model, $h = \delta\phi - \delta\phi_{\text{fit}}$, as discussed in Sec. II A and derived in Appendix B. We quote the result here,

$$h(t) = \delta\phi(t) - \sum_{n=0}^{2} \left[ \frac{1}{T} \int_0^T dt' \, \delta\phi(t') f_n(t') \right] f_n(t) , \tag{27}$$

where $f_n(t) = \sqrt{2n+1} P_n(2t/T - 1)$, and $P_n$ are the Legendre polynomials. The sum is from zero to two in order to include $\phi^0$, $\nu$, and $\dot{\nu}$ in the timing model fit. It follows that correlators of $h$, $R_{\text{sub}}(t,t') \equiv \langle h(t)h(t') \rangle$, are related to correlators of $\delta\phi$ and $R$ by (see details in Appendix B):

$$R_{\text{sub}}(t,t') = R(t,t') - \sum_{n=0}^{2} f_n(t)\mathcal{R}_n(t') - \sum_{n=0}^{2} f_n(t')\mathcal{R}_n(t) + \sum_{n=0}^{2}\sum_{m=0}^{2} \mathcal{R}_{nm} f_n(t) f_m(t') \tag{28}$$

$$\mathcal{R}_n(t') \equiv \frac{1}{T} \int_0^T R(t,t') f_n(t) \, dt \tag{29}$$

$$\mathcal{R}_{nm} \equiv \frac{1}{T^2} \int_0^T \int_0^T dt dt' \, R(t,t') \, f_n(t) f_m(t') . \tag{30}$$

Having defined the subtracted signal we construct the SNR using a matched filter procedure as in Refs. [43, 44]. We begin with the deterministic signal, as in Ref. [30], where we study the SNR, $\text{SNR}_{\text{det}}$, from the subhalo which imprints the largest signal. The expressions for the SNR are derived in Appendix A 1, and depend on whether the signal is uncorrelated across pulsars (the 'pulsar' term, $\text{SNR}_{\text{det},P}$) or correlated (the 'Earth' term, $\text{SNR}_{\text{det},E}$):

$$\text{SNR}_{\text{det},P}^2 = \frac{1}{\widetilde{N}} \max_{\{I\}} \left[ \int_0^T dt \, h_I^2(t) \right] \tag{31}$$

$$\text{SNR}_{\text{det},E}^2 = \frac{N_P}{\widetilde{N}} \int_0^T dt \, \langle h^2(t) \rangle_{\mathcal{P}}, \tag{32}$$

where the maximum over $I$ denotes the maximum signal across all pulsars in the array. $\langle\rangle_{\mathcal{P}}$ denotes averaging over the pulsar positions, $\widetilde{N} = \nu^2 t_{\text{rms}}^2 \Delta t$, $T$ is the observing time, $\Delta t$ the cadence, and the residual timing noise is $t_{\text{rms}}$.

There are two key differences between the SNR here and in Ref. [30]. The first is that, in order to unify the formalism presented here with the analysis in Ref. [30], we cast the signal in the residual phase and not residual frequency shift, $\delta\nu$ (related to each other by Eq. (5)). Second, and more importantly, the signal in the SNR is the subtracted signal, $h$, as opposed to $\delta\phi$, which was neglected in Ref. [30]. This causes an $\mathcal{O}(1)$ difference in the overall SNR and is discussed in more detail in Sec. III A.

The pulsar and Earth term SNR for the stochastic signal ($\text{SNR}_P$ and $\text{SNR}_E$ respectively)

are derived in Appendices A 2, A 3, and we quote the result here,

$$\text{SNR}_P^2 = \frac{N_P}{2\widetilde{N}^2} \int dt dt' \langle R_I^{\text{sub}}(t,t')^2 \rangle_{\mathcal{P}} \tag{33}$$

$$\text{SNR}_E^2 = \frac{N_P(N_P - 1)}{2\widetilde{N}^2} \int dt dt' \left\langle R_{IJ}^{\text{sub}}(t,t')^2 \right\rangle_{\mathcal{P}} . \tag{34}$$

The indices $I, J$ run over the pulsars in the array and $\langle \rangle_{\mathcal{P}}$ averages over the pulsar positions. The pulsar timing array parameters, $N_P$, $T$, $\Delta t$, $t_{\text{rms}}$ are drawn based on the capabilities of current PTAs, extrapolated to the potential of future PTAs. The currently operating PTAs are European Pulsar Timing Array (EPTA) [45], Parkes Pulsar Timing Array (PPTA) [46], North American Nanohertz Observatory for Gravitational Waves (NANOGrav) [47]; the MeerKAT telescope has a pulsar timing program (MeerTime [48]), along with the Five-hundred-meter Aperture Spherical Telescope (FAST) [49]. The International Pulsar Timing Array (IPTA) [50] is comprised of EPTA [45], PPTA [51], and NANOGrav [52], and between the three collaborations has measured $N_P = 73$ unique millisecond pulsars for $T \sim 10 - 30$ years, with timing residuals in the range $t_{\text{rms}} = 50$ ns $- 10^4$ ns, at a distance of $z_0 \sim 1 - 5$ kpc, and a cadence of $\Delta t \sim 1 - 4$ week. The future Square Kilometer Array (SKA) [40] could increase the number of pulsars to $N_P = 200$, with timing residuals of $t_{\text{rms}} \sim 50$ ns. The FAST telescope could optimistically reduce the timing residuals to $t_{\text{rms}} \sim 1 - 10$ ns, assuming the current limitation is statistics [49]. Our baseline PTA parameters, which we utilize in the next section, and assume when simplifying analytic results, are based on the estimated capability of the future SKA PTA: $N_P = 200$, $t_{\text{rms}} = 50$ ns, $\Delta t = 2$ week, $T = 20$ years, $z_0 = 5$ kpc.

## III. OBSERVABILITY OF DARK MATTER SUBSTRUCTURE WITH PULSAR TIMING ARRAYS

We can now determine the observability of dark matter substructure. In this Section, we consider both monochromatic and CDM-like HMFs, with constituent subhalos having either PBH or NFW density profiles. We determine the constraints on the mass fraction, $f \equiv \Omega/\Omega_{\text{DM}}$, in these models with a future PTA with SKA-like capabilities (defined in Sec. II C).

### A. Monochromatic Mass Distribution

To gain intuition for how PTAs derive the power of their constraints, we begin with the simplest case of a monochromatic mass distribution of subhalos of mass $M$, $F(M') = \overline{n}M'\delta(M - M')$. Before deriving these constraints in detail it will be important to understand the length
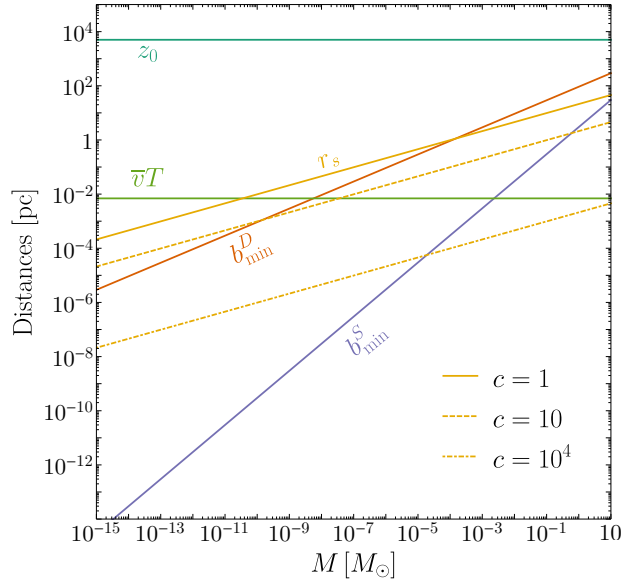
FIG. 1: Relevant length scales for measurement of dark matter subhalos with PTAs, as a function of the subhalo mass $M$ (assuming $f = 1$). $z_0$, the largest scale, is the Earth-pulsar distance. $b_{\min}^{D(S)}$ denote subhalo impact parameters for the Doppler (Shapiro) searches (see text for details); to have sensitivity they must be smaller than $\overline{v}_{(\perp)}T$, showing why the Doppler search is sensitive to lower mass subhalos than the Shapiro search. The yellow lines are the scale radius, $r_s$, for $c = 1, 10, 10^4$. A subhalo with $r_s > b_{\min}$ has weakened constraints, and when $r_s > vT$ the constraints are negligible, indicating the subhalo concentration parameters to which PTAs are sensitive.

scales which shape our results, summarized in Fig. (1). The first important length scales are the range of impact parameters of transiting subhalos to which the PTA measurement is sensitive. The distance a subhalo can travel over the observation time $T$ is

$$b_{\max} \sim 10^{-8} \text{ Mpc} \left( \frac{v}{10^{-3} c} \right) \left( \frac{T}{20 \text{ yr}} \right), \tag{35}$$

shown in Fig. (1). A subhalo must have an impact parameter smaller than $vT$ to be observable in a stochastic or dynamic search. On the lower end of the impact parameter range, the smallest impact parameter over an ensemble of $N$ events can be derived (from evenly spatially distributed subhalos), as in Ref. [30], for the Shapiro and Doppler delays,

$$b_{\min}^S \sim \frac{\overline{v}_\perp T}{N} \quad , \quad b_{\min}^D \sim \frac{3}{2} \frac{\overline{v}T}{\sqrt{N}}, \tag{36}$$

where we have quoted the $90^{\text{th}}$ percentile result. The number of subhalos in the observing volume is $N = f \rho_{\text{DM}} V / M$. The observing volume (as appears in Eq. (14)) is a cylinder for

the Doppler effect, $V_D = \pi \overline{v}^3 T^3$, and a rectangular box for the Shapiro effect, $V_S = \overline{v}_\perp^2 T^2 z_0$. We have the minimum impact parameters labeled $b_{\min}^S$, $b_{\min}^D$ in Fig. (1) for $f = 1$. In order to constrain a subhalo of mass $M$, this minimum impact parameter must be less than the maximum in Eq. (35), and by examining Fig. (1), we see that a Shapiro search will have greater sensitivity to larger mass subhalos than a Doppler signal.

For subhalos with NFW profile the size of the subhalo is also a relevant scale. There are two important sizes, the virial radius, $r_v$, which contains all the mass, and the scale radius, $r_s$,

$$r_s \sim \frac{10^{-5} \text{ Mpc}}{c} \left( \frac{M}{M_\odot} \right)^{\frac{1}{3}}, \tag{37}$$

which quantifies the compactness of a subhalo, parameterized by the concentration parameter, $c \equiv r_v/r_s$. If $b > r_v$ then the subhalo can be treated as point-like, and if $b > r_s$, there is only a modest loss in sensitivity (as we will discuss in more detail below, Sec. (III A 2)). From Fig. (1), we see that the Doppler search in particular has strong sensitivity to low concentration subhalos, while the Shapiro search will more rapidly lose its reach for extended subhalos in comparison to PBHs.

We now discuss constraints for point-like (PBH) and NFW density profiles in more detail.

### 1. Point-like subhalo (PBH)

We begin by laying out the constraints in the simplest case: a point-like subhalo (or PBH), where the impact form factor in Eq. (18) simplifies to $\mathcal{F} = 1$. The projected constraints are shown in Fig. (2) for the PTA parameters discussed in Sec. II C, where the new results, labeled by 'DopStoch' and 'ShapStoch', are from the stochastic signal caused by the Doppler and Shapiro effects respectively. Shown in orange and green are the reach curves derived in Ref. [30] from deterministic ('Det') or static ('Static') events, corrected to include the subtraction effects. Fig. (3) shows the same results as Fig. (2) but for more futuristic PTA parameters described in the caption.

We first note that the deterministic signal constraints shown in Fig. (2) differ from those in Ref. [30] even for the same sets of PTA parameters. This is because subtraction of fitted pulsar parameters was neglected previously. To better understand the effect of the subtraction, we look at both the signal and noise power in frequency space.

In Fig. (4) we plot the signal and noise strain ($\mathcal{S}$ and $\mathcal{N}$ respectively) for the deterministic and stochastic signals, following the conventions in Ref. [43], and show the effects of subtracting different terms in the timing model, as in Eq. (1). We define the signal strains for the
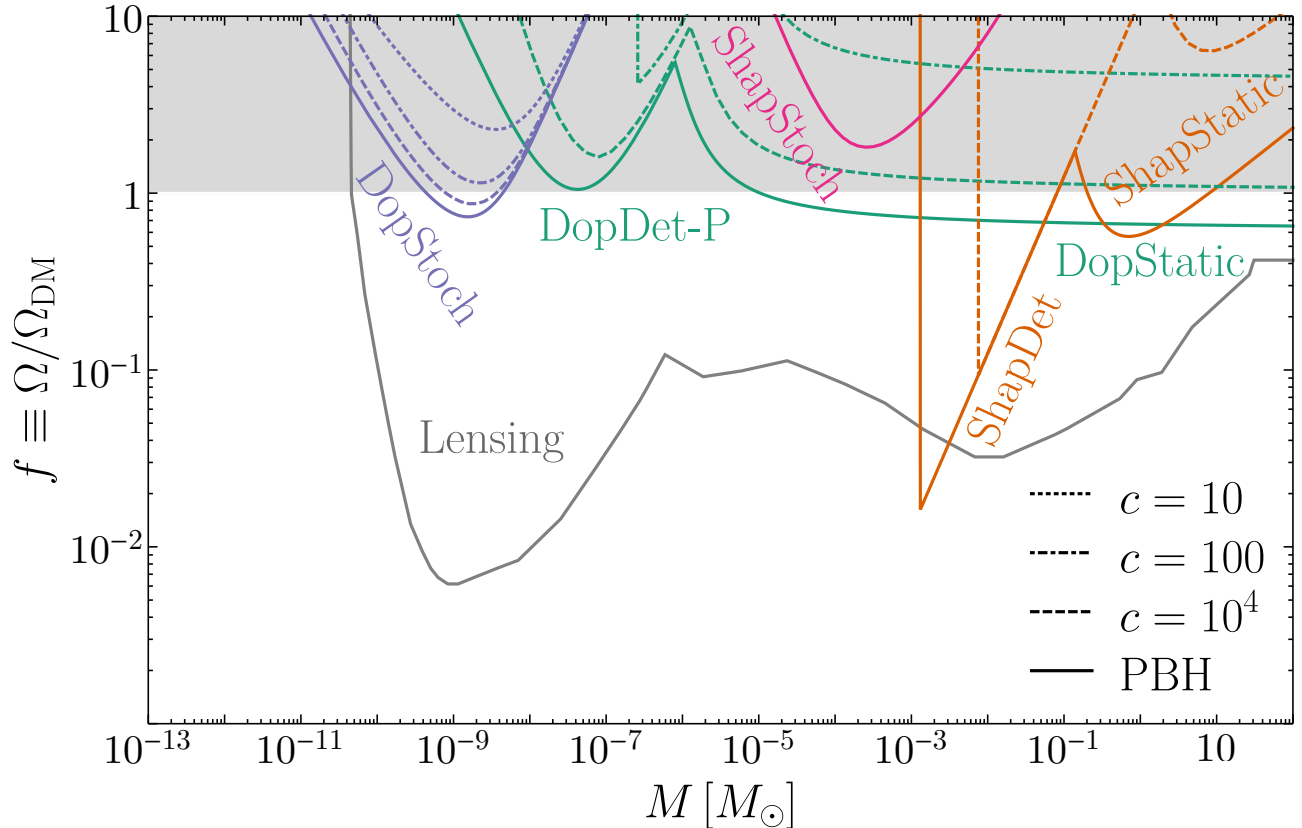
FIG. 2: Limits from PTAs on the dark matter mass fraction $f = \Omega/\Omega_{\mathrm{DM}}$ in subhalos of mass $M$ for different subhalo concentration parameters, $c = 10, 100, 10^4$, and the PBH limit, $c \to \infty$. Results derived in Ref. [30] from deterministic single transiting objects and static signals are labeled 'DopDet-P', 'DopStatic', 'ShapDet', and 'ShapStatic' and shown in green and orange. The 'DopDet-P' and 'ShapDet' constraints have been weakened relative to Ref. [30] due to the subtraction procedure discussed in Appendix B. New results of this paper utilizing a stochastic signal induced by multiple transiting subhalos are labeled 'DopStoch' and 'ShapStoch', and shown in blue and pink, respectively. An SKA-like PTA, described in Sec. II C, with identical pulsars was assumed. Lensing constraints in gray are from Refs. [22–26, 53], and disappear for $c < 10^7$.

deterministic and stochastic signals, $\mathcal{S}_{\mathrm{det}}, \mathcal{S}_{\mathrm{stoch}}$ such that

$$\mathrm{SNR}_{\mathrm{det}}^2 = \int d\log\mathfrak{f}\, \frac{\mathcal{S}_{\mathrm{det}}^2(\mathfrak{f})}{\mathcal{N}_{\mathrm{det}}^2(\mathfrak{f})} \tag{38}$$

$$\mathrm{SNR}_{\mathrm{stoch}}^2 = \int d\log\mathfrak{f}\, d\log\mathfrak{f}'\, \frac{\mathcal{S}_{\mathrm{stoch}}^2(\mathfrak{f}, \mathfrak{f}')}{\mathcal{N}_{\mathrm{stoch}}(\mathfrak{f})\mathcal{N}_{\mathrm{stoch}}(\mathfrak{f}')}\,, \tag{39}$$

where the overall $N_P$ dependence for pulsar and Earth terms, is absorbed in $\mathcal{S}_{\mathrm{det}}, \mathcal{S}_{\mathrm{stoch}}$, and $\mathfrak{f}$ is
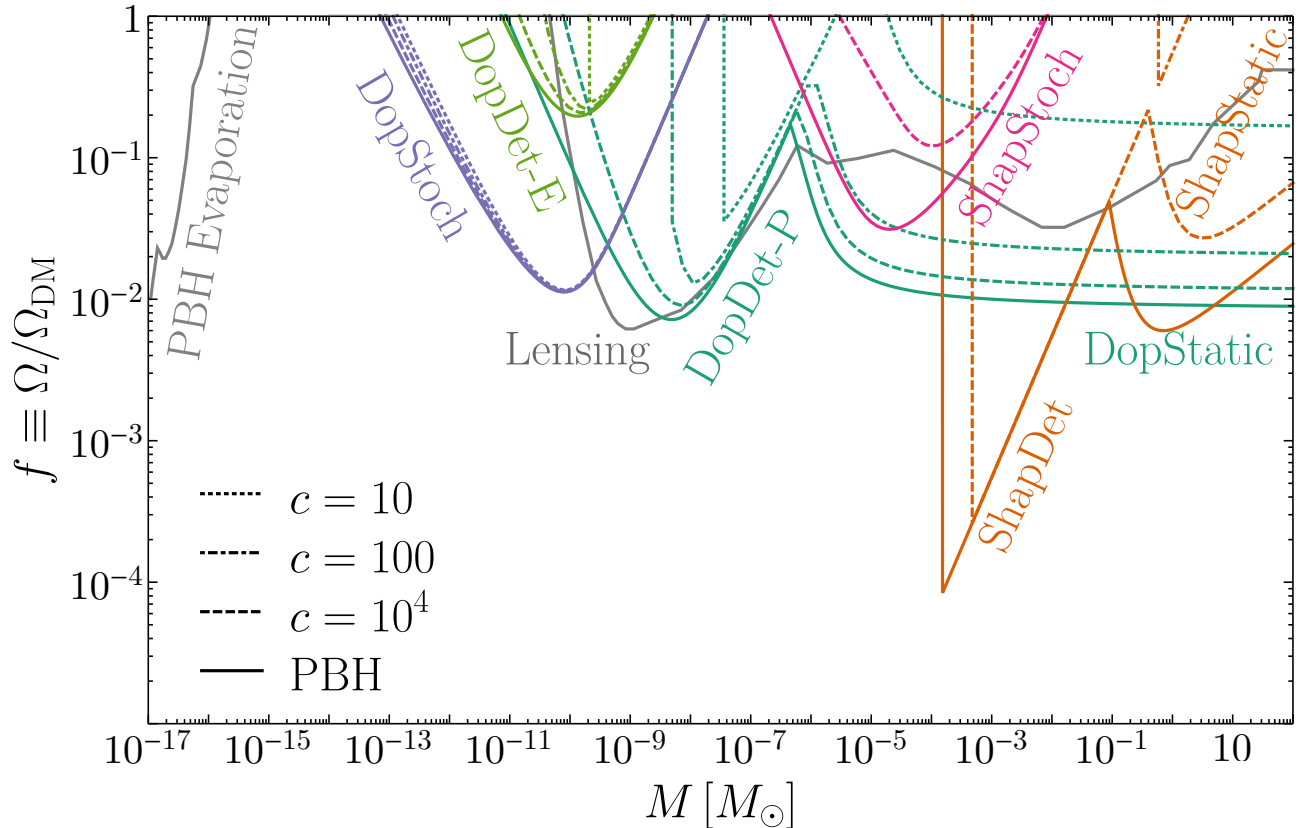
FIG. 3: Similar to Fig. (2) we show limits from PTAs on the dark matter mass fraction $f = \Omega/\Omega_{\rm DM}$ for $c = 10, 100, 10^4$ and the PBH limit. Here we have assumed a more optimistic set of PTA parameters compared to Fig. (2): $N_P = 1000$, $T = 30$ yr, $t_{\rm rms} = 10$ ns, $\Delta t = 1$ week, and $z_0 = 10$ kpc. One additional constraint 'DopDet-E' is visible here and corresponds to the deterministic Doppler Earth term. Lensing constraints are again from Refs. [22–26, 53], and disappear for $c < 10^7$. The PBH evaporation constraint ('PBH Evaporation') is from Ref. [54].

used for frequency, to avoid confusion with the dark matter mass fraction, $f$. These definitions allow one to estimate the contribution from different decades in frequency to the SNR using Fig. (4). For comparison with white noise we show the white noise strain, $\mathcal{N}_{\rm det}(\mathfrak{f}) = \sqrt{\mathfrak{f}\,\nu^2 t_{\rm rms}^2 \Delta t}$, and $\mathcal{N}_{\rm stoch}(\mathfrak{f}) = \mathfrak{f}\,\nu^2 t_{\rm rms}^2 \Delta t$ for an SKA-like PTA.

The deterministic signal strain, shown in the top row of Fig. (4), is from the pulsar term SNR and therefore

$$\mathcal{S}_{\rm det} = \mathfrak{f} \left| \int dt\, e^{2\pi i \mathfrak{f} t} h(t) \right| , \tag{40}$$

where $h$ is the subtracted signal in Eq. (27). The stochastic signal SNR depends on a two

dimensional integral, as seen in Eq. (39), and therefore the signal strain cannot be plotted as simply as the deterministic signal strain. Instead we show a one-dimensional slice $\mathfrak{f} = \mathfrak{f}'$, where the signal strains for the Doppler (Shapiro) delays, $\mathcal{S}_{D(S)}^{\text{stoch}}$, can be written in terms of a power, $P(\mathfrak{f})$,

$$P(\mathfrak{f}) \equiv \int dt \, dt' \, e^{2\pi i \mathfrak{f}(t+t')} R_{\text{sub}}(t, t') \,. \tag{41}$$

The signal strains are then

$$\mathcal{S}_{D(S)}^{\text{stoch}}(\mathfrak{f}, \mathfrak{f}) = \sqrt{\frac{N_P}{2}} \mathfrak{f}^2 \left| P_{D(S)}(\mathfrak{f}) \right| \,. \tag{42}$$

In each panel of Fig. (4) the strain corresponding to the unsubtracted signal is contrasted with the strain from signals with increasingly higher order subtractions corresponding to $\phi^0, \nu$, and $\dot{\nu}$. Subtraction has the largest effect at frequencies $\lesssim 1/T$, and increasing the number of terms subtracted increases the power law scaling at low frequencies. However, there is also substantial reduction in the strain at large frequencies, although the noise strain is larger and therefore these decades in frequency contribute less to the SNR than the frequencies $\sim 1/T$. For a rough estimate of the SNR one simply needs to estimate the area between the signal and noise strain curves in Fig. (4) near $\mathfrak{f} \sim 1/T$.

We now discuss in more detail how the subtracted signals shown in Fig. (4) were obtained. Consider first the deterministic Doppler and Shapiro signals. The raw signal, $\delta\phi$, is found by integrating Eqs. (19), (24), and the subtracted signal, $h$, is subsequently computed by Eq. (B6). The pulsar term SNR from Eq. (31) for a subtracted and centered ($t_0 = T/2$) Doppler signal ($\text{SNR}_{\text{det, D, P}}$), in the $b \ll vT$ limit, is given by[4]

$$\text{SNR}_{\text{det},D,P} \left( t_0 = \frac{T}{2} \right) = \frac{GM}{16\sqrt{3}} \frac{T^{\frac{3}{2}}}{t_{\text{rms}} \sqrt{\Delta t}} \max_{\{I, i\}} \left[ \frac{\left| \hat{\mathbf{d}}_I \cdot \hat{\mathbf{b}}_i \right|}{b_i v_i} \right] \,, \tag{43}$$

where the maximum is taken over all events, $i$, in all the pulsars, $I$. This subtracted result is a factor of 8 smaller compared to Ref. [30] where no subtraction was done. Likewise, the subtracted pulsar term SNR for a subtracted and centered ($t_0 = T/2$) Shapiro signal ($\text{SNR}_{\text{det},S,P}$), in the $b \ll \overline{v}_\perp T$ limit, is given by

$$\text{SNR}_{\text{det},S,P} \left( t_0 = \frac{T}{2} \right) \approx 1.33 \times \frac{GM}{t_{\text{rms}}} \sqrt{\frac{T}{\Delta t}} \,, \tag{44}$$

---

[4] In the $b \ll vT$ limit the Doppler signal is a step function, as seen in Eq. (24), similar to the gravitational wave memory effect discussed in Ref. [55]. We also observe that the deterministic SNR is peaked at a signal offset of $t_0/T = 1/2 \pm 1/(2\sqrt{5})$.

which is a factor of 0.24 smaller in Ref. [30]. Subtraction has a larger effect on the Doppler delay because the Shapiro signal is much more peaked, and therefore less susceptible to subtraction. The projected constraints for the deterministic signals shown in our reach plots have been appropriately rescaled to account for this $\mathcal{O}(1)$ change in the SNR. Also note that we have considered a centered signal, $t_0 = T/2$, to mirror the analysis done in [30]. A more accurate analysis would include the $t_0$ dependence when computing the SNR via a Monte Carlo (MC) simulation; we expect this effect to be small, however, as the SNR only decreases rapidly when $t_0$ is near the observation edge: $t_0 = 0$ or $t_0 = T$.

Lastly, in Ref. [30] the Doppler Earth term was considered subdominant compared to the pulsar term and ignored. This is true for PBHs, as the PBH closest to the Earth is farther than the one closest to any pulsar. However, this effect is compensated by a factor of $N_P$ in the Earth term SNR, and for more diffuse subhalos this increase in the minimum impact parameter extends the reach, as more subhalo mass is contained within the impact parameter. The Earth term will therefore be more sensitive to diffuse subhalos relative to the pulsar term, as seen in Fig. (2). The Earth term SNR, for a centered Doppler signal, is given by

$$
\mathrm{SNR}_{\mathrm{det,\ D,\ E}}\left(t_0 = \frac{T}{2}\right) = \frac{\sqrt{N_P}}{48}\frac{GMT^{\frac{3}{2}}}{t_{\mathrm{rms}}\sqrt{\Delta t}}\max_{\{i\}}\left[\frac{1}{b_i v_i}\right]\,,
\tag{45}
$$

where $i$ is over all of the events near the Earth, and the average over the pulsar positions gives a factor of $1/\sqrt{3}$ relative to Eq. (43).

Next we discuss and derive in detail the features from the new analysis of the stochastic signal (curves labeled 'stoch' in Fig. (2)), which can be understood from the distance scales discussed in the previous subsection. First, the right-hand side of the 'stoch' reach curves in Fig. (2), $f_{\mathrm{DM}}^R$, is derived from the requirement that events transit the Earth-pulsar system during the observation time, $i.e.$ $b_{\mathrm{min}}^{D(S)} < \overline{v}_{(\perp)}T$. Utilizing Eq. (36) we obtain:

$$
f_S^R \sim 0.5\left(\frac{M}{10^{-3}M_\odot}\right)\left(\frac{20\ \mathrm{year}}{T}\right)^2\left(\frac{5\ \mathrm{kpc}}{z_0}\right)
\tag{46}
$$

$$
f_D^R \sim 0.2\left(\frac{M}{10^{-9}\,M_\odot}\right)\left(\frac{20\ \mathrm{year}}{T}\right)^3.
\tag{47}
$$

The opposite (left-hand) side of the 'Stoch' constraints are derived by the strength of the
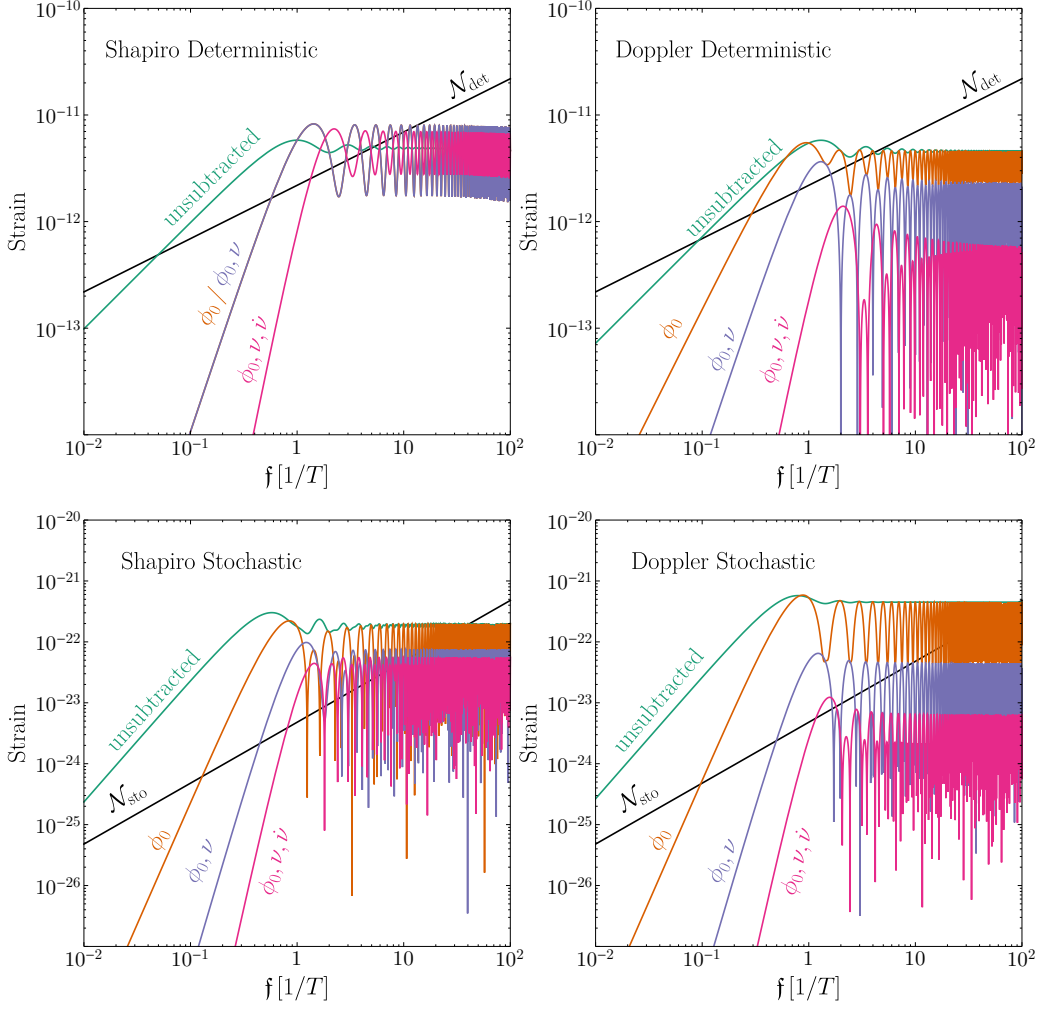
FIG. 4: Signal strain for the deterministic and stochastic signals from monochromatic PBHs compared to timing noise, $\mathcal{N}$, with parameters $N_P = 200$, $t_{\rm rms} = 50$ ns, $\Delta t = 2$ weeks and $T = 20$ years. We illustrate the effect of subtraction due to fitting the terms in the timing model, $\phi_0, \nu, \dot{\nu}$. $M = 10^{-3} M_\odot$, $b = 10^{-4}\, \bar{v}_\perp T$ are assumed for the Shapiro signals, $M = 10^{-9} M_\odot$ for the stochastic Doppler signal, and $M = 10^{-8} M_\odot$ for the deterministic Doppler signal.

SNR. Evaluating Eqs. (20), (25),

$$R_S(t,t') = 4\pi f G^2 \rho_{\rm DM} M \nu^2 \overline{v}_\perp^2 z_0 T^2 \int_0^{\frac{t}{T}} \int_0^{\frac{t'}{T}} dx_1 dx_2 \log\left(\frac{4+(x_1-x_2)^2}{4\left(\frac{b_{S,\min}}{\overline{v}_\perp T}\right)^2 + (x_1-x_2)^2}\right)$$

$$\equiv 4\pi f G^2 \rho_{\rm DM} M \nu^2 \overline{v}_\perp^2 z_0 T^2 \, C\left(\frac{t}{T}, \frac{t'}{T}, \frac{b_{\min}^S}{\overline{v}_\perp T}\right) \tag{48}$$

$$R_D^{IJ}(t,t') = \frac{4\pi f G^2 \rho_{\rm DM} M \nu^2}{\overline{v}} \left(\hat{\mathbf{d}}_I \cdot \hat{\mathbf{d}}_J\right) B(t,t') \log\left(\frac{vT}{b_{\min}^D}\right), \tag{49}$$

which must then be substituted in to Eq. (30) before computing the SNR with Eqs. (33), (34). Beginning with the Shapiro stochastic signal, which has an unsubtracted correlator given in Eq. (48), the SNR in Eq. (33) is approximately,

$$\text{SNR}_S \approx \left(9.4 \times 10^{-2}\right) f \frac{\sqrt{N_P} G^2 \rho_{\rm DM} M \overline{v}_\perp^2 z_0 T^3}{t_{\rm rms}^2 \Delta t}, \tag{50}$$

in the $b_{\min} \ll \overline{v}_\perp T$ limit. Setting $\text{SNR}_S = 2$, gives the left-hand side of the 'ShapStoch' constraint, $f_{\rm DM(S)}^L$,

$$f_S^L \approx 2.3 \left(\frac{200}{N_P}\right)^{\frac{1}{2}} \left(\frac{10^{-4} M_\odot}{M}\right) \left(\frac{t_{\rm rms}}{50 \text{ ns}}\right)^2 \left(\frac{5 \text{ kpc}}{z_0}\right) \left(\frac{20 \text{ year}}{T}\right)^3. \tag{51}$$

The left-hand side of the 'DopStoch' curves in Fig. (2) are derived similarly. The average over the pulsar positions contributes a factor of

$$\left\langle \left(\hat{\mathbf{d}}_I \cdot \hat{\mathbf{d}}_J\right)^2 \right\rangle_{\mathcal{P}} = \frac{1}{3}, \tag{52}$$

and the SNR in Eq. (34) is then,

$$\text{SNR}_D \approx \left(1.4 \times 10^{-3}\right) f \rho_{\rm DM} \frac{N_P G^2 T^4 M}{\Delta t \, t_{\rm rms}^2 \overline{v}} \log\left(\frac{\overline{v} T}{b_{\min}^D}\right). \tag{53}$$

The constraint is again derived from $\text{SNR}_D = 2$. There is no simple scaling law as there is $f$ dependence inside the logarithm.

In deriving analytic results for the stochastic signals, we are using the expressions for $b_{\min}$ from Eq. (36), and we justify their use here. The SNRs in Eqs. (33), (34) have been calculated by averaging over the subhalo random variables, $e.g.$ $\mathbf{b}_i$. The subtlety is that the average can be skewed by unlikely values. For example, the Doppler delay $\text{SNR}^2 \propto \langle b^{-2} \rangle$, diverging with the lower cut-off on the $b$ integral; even though small values of $b$ are unlikely, their effect on the SNR is large enough to skew the average. This calls into question the robustness of our analytic prescription with $b_{\min}$ calculated from Eq. (36).
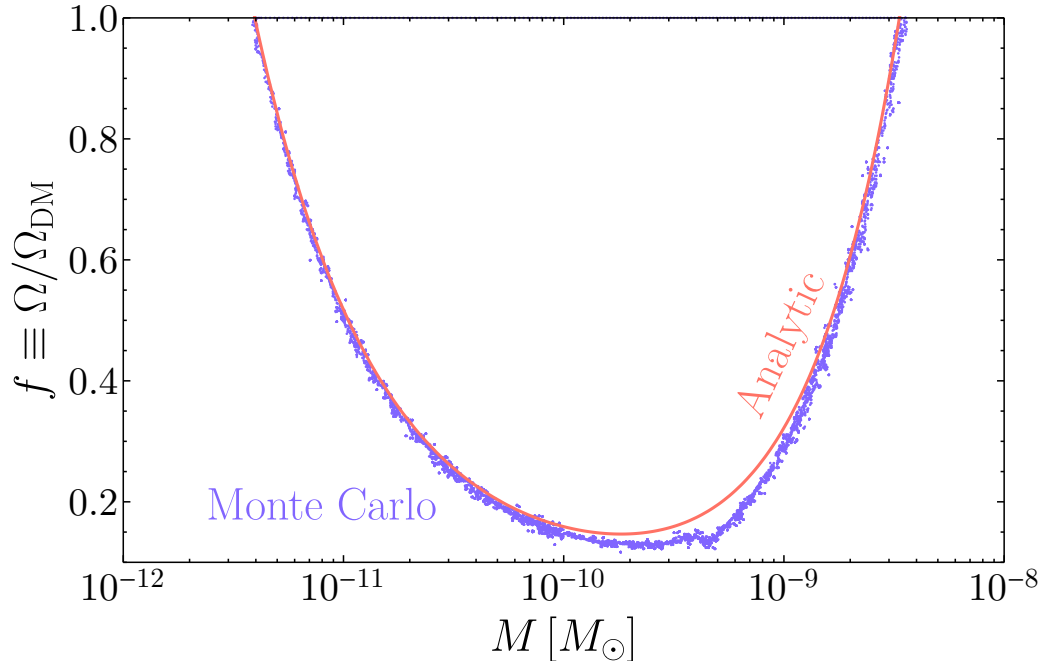
FIG. 5: Comparison of constraints for the Doppler stochastic signal using the Monte Carlo (MC) and analytic approaches in the PBH limit. The MC derives constraints from the simulated $10^{\text{th}}$ percentile SNR, whereas the analytic constraint imposes $b > b_{\text{min}}^D$, where $b_{\text{min}}^D$ is the $90^{\text{th}}$ percentile minimum impact parameter. Constraints are created assuming $v = 10^{-3}$ and PTA parameters of $N_P = 200, t_{\text{rms}} = 10$ ns, $\Delta t = 1$ week.

The solution is to calculate a skew-independent statistic of the signal, such as a percentile, which we obtain via a Monte Carlo (MC) simulation. We find good agreement between the reach calculated analytically, with $b_{\text{min}}$ from Eq. (36), and the $10^{\text{th}}$ percentile SNR computed from the MC. The results are compared in Fig. (5). Because of the strong agreement, we will proceed to use the analytic results in the remainder of this paper.

### 2. NFW Subhalo

We now turn to less concentrated subhalos and, for concreteness, consider an NFW density profile,

$$\rho(r, c) = \frac{4\rho_s(c)}{(r/r_s)(1 + r/r_s)^2}, \tag{54}$$

where $r_s$ is the scale radius, and the scale density $\rho_s(c) = \rho(r_s, c)$ is given by,

$$\rho_s(c) = \frac{50c^3\rho_c}{3(\log(c+1) - c/(1+c))}, \tag{55}$$

where $\rho_c$ is the critical density and $c$ is the concentration parameter. As emphasized in Eq. (37), for a given subhalo mass $M$, larger concentration parameters lead to more compact subhalos, with the PBH limit $c \to \infty$. $N$-body simulations of CDM subhalos indicate that $10 \lesssim c \lesssim 100$, but more concentrated subhalos can be formed from earlier collapse, as $c \propto 1 + z_{\text{col}}$, where $z_{\text{col}}$ is the collapse redshift.

With the density profile defined, the steps to calculating the constraints are the same as that of a point-like subhalo but now the form factor $\mathcal{F}$ appears in Eqs. (20), (25); the results are similar to those in Eqs. (48), (49), except now we do not take $\mathcal{F} \to 1$:

$$R_S(t, t') = 32\pi G^2 M \, f \, \rho_{\text{DM}} \nu^2 z_0 \int db \, b \, A\left(t, t', \frac{b}{\overline{v}_\perp}\right) \left(\mathcal{F}\left(\frac{b}{r_v}, c\right)\right)^2 \tag{56}$$

$$R_D^{IJ}(t, t') = \frac{4\pi G^2 M \, f \, \rho_{\text{DM}} \nu^2}{\overline{v}} \left(\hat{\mathbf{d}}^I \cdot \hat{\mathbf{d}}^J\right) B(t, t') \int \frac{db}{b} \left(\mathcal{F}\left(\frac{b}{r_v}, c\right)\right)^2. \tag{57}$$

The result of computing the constraints on the dark matter mass fraction, $f$, are shown in Fig. (2) for $c = 10, 100, 10^4$. Because the signal depends on the integral from $b_{\min}$ to $b_{\max}$, the difference in constraints between finite $c$ and $c \to \infty$ can be understood from the behavior of the form factor $\mathcal{F}$ over this range of impact parameters. This is shown in Fig. (6) (obtained from Eq. (18) and simplified analytically in Appendix D) as a function of $b/r_v$ and $c$. $\mathcal{F} \to 1$ when $b/r_v > 1$ and the masses can be treated as point-like. As long as $b/r_v > 1/c$, $\mathcal{F}$ remains relatively large. Only once $b/r_v < 1/c$ does $\mathcal{F}$ drop rapidly and the signal becomes very weak. This relative insensitivity to the subhalo radius allows PTAs to constrain a wide range of concentration parameters.

For example, the 'DopStoch' constraints are relatively $c$-independent because, even for $c \gtrsim 10$, $r_v/c \lesssim b_{\min}^D$ (see Fig. (1)), such that $\mathcal{F} \sim 1$ over the integration region. On the other hand, the Shapiro search is only sensitive to larger mass subhalos (as seen in Fig. (2)) that have larger radii. $b/r_v$ is typically thus much smaller, and $\mathcal{F}$ is rapidly suppressed, as shown in Fig. (6), so that the Shapiro search has much less reach to low concentration subhalos.

In summary, in order for subhalos to be sufficiently compact to be observable by PTAs, we require their scale radius be smaller than the radius of the observing volume, $r_s < b_{\max}$, which is only satisfied for large concentration parameters:

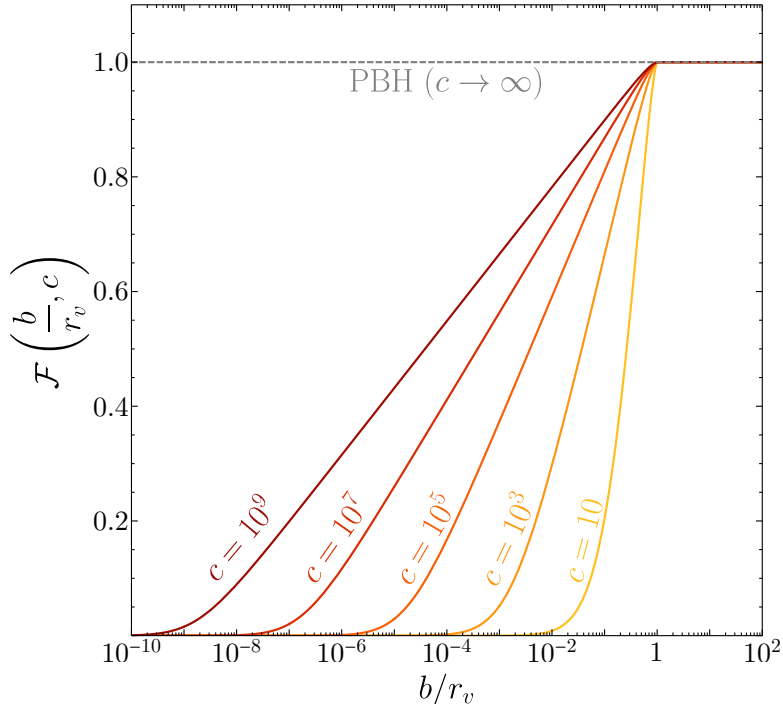$$c \gg 4 \times 10^3 \left(\frac{20 \text{ yr}}{T}\right) \left(\frac{M}{M_\odot}\right)^{\frac{1}{3}}. \tag{58}$$

FIG. 6: Subhalo form factor $\mathcal{F}$ for various concentration parameters $c$, as a function of the subhalo impact parameter $b$ and virial radius $r_v$. When $b/r_v > 1$, the subhalo is point-like and $\mathcal{F} \to 1$. As $b/r_v$ drops below 1, $\mathcal{F}$ drops only logarithmically with $b/r_v$ until $b/r_v < 1/c$, where the form factor rapidly goes to zero. A PTA can only observe subhalos with $b < b_{\max} = vT$, set by the PTA observing time $T$. Therefore more massive subhalos, having larger virial radii, can only be constrained if they have sufficiently large concentration parameter.

Overall, this means that PTAs are particularly powerful probes for low concentration subhalos with $M \lesssim 10^{-6}\ M_\odot$.

In Refs. [7, 12] the effect of tidal stripping on diffuse subhalos was incorporated by assuming that only cores survive until late times. This was modeled by an abrupt fall-off in density outside the scale radius $r_s$ in Eq. (54). For these subhalos, $M = M_s$ where $M_s$ is the mass contained inside the radius $r_s$. Constraints from different probes were projected in the $\rho_s$ versus $M_s$ plane, where $\rho_s = \rho(r_s)$ is the scale density, for monochromatic subhalos which make up a fraction $f$ of the dark matter. For a direct comparison with other futuristic proposals, we show constraints from PTAs with optimistic pulsar parameters in Fig. (7). For each probe, we show two contours corresponding to the minimum scale density that can be probed for a particular core mass, $M_s$, for $f = 1$ and $f = 0.3$. Also shown are projections from photometric lensing [32], as computed in [12], $f = 1$ constraints from astrometric $\alpha$ and $\mu$ lensing from Gaia data [18–20], and diffraction of gravitational wave from BH mergers observable at aLIGO [7, 56]

in dashed gray. The hatched region corresponds to subhalos with $10 \leq c \leq 100$, to account for $\Lambda$CDM-like subhalos at masses much lower than those typically found in simulations [57, 58]. The black dot-dashed line corresponds to the local DM density; subhalos cannot make up all of dark matter, without sufficiently overlapping, below this line. The robustness of constraints with respect to the concentration parameter, as seen in Figs. (2), (3), translates to sensitivities to very small $\rho_s$ in Fig. (7). Remarkably, with optimistic PTA parameters, a combination of the different Doppler constraints will be sensitive to a monochromatic mass distribution with even the most diffuse cores, in a mass window spanning as low as $10^{-13} M_\odot$ to well above a solar mass.

## B.   Extended Halo Mass Functions

We now consider a mass distribution that is not simply monochromatic, focusing on the highly relevant case of the Cold Dark Matter (CDM) paradigm where scale invariant perturbations are seeded by inflation. The (nearly) scale invariant nature of the perturbations fairly firmly fixes the spectrum of the Halo Mass Function (HMF), which can be parameterized as

$$F(M) \equiv \frac{dn}{d \log M} = \frac{f \rho_{\mathrm{DM}}}{\mathcal{N}(M_{\mathrm{min}}, M_{\mathrm{max}}, 1-\beta)} M^{1-\beta} \Theta(M_{\mathrm{max}} - M) \Theta(M - M_{\mathrm{min}}) \qquad (59)$$

$$\mathcal{N}(M_{\mathrm{min}}, M_{\mathrm{max}}, \alpha) \equiv \int_{M_{\mathrm{min}}}^{M_{\mathrm{max}}} M^\alpha dM = \begin{cases} \log\left(\frac{M_{\mathrm{max}}}{M_{\mathrm{min}}}\right) & \alpha = -1 \\ \frac{1}{\alpha+1}\left(M_{\mathrm{max}}^{1+\alpha} - M_{\mathrm{min}}^{1+\alpha}\right) & \alpha \neq -1 \end{cases}, \qquad (60)$$

where the overall normalization, $\mathcal{N}$, is found by requiring that the expected matter density is equal to its measured value, $f \rho_{\mathrm{DM}} = \int dM\, F(M)$. The total number of subhalos within a volume $V$ is given by,

$$N = V \int_{M_{\mathrm{min}}}^{M_{\mathrm{max}}} \frac{dn}{dM} dM = f \rho_{\mathrm{DM}} V \frac{\mathcal{N}(M_{\mathrm{min}}, M_{\mathrm{max}}, -\beta)}{\mathcal{N}(M_{\mathrm{min}}, M_{\mathrm{max}}, -\beta+1)}. \qquad (61)$$

One can estimate $\beta$ with the standard Press-Schechter theory [59], assuming a scale invariant primordial power spectrum. For large $k$ the power spectrum today scales as $\sim k^{-3}$, up to the free-streaming scale (which sets $M_{\mathrm{min}}$), corresponding to $\beta = 2$. This scale invariant spectrum gives equal mass density in equal logarithmic intervals.

Cosmological $\Lambda$-CDM $N$-body simulations have, however, made more precise estimates which indicate that $\beta \approx 1.9$ [60, 61]. As innocent as this difference seems, it has a large impact on the constraints that can be placed at low subhalo masses where PTA constraints are most powerful. This is because, for $\beta = 1.9$, the mass density is dominated by the large mass subhalos. Since
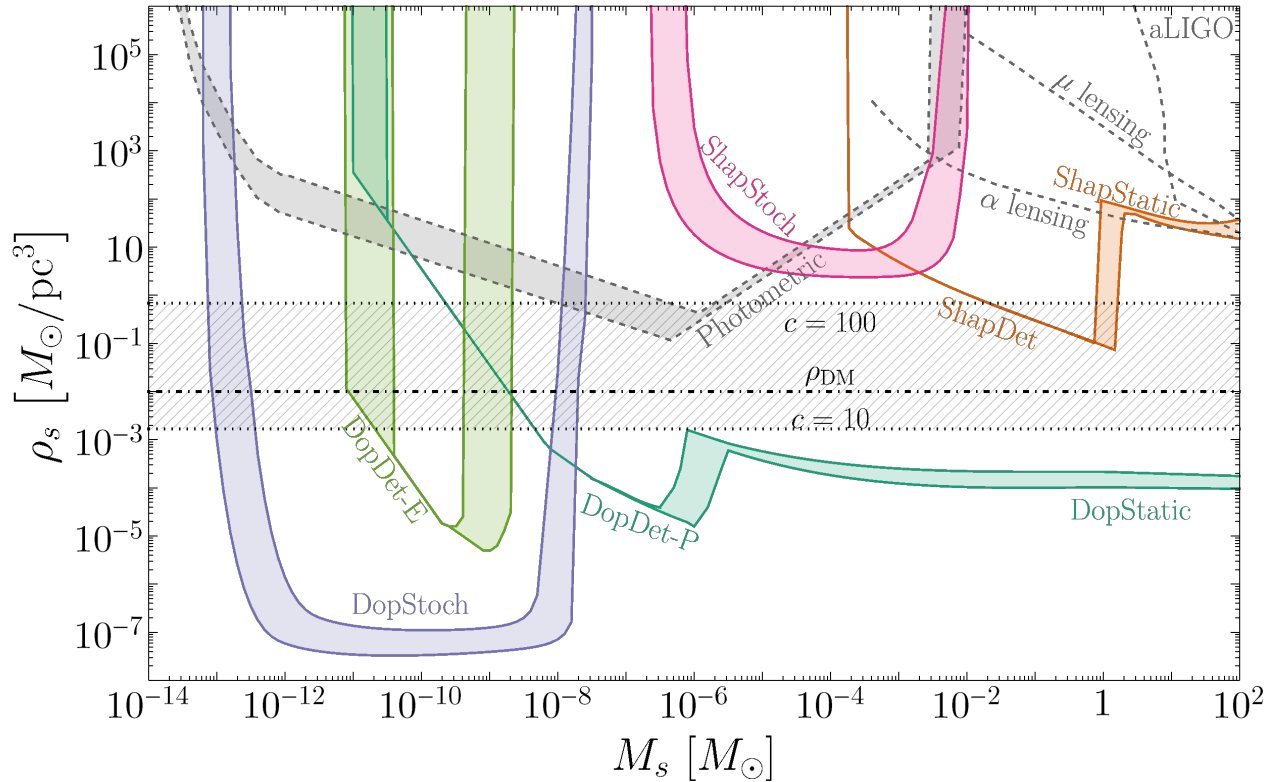
FIG. 7: PTA constraints on a monochromatic mass distribution of 'core-only' subhalos with scale density $\rho_s$, as a function of core mass $M_s$ arising from optimistic PTA parameters. PTA constraints have the same color scheme as in Fig. (3). To show the dependence on the dark matter fraction $f$ in such subhalos, we show a band with $0.3 \leq f \leq 1$. Scale densities corresponding to $10 \leq c \leq 1$ in the $\Lambda$CDM range are shown as a hatched region (though note this is not the $\Lambda$CDM model, which features a broad spectrum of subhalo masses). Halos below the dot-dashed-black line corresponding to $\rho_s = \rho_{\mathrm{DM}}$ cannot make up $f = 1$ . Also shown in dashed-gray are projections from photometric lensing [12, 32], $\alpha$ and $\mu$ lensing from astrometric lensing in Gaia data [18–20] and diffraction of BH mergers observable in aLIGO[7, 56].

$M_{\mathrm{min}}, M_{\mathrm{max}}$ and $\beta$ are largely model-dependent, we allow them to vary, but choose parameters that do not dramatically vary from a scale invariant spectrum.

$N$-body simulations, having dark matter only, favor constituent subhalos with an NFW profile.[5] Furthermore, CDM subhalos, obtained from galactic simulations [58, 60, 62, 63], are typically quoted to have concentration parameters with $c \gtrsim 50$ for subhalos below $M_\odot$.

---

[5] Baryons tend to change these profiles near the core of the subhalo, but the small subhalos that we consider here do not hold baryons

These models, however, are usually obtained from simulation data with $M \gtrsim M_\odot$. More recent simulations which study lower mass subhalos suggest that, below $M \sim 10^{-3} M_\odot$, the concentration parameter decreases with decreasing mass [58]. While the 'Doppler-stoch' search will have some reach for $c \sim 10$, it will only be at masses much smaller than the typical minimum mass of CDM subhalos from WIMP dark matter which have $M_{\min} \sim 10^{-6} M_\odot$. Given the uncertainty on the concentration parameters of low mass subhalos, we show our results in Figs. (8), (9) for $c = 10, 100, 10^4$, and the $c \to \infty$, PBH limit for comparison.

The constraints shown in Figs. (8), (9) were derived from the monochromatic mass distribution, following Ref. [64]. The method advocated there re-weights the monochromatic distribution constraints shown in Fig. (2), $f_{\mathrm{mono}}(M)$, according to the relation

$$\frac{\rho_{\mathrm{DM}}}{f} = \int dM \frac{F(M)}{f_{\mathrm{mono}}(M)} \,. \tag{62}$$

To check the validity of this approximation, we compare the results obtained utilizing this analytic prescription with a Monte Carlo, in Fig. (11). We see good agreement for more than one HMF, and proceed to use this analytic formula in our main results, Figs. (8), (9).

The first row in Fig. (8) shows constraints for $\beta = 2$, and the second $\beta = 1.9$. We take $M_{\max} = 10^{12} M_\odot$ (the Milky Way galaxy mass) and $10^8 M_\odot$ to show the dependence on $M_{\max}$ in the left and middle columns, and lastly $M_{\max} = 10^3 M_{\min}$, in the right column. Such narrow HMFs can be produced in theories with peaks in the primordial power spectrum [2–11]. We will use a shorthand for the variety of search types: 'DopDet-P+Static' for the combined deterministic and static Doppler searches that involve the pulsar term, 'ShapDet+Static' for the combined deterministic and static Shapiro searches, 'DopStoch' for the stochastic Doppler signal, and 'ShapStoch' for stochastic Shapiro signal. The 'DopDet-P+Static' and 'ShapDet+Static' curves were derived in Ref. [30] and corrected due to the subtraction procedure here, while 'DopStoch' and 'ShapStoch' were derived here.

It is clear from Eq. (62) that there is enhanced sensitivity to Halo Mass Functions $F(M)$ with large support to masses for which a particular type of PTA search is sensitive. We show this in Fig. (10) by comparing PTA search constraints from Fig. (3) and the mass fraction in $[M, 10M]$, $\int_M^{10M} (F(M')/\rho_{\mathrm{DM}}) \, dM'$, for a few different HMFs. This translates to sensitivity in Fig. (8) when $M_{\min}$ and $M_{\max}$ encapsulate the entire mass sensitivity range of a particular type of search. In addition, if $M_{\max} - M_{\min}$ is substantially larger than the sensitivity range of a particular search, this results in the reduction of reach to such dark matter masses; this is simply because the fraction of dark matter in the sensitivity range is diluted. Across the board, this is seen in weaker limits for $M_{\max} = 10^{12} \, M_\odot$ compared to $M_{\max} = 10^8 \, M_\odot$. However this reduction scales only as $f \sim 1/\log(M_{\max}/M_{\min})$ for $\beta = 2$, such that there is only a logarithmic decline for small enough $M_{\min}$ in the top row, left and middle panels of Fig. (8). For $\beta = 1.9$,
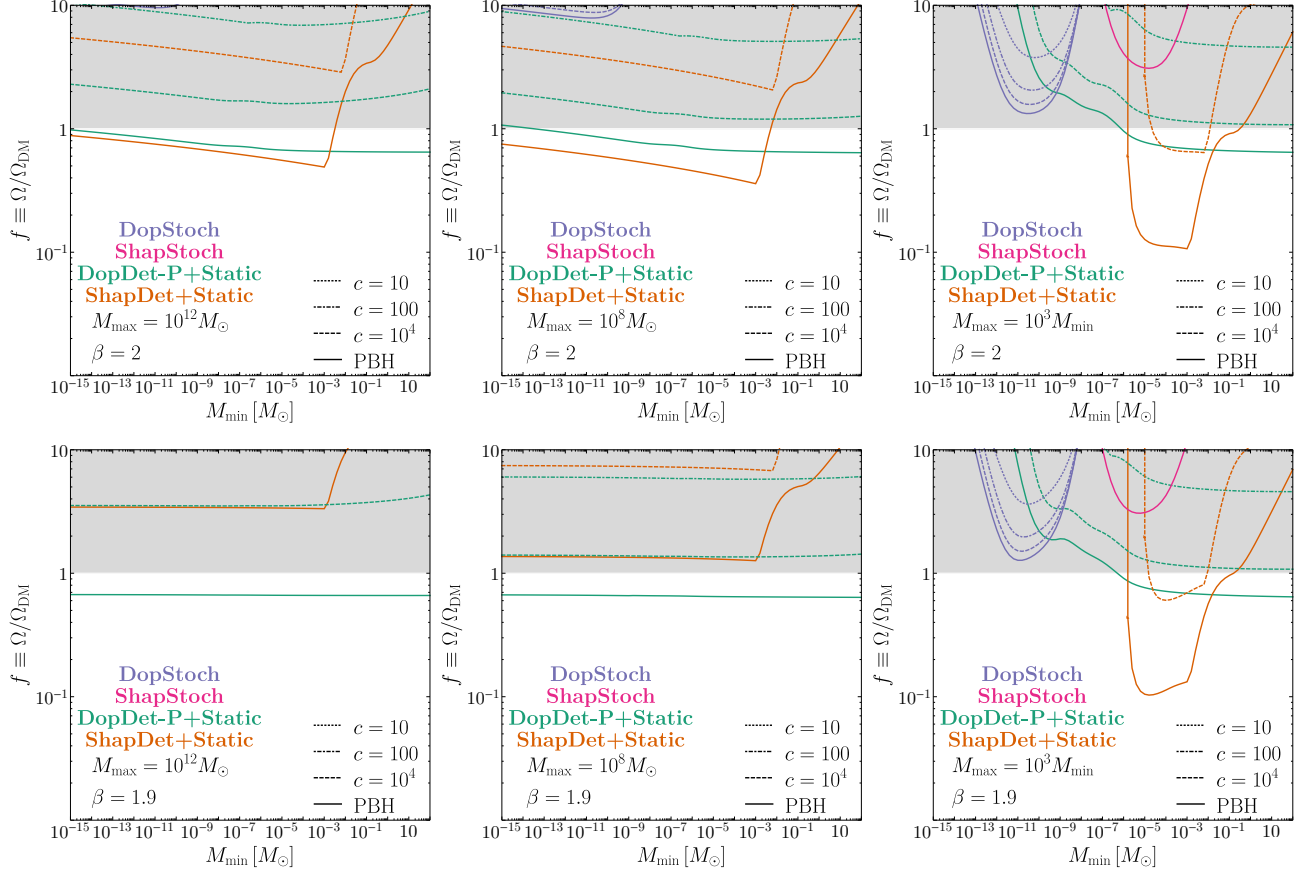
FIG. 8: Constraint on the fraction of dark matter $f = \Omega/\Omega_{\mathrm{DM}}$ in a Halo Mass Function (HMF) parameterized by $M_{\min}, M_{\max}$ and $\beta$, as in Eq. (60). We assume SKA-like PTA parameters of $N_P = 200$, $T = 20$ yr, $t_{\mathrm{rms}} = 50$ ns, $\Delta t = 2$ weeks, $z_0 = 5$ kpc. We show constraints from four different signal regimes: combined pulsar term deterministic and static Doppler (DopDet-P+Static), combined deterministic and static Shapiro (ShapDet+Static), stochastic Doppler (DopStoch), stochastic Shapiro (ShapStoch), with four different concentrations $c = 10, 100, 10^4$, and the PBH-like $c \to \infty$.

$f \sim 1/M_{\max}^{0.1}$ so that the reach curves flatten out for small enough $M_{\min}$ in the left and middle panels.

For $M_{\max} = 10^{12}\ M_\odot$ and $M_{\max} = 10^8\ M_\odot$ with $\beta = 2$, non-trivial constraints will be set with an SKA-like PTA for $c = 10^4$ and above, from the deterministic Doppler and Shapiro constraints derived in Ref. [30], primarily due to the sensitivity to a wide range in mass. There is less reach for $\beta = 1.9$ due to the skew of the HMF towards larger mass subhalos, as can be seen in Fig. (10). There are more optimistic projections when $M_{\max} = 10^3 M_{\min}$, even for small concentration parameters.
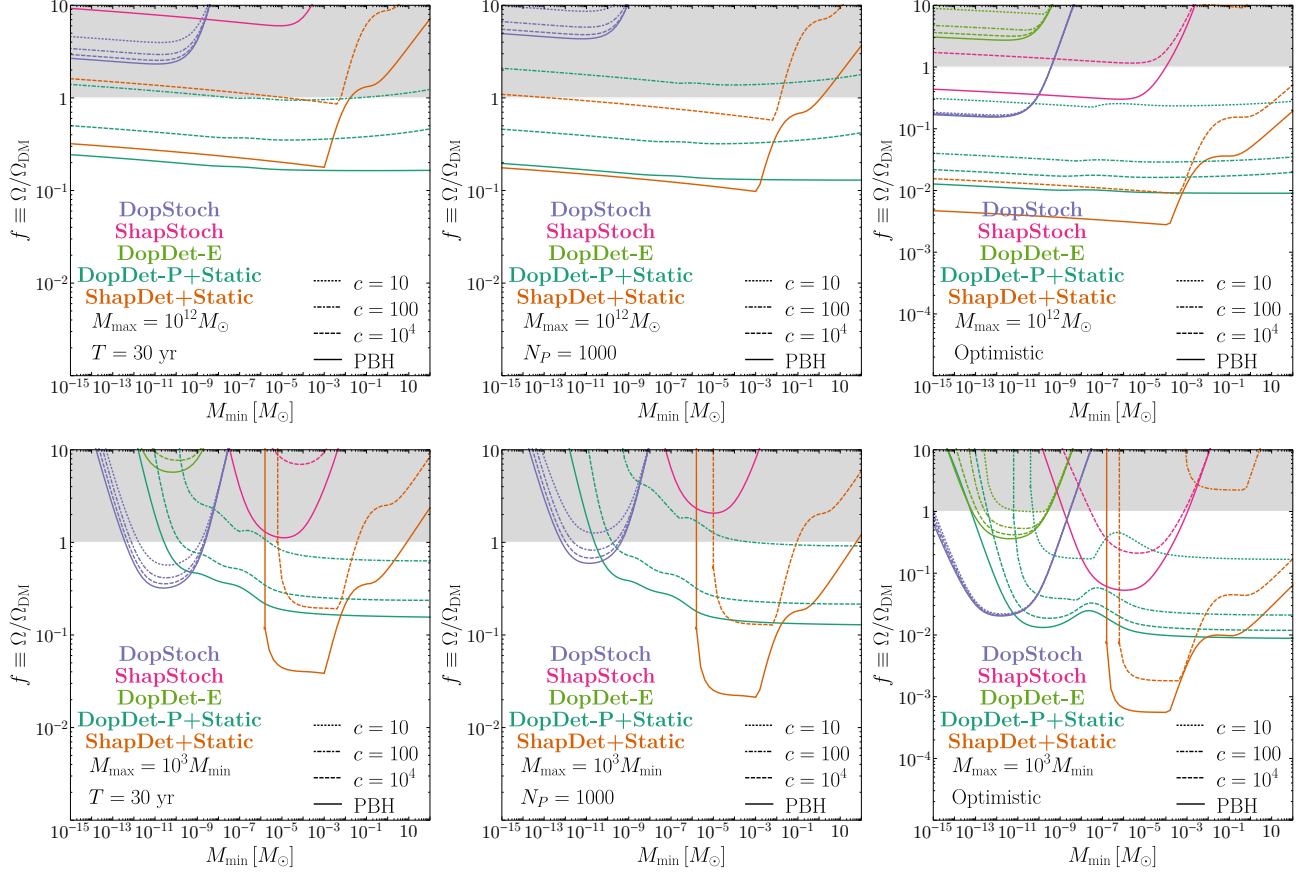
FIG. 9: Constraint on the fraction of dark matter $f = \Omega/\Omega_{\rm DM}$ in a Halo Mass Function (HMF) parameterized by $M_{\rm min}, M_{\rm max}$ and $\beta = 2$, as in Eq. (60). In the three columns we vary the PTA capability parameters to show what will be necessary to reach a CDM-like HMF. The left column assumes an SKA-like PTA with $T = 30$ year; the middle column an SKA-like PTA with $N_P = 1000$; the right column shows a futuristic PTA with optimistic parameters: $N_P = 1000$, $T = 30$ yr, $t_{\rm rms} = 10$ ns, $\Delta t = 1$ week, $z_0 = 10$ kpc. As in Fig. (8) we show constraints from four different signal regimes: combined deterministic and static Doppler pulsar term (DopDet-P+Static), combined deterministic and static Shapiro (ShapDet+Static), stochastic Doppler (DopStoch), stochastic Shapiro (ShapStoch). Additionally, the deterministic Earth term Doppler DopDet-E is shown. Four choices of concentration parameter $c = 10, 100, 10^4$, and the PBH limit $c \to \infty$ are shown.

In order to demonstrate what will be necessary from PTAs to probe CDM-like substructure, we also show improvements in projected reach as some PTA parameters are dialed to more optimistic values in Fig. (9). We take the SKA-like parameters discussed in Sec. II C and increase the observing time to 30 years in the left panel, the number of pulsars to 1000 in
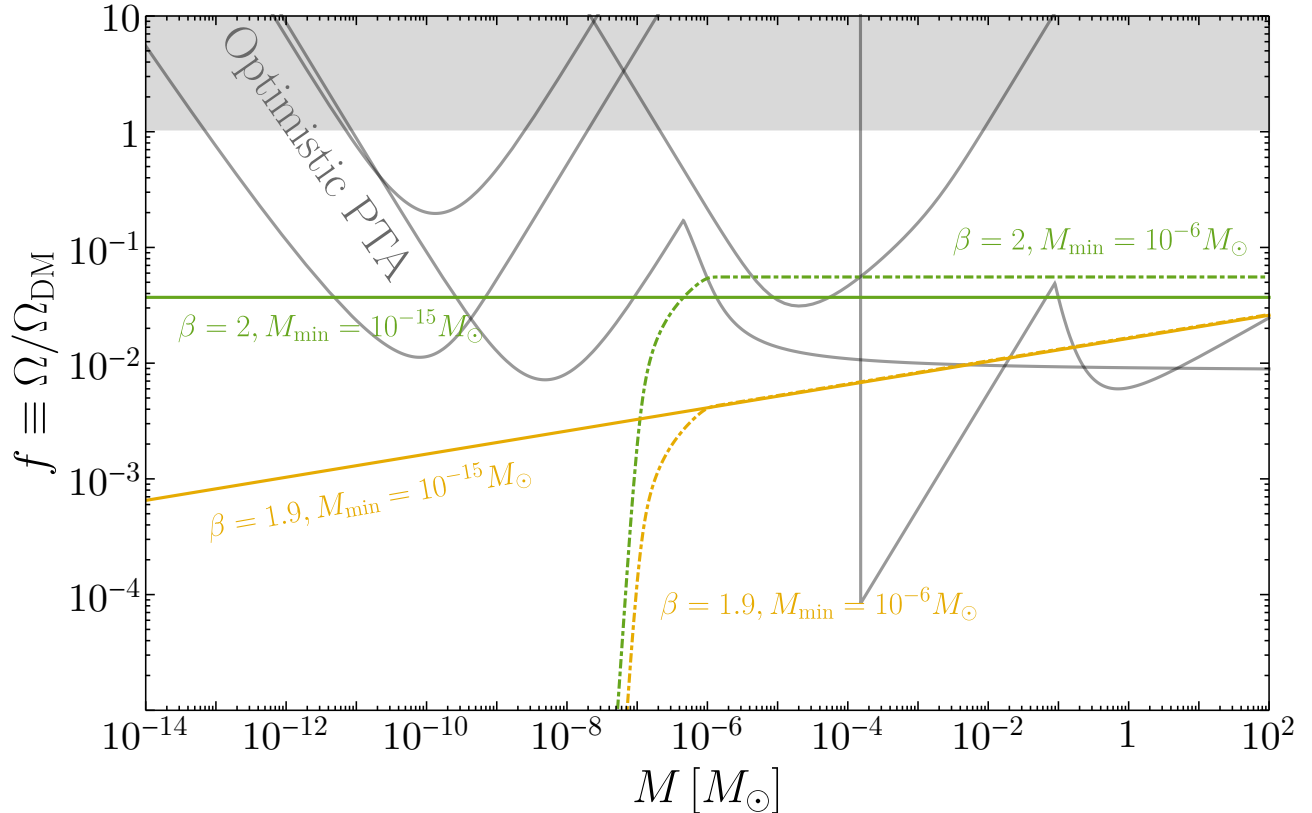
FIG. 10: Comparison of the mass fraction of dark matter in $[M, 10M]$, $\int_M^{10M}(F(M')/\rho_{DM})\,dM'$, and monochromatic PBH constraints (gray) from Fig. (3). The Halo Mass Function is labeled by a spectral index $\beta$ and minimum subhalo mass $M_{min}$ where the drop-off in $M_{min}$ has been smoothed from Eq. (60). The values of $\beta$ shown are motivated by scale invariant density perturbations seeded during inflation (*i.e.* CDM HMF). Comparing the gray PTA reach curves against the prevalence of subhalos in a particular mass bin (green or yellow curves) roughly shows for which mass subhalos PTA reach, via the procedure summarized in Eq. (62), is viable.

the middle panel and finally a combined set of optimistic parameters ($N_P = 1000$, $T = 30$ yr, $t_{rms} = 10$ ns, $\Delta t = 1$ week, $z_0 = 10$ kpc) in the right panel. 30 years of observation time will allow reach to HMFs with $\beta = 2$ and subhalos with $c = 100$ via the deterministic event Doppler constraints from Ref. [30]. With the optimistic set of PTA parameters shown in the right-hand panel, the stochastic Doppler signal can reach subhalos with concentration parameters $c = 10$ composing less than two percent of dark matter.
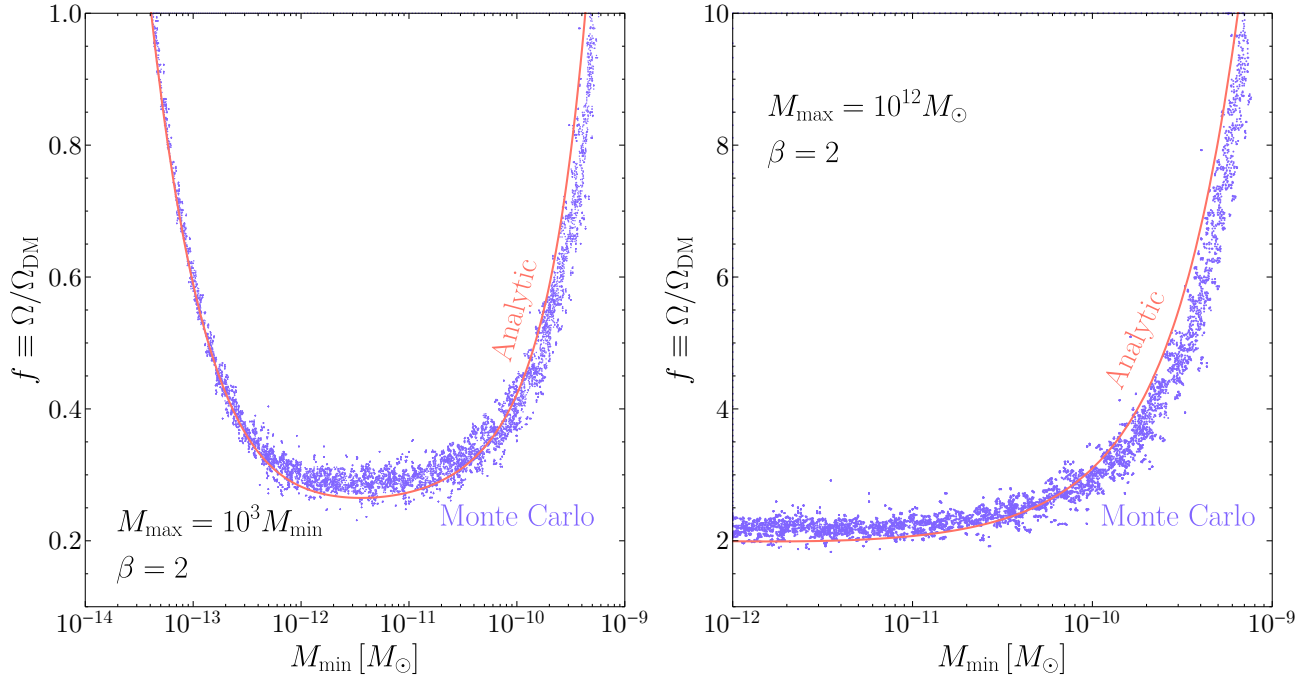
FIG. 11: Comparison of constraints for the 'DopStoch' signal using the Monte Carlo (MC) and analytic approach from Ref. [64]. The MC calculates constraints from the the $10^{\text{th}}$ percentile SNR, while the analytic approach re-integrates the monochromatic results using Eq. (62). The two methods are nevertheless in good agreement. Constraints are again created assuming $v = 10^{-3}$ and PTA parameters of $N_P = 200, t_{\text{rms}} = 10$ ns, $\Delta t = 1$ week.

## C.   Backgrounds

Similar to direct detection of dark matter, non-observation of a signal allows one to set constraints, but a claim for discovery requires careful noise discrimination. While the noise modeling adopted in this work assumes only white noise, red-noise has been observed in some pulsars primarily due to intrinsic deviations from the spin-down timing model [65] and fluctuations in the dispersion measure [66]. Similar to stochastic gravitational waves, the signal can be differentiated from the background exploiting the differences in the power spectral indices [65]. The contributions from the interstellar medium are dependent on the pulsar light frequency and could be corrected for without removing the pulsar-frequency-independent dark matter signal. Furthermore in the case of the Earth term, variations unique to individual pulsars are suppressed when considering cross-correlations between pulsars.

Gravitational waves from supermassive black hole merger events are expected to be detected before a dark matter signal and hence will constitute another background to a dark matter signal [67]. A single merger event can be differentiated from the deterministic dark matter subhalo

event via the characteristic signal shape, as outlined in Ref. [30]. The stochastic gravitational wave background will also have a different power spectral index compared to the stochastic dark matter signal analyzed in this work. Furthermore, dark matter signals are dipolar in nature and exhibit characteristically different angular correlations compared to gravitational waves which are quadrupolar signal. See *e.g.* Ref. [68] for the analog of the Hellings and Downs analysis [69] for other signal patterns.

Finally, baryonic objects could cause Doppler and Shapiro delays identical to dark matter subhalos in the mass range of sensitivity. However the baryonic matter is dominated by stars in the solar mass range [70], which would only affect the static signals. A discovery here would necessarily involve supplemental analysis with luminosity discrimination. At lower masses ($10^{-2} - 10^{-1} M_\odot$), sub-stellar objects including brown dwarfs make up less than one percent of the total baryon density. At even lower masses, planets make up less than 0.005% of all the baryons. Furthermore these objects have to be transiting and not bound to the pulsar/solar system in order to mimic a dark matter signal. Nonetheless, pulsars in baryon rich environments might very well display irreducible backgrounds from these loose baryonic objects in certain mass ranges. However, for high latitude pulsars which have been discovered recently [71], the baryonic background will not be limiting in the near future due to the strong limits set on by gravitational MACHO lensing surveys.

## IV. CONCLUSIONS

Pulsar Timing Arrays offer a unique glimpse of dark matter substructure at previously inaccessible small scales. Current constraints from measurements of large scale structure, for diffuse CDM-like subhalos, become weak below $\sim 10^7 \, M_\odot$. For objects as dense as primordial black holes, observational limits from disruption of structure in conjunction with lensing extend the reach down to $\sim 10^{-10} \, M_\odot$, though these limits are rapidly lifted as soon as the subhalos become become even modestly less dense. By contrast, the methods presented here, and in our companion paper [30], reach subhalos having a concentration parameter as small as $c = 10$ and as light as $\sim 10^{-13} \, M_\odot$. PTAs are thus potentially more powerful than other existing or proposed probes of dark matter substructure.

As summarized in Figs. (2), (3), we showed that future PTAs will be able to place strong constraints on the fraction of dark matter in such substructure, even for NFW subhalos having a concentration as small as $c \sim 10$ where lensing cannot reach. We also found that our analysis utilizing multiple transiting subhalos causing a stochastic signal can extend constraints for both Doppler and Shapiro signals by more than an order of magnitude to smaller mass relative to the constraint obtained from a deterministic event of a single transiting subhalo [30], as seen

by comparing the curves labeled 'stoch' to the other curves in Fig. (2).

A definitive goal is to observe substructure consistent with ordinary CDM at as small of a mass scale as possible, to see to what mass scale such substructures persist. For the SKA-like PTA parameters described at the end of Sec. II C, we find this will be difficult unless the power spectrum is significantly more skewed towards low mass subhalos than we expect from standard CDM, as shown in Fig. (8). We showed the requirements from a PTA in Fig. (9) to reach standard CDM. While in principle possible, standard CDM offers a challenge, though one that would offer an unprecedented view if reached.

On the other hand, dark matter often has dynamics that, on small scales, enhance the density perturbations seeded by adiabatic, scale-invariant inflation. The formalism developed here is sufficiently general to account for any HMF. In future work we will apply this to other well motivated forms of substructure, such as axion miniclusters, cosmologies with a period of matter domination, and vector bosons produced during inflation. Observing dark matter substructure on small scales gives a unique window into post-inflationary dynamics and the nature of the dark matter.

## Appendix A: Derivation of the Optimal SNR

We derive the optimal signal-to-noise ratio (SNR) for different signals using a matched filter procedure, similar to Refs. [43, 44, 72]. To generalize our discussion we define the measured signal, $s_I(t)$, the dark matter signal $h_I(t)$, and the noise in the detector, $n_I(t)$, which satisfies the relation $s_I = h_I + n_I$ for pulsar $I$. We will derive the optimal SNR for a deterministic signal, where $h_I$ is known, as well as a stochastic signal where $\langle h_I(t) h_J(t') \rangle \equiv R_{IJ}(t, t')$ is known. We will consider both a *pulsar* term, which assumes an independent signal in each pulsar, and an *Earth* term, where correlations between pulsars can boost the SNR. For most of the derivation we assume that the detector noise is stationary and independent across pulsars,

$$\langle n_I(t) n_J(t') \rangle = \delta_{IJ} N_I(t - t') . \tag{A1}$$

Finally we will simplify the SNR in the limit where the timing residual noise is white and identical in each pulsar: $\langle \delta t_I(t) \delta t_J(t') \rangle = \delta_{IJ} t_{\mathrm{rms}}^2 \Delta t \, \delta(t - t')$, where $\Delta t$ is the measurement cadence. The residual noise is related to the timing residual by a factor of the pulsar frequency, $n_I = \nu_I \delta t_I$ and therefore[6]

$$N_I(t - t') = \nu_I^2 t_{\mathrm{rms}}^2 \Delta t \, \delta(t - t') \tag{A2}$$

$$\widetilde{N}_I = \nu_I^2 t_{\mathrm{rms}}^2 \Delta t . \tag{A3}$$

### 1. Deterministic Signal SNR

We begin with computing the optimal SNR for a deterministic signal. This derivation will closely follow the discussion given in Ref. [43]. We begin by defining a test statistic,

$$\mathcal{T} = \sum_{I=1}^{N_P} \int dt \, s_I(t) \, Q_I(t) , \tag{A4}$$

where $Q_I(t)$ is a filter function chosen to maximize the SNR,

$$\mathrm{SNR}^2 = \frac{\langle \mathcal{T} \rangle_{s=h+n}^2}{\langle \mathcal{T}^2 \rangle_{s=n} - \langle \mathcal{T} \rangle_{s=n}^2} . \tag{A5}$$

---

[6] The signals discussed in the main text also have an accompanying factor of $\nu_I$, and, as we will show, the SNR is independent of this factor.

The subscripts on $\langle \rangle$ indicate what $s$ is assumed to be. We can compute the expectation values,

$$\langle \mathcal{T} \rangle_{s=n} = 0 \tag{A6}$$

$$\langle \mathcal{T} \rangle_{s=h+n} = \sum_{I=1}^{N_P} \int dt\, h_I(t) Q_I(t)$$

$$= \sum_{I=1}^{N_P} \int dt d\mathfrak{f}_1 d\mathfrak{f}_2\, e^{2\pi i(\mathfrak{f}_1+\mathfrak{f}_2)t}\, \widetilde{h}_I(\mathfrak{f}_1) \widetilde{Q}_I(\mathfrak{f}_2)$$

$$= \sum_{I=1}^{N_P} \int d\mathfrak{f}\, \widetilde{h}_I(\mathfrak{f}) \widetilde{Q}_I^*(\mathfrak{f}) \tag{A7}$$

$$\langle \mathcal{T}^2 \rangle_{s=n} = \sum_{I=1}^{N_P} \sum_{J=1}^{N_P} \int dt dt'\, \langle n_I(t) n_J(t') \rangle Q_I(t) Q_J(t')$$

$$= \sum_{I=1}^{N_P} \int dt dt'\, N_I(t-t') Q_I(t) Q_I(t')$$

$$= \sum_{I=1}^{N_P} \int dt dt' d\mathfrak{f}_1 d\mathfrak{f}_2 d\mathfrak{f}_3\, e^{2\pi i(\mathfrak{f}_1+\mathfrak{f}_2)t} e^{2\pi i(\mathfrak{f}_3-\mathfrak{f}_1)t'}\, \widetilde{N}_I(\mathfrak{f}_1) \widetilde{Q}_I(\mathfrak{f}_2) \widetilde{Q}_I(\mathfrak{f}_3)$$

$$= \sum_{I=1}^{N_P} \int d\mathfrak{f}\, \widetilde{N}_I(\mathfrak{f}) \widetilde{Q}_I^*(\mathfrak{f}) \widetilde{Q}_I(\mathfrak{f})\,, \tag{A8}$$

The SNR is then,

$$\mathrm{SNR}^2 = \frac{\left| \sum_{I=1}^{N_P} \int d\mathfrak{f}\, \widetilde{h}_I(\mathfrak{f}) \widetilde{Q}_I^*(\mathfrak{f}) \right|^2}{\sum_{I=1}^{N_P} \int d\mathfrak{f}\, \widetilde{N}_I(\mathfrak{f}) \widetilde{Q}_I^*(\mathfrak{f}) \widetilde{Q}_I(\mathfrak{f})}\,. \tag{A9}$$

The $Q(\mathfrak{f})$ which optimizes this SNR is $Q_I(\mathfrak{f}) = \widetilde{h}_I(\mathfrak{f})/\widetilde{N}_I$, and therefore the optimal SNR is

$$\mathrm{SNR}^2 = \sum_{I=1}^{N_P} \int d\mathfrak{f}\, \frac{\left| \widetilde{h}_I(\mathfrak{f}) \right|^2}{\widetilde{N}_I(\mathfrak{f})}\,. \tag{A10}$$

Assuming the noise is white and pulsar independent we can further simplify,

$$\mathrm{SNR}^2 = \frac{1}{\widetilde{N}} \sum_{I=1}^{N_P} \int dt\, h_I^2(t)\,. \tag{A11}$$

If the signal is independent in each pulsar and the SNR is dominated by the largest signal across the array, then the SNR can be approximated as

$$\mathrm{SNR}^2 = \frac{1}{\widetilde{N}} \max_{\{I\}} \left[ \int dt\, h_I^2(t) \right]\,, \tag{A12}$$

which we define as the *pulsar* term SNR. If the signal has a similar amplitude, but not independent across the pulsars, then Eq. (A11) will be parametrically larger by a factor of $N_P$. We define this as the *Earth* term SNR, because it's the SNR used when the dark matter interacts with the Earth. We can simplify this further by replacing the sum with an average over the pulsar positions, defined by $\langle\rangle_\mathcal{P}$

$$\text{SNR}^2 = \frac{N_P}{\widetilde{N}} \int dt \, \langle h_I^2(t) \rangle_\mathcal{P} \,. \tag{A13}$$

### 2. Stochastic Pulsar Term SNR

We begin by defining a test statistic, $\mathcal{T}$,

$$\mathcal{T} = \sum_{I=1}^{N_P} \int dt dt' \, (s_I(t)s_I(t') - \langle n_I(t)n_I(t') \rangle) \, Q_I(t,t') \tag{A14}$$

where $Q_I(t,t')$ is a filter function applied to the time series of the $I^{\text{th}}$ pulsar. Our goal is to find the $Q_I$'s which maximize the SNR,

$$\text{SNR}^2 = \frac{\langle \mathcal{T} \rangle_{s=h+n}^2}{\langle \mathcal{T}^2 \rangle_{s=n} - \langle \mathcal{T} \rangle_{s=n}^2}, \tag{A15}$$

where the subscripts denote the assumptions under which we should evaluate the expectation values. We have

$$\langle \mathcal{T} \rangle_{s=n} = 0 \tag{A16}$$

$$\langle \mathcal{T} \rangle_{s=h+n} = \sum_{I=1}^{N_P} \int dt dt' \langle h_I(t) h_I(t') \rangle Q_I(t,t')$$

$$= \sum_{I=1}^{N_P} \int dt dt' \prod_{k=1}^{4} d\mathfrak{f}_k \, e^{2\pi i (\mathfrak{f}_1 + \mathfrak{f}_3)t} e^{2\pi i (\mathfrak{f}_2 + \mathfrak{f}_4)t'} \langle \widetilde{h}_I(\mathfrak{f}_1) \widetilde{h}_I(\mathfrak{f}_2) \rangle \widetilde{Q}_I(\mathfrak{f}_3, \mathfrak{f}_4)$$

$$= \sum_{i=1}^{N_P} \int d\mathfrak{f} d\mathfrak{f}' \, S_I(\mathfrak{f}, \mathfrak{f}') \widetilde{Q}_I^*(\mathfrak{f}, \mathfrak{f}') \tag{A17}$$

$$\langle \mathcal{T}^2 \rangle_{s=n} = \sum_{I=1}^{N_P} \sum_{J=1}^{N_P} \int \prod_{k=1}^{4} dt_k \, (\langle n_I(t_1) n_I(t_2) n_J(t_3) n_J(t_4) \rangle - N_I(t_1 - t_2) N_J(t_3 - t_4))$$

$$\times Q_I(t_1, t_2) Q_J(t_3, t_4) \,, \tag{A18}$$

where $S_I(\mathfrak{f}, \mathfrak{f}') \equiv \langle \widetilde{h}_I(\mathfrak{f}) \widetilde{h}_I(\mathfrak{f}') \rangle$. Evaluating the four point function of the noise,

$$\langle n_I(t_1) n_I(t_2) n_J(t_3) n_J(t_4) \rangle = N_I(t_1 - t_2) N_J(t_3 - t_4)$$

$$+ \delta_{IJ} \left( N_I(t_1 - t_3) N_J(t_2 - t_4) + N_I(t_1 - t_4) N_J(t_2 - t_3) \right), \tag{A19}$$

which allows us to simplify Eq. (A18). The two remaining terms are identical, and the whole expression can be simplified to

$$\langle \mathcal{T}^2 \rangle_{s=n} = 2 \sum_{I=1}^{N_P} \int d\mathfrak{f} d\mathfrak{f}' \, \widetilde{N}_I(\mathfrak{f}) \widetilde{N}_I(\mathfrak{f}') \widetilde{Q}_I(\mathfrak{f}) \widetilde{Q}_I^*(\mathfrak{f}'). \tag{A20}$$

The $\widetilde{Q}$ which maximizes the SNR is $Q_I(\mathfrak{f}, \mathfrak{f}') = S_I(\mathfrak{f}, \mathfrak{f}')/(\widetilde{N}_I(\mathfrak{f}) \widetilde{N}_I(\mathfrak{f}'))$ and therefore the optimal SNR is

$$\mathrm{SNR}^2 = \frac{1}{2} \sum_{I=1}^{N_P} \int d\mathfrak{f} d\mathfrak{f}' \, \frac{|S_I(\mathfrak{f}, \mathfrak{f}')|^2}{\widetilde{N}_I(\mathfrak{f}) \widetilde{N}_I(\mathfrak{f}')} \, . \tag{A21}$$

We can simplify further by assuming that the signal and noise are independent of the pulsar, and that the noise is white,

$$\mathrm{SNR}^2 = \frac{N_P}{2\widetilde{N}^2} \int dt dt' R(t, t')^2 \, . \tag{A22}$$

### 3. Stochastic Earth Term SNR

The derivation of the optimal Earth term SNR is similar to the pulsar term, except the test statistic is slightly different,

$$\mathcal{T} = \sum_{I \neq J}^{N_P(N_P-1)} \int dt dt' s_I(t) s_J(t') Q_{IJ}(t, t'), \tag{A23}$$

where the sum is over pairs of pulsars. There is no subtracted piece, as there is in Eq. (A14), because the second term evaluates to zero when $I \neq J$. Computing the terms in the SNR gives,

$$\langle \mathcal{T} \rangle_{s=n} = 0 \tag{A24}$$

$$\langle \mathcal{T} \rangle_{s=h+n} = \sum_{I \neq J} \int dt dt' \, \langle h_I(t) h_J(t') \rangle Q_{IJ}(t, t')$$

$$\langle \mathcal{T} \rangle_{s=h+n} = \sum_{I \neq J} \int d\mathfrak{f} d\mathfrak{f}' \, S_{IJ}(\mathfrak{f}, \mathfrak{f}') \widetilde{Q}_{IJ}^*(\mathfrak{f}, \mathfrak{f}') \tag{A25}$$

$$\langle \mathcal{T}^2 \rangle_{s=n} = \sum_{I \neq J} \sum_{K \neq L} \int \prod_{k=1}^{4} dt_k \, \langle n_I(t_1) n_J(t_2) n_K(t_3) n_L(t_4) \rangle Q_{IJ}(t_1, t_2) Q_{KL}(t_3, t_4) \,, \tag{A26}$$

where $S_{IJ}(\mathfrak{f}, \mathfrak{f}') \equiv \langle \widetilde{h}_I(\mathfrak{f}) \widetilde{h}_J(\mathfrak{f}') \rangle$ and we use a finite time delta function to remove the $t$ integrals. Computing the four point function, noting that $I \neq J$, $K \neq L$ by definition gives,

$$\langle n_I(t_1) n_J(t_2) n_K(t_3) n_L(t_4) \rangle = \delta_{IK} \delta_{JL} N_I(t_1 - t_3) N_J(t_2 - t_4) + \delta_{IL} \delta_{JK} N_I(t_1 - t_4) N_J(t_2 - t_3) \,. \tag{A27}$$

Substituting this expression in $\langle \mathcal{T}^2 \rangle_{s=n}$ gives two identical terms which we can simplify to

$$\langle \mathcal{T}^2 \rangle_{s=n} = 2 \sum_{I \neq J} \int d\mathfrak{f} d\mathfrak{f}' \, \widetilde{N}_I(\mathfrak{f}) \widetilde{N}_J(\mathfrak{f}') \widetilde{Q}_{IJ}^*(\mathfrak{f}) \widetilde{Q}_{IJ}(\mathfrak{f}') \,. \tag{A28}$$

Again one can show that the optimal $\widetilde{Q}_{IJ}(\mathfrak{f}, \mathfrak{f}')$ is $S_{IJ}(\mathfrak{f}, \mathfrak{f}')/\left( \widetilde{N}_I(\mathfrak{f}) \widetilde{N}_J(\mathfrak{f}') \right)$ and therefore the optimal SNR is given by

$$\mathrm{SNR}^2 = \frac{1}{2} \sum_{I \neq J} \int d\mathfrak{f} d\mathfrak{f}' \, \frac{|S_{IJ}(\mathfrak{f}, \mathfrak{f}')|^2}{\widetilde{N}_I(\mathfrak{f}) \widetilde{N}_J(\mathfrak{f}')} \,. \tag{A29}$$

Lastly, we assume that the noise is white and identical across pulsars, and replace the sum by taking an average over the pulsar positions, denoted by $\langle \rangle_{\mathcal{P}}$,

$$\mathrm{SNR}^2 = \frac{N_P(N_P - 1)}{2\widetilde{N}^2} \int dt dt' \, \langle R_{IJ}(t, t')^2 \rangle_{\mathcal{P}} \tag{A30}$$

## Appendix B: Subtraction of Best Fit Parameters in PTA Signal

We discuss how the parameters of the pulsar timing model, *e.g.* $\phi_{\mathrm{fit}}^0, \nu_{\mathrm{fit}}, \dot{\nu}_{\mathrm{fit}}, ...,$ impacts the inferred dark matter signal. Let $\phi_m(t)$ denote the phase that is measured at time $t$, and therefore the goodness of fit is characterized by [7]

$$\chi^2 = \frac{1}{T} \int \left( \phi_m(t) - \phi^0 - \nu t - \frac{1}{2}\dot{\nu}t^2 \right)^2 dt \,. \tag{B1}$$

The generalization to a timing model with higher order terms is straightforward. In order to find the best-fit parameters we minimize $\chi^2$ with respect to $\phi^0, \nu, \dot{\nu}$, which is more easily done by defining an inner product,

$$(a, b) \equiv \frac{1}{T} \int a(t)b(t) \, dt \,, \tag{B2}$$

along with a set of polynomial basis functions, $f_i$, with respect to this inner product, [8]

$$(f_i, f_j) = \delta_{ij} \,. \tag{B3}$$

Therefore $\chi^2$ in Eq. (B1) can be written as

$$\chi^2 = \left( \phi_m(t) - \sum_{k=0}^{2} c_k f_k(t), \phi_m(t) - \sum_{k=0}^{2} c_k f_k(t) \right) \,, \tag{B4}$$

---

[7] In a PTA measurement, $\phi_M$ is measured at an arrival time of the $n^{\mathrm{th}}$ pulse such that $\phi(t_n)$ is subtracted in the standard discrete time formulation of $\chi^2$. The analysis here only differs by working in continuous time.

[8] These are related to the standard Legendre polynomials, $P_n$, by a scaling and shift: $f_n(t) = \sqrt{2n+1}\, P_n(2t/T - 1)$.

where $c_k = (\phi_0 + \nu t + \dot{\nu}t^2/2, f_k)$. We can now minimize with respect to $c_k$ (and later work out $\phi_{\rm fit}^0, \nu_{\rm fit}, \dot{\nu}_{\rm fit}$ if necessary). The minimization condition is

$$\frac{\partial \chi^2}{\partial c_k}(c_k^{\rm fit}) = -2(f_k, \phi_m) + 2\sum_l c_l^{\rm fit}(f_k, f_l) = -2(f_k, \phi_m) + 2c_k^{\rm fit} = 0, \qquad (B5)$$

and therefore $c_k^{\rm fit} = (\phi_m, f_k)$. The residual, $s$, is then given by $s = \phi_m(t) - \sum_k c_k^{\rm fit} f_k(t)$, and the subtracted DM signal, $h$, is related to the unsubtracted signal, $\delta\phi$ by,

$$h(t) = \delta\phi(t) - \sum_{k=0}^{2}(\delta\phi, f_k)f_k(t). \qquad (B6)$$

This is how the subtraction procedure effects the single deterministic event analysis, as the SNR is only a function of $h$. However for the stochastic signal the subtraction procedure enters through a correlator of subtracted signals. We can relate the unsubtracted correlator, $R$, to the subtracted correlator, $R_{\rm sub}$, by

$$R_{\rm sub}(t, t') \equiv \langle h(t)h(t') \rangle$$

$$= \left\langle \left( \delta\phi(t) - \sum_{n=0}^{2}(\delta\phi, f_n)f_n(t) \right) \left( \delta\phi(t') - \sum_{n=0}^{2}(\delta\phi, f_n)f_n(t') \right) \right\rangle$$

$$= R(t, t') - \sum_{n=0}^{2} f_n(t)\mathcal{R}_n(t') - \sum_{n=0}^{2} f_n(t')\mathcal{R}_n(t) + \sum_{n=0}^{2}\sum_{m=0}^{2} f_n(t)f_m(t')\mathcal{R}_{nm}, \qquad (B7)$$

where

$$\mathcal{R}_n(t') \equiv (R(t, t'), f_n(t)) \qquad (B8)$$

$$\mathcal{R}_{nm} \equiv \frac{1}{T^2}\int dt dt'\, R(t, t')f_n(t)f_m(t'). \qquad (B9)$$

The results of this procedure are quoted in Eq. (30). The effect of this subtraction procedure on the reach to monochromatic PBHs can be seen in Fig. (12). Note that in this figure we have not included 'Static' constraints, as these are derived with only $\ddot{\nu}$ and higher order terms, such that subtraction cannot be meaningfully applied to this analysis.

## Appendix C: Relationship with the Power Spectrum Approach

In this section, we explain the difference between this work and the power spectrum procedure adopted in Ref. [37]. If the observing volume is all of space then $R$ in Eq. (14), for the Doppler and Shapiro delays, can be written in terms of the matter power spectrum, as done
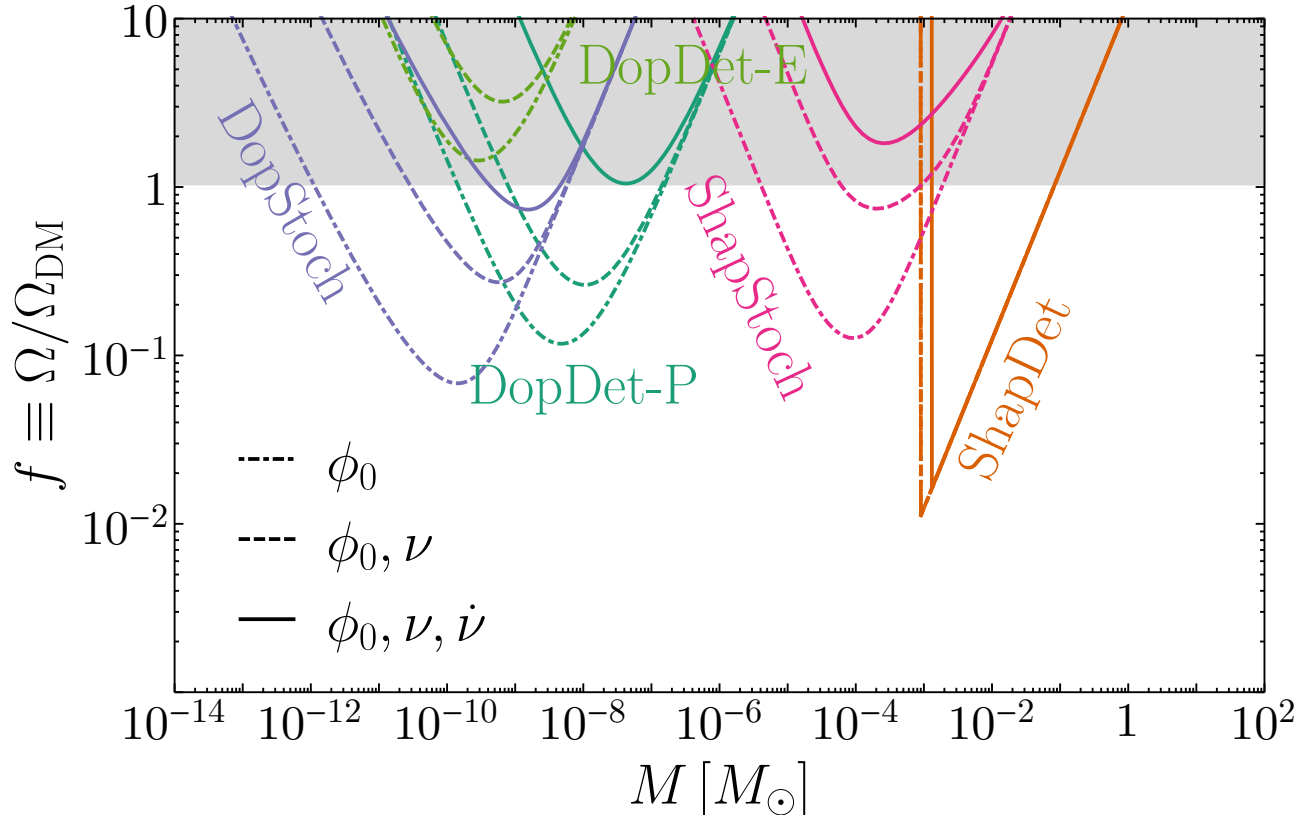
FIG. 12: Constraints of the fraction of dark matter, $f = \Omega/\Omega_{\text{DM}}$ in monochromatic PBHs (similar to Fig. (2)) when different number of parameters are included in the timing model for deterministic and stochastic constraints. Static constraints are not shown since the fitting procedure trivially picks out only the second derivative.

in Ref. [37].[9] To show this, we simplify the gravitational potential correlator, from which both the Doppler and Shapiro signal correlators can be derived. For example, the Shapiro delay frequency shift correlator is

$$\left\langle \frac{\delta\nu}{\nu}(t)\frac{\delta\nu}{\nu}(t') \right\rangle = 4\frac{d}{dt}\frac{d}{dt'}\int_0^{z_0}\int_0^{z_0} dz_1 dz_2 \int \frac{d^3\mathbf{k}d^3\mathbf{k}'}{(2\pi)^6} e^{i(\mathbf{k}+\mathbf{k}')\cdot\mathbf{r}} \langle \widetilde{\Phi}(\mathbf{k}, t; Q)\widetilde{\Phi}(\mathbf{k}', t'; Q) \rangle \qquad \text{(C1)}$$

where $\widetilde{\Phi}(\mathbf{k}; Q)$ is the Fourier transform of the gravitational potential at position $\mathbf{r}$, and $Q$ represents all of the random variables, $\mathbf{r}_1^0, \mathbf{r}_2^0, M_1, M_2$, etc.

Since the object's individual potentials only depend on the distance from the center of mass,

---

[9] The $\langle \delta\phi(t)\delta\phi(t') \rangle$ signal correlator we consider here and $\langle \frac{\delta\nu}{\nu}(t)\frac{\delta\nu}{\nu}(t') \rangle$ considered in Ref. [37] are related by $\langle \delta\phi(t)\delta\phi(t') \rangle = \nu^2 \int_0^t \int_0^{t'} dt_1 dt_2 \langle \frac{\delta\nu}{\nu}(t_1)\frac{\delta\nu}{\nu}(t_2) \rangle$.

$$\mathbf{r}_i(t) = \mathbf{v}t + \mathbf{r}_i^0,$$

$$\widetilde{\Phi}(\mathbf{k}, t; Q) = \sum_i e^{i\mathbf{k}\cdot\mathbf{r}_i(t)}\widetilde{\Phi}_i(\mathbf{k}; M_i)\,. \tag{C2}$$

The potential correlator is then a sum over individual contributions as,

$$\langle\widetilde{\Phi}(\mathbf{k}, t; Q)\widetilde{\Phi}(\mathbf{k}', t'; Q)\rangle = \sum_{i,j}\left\langle e^{i(\mathbf{k}\cdot\mathbf{r}_i + \mathbf{k}'\cdot\mathbf{r}_j)}\right\rangle_{\mathbf{r}^0}\langle\widetilde{\Phi}_i(\mathbf{k}, M_i)\widetilde{\Phi}_j(\mathbf{k}', M_j)\rangle_M\,, \tag{C3}$$

where the subscripted $\langle\rangle$ denotes averaging over only the subscripted random variable. In the limit where the $\mathbf{r}^0$ integral is over all space, the exponential in the $i = j$ term becomes a delta function, and the potential correlator from a single subhalo can be simplified,

$$\left\langle e^{i\left(\mathbf{k}\cdot\mathbf{r}_i^0 + \mathbf{k}'\cdot\mathbf{r}_i^0\right)}\right\rangle_{\mathbf{r}^0} = \frac{(2\pi)^3}{V}\delta^3(\mathbf{k} + \mathbf{k}') \tag{C4}$$

$$\langle\widetilde{\Phi}_i(\mathbf{k}, M_i)\widetilde{\Phi}_i(\mathbf{k}, M_i)\rangle_M = \frac{16\pi^2 G^2\overline{n}^2}{k^4}P_m^{1h}(k)\,, \tag{C5}$$

where $P_m^{1h}$ is the 1-subhalo matter power spectrum. It can be shown that the $i \neq j$ term in Eq. (C3) simplifies similarly with $P_m^{1h}$ being replaced by $P_m^{2h}$. Finally,

$$\langle\widetilde{\Phi}(\mathbf{k}, t; Q)\widetilde{\Phi}(\mathbf{k}', t'; Q)\rangle = (2\pi)^3\delta^3\left(\mathbf{k} + \mathbf{k}'\right)\frac{16\pi^2 G^2\overline{n}^2}{k^4}e^{i\mathbf{k}\cdot\mathbf{v}(t-t')}\left(P_m^{1h}(k) + P_m^{2h}(k)\right)\,, \tag{C6}$$

and the rest of the derivation for the signal correlator proceeds as in Ref. [37]. However the Earth-pulsar system samples a finite volume of space, since subhalos at a larger distance are static, and hence susceptible to being absorbed in the fit. To obtain a physical result from a power spectrum approach one would need to introduce cuts on the $\mathbf{k}$ integral to incorporate these finite volume effects. Here we simplify the signal correlators differently than Ref. [37] by incorporating finite volume effects directly in the subhalo position average.

## Appendix D: $\mathcal{F}$ Form Factor Integral

We derive an analytic expression for the form factor in Eq. (18),

$$\mathcal{F}(M, b) = b\int dk\, W(k, M)J_1(kb)\,, \tag{D1}$$

with the definition of $W$,

$$W(k, M) = \frac{4\pi}{M}\int_0^{r_v} r^2\text{sinc}(kr)\rho(r, M)dr\,. \tag{D2}$$

Substituting this in to Eq. (D1),

$$\mathcal{F}(M,b) = \frac{4\pi b}{M} \int_0^{r_v} r^2 \rho(r,M) dr \int dk\, J_1(kb)\text{sinc}(kr)\,, \tag{D3}$$

and using

$$\int dk\, J_1(kb)\text{sinc}(kr) = \frac{1}{b}\left(1 - \Theta(r-b)\sqrt{1 - \frac{b^2}{r^2}}\right)\,, \tag{D4}$$

gives

$$\mathcal{F}(M,b) = 1 - \frac{4\pi}{M}\int_b^{r_v}\sqrt{1 - \frac{b^2}{r^2}}r^2\rho(r,M)dr\,. \tag{D5}$$

Assuming $\rho$ is given by an NFW profile then

$$\rho = \rho_s\frac{1}{\left(\frac{r}{r_s}\right)\left(1 + \frac{r}{r_s}\right)^2} \tag{D6}$$

$$\rho_s = \frac{1}{r_s^3}\frac{M}{4\pi}\frac{1}{\left(\log\left(1+c\right) - \frac{c}{1+c}\right)}\,, \tag{D7}$$

and therefore

$$\mathcal{F}(M,b) = 1 - \frac{1}{r_v^3}\frac{c^3}{\left(\log\left(1+c\right) - \frac{c}{1+c}\right)}\int_b^{r_v}\sqrt{1 - \frac{b^2}{r^2}}r^2\frac{1}{\left(\frac{r}{r_s}\right)\left(1 + \frac{r}{r_s}\right)^2}dr\,. \tag{D8}$$

We can simplify further by changing variables. Defining $y \equiv r/r_v$, and $x \equiv b/r_v$

$$\mathcal{F}(x,c) = 1 - \frac{c^2}{\left(\log\left(1+c\right) - \frac{c}{1+c}\right)}\int_x^1\sqrt{1 - \frac{x^2}{y^2}}\frac{y}{\left(1 + cy\right)^2}dy \tag{D9}$$

We see that the form factor $\mathcal{F}$ is only a function of $c$ and $b/r_v$, as shown in Fig. 6.

---

[1] A. M. Green, S. Hofmann,  and D. J. Schwarz, JCAP **0508**, 003 (2005), arXiv:astro-ph/0503387 [astro-ph].

[2] C. Hogan and M. Rees, Physics Letters B **205**, 228 (1988).

[3] E. W. Kolb and I. I. Tkachev, Phys. Rev. Lett. **71**, 3051 (1993), arXiv:hep-ph/9303313 [hep-ph].

[4] E. W. Kolb and I. I. Tkachev, Phys. Rev. **D49**, 5040 (1994), arXiv:astro-ph/9311037 [astro-ph].

[5] K. M. Zurek, C. J. Hogan,  and T. R. Quinn, Phys. Rev. **D75**, 043511 (2007), arXiv:astro-ph/0607341 [astro-ph].

[6] M. Buschmann, J. W. Foster, and B. R. Safdi, Phys. Rev. Lett. **124**, 161103 (2020), arXiv:1906.00967 [astro-ph.CO].

[7] A. Arvanitaki, S. Dimopoulos, M. Galanis, L. Lehner, J. O. Thompson, and K. Van Tilburg, Phys. Rev. D **101**, 083014 (2020), arXiv:1909.11665 [astro-ph.CO].

[8] P. W. Graham, J. Mardon, and S. Rajendran, Phys. Rev. **D93**, 103520 (2016), arXiv:1504.02102 [hep-ph].

[9] A. L. Erickcek and K. Sigurdson, Phys. Rev. **D84**, 083503 (2011), arXiv:1106.0536 [astro-ph.CO].

[10] G. Barenboim and J. Rasero, JHEP **04**, 138 (2014), arXiv:1311.4034 [hep-ph].

[11] J. Fan, O. Özsoy, and S. Watson, Phys. Rev. **D90**, 043536 (2014), arXiv:1405.7373 [hep-ph].

[12] N. Blinov, M. J. Dolan, and P. Draper, Phys. Rev. D **101**, 035002 (2020), arXiv:1911.07853 [astro-ph.CO].

[13] D. Gilman, S. Birrer, A. Nierenberg, T. Treu, X. Du, and A. Benson, Mon. Not. Roy. Astron. Soc. **491**, 6077 (2020), arXiv:1908.06983 [astro-ph.CO].

[14] A. Bonaca, D. W. Hogg, A. M. Price-Whelan, and C. Conroy, The Astrophysical Journal **880**, 38 (2019).

[15] J. B. Muñoz, C. Dvorkin, and F.-Y. Cyr-Racine, Phys. Rev. D **101**, 063526 (2020), arXiv:1911.11144 [astro-ph.CO].

[16] A. Diaz Rivero, F.-Y. Cyr-Racine, and C. Dvorkin, Phys. Rev. **D97**, 023001 (2018), arXiv:1707.04590 [astro-ph.CO].

[17] M. Buschmann, J. Kopp, B. R. Safdi, and C.-L. Wu, Phys. Rev. Lett. **120**, 211101 (2018), arXiv:1711.03554 [astro-ph.GA].

[18] K. Van Tilburg, A.-M. Taki, and N. Weiner, JCAP **1807**, 041 (2018), arXiv:1804.01991 [astro-ph.CO].

[19] C. Mondino, A.-M. Taki, K. Van Tilburg, and N. Weiner, (2020), arXiv:2002.01938 [astro-ph.CO].

[20] S. Mishra-Sharma, K. Van Tilburg, and N. Weiner, (2020), arXiv:2003.02264 [astro-ph.CO].

[21] T. D. Brandt, Astrophys. J. **824**, L31 (2016), arXiv:1605.03665 [astro-ph.GA].

[22] R. A. Allsman *et al.* (Macho), The Astrophysical Journal **550**, L169 (2001), arXiv:astro-ph/0011506 [astro-ph].

[23] P. Tisserand *et al.* (EROS-2), Astronomy & Astrophysics **469**, 387 (2007), arXiv:astro-ph/0607207 [astro-ph].

[24] L. Wyrzykowski *et al.*, Monthly Notices of the Royal Astronomical Society **416**, 2949 (2011), arXiv:1106.2925 [astro-ph.GA].

[25] H. Niikura, M. Takada, N. Yasuda, R. H. Lupton, T. Sumi, S. More, A. More, M. Oguri, and M. Chiba, Nature Astronomy **3**, 524 (2019), arXiv:1701.02151 [astro-ph.CO].

[26] N. Smyth, S. Profumo, S. English, T. Jeltema, K. McKinnon, and P. Guhathakurta, (2019), arXiv:1910.01285 [astro-ph.CO].

[27] K. Griest, A. M. Cieplak, and M. J. Lehner, Astrophys. J. **786**, 158 (2014), arXiv:1307.5798 [astro-ph.CO].

[28] A. Katz, J. Kopp, S. Sibiryakov, and W. Xue, JCAP **1812**, 005 (2018), arXiv:1807.11495 [astro-ph.CO].

[29] A. Katz, J. Kopp, S. Sibiryakov, and W. Xue, (2019), arXiv:1912.07620 [astro-ph.CO].

[30] J. A. Dror, H. Ramani, T. Trickle, and K. M. Zurek, Phys. Rev. **D100**, 023003 (2019), arXiv:1901.04490 [astro-ph.CO].

[31] D. Croon, D. McKeen, and N. Raj, Phys. Rev. D **101**, 083013 (2020), arXiv:2002.08962 [astro-ph.CO].

[32] L. Dai and J. Miralda-Escudé, Astron. J. **159**, 49 (2020), arXiv:1908.01773 [astro-ph.CO].

[33] E. R. Siegel, M. Hertzberg, and J. Fry, Mon. Not. Roy. Astron. Soc. **382**, 879 (2007), arXiv:astro-ph/0702546.

[34] N. Seto and A. Cooray, Astrophys. J. **659**, L33 (2007), arXiv:astro-ph/0702586.

[35] H. A. Clark, G. F. Lewis, and P. Scott, Mon. Not. Roy. Astron. Soc. **456**, 1394 (2016), [Erratum: Mon.Not.Roy.Astron.Soc. 464, 2468 (2017)], arXiv:1509.02938 [astro-ph.CO].

[36] K. Schutz and A. Liu, Phys. Rev. D **95**, 023002 (2017), arXiv:1610.04234 [astro-ph.CO].

[37] S. Baghram, N. Afshordi, and K. M. Zurek, Phys. Rev. **D84**, 043511 (2011), arXiv:1101.5487 [astro-ph.CO].

[38] K. Kashiyama and N. Seto, MNRAS **426**, 1369 (2012), arXiv:1208.4101.

[39] K. Kashiyama and M. Oguri, (2018), arXiv:1801.07847 [astro-ph.CO].

[40] P. A. Rosado, A. Sesana, and J. Gair, Mon. Not. Roy. Astron. Soc. **451**, 2417 (2015), arXiv:1503.04803 [astro-ph.HE].

[41] E. R. Siegel, M. P. Hertzberg, and J. N. Fry, Monthly Notices of the Royal Astronomical Society **382**, 879 (2007), arXiv:astro-ph/0702546 [astro-ph].

[42] N. Seto and A. Cooray, Astrophys. J. **659**, L33 (2007), arXiv:astro-ph/0702586 [astro-ph].

[43] C. J. Moore, R. H. Cole, and C. P. L. Berry, Class. Quant. Grav. **32**, 015014 (2015), arXiv:1408.0740 [gr-qc].

[44] T. L. Smith and R. Caldwell, Phys. Rev. D 100, 104055 (2019) (2019), 10.5281/zenodo.3341817, arXiv:1908.00546v1 [astro-ph.CO].

[45] G. Desvignes *et al.*, Mon. Not. Roy. Astron. Soc. **458**, 3341 (2016), arXiv:1602.08511 [astro-ph.HE].

[46] R. N. Manchester *et al.*, Publ. Astron. Soc. Austral. **30**, 17 (2013), arXiv:1210.6130 [astro-ph.IM].

[47] Z. Arzoumanian *et al.* (NANOGrav), Astrophys. J. **821**, 13 (2016), arXiv:1508.03024 [astro-ph.GA].

[48] M. Bailes *et al.*, *Proceedings, MeerKAT Science: On the Pathway to the SKA (MeerKAT2016): Stellenbosch, South Africa, May 25-27, 2016*, PoS **MeerKAT2016**, 011 (2018), arXiv:1803.07424 [astro-ph.IM].

[49] G. Hobbs, S. Dai, R. N. Manchester, R. M. Shannon, M. Kerr, K. J. Lee, and R. Xu, Res. Astron. Astrophys. **19**, 020 (2019), arXiv:1407.0435 [astro-ph.IM].

[50] B. B. P. Perera *et al.*, Mon. Not. Roy. Astron. Soc. **490**, 4666 (2019), arXiv:1909.04534 [astro-ph.HE].

[51] D. Reardon, G. Hobbs, W. Coles, Y. Levin, M. Keith, M. Bailes, N. Bhat, S. Burke-Spolaor, S. Dai, M. Kerr, *et al.*, Monthly Notices of the Royal Astronomical Society **455**, 1751 (2016).

[52] Z. Arzoumanian *et al.* (NANOGrav), Astrophys. J. **813**, 65 (2015), arXiv:1505.07540 [astro-ph.IM].

[53] M. Zumalacárregui and U. Seljak, Physical Review Letters **121** (2018), 10.1103/physrevlett.121.141101.

[54] R. Laha, J. B. Muñoz, and T. R. Slatyer, (2020), arXiv:2004.00627 [astro-ph.CO].

[55] R. Van Haasteren and Y. Levin, *Mon. Not. R. Astron. Soc*, Tech. Rep. 0000 (2018) arXiv:0909.0954v1.

[56] L. Dai, S.-S. Li, B. Zackay, S. Mao, and Y. Lu, Phys. Rev. D **98**, 104029 (2018), arXiv:1810.00003 [gr-qc].

[57] A. Moline, M. A. Sanchez-Conde, S. Palomares-Ruiz, and F. Prada, Mon. Not. Roy. Astron. Soc. **466**, 4974 (2017), arXiv:1603.04057 [astro-ph.CO].

[58] J. Wang, S. Bose, C. S. Frenk, L. Gao, A. Jenkins, V. Springel, and S. D. M. White, (2019), arXiv:1911.09720 [astro-ph.CO].

[59] W. H. Press and P. Schechter, Astrophys. J. **187**, 425 (1974).

[60] V. Springel, J. Wang, M. Vogelsberger, A. Ludlow, A. Jenkins, A. Helmi, J. F. Navarro, C. S. Frenk, and S. D. M. White, Mon. Not. Roy. Astron. Soc. **391**, 1685 (2008), arXiv:0809.0898 [astro-ph].

[61] D. Fiacconi, P. Madau, D. Potter, and J. Stadel, Astrophys. J. **824**, 144 (2016), arXiv:1602.03526 [astro-ph.GA].

[62] A. D. Ludlow, J. F. Navarro, R. E. Angulo, M. Boylan-Kolchin, V. Springel, C. Frenk, and S. D. M. White, Mon. Not. Roy. Astron. Soc. **441**, 378 (2014), arXiv:1312.0945 [astro-ph.CO].

[63] B. Diemer and M. Joyce, Astrophys. J. **871**, 168 (2019), arXiv:1809.07326 [astro-ph.CO].

[64] B. Carr, M. Raidal, T. Tenkanen, V. Vaskonen, and H. Veermäe, Phys. Rev. **D96**, 023514 (2017), arXiv:1705.05567 [astro-ph.CO].

[65] Y. Wang, J. Phys. Conf. Ser. **610**, 012019 (2015), arXiv:1505.00402 [astro-ph.IM].

[66] W. Becker, M. Kramer, and A. Sesana, Space Sci. Rev. **214**, 30 (2018), arXiv:1705.11022 [astro-ph.IM].

[67] A. N. Lommen, Reports on Progress in Physics **78**, 124901 (2015).

[68] F. A. Jenet and J. D. Romano, Am. J. Phys. **83**, 635 (2015), arXiv:1412.1142 [gr-qc].

[69] R. Hellings and G. Downs, Astrophys. J. **265**, L39 (1983).

[70] M. Fukugita and P. E. Peebles, Astrophys. J. **616**, 643 (2004), arXiv:astro-ph/0406095.

[71] D. R. Lorimer, Living reviews in relativity **11**, 8 (2008).

[72] B. Allen and J. D. Romano, Physical Review D **59** (1999), 10.1103/physrevd.59.102001.