Little Red Dots from Small-Scale Primordial Black Hole Clustering

Borui Zhang,[1] Wei-Xiang Feng,[1,*] and Haipeng An[1,2,†]

[1] *Department of Physics, Tsinghua University, Beijing 100084, China*
[2] *Center for High Energy Physics, Tsinghua University, Beijing 100084, China*

The James Webb Space Telescope (JWST) observations have identified a class of compact galaxies at high redshifts ($4 \lesssim z \lesssim 11$), dubbed "little red dots" (LRDs). The supermassive black holes (SMBHs) of $10^{5-8}\,\mathrm{M}_\odot$ in LRDs favor a heavy-seed origin. We propose a mechanism for their formation: Clusters of primordial black holes, formed through long-short mode coupling on small scales in the early Universe, undergo sequential mergers over extended timescales. This mechanism can evade cosmic microwave background distortions and result in heavy-seed SMBHs via runaway mergers. We employ Monte Carlo simulations to solve the Smoluchowski coagulation equation and determine the runaway merging timescale. The resulting stochastic gravitational wave background offers a distinct signature of this process, and the forming SMBHs can be highly spinning at their formation due to the spin residual of the cluster from tidal fields. This mechanism may explain the rapidly spinning SMBHs in LRDs under the assumption of obscured active galactic nuclei.

Introduction and Summary.— The James Webb Space Telescope (JWST) has recently discovered a class of compact, red-hued galaxies at high redshifts ($4 \lesssim z \lesssim 11$), dubbed "little red dots" (LRDs) [1–5], for which the origin remains open. Moreover, around 80% of these galaxies display broad Balmer emission lines, suggesting that they host active galactic nuclei (AGNs) with central supermassive black holes (SMBHs) of $10^{5-8}\,\mathrm{M}_\odot$ [6–15].

SMBHs in LRDs could be of primordial origin from the high-density perturbations in the early Universe. While primordial black holes (PBHs) produced from direct collapse generally harbor small spin [16, 17] and are excluded for heavy seeds $\gtrsim 10^4\,\mathrm{M}_\odot$ by cosmic microwave background (CMB) distortions given the nearly Gaussian primordial density perturbations [18, 19].

In this *Letter*, we propose a mechanism wherein a high-redshift spinning SMBH can originate from successive mergers of PBHs in an initial small-scale cluster, which avoids the CMB constraints. Although the high spin, resulting in high radiative efficiency, limits their growth by gas accretion, a runaway merger of PBHs in the cluster can give rise to heavy-seed SMBHs favored in LRDs. Besides, based on the Sołtan argument [20], slower accretion growth could reconcile the SMBH mass density at $z \simeq 4$–5 under the obscured AGN hypothesis [21].

As a proof of concept, we consider the primordial power spectrum consisting of two components: short (k_s) and long (k_l) wavelength modes, where the characteristic length scale $r_{\rm pbh} \sim 1/k_s$ of a PBH, and that of a cluster is $r_{\rm cl} \sim 1/k_l$. Here we take $k_s \simeq 2.5\,\mathrm{pc}^{-1}$ at the horizon-exit for PBHs of mass $m_{\rm pbh} \simeq 30\,\mathrm{M}_\odot$ [22]. Single PBHs virialized in a cluster with number density $n_{\rm cl}$ will form binaries via the emission of gravitational radiation during close encounters and eventually merge. The fraction of PBHs that will coalesce over a time period Δt can be computed as:

$$\frac{\Delta n_{\rm cl}}{n_{\rm cl}} = \frac{1}{2}\sigma_{\rm merg}\, v_{\rm rel}\, n_{\rm cl}\, \Delta t\,, \tag{1}$$

where

$$\sigma_{\rm merg} = 2\pi\left(\frac{85\pi}{6\sqrt{2}}\right)^{2/7}\frac{G^2(m_1+m_2)^{10/7}m_1^{2/7}m_2^{2/7}}{c^{10/7}v_{\rm rel}^{18/7}} \tag{2}$$

is the merging cross section [23, 24]. Here, G is the Newton's constant, c is the speed of light, m_1 and m_2 are the masses of the two PBHs in a close encounter, and their relative velocity is $v_{\rm rel} \approx v_{\rm vir} = \sqrt{GM_{\rm cl}/r_{\rm vir}}$, the virial velocity of the cluster, where $M_{\rm cl}$ is the cluster mass within the virial radius $r_{\rm vir}$. To form a SMBH via mergers of an initial population of $N_{\rm cl}$ PBHs of equal mass $m_{\rm pbh}$, we can estimate the timescale over which half of the PBHs in the cluster coalesce into larger ones, by

$$t_{\rm merg} \sim \frac{1}{n_{\rm cl}\,\sigma_{\rm merg}\,v_{\rm vir}} \simeq 1.07\,\mathrm{Gyr}\left(\frac{N_{\rm cl}}{10^5}\right)^{-1} \times \left(\frac{m_{\rm pbh}}{30\,\mathrm{M}_\odot}\right)^{-2}\left(\frac{M_{\rm cl}}{3\times10^6\,\mathrm{M}_\odot}\right)^{11/14}\left(\frac{r_{\rm vir}}{0.02\,\mathrm{pc}}\right)^{31/14}, \tag{3}$$

which indicates that $n_{\rm cl} = 3N_{\rm cl}/4\pi r_{\rm vir}^3 \gtrsim 3\times10^9\,\mathrm{pc}^{-3}$ is required for the emergence of a $\sim 10^6\,\mathrm{M}_\odot$ SMBH within the first billion years after the Big Bang. This estimate is conservative, as during mergers the merging cross section increases as PBHs coalesce into bigger ones, until the *runaway merger* results. Then the question is: how can such a dense cluster be generated in the early Universe?

Clusters of this kind may arise from the modulation of the overdensity field at the PBH formation scale (k_s) by long-wavelength (k_l) modes on larger scales, through various mechanisms such as local-type non-Gaussianity [22, 25–37], closed domain walls [38–40], multistream inflation [41, 42], long-range scalar forces [43], quantum diffusion [44–46], or correlated bubble collisions [47]. A simple parametrization of the long-mode spatial modulation is given by [22]:

$$\nu(\boldsymbol{x}) = \nu_g\,(1+\eta\,\Phi_l(\boldsymbol{x}))\,, \tag{4}$$

* wxfeng@mail.tsinghua.edu.cn
† anhp@mail.tsinghua.edu.cn

where η is the coupling strength between long and short modes, $\Phi_l(\boldsymbol{x})$ is the long-mode field, and $\nu_g \equiv \delta_c/\bar{\sigma}_s$ is the reduced threshold of the short mode without modulation, with $\delta_c \simeq 0.414$ during the radiation era [48], and $\bar{\sigma}_s^2 = \langle \delta_s^2 \rangle$ the variance of δ_s averaged over the whole space. Both the long- and short-mode density fields are assumed to follow Gaussian distributions. In this study, we fix $\nu_g = 8.5$ to generate a sufficiently large initial clustering amplitude and to achieve an adequate PBH abundance (see Eqs. 6 and 7), resulting in $\bar{\sigma}_s \simeq 0.0487$. The variance of the long-mode, σ_l^2, is set to be $\sigma_l \simeq 0.0064$ in order to evade the CMB μ-distortion constraint [18, 19].

After the cluster decouples from the Hubble flow and undergoes virialization, we use the *Smoluchowski coagulation equation* [49–51] to model the PBH population evolution during successive mergers and to calculate the associated gravitational wave (GW) signals. The merger process may last from millions to billions of years, during which the cluster can acquire a significant amount of angular momentum from tidal fields, such that the resulting SMBH inherits a large spin.

Fig. 1 summarizes the parameter space in terms of the modulation strength $|\eta|$ and the comoving wavenumber ratio k_l/k_s, where PBH clusters can form and potentially explain the origin of LRDs (light-gray region) via runaway mergers. The remaining regions represent parameter ranges where SMBHs at high redshift cannot form through mergers, as will be discussed in the following sections.

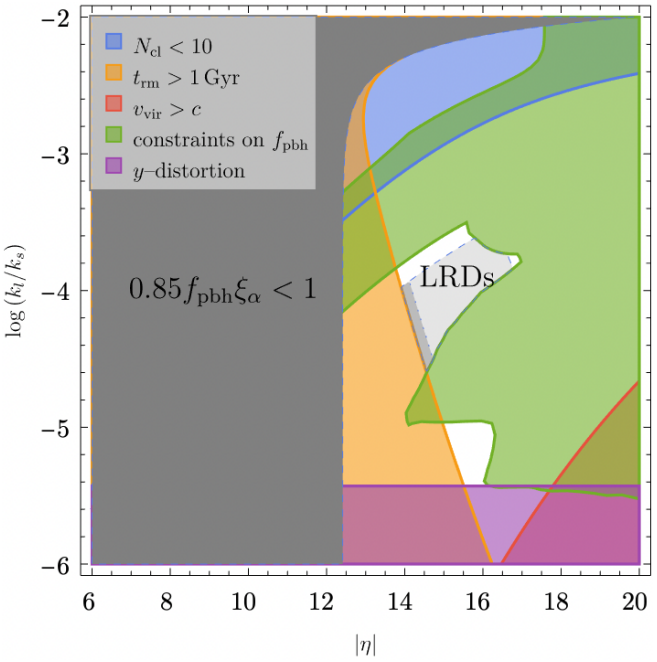

FIG. 1: The $|\eta|$–(k_l/k_s) parameter plane for $\nu_g = 8.5$, $\kappa \equiv \sigma_l/\bar{\sigma}_s = 0.1313$, chosen to satisfy the stringent CMB μ-distortion constraint. The light-gray region denotes the parameter space consistent with observational bounds that permits the formation of high-spin SMBHs ($10^{5\text{–}8}\,{\rm M}_\odot$) in LRDs at high redshifts.

Initial Small-Scale PBH Clustering.— According to Eq. 4, for $|\eta| \neq 0$, the long-mode-induced modulation can generate a significant two-point correlation that remains nearly constant within the clustering scale $r_{\rm cl} \sim 1/k_l$:

$$\xi_{\rm pbh}(r) \simeq \begin{cases} \xi_\alpha, & r \lesssim r_{\rm cl}; \\ 0, & r > r_{\rm cl}\,, \end{cases} \tag{5}$$

where the subscript denotes coupling strength $\alpha(\eta,\kappa) \equiv \delta_c \eta \kappa$ with $\kappa \equiv \sigma_l/\bar{\sigma}_s \simeq 0.1313$ the ratio of variance of long to short mode. In our setting, $\xi_\alpha \sim 10^6$ given $|\eta| \sim 14$. The fraction of PBHs relative to the total dark matter abundance, defined as $f_{\rm pbh} \equiv \Omega_{\rm pbh}/\Omega_{\rm dm}$, also depends on the long-short mode coupling $|\eta|$, and is given by [18, 19]:

$$f_{\rm pbh} \simeq 2.8\times10^7 \left(\frac{g_\star}{106.75}\right)^{-1/4} \left(\frac{30\,{\rm M}_\odot}{m_{\rm pbh}}\right)^{1/2} \beta(m_{\rm pbh}), \tag{6}$$

where $g_\star$ is the number of relativistic degrees of freedom at the time of PBH formation, and [22]

$$\beta(m_{\rm pbh}) \approx 0.1\,{\rm erfc}\left(\frac{\nu_g}{\sqrt{2}}\frac{1}{\sqrt{1+\alpha^2}}\right), \tag{7}$$

so the coupling strength enhances the PBH abundance.

The properties of the PBH cluster depend on the combination $f_{\rm pbh}\xi_\alpha$, which is a function of $\alpha(\eta,\kappa)$ and k_l/k_s. In particular, the condition $0.85 f_{\rm pbh}\xi_\alpha > 1$ (as shown in Fig. 1) must be satisfied for decoupling before the matter-radiation equality. It turns out $f_{\rm pbh} \sim 10^{-5}$ is sufficient to explain the SMBH abundance given the strong correlation $\xi_\alpha \sim 10^6$. The cluster parameters are summarized below [36, 37] (see also Supplemental Material).

- *The Number Density:* After the cluster decouples from the Universe's expansion, its number density can be treated as fixed [36, 37]:

$$n_{\rm cl} \simeq 2\times10^8\,{\rm pc}^{-3} \left(\frac{f_{\rm pbh}\xi_\alpha}{25}\right)^4 \left(\frac{C}{20}\right) \left(\frac{30\,{\rm M}_\odot}{m_{\rm pbh}}\right), \tag{8}$$

  where the compactness parameter $C = 20$ [52] is adopted to characterize the cluster's overdensity after virialization. The number density $n_{\rm cl}$ is only determined by the combination $f_{\rm pbh}\xi_\alpha$, and therefore remains constant after decoupling regardless of the cluster size $r_{\rm cl} \sim 1/k_l$ at formation. A strong coupling η boosts the clustering via $n_{\rm cl} \propto (f_{\rm pbh}\xi_\alpha)^4$.

- *The Cluster Mass:* The total mass of the cluster is given by:

$$M_{\rm cl} \simeq 3\times10^6\,{\rm M}_\odot \left(\frac{f_{\rm pbh}\xi_\alpha}{25}\right) \left(\frac{r_{\rm cl}}{10\,{\rm kpc}}\right)^3, \tag{9}$$

  which scales as $r_{\rm cl}^3 \sim 1/k_l^3$, implying that the wavelength of the long mode primarily determines the total mass of the cluster and ultimately sets the mass scale of the final SMBH. As $n_{\rm cl}$ is nearly independent of $r_{\rm cl}$, a longer-wavelength mode (smaller

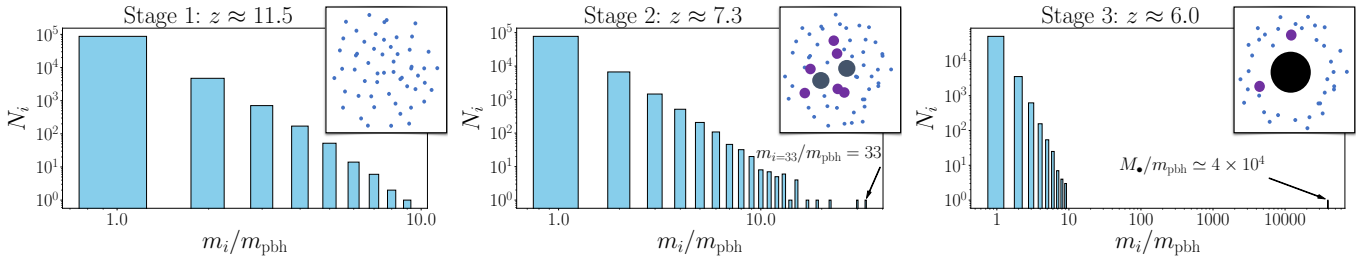

FIG. 2: PBH mass population evolution in a cluster of $N_{\rm cl} = 10^5$ from redshift $z \simeq 11.5$–6.0 with $n_{\rm cl} = 2.0 \times 10^8\,{\rm pc}^{-3}$ and $v_{\rm vir} = 512\,{\rm km\,s}^{-1}$ corresponding to $|\eta| \simeq 14.34$ and $\log(k_l/k_s) \simeq -4.38$ in Fig. 1.

k_l) results in a larger cluster volume and hence a greater number of PBHs, $N_{\rm cl} = M_{\rm cl}/m_{\rm pbh}$. Therefore, for a fixed $M_{\rm cl}$, the number of PBHs in the cluster is inversely proportional to their individual masses $m_{\rm pbh}$.

- *The Virial Radius and Velocity:* The virial radius of the cluster after decoupling is computed as $r_{\rm vir} = (3M_{\rm cl}/4\pi m_{\rm pbh} n_{\rm cl})^{1/3} \propto (f_{\rm pbh}\xi_\alpha)^{-1} C^{-1/3} r_{\rm cl}$, which decreases with the coupling $|\eta|$. The virial velocity of the cluster is then given by $v_{\rm vir} = \sqrt{GM_{\rm cl}/r_{\rm vir}} \propto (f_{\rm pbh}\xi_\alpha) C^{1/6} r_{\rm cl}$, which increases with the coupling $|\eta|$. As both $r_{\rm vir}$ and $v_{\rm vir} \propto r_{\rm cl} \sim 1/k_l$, the smaller the ratio k_l/k_s, the larger the $r_{\rm vir}$ and $v_{\rm vir}$. Given the above setting and $k_l/k_s \sim 10^{-4}$, we will show that a cluster with $r_{\rm vir} \sim 0.05\,{\rm pc}$ and $v_{\rm vir} \sim 500\,{\rm km\,s}^{-1}$ can explain the SMBHs in LRDs through runaway mergers.

Runaway Mergers and the Gravitational Wave.— In the PBH-merger scenario, clusters at formation must not undergo direct collapse into an SMBH immediately after decoupling from the Hubble flow, such that $r_{\rm vir} > 2GM_{\rm cl}/c^2$ or $v_{\rm vir} < c$ based on the hoop conjecture [53]. For a virialized homogeneous cluster with gravitational potential energy $U = -0.6\,GM_{\rm cl}^2/r_{\rm vir}$, PBHs generally reach stable equilibrium with $r_{\rm vir} < -0.335\,GM_{\rm cl}^2/E$, where the total energy is $E = U/2$. Thus, no gravothermal catastrophe occurs, according to the Antonov instability [54–57], and the emergence of an SMBH can result only from successive mergers of PBHs.

The initial cluster decouples with number density $n_{\rm cl}$ and PBHs of monochromatic mass $m_{\rm pbh}$. As PBH mergers are predominantly governed by two-body interactions, and binaries coalesce via the emission of gravitational radiation, we employ the Smoluchowski coagulation equation [49–51] to study the population evolution and the timescale leading to central massive black holes:

$$\frac{{\rm d}}{{\rm d}t} n_i = \frac{1}{2} \sum_{j+k=i} n_j n_k \mathcal{K}_{jk} - n_i \sum_{j=1}^{N_{\rm cl}-i} n_j \mathcal{K}_{ij}\,, \tag{10}$$

where n_i is the number density of PBH with mass $m_i = i m_{\rm pbh}$ $(i = 1, 2, ..., N_{\rm cl})$. The first sum accounts for the gain in n_i due to mergers of PBHs with $m_j + m_k = m_i$, while the second represents the loss in n_i due to the mergers of PBHs of mass m_i with any other PBH mass. The merger kernel $\mathcal{K}_{ij} = \langle \sigma_{\rm merg}(i,j)\, v_{\rm rel} \rangle$, capturing the rate of PBH mergers, is the velocity-weighted average of Eq. 2. Assuming the relative velocities of PBHs in the initial virialized cluster follow a Maxwell–Boltzmann distribution, the kernel is given by [23]

$$\mathcal{K}_{ij} = \mathcal{A} \frac{G^2 m_{\rm pbh}^2}{c^3} \left(\frac{v_0}{c}\right)^{-11/7} \left(\frac{m_i m_j}{m_{\rm pbh}^2}\right)^p \left(\frac{m_i + m_j}{m_{\rm pbh}}\right)^q \tag{11}$$

with $\mathcal{A} = 85^{2/7}(2\pi)^{11/14} 3^{1/2} \Gamma(5/7)$, $v_0 = \sqrt{3/5}\, v_{\rm vir}$ the initial (root-mean-square) velocity of PBHs, and $(p, q) = (15/14, 9/14)$ without considering mass segregation.

Fig. 2 shows the PBH mass population evolution in an initial cluster of $N_{\rm cl} = 10^5$ monochromatic PBHs from redshift $z \simeq 11.5$ to 6.0 with $n_{\rm cl} = 2 \times 10^8\,{\rm pc}^{-3}$ and $v_{\rm vir} = 512\,{\rm km\,s}^{-1}$, where we utilize the Monte Carlo scheme [58–65] to solve Eq. 10 using a full-conditioning method for sampling [59]. For the simulation to be valid, it can be justified that the coalescence time of two PBHs is much shorter than the disruption timescale caused by a third PBH, so that three-body interactions can be ignored during mergers. A SMBH of mass $M_\bullet \simeq 10^6\,{\rm M}_\odot$ emerges by $z \sim 6$ if the initial PBHs are of $m_{\rm pbh} = 30\,{\rm M}_\odot$, which can mainly explain the LRDs that host $M_\bullet \simeq 10^{5\text{–}8}\,{\rm M}_\odot$, ranging from $z \simeq 4$–11, with or without the accretion of surrounding gas. In the simulation, the runaway merging timescale $t_{\rm rm} \approx 0.2/n_{\rm cl}\mathcal{K}_{00}$ for completing the merger is smaller if $n_{\rm cl}$ is larger. This is consistent with the given upper bound, $1/n_{\rm cl}\mathcal{K}_{00}$, shown in Ref. [23].

The merger of a PBH cluster can be divided into three stages. The first stage involves only small-mass PBH mergers, initially at high redshifts. In the second stage, PBHs with intermediate masses begin to form and boost the merger process due to the increased merger cross section. The third stage eventually arises when the number density of these intermediate-mass PBHs grows beyond a certain threshold, and an SMBH will emerge with the surrounding small-mass PBHs, forming an *extreme mass ratio inspiral system*. The overall merger processes will provide a distinct GW signature of this mechanism.

Fig. 3 shows the stochastic GW background associated with the PBH merging processes. A key feature of the GW spectrum is the presence of two peaks: one at low

frequencies and the other at relatively high frequencies. The first peak arises due to the initially large abundance of small-mass PBHs in the first stage in Fig. 2. The third stage explains the emergence of the peak at low frequencies: The strength of the GW signal is proportional to a positive power of the chirp mass, while the maximum frequency is inversely proportional to the total mass. After the binary formation at relatively low redshifts, mergers of PBHs with their central SMBH contribute significantly at low frequencies, thereby enhancing the GW spectrum. The infrared behavior of the gravitational wave spectrum follows a $\sim f^3$ scaling, which is much steeper than the $f^{2/3}$ scaling of a single binary inspiral event. This is because orbital eccentricity transfers power from the fundamental harmonic to higher harmonics, significantly suppressing the GW background at frequencies below the source frequency while enhancing it at intermediate frequencies, see Ref. [66] for details.

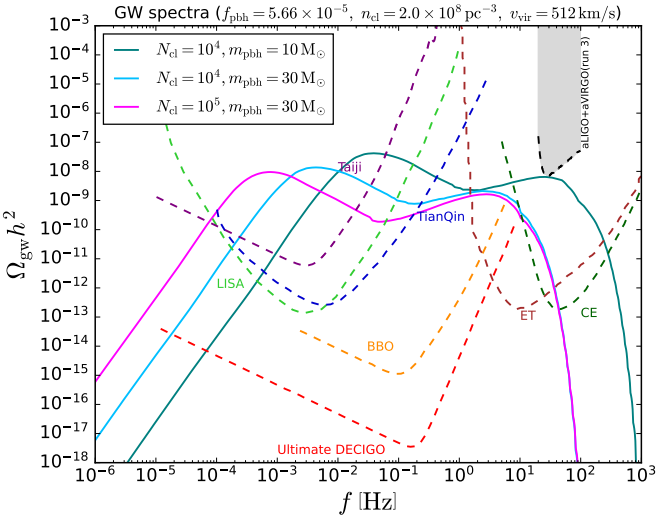

FIG. 3: Stochastic GW background from PBH cluster merger processes. Sensitivities of future GW detectors are also shown [67–73].

Accretion Growth and the Spin from Tidal Fields.— The mass $M_\bullet$ and spin $J_\bullet$ of the most massive black hole are determined mainly by mergers and mass accretion. The accretion efficiency is sensitive to the spin of accreting black holes through the radiative efficiency $\epsilon_M(a_\bullet) = 1 - \tilde{E}_{\rm ISCO}(a_\bullet)$, such that ϵ_M increases monotonically with $a_\bullet \equiv J_\bullet/J_{\rm max}$, the dimensionless spin, where $J_{\rm max} \equiv GM_\bullet^2/c$ is the maximal spin of a Kerr black hole, and $\tilde{E}_{\rm ISCO}$ is the unit mass energy at the innermost stable circular orbit [74, 75]. For example, $\epsilon_M(1) = 0.42$, while $\epsilon_M(0.95) = 0.19$. A small difference in spin results in a substantial difference in radiative efficiency. The enhancement of mass growth through Eddington accretion is given by [75]

$$\frac{M_\bullet}{m_{\rm pbh}} = \mathcal{F}_{\rm merg} \exp\left[\frac{\epsilon_L\,(1-\epsilon_M)}{\epsilon_M\,\tau_{\rm acc}} t_{\rm acc}\right], \qquad (12)$$

where $\mathcal{F}_{\rm merg} \propto N_{\rm cl}$ is the merger enhancement factor, ϵ_L the accretion luminosity efficiency, ϵ_M the radiative efficiency, $\tau_{\rm acc} \approx 0.4\,{\rm Gyr}$ the characteristic accretion timescale, and $t_{\rm acc}$ the accretion time of baryon gas.

For Eddington-limited accretion ($\epsilon_L = 1$), the spin can be driven to its maximal equilibrium value ($a_\bullet \simeq 1$ for a thin disk; $a_\bullet \simeq 0.95$ for an MHD disk) through gas accretion in $t_{\rm acc} \sim 0.1\,\tau_{\rm acc} \approx 40\,{\rm Myr}$, and the mass growth is limited to at most 10^4 in $t_{\rm acc} \sim 2\,\tau_{\rm acc} \approx 800\,{\rm Myr}$ for the MHD disk model ($\epsilon_M = 0.19$). Thus, a sufficiently heavy seed is favored at the beginning of the accretion phase to explain those LRDs hosting $10^{5-8}\,{\rm M}_\odot$ SMBHs in the first billion years. Crucially, a slower mass growth rate can also reconcile the cumulative mass density accreted onto black holes with the SMBH mass density at $z \simeq 4$–5, based on the Sołtan argument [20], if their bolometric luminosities are estimated via spectral energy distribution fitting employing dust-obscured AGN models [21].

Nevertheless, accretion strongly depends on the abundance of surrounding matter and the specific accretion model. For example, chaotic accretion of short-lived episodes with random orientations tends to damp the black hole spin down to average values of $a_\bullet \simeq 0.2$ [76]. Moreover, cosmological and hydrodynamical simulations have shown that the gas angular momentum directions of subsequent accretion episodes become uncorrelated with the black hole spin once $M_\bullet \gtrsim 2\times10^7\,{\rm M}_\odot$ [77].

Regardless of accretion, our model can generate SMBHs with large spin, even if the PBHs had zero spin at their formation [17, 78]. The PBH cluster acquires significant spin through interactions with tidal fields induced by other clusters, due to its extended merger history and non-sphericity. As a result, the spin of the SMBH reflects the residual spin of the cluster after mergers. To estimate the magnitude, we consider the quadrupole interaction model [79] and a homogeneous ellipsoid model for clusters. The dimensionless spin parameter of the cluster is estimated as (see Supplemental Material):

$$a_{\rm cl} \equiv cJ_{\rm cl}/GM_{\rm cl}^2 \simeq 5.6\,\mathcal{H}(e,p) \times \left(\frac{f_{\rm pbh}}{10^{-5}}\right)^{-1/2}\left(\frac{\xi_\alpha}{10^6}\right)^{-1}\left(\frac{C}{20}\right)^{-1/6}\left(\frac{r_{\rm cl}}{10\,{\rm kpc}}\right)^{-1}, \qquad (13)$$

where $\mathcal{H}(e,p)$, with the ellipticity e and prolateness p, characterizes the non-sphericity ($\mathcal{H}(0,0) = 0$ if spherical) of the cluster according to peak theory [17, 78, 80–82], $\mathcal{H}(\bar{e} \simeq 0.09, \bar{p} \simeq 0) \simeq 0.53$ on average.

To explain SMBHs of $10^{5-8}\,{\rm M}_\odot$ in the LRDs (light-gray in Fig. 1), the runaway timescale $t_{\rm rm}$ must be sufficiently short, $t_{\rm rm} \lesssim 1\,{\rm Gyr}$ ($z \gtrsim 5$). Remarkably, the cluster spin acquired through tidal interactions in this region exceeds the Kerr limit, implying a nearly maximal residual spin ($a_\bullet \simeq 1$) of the SMBH after mergers. The remaining PBHs, forming extreme mass ratio inspiral systems with the central SMBH, cannot merge until the GW carries away all the spin remnant of cluster. Besides, the darker band in this region exhibits $t_{\rm rm} \gtrsim 400\,{\rm Myr}$ for rapid SMBH emergence up to the LRD redshifts ($z \lesssim 11$).

Discussion.— In addition to the requirements discussed in the preceding sections, the green region in Fig. 1 shows the experimental constraints on PBH abundance, $f_{\rm pbh}$ [18, 19, 83], derived from gravitational lensing, Lyman-α measurements, X-ray observations [84], and dynamical friction [85]. As we focus on PBHs with an initial mass of $\mathcal{O}(10)\,{\rm M}_\odot$, their abundance is only weakly constrained, with $f_{\rm pbh} \lesssim 10^{-2}$. Especially, following successive mergers of PBHs in the cluster, their number density declines significantly over time, making them less subject to these observational constraints. The primary constraint on the PBH abundance then arises only from the SMBH mass density of $\rho_\bullet \sim 10^5\,{\rm M}_\odot\,{\rm Mpc}^{-3}$ in the local universe, currently accounting for $f_{\rm pbh} \sim 10^{-5}$ dark matter density – consistent with our parameter choice.

The most stringent constraints on long-wavelength perturbations Φ_l arise from CMB spectral distortions [86–90], see also Refs. [18, 19]. COBE/FIRAS observations [91] impose upper limits on the μ and y distortion parameters, with $|\mu| \lesssim 9\times10^{-5}$ and $|y| \lesssim 1.5\times10^{-5}$. These constraints can be evaded through the long-short mode coupling. In our model, SMBHs do not form from the direct collapse of long-wavelength perturbations but instead emerge from successive mergers of small-mass PBHs. Crucially, the variance of long modes in our scenario is sufficiently small, $\sigma_l \lesssim 0.0064$, allowing the μ-distortion constraint to be satisfied while maintaining a sufficient PBH abundance $f_{\rm pbh} \sim 10^{-5}$, which will ultimately result in the SMBH population. The y-distortion constraint is weaker than that from μ-distortions in the mass range $10^{5-8}\,{\rm M}_\odot$ for LRDs, while it primarily excludes PBH clusters that would result in ultra-massive black holes, $M_\bullet \gtrsim 10^{11}\,{\rm M}_\odot$, the purple region in Fig. 1.

In the simulation, the runaway merger time of the PBH cluster is conservative, as we have neglected the effect of mass segregation due to dynamical friction. When mass segregation is included [66], the runaway merger can be even faster, as low-velocity massive PBHs will sink toward the center, and increase the merging cross section. This indeed prevents the relativistic instability of core collapse [92], while it facilitates the merger and further relaxes the parameter space constraints in Fig.1.

Conclusions.— Initial small-scale PBH clustering could potentially explain the rapid formation of SMBHs within the first billion years after the Big Bang, particularly the LRDs observed in recent JWST data. We have considered a simplified model in which small-scale clustering arises from long-wavelength mode modulation at the PBH formation scale. In this framework, a central SMBH emerges through successive mergers of light PBHs within the cluster. Due to tidal interactions from other clusters, the resulting SMBH is expected to be highly spinning, and extreme mass ratio inspiral systems with the central black hole will naturally form [24]. The GW signature associated with this mechanism may be detected by next-generation GW observatories such as the Einstein Telescope [93–95], and are distinguishable from alternative SMBH formation models, e.g., heavy PBH seeds [36, 96] or gravothermal core collapse [97–102].

We thank Simeon Bird and Shi Pi for correspondence. This work is supported in part by the National Key R&D Program of China under Grand Nos 2023YFA1607104 and 2021YFC2203100, the National Science Foundation of China under Grant No. 12475107, the China Postdoctoral Science Foundation under Grant No. 2024M761594, and the Tsinghua University Dushi program.

[1] J. Matthee *et al.*, "Little Red Dots: An Abundant Population of Faint Active Galactic Nuclei at z ∼ 5 Revealed by the EIGER and FRESCO JWST Surveys," *Astrophys. J.* **963** no. 2, (2024) 129, arXiv:2306.05448 [astro-ph.GA].

[2] F. Pacucci, B. Nguyen, S. Carniani, R. Maiolino, and X. Fan, "JWST CEERS and JADES Active Galaxies at z = 4–7 Violate the Local $M_\bullet$–$M_\star$ Relation at $>3\sigma$: Implications for Low-mass Black Holes and Seeding Models," *Astrophys. J. Lett.* **957** no. 1, (2023) L3, arXiv:2308.12331 [astro-ph.GA].

[3] F. Gentile *et al.*, "Not-so-little Red Dots: Two Massive and Dusty Starbursts at z ∼ 5–7 Pushing the Limits of Star Formation Discovered by JWST in the COSMOS-Web Survey," *Astrophys. J. Lett.* **973** no. 1, (2024) L2, arXiv:2408.10305 [astro-ph.GA].

[4] H. B. Akins, C. M. Casey, E. Lambrides, N. Allen, I. T. Andika, M. Brinch, J. B. Champagne, O. Cooper, X. Ding, N. E. Drakos, *et al.*, "Cosmos-web: The over-abundance and physical nature of" little red dots"–implications for early galaxy and smbh assembly," arXiv:2406.10341 [astro-ph.GA].

[5] C. C. Williams, S. Alberts, Z. Ji, K. N. Hainline, J. Lyu, G. Rieke, R. Endsley, K. A. Suess, F. Sun, B. D. Johnson, *et al.*, "The galaxies missed by hubble and alma: the contribution of extremely red galaxies to the cosmic census at 3< z< 8," *Astrophys. J.* **968** no. 1, (2024) 34, arXiv:2311.07483.

[6] Y. Harikane, Y. Zhang, K. Nakajima, M. Ouchi, Y. Isobe, Y. Ono, S. Hatano, Y. Xu, and H. Umeda, "A JWST/NIRSpec First Census of Broad-Line AGNs at $z = 4$–7: Detection of 10 Faint AGNs with $M_{BH} \sim 10^6$–$10^8 M_{sun}$ and Their Host Galaxy Properties," *Astrophys. J.* **959** no. 1, (2023) 39, arXiv:2303.11946 [astro-ph.GA].

[7] R. Maiolino *et al.*, "JADES - The diverse population of infant black holes at 4 < z < 11: Merging, tiny, poor, but mighty," *Astron. Astrophys.* **691** (2024) A145, arXiv:2308.01230 [astro-ph.GA].

[8] D. D. Kocevski, M. Onoue, K. Inayoshi, J. R. Trump, P. A. Haro, A. Grazian, M. Dickinson, S. L. Finkelstein, J. S. Kartaltepe, M. Hirschmann, *et al.*, "Hidden little monsters: spectroscopic identification of low-mass, broad-line agns at z> 5 with ceers,"

Astrophys.J.Lett. **954** no. 1, (2023) L4, arXiv:2302.00012 [astro-ph.GA].

[9] V. Kokorev, S. Fujimoto, I. Labbe, J. E. Greene, R. Bezanson, P. Dayal, E. J. Nelson, H. Atek, G. Brammer, K. I. Caputi, *et al.*, "UNCOVER: A NIRSpec Identification of a Broad-line AGN at z = 8.50," *Astrophys. J. Lett.* **957** no. 1, (2023) L7, arXiv:2308.11610 [astro-ph.GA].

[10] M. Killi, D. Watson, G. Brammer, C. McPartland, J. Antwi-Danso, R. Newshore, D. Coe, N. Allen, J. P. Fynbo, K. Gould, *et al.*, "Deciphering the jwst spectrum of a 'little red dot'at z 4.53: An obscured agn and its star-forming host," *Astron. Astrophys.* **691** (2024) A52, arXiv:2312.03065 [astro-ph.GA].

[11] V. Kokorev, K. I. Caputi, J. E. Greene, P. Dayal, M. Trebitsch, S. E. Cutler, S. Fujimoto, I. Labbé, T. B. Miller, E. Iani, *et al.*, "A Census of Photometrically Selected Little Red Dots at $4 < z < 9$ in JWST Blank Fields," *Astrophys. J.* **968** no. 1, (2024) 38, arXiv:2401.09981 [astro-ph.GA].

[12] B. Wang, A. de Graaff, R. L. Davies, J. E. Greene, J. Leja, G. B. Brammer, A. D. Goulding, T. B. Miller, K. A. Suess, A. Weibel, *et al.*, "RUBIES: JWST/NIRSpec Confirmation of an Infrared-luminous, Broad-line Little Red Dot with an Ionized Outflow,". http://arxiv.org/abs/2403.02304.

[13] E. Durodola, F. Pacucci, and R. C. Hickox, "Exploring the Active Galactic Nucleus Fraction of a Sample of JWST's Little Red Dots at $4 < z < 8$: Overmassive Black Holes Are Strongly Favored," *Astrophys. J.* **985** no. 2, (2025) 169, arXiv:2406.10329 [astro-ph.GA].

[14] T. T. Ananna, A. Bogdán, O. E. Kovács, P. Natarajan, and R. C. Hickox, "X-Ray View of Little Red Dots: Do They Host Supermassive Black Holes?," *Astrophys. J. Lett.* **969** no. 1, (2024) L18, arXiv:2404.19010 [astro-ph.GA].

[15] D. D. Kocevski, S. L. Finkelstein, G. Barro, A. J. Taylor, A. Calabrò, B. Laloux, J. Buchner, J. R. Trump, G. C. Leung, G. Yang, *et al.*, "The Rise of Faint, Red AGN at $z>4$: A Sample of Little Red Dots in the JWST Extragalactic Legacy Fields,". http://arxiv.org/abs/2404.03576.

[16] M. Mirbabayi, A. Gruzinov, and J. Noreña, "Spin of Primordial Black Holes," *JCAP* **03** (2020) 017, arXiv:1901.05963 [astro-ph.CO].

[17] V. De Luca, V. Desjacques, G. Franciolini, A. Malhotra, and A. Riotto, "The initial spin probability distribution of primordial black holes," *JCAP* **05** (2019) 018, arXiv:1903.01179 [astro-ph.CO].

[18] M. Sasaki, T. Suyama, T. Tanaka, and S. Yokoyama, "Primordial black holes—perspectives in gravitational wave astronomy," *Class. Quant. Grav.* **35** no. 6, (2018) 063001, arXiv:1801.05235 [astro-ph.CO].

[19] B. Carr, K. Kohri, Y. Sendouda, and J. Yokoyama, "Constraints on primordial black holes," *Rept. Prog. Phys.* **84** no. 11, (2021) 116902, arXiv:2002.12778 [astro-ph.CO].

[20] A. Soltan, "Masses of quasars," *Mon. Not. Roy. Astron. Soc.* **200** (1982) 115–122.

[21] K. Inayoshi and K. Ichikawa, "Birth of Rapidly Spinning, Overmassive Black Holes in the Early Universe," *Astrophys. J. Lett.* **973** no. 2, (2024) L49, arXiv:2402.14706 [astro-ph.GA].

[22] V. Atal, A. Sanglas, and N. Triantafyllou, "LIGO/Virgo black holes and dark matter: The effect of spatial clustering," *JCAP* **11** (2020) 036, arXiv:2007.07212 [astro-ph.CO].

[23] H. Mouri and Y. Taniguchi, "Runaway merging of black holes: analytical constraint on the timescale," *Astrophys. J. Lett.* **566** (2002) L17–L20, arXiv:astro-ph/0201102.

[24] W.-X. Feng, S. Bird, and H.-B. Yu, "Gravitational Waves from Primordial Black Hole Dark Matter Spikes," *Astrophys. J.* **986** no. 2, (2025) 151, arXiv:2411.05065 [astro-ph.CO].

[25] C. T. Byrnes, S. Nurmi, G. Tasinato, and D. Wands, "Inhomogeneous non-Gaussianity," *JCAP* **03** (2012) 012, arXiv:1111.2721 [astro-ph.CO].

[26] C. T. Byrnes, E. J. Copeland, and A. M. Green, "Primordial black holes as a tool for constraining non-Gaussianity," *Phys. Rev. D* **86** (2012) 043512, arXiv:1206.4188 [astro-ph.CO].

[27] S. Young and C. T. Byrnes, "Primordial black holes in non-Gaussian regimes," *JCAP* **08** (2013) 052, arXiv:1307.4995 [astro-ph.CO].

[28] S. Young and C. T. Byrnes, "Long-short wavelength mode coupling tightens primordial black hole constraints," *Phys. Rev. D* **91** no. 8, (2015) 083521, arXiv:1411.4620 [astro-ph.CO].

[29] S. Young and C. T. Byrnes, "Signatures of non-gaussianity in the isocurvature modes of primordial black hole dark matter," *JCAP* **04** (2015) 034, arXiv:1503.01505 [astro-ph.CO].

[30] Y. Tada and S. Yokoyama, "Primordial black holes as biased tracers," *Phys. Rev. D* **91** no. 12, (2015) 123534, arXiv:1502.01124 [astro-ph.CO].

[31] G. Franciolini, A. Kehagias, S. Matarrese, and A. Riotto, "Primordial Black Holes from Inflation and non-Gaussianity," *JCAP* **03** (2018) 016, arXiv:1801.09415 [astro-ph.CO].

[32] V. Desjacques and A. Riotto, "Spatial clustering of primordial black holes," *Phys. Rev. D* **98** no. 12, (2018) 123533, arXiv:1806.10414 [astro-ph.CO].

[33] Y. Ali-Haïmoud, "Correlation Function of High-Threshold Regions and Application to the Initial Small-Scale Clustering of Primordial Black Holes," *Phys. Rev. Lett.* **121** no. 8, (2018) 081304, arXiv:1805.05912 [astro-ph.CO].

[34] S. Young and C. T. Byrnes, "Initial clustering and the primordial black hole merger rate," *JCAP* **03** (2020) 004, arXiv:1910.06077 [astro-ph.CO].

[35] T. Suyama and S. Yokoyama, "Clustering of primordial black holes with non-Gaussian initial fluctuations," *PTEP* **2019** no. 10, (2019) 103E02, arXiv:1906.04958 [astro-ph.CO].

[36] V. De Luca, G. Franciolini, and A. Riotto, "Heavy Primordial Black Holes from Strongly Clustered Light Black Holes," *Phys. Rev. Lett.* **130** no. 17, (2023) 171401, arXiv:2210.14171 [astro-ph.CO].

[37] V. De Luca, G. Franciolini, A. Riotto, and H. Veermäe, "Ruling Out Initially Clustered Primordial Black Holes as Dark Matter," *Phys. Rev. Lett.* **129** no. 19, (2022) 191302, arXiv:2208.01683 [astro-ph.CO].

[38] M. Y. Khlopov, S. G. Rubin, and A. S. Sakharov, "Primordial structure of massive black hole clusters," *Astropart. Phys.* **23** (2005) 265, arXiv:astro-ph/0401532.

[39] V. Dokuchaev, Y. Eroshenko, and S. Rubin, "Quasars formation around clusters of primordial black holes," *Grav. Cosmol.* **11** (2005) 99–104, arXiv:astro-ph/0412418.

[40] K. M. Belotsky, V. I. Dokuchaev, Y. N. Eroshenko, E. A. Esipova, M. Y. Khlopov, L. A. Khromykh, A. A. Kirillov, V. V. Nikulin, S. G. Rubin, and I. V. Svadkovsky, "Clusters of primordial black holes," *Eur. Phys. J. C* **79** no. 3, (2019) 246, arXiv:1807.06590 [astro-ph.CO].

[41] Q. Ding, T. Nakama, J. Silk, and Y. Wang, "Detectability of Gravitational Waves from the Coalescence of Massive Primordial Black Holes with Initial Clustering," *Phys. Rev. D* **100** no. 10, (2019) 103003, arXiv:1903.07337 [astro-ph.CO].

[42] H.-L. Huang and Y.-S. Piao, "Toward supermassive primordial black holes from inflationary bubbles," *Phys. Rev. D* **110** no. 2, (2024) 023501, arXiv:2312.11982 [astro-ph.CO].

[43] M. M. Flores and A. Kusenko, "Primordial Black Holes from Long-Range Scalar Forces and Scalar Radiative Cooling," *Phys. Rev. Lett.* **126** no. 4, (2021) 041101, arXiv:2008.12456 [astro-ph.CO].

[44] J. M. Ezquiaga, J. García-Bellido, and V. Vennin, "The exponential tail of inflationary fluctuations: consequences for primordial black holes," *JCAP* **03** (2020) 029, arXiv:1912.05399 [astro-ph.CO].

[45] J. M. Ezquiaga, J. García-Bellido, and V. Vennin, "Massive Galaxy Clusters Like El Gordo Hint at Primordial Quantum Diffusion," *Phys. Rev. Lett.* **130** no. 12, (2023) 121003, arXiv:2207.06317 [astro-ph.CO].

[46] C. Animali and V. Vennin, "Clustering of primordial black holes from quantum diffusion during inflation," *JCAP* **08** (2024) 026, arXiv:2402.08642 [astro-ph.CO].

[47] V. De Luca, G. Franciolini, and A. Riotto, "Bubble correlation in first-order phase transitions," *Phys. Rev. D* **104** no. 12, (2021) 123539, arXiv:2110.04229 [hep-ph].

[48] T. Harada, C.-M. Yoo, and K. Kohri, "Threshold of primordial black hole formation," *Phys. Rev. D* **88** no. 8, (2013) 084051, arXiv:1309.4201 [astro-ph.CO]. [Erratum: Phys.Rev.D 89, 029903 (2014)].

[49] M. V. Smoluchowski, "Drei Vortrage uber Diffusion, Brownsche Bewegung und Koagulation von Kolloidteilchen," *Zeitschrift für Physik* **17** (1916) 557–585.

[50] M. V. Smoluchowski, "Versuch einer mathematischen theorie der koagulationskinetik kolloider lösungen," *Zeitschrift für Physikalische Chemie* **92** no. 1, (1917) 129–168. http://dx.doi.org/10.1515/zpch-1918-9209.

[51] S. Chandrasekhar, "Stochastic problems in physics and astronomy," *Rev. Mod. Phys.* **15** (1943) 1–89.

[52] E. W. Kolb and I. I. Tkachev, "Large amplitude isothermal fluctuations and high density dark matter clumps," *Phys. Rev. D* **50** (1994) 769–773, arXiv:astro-ph/9403011.

[53] E. Flanagan, "Hoop conjecture for black-hole horizon formation," *Phys. Rev. D* **44** no. 8, (1991) 2409–2420.

[54] V. A. Antonov, "The Most Probable Phase Distribution in Spherical Stellar Systems and the Conditions Governing Its Existence,".

[55] D. Lynden-Bell, R. Wood, and A. Royal, "The Gravo-Thermal Catastrophe in Isothermal Spheres and the Onset of Red-Giant Structure for Stellar Systems," *Mon. Not. Roy. Astron. Soc.* **138** no. 4, (1968) 495–525.

[56] T. Padmanabhan, "Antonov instability and gravothermal catastrophe-revisited," *Astrophys. J. Suppl.* **71** (1989) 651–664.

[57] P.-H. Chavanis, "Gravitational instability of finite isothermal spheres," *Astron. Astrophys.* **381** (2002) 340, arXiv:astro-ph/0103159.

[58] D. T. Gillespie, "The stochastic coalescence model for cloud droplet growth," *J. Atmos. Sci.* **29** no. 8, (1972) 1496–1510.

[59] D. T. Gillespie, "An exact method for numerically simulating the stochastic coalescence process in a cloud," *J. Atmos. Sci.* **32** no. 10, (1975) 1977–1989.

[60] A. L. Garcia, C. Van Den Broeck, M. Aertsens, and R. Serneels, "A Monte Carlo simulation of coagulation," *Phys. A: Stat. Mech. Appl.* **143** no. 3, (1987) 535–546.

[61] K. Liffman, "A direct simulation Monte-Carlo method for cluster coagulation," *J. Comput. Phys.* **100** no. 1, (1992) 116–127.

[62] H. Babovsky, "On a Monte Carlo scheme for Smoluchowski's coagulation equation," *Monte Carlo Methods Appl.* **5** no. 1, (1999) 1–18.

[63] R. I. Patterson, W. Wagner, and M. Kraft, "Stochastic weighted particle methods for population balance equations," *J. Comput. Phys.* **230** no. 19, (2011) 7456–7472.

[64] G. Kotalczyk and F. Kruis, "A Monte Carlo method for the simulation of coagulation and nucleation based on weighted particles and the concepts of stochastic resolution and merging," *J. Comput. Phys.* **340** (2017) 276–296.

[65] Q. D. Tran, V. Sorichetti, G. Pehau-Arnaudet, M. Lenz, and C. Leduc, "Fragmentation and entanglement limit vimentin intermediate filament assembly," *Phys. Rev. X* **13** no. 1, (2023) 011014.

[66] B. Zhang, W.-X. Feng, and H. An. In preparation, 2025.

[67] M. Colpi *et al.*, "LISA Definition Study Report," arXiv:2402.07571 [astro-ph.CO].

[68] W.-H. Ruan, Z.-K. Guo, R.-G. Cai, and Y.-Z. Zhang, "Taiji program: Gravitational-wave sources," *Int. J. Mod. Phys. A* **35** no. 17, (2020) 2050075, arXiv:1807.09495 [gr-qc].

[69] **TianQin** Collaboration, J. Mei *et al.*, "The TianQin project: current progress on science and technology," *PTEP* **2021** no. 5, (2021) 05A107, arXiv:2008.10332 [gr-qc].

[70] H. Kudoh, A. Taruya, T. Hiramatsu, and Y. Himemoto, "Detecting a gravitational-wave background with next-generation space interferometers," *Phys. Rev. D* **73** (2006) 064006, arXiv:gr-qc/0511145.

[71] S. Kawamura *et al.*, "Current status of space gravitational wave antenna DECIGO and B-DECIGO," *PTEP* **2021** no. 5, (2021) 05A105, arXiv:2006.13545 [gr-qc].

[72] A. Abac *et al.*, "The Science of the Einstein Telescope," arXiv:2503.12263 [gr-qc].

[73] **KAGRA, Virgo, LIGO Scientific** Collaboration, R. Abbott *et al.*, "Upper limits on the isotropic gravitational-wave background from Advanced LIGO and Advanced Virgo's third observing run," *Phys. Rev. D* **104** no. 2, (2021) 022004, arXiv:2101.12130 [gr-qc].

[74] J. M. Bardeen, W. H. Press, and S. A. Teukolsky, "Rotating black holes: Locally nonrotating frames, energy extraction, and scalar synchrotron radiation," *Astrophys. J.* **178** (1972) 347.

[75] S. L. Shapiro, "Spin, accretion and the cosmological growth of supermassive black holes," *Astrophys. J.* **620** (2005) 59–68, arXiv:astro-ph/0411156.

[76] A. R. King, J. E. Pringle, and J. A. Hofmann, "The Evolution of Black Hole Mass and Spin in Active Galactic Nuclei," *Mon. Not. Roy. Astron. Soc.* **385** (2008) 1621, arXiv:0801.1564 [astro-ph].

[77] L. Sala, M. Valentini, V. Biffi, and K. Dolag, "Supermassive black hole spin evolution in cosmological simulations with OPENGADGET3," *Astron. Astrophys.* **685** (2024) A92, arXiv:2312.07657 [astro-ph.GA].

[78] C.-M. Yoo, "Primordial black hole formation from a nonspherical density profile with a misaligned deformation tensor," *Phys. Rev. D* **110** no. 4, (2024) 043526, arXiv:2403.11147 [gr-qc].

[79] P. J. E. Peebles, "Origin of the Angular Momentum of Galaxies," *Astrophys. J.* **155** (1969) 393.

[80] J. M. Bardeen, J. R. Bond, N. Kaiser, and A. S. Szalay, "The Statistics of Peaks of Gaussian Random Fields," *Astrophys. J.* **304** (1986) 15–61.

[81] R. K. Sheth, H. J. Mo, and G. Tormen, "Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes," *Mon. Not. Roy. Astron. Soc.* **323** (2001) 1, arXiv:astro-ph/9907024.

[82] A. Escrivà and C.-M. Yoo, "Simulations of Ellipsoidal Primordial Black Hole Formation," arXiv:2410.03452 [gr-qc].

[83] B. Carr, F. Kuhnel, and M. Sandstad, "Primordial Black Holes as Dark Matter," *Phys. Rev. D* **94** no. 8, (2016) 083504, arXiv:1607.06077 [astro-ph.CO].

[84] Y. Inoue and A. Kusenko, "New X-ray bound on density of primordial black holes," *JCAP* **10** (2017) 034, arXiv:1705.00791 [astro-ph.CO].

[85] B. J. Carr and M. Sakellariadou, "Dynamical constraints on dark compact objects," *Astrophys. J.* **516** (1999) 195–220.

[86] J. Chluba, A. L. Erickcek, and I. Ben-Dayan, "Probing the inflaton: Small-scale power spectrum constraints from measurements of the CMB energy spectrum," *Astrophys. J.* **758** (2012) 76, arXiv:1203.2681 [astro-ph.CO].

[87] J. Chluba, R. Khatri, and R. A. Sunyaev, "CMB at 2x2 order: The dissipation of primordial acoustic waves and the observable part of the associated energy release," *Mon. Not. Roy. Astron. Soc.* **425** (2012) 1129–1169, arXiv:1202.0057 [astro-ph.CO].

[88] M. Kawasaki, A. Kusenko, and T. T. Yanagida, "Primordial seeds of supermassive black holes," *Phys. Lett. B* **711** (2012) 1–5, arXiv:1202.3848 [astro-ph.CO].

[89] K. Kohri, T. Nakama, and T. Suyama, "Testing scenarios of primordial black holes being the seeds of supermassive black holes by ultracompact minihalos and CMB μ-distortions," *Phys. Rev. D* **90** no. 8, (2014) 083514, arXiv:1405.5999 [astro-ph.CO].

[90] T. Nakama, B. Carr, and J. Silk, "Limits on primordial black holes from μ distortions in cosmic microwave background," *Phys. Rev. D* **97** no. 4, (2018) 043525, arXiv:1710.06945 [astro-ph.CO].

[91] D. J. Fixsen, E. S. Cheng, J. M. Gales, J. C. Mather, R. A. Shafer, and E. L. Wright, "The Cosmic Microwave Background spectrum from the full COBE FIRAS data set," *Astrophys. J.* **473** (1996) 576, arXiv:astro-ph/9605054.

[92] W.-X. Feng, H.-B. Yu, and Y.-M. Zhong, "Dynamical instability of collapsed dark matter halos," *JCAP* **05** no. 05, (2022) 036, arXiv:2108.11967 [astro-ph.CO].

[93] S. Hild, S. Chelkowski, and A. Freise, "Pushing towards the ET sensitivity using 'conventional' technology," arXiv:0810.0604 [gr-qc].

[94] S. Hild, "Beyond the Second Generation of Laser-Interferometric Gravitational Wave Observatories," *Class. Quant. Grav.* **29** (2012) 124006, arXiv:1111.6277 [gr-qc].

[95] N. Singh, T. Bulik, K. Belczynski, and A. Askar, "Exploring compact binary populations with the Einstein Telescope," *Astron. Astrophys.* **667** (2022) A2, arXiv:2112.04058 [astro-ph.HE].

[96] D. Hooper, A. Ireland, G. Krnjaic, and A. Stebbins, "Supermassive primordial black holes from inflation," *JCAP* **04** (2024) 021, arXiv:2308.00756 [astro-ph.CO].

[97] W.-X. Feng, H.-B. Yu, and Y.-M. Zhong, "Seeding Supermassive Black Holes with Self-interacting Dark Matter: A Unified Scenario with Baryons," *Astrophys. J. Lett.* **914** no. 2, (2021) L26, arXiv:2010.15132 [astro-ph.CO].

[98] M. G. Roberts, L. Braff, A. Garg, S. Profumo, T. Jeltema, and J. O'Donnell, "Early formation of supermassive black holes from the collapse of strongly self-interacting dark matter," *JCAP* **01** (2025) 060, arXiv:2410.17480 [astro-ph.GA].

[99] F. Jiang, Z. Jia, H. Zheng, L. C. Ho, K. Inayoshi, X. Shen, M. Vogelsberger, and W.-X. Feng, "Formation of the Little Red Dots from the Core-collapse of Self-interacting Dark Matter Halos," arXiv:2503.23710 [astro-ph.GA].

[100] T. Shen, X. Shen, H. Xiao, M. Vogelsberger, and F. Jiang, "Massive Black Holes Seeded by Dark Matter – Implications for Little Red Dots and Gravitational Wave Signatures," arXiv:2504.00075 [astro-ph.GA].

[101] W.-X. Feng, H.-B. Yu, and Y.-M. Zhong, "Dark Bondi Accretion Aided by Baryons and the Origin of JWST Little Red Dots," arXiv:2506.17641 [astro-ph.GA].

[102] M. G. Roberts, L. Braff, A. Garg, S. Profumo, and T. Jeltema, "Little Red Dots from Ultra-Strongly Self-Interacting Dark Matter," arXiv:2507.03230 [astro-ph.GA].

[103] B. Carr and J. Silk, "Primordial Black Holes as Generators of Cosmic Structures," *Mon. Not. Roy. Astron. Soc.* **478** no. 3, (2018) 3756–3775, arXiv:1801.00672 [astro-ph.CO].

[104] B.-Y. Su, N. Li, and L. Feng, "An inflation model for massive primordial black holes to interpret the JWST observations," arXiv:2306.05364 [astro-ph.CO].

— Supplemental Material —

Little Red Dots from Small-Scale Primordial Black Hole Clustering

Borui Zhang, Wei-Xiang Feng, and Haipeng An

Appendix S.1: The two-point correlation function of PBHs

A simple parametrization of the long-mode modulation is given by [22]:

$$\nu(\boldsymbol{x}) = \nu_g \left(1 + \eta\, \Phi_l(\boldsymbol{x})\right) , \tag{S.1–1}$$

where η denotes the coupling strength between long and short modes, and $\Phi_l(\boldsymbol{x})$ is the long-wavelength density perturbation field. The unmodulated reduced threshold is defined as $\nu_g \equiv \delta_c/\bar{\sigma}_s$ with $\delta_c \simeq 0.414$ the critical density contrast for PBH formation and $\bar{\sigma}_s^2 = \langle \delta_s^2 \rangle$ the spatially averaged variance of the short-mode density perturbations. Both Φ_l and δ_s are assumed to follow Gaussian statistics. The probability of forming a PBH and the joint probability of finding two PBHs separated by a distance r, after integrating over the configuration of the long mode are, respectively,

$$\begin{aligned} P_1 = \frac{1}{2}\mathrm{erfc}\left(\frac{\nu(\boldsymbol{x})}{\sqrt{2}}\right) \quad \text{and} \quad P_2 = \frac{1}{4}\left[\mathrm{erfc}\left(\frac{\nu_1}{\sqrt{2}}\right) + \mathrm{erfc}\left(\frac{\nu_2}{\sqrt{2}}\right) + \mathrm{sgn}(\nu_1)\mathrm{sgn}(\nu_2) - 1\right] \\ - T\left(\nu_1, \frac{\nu_2 - \omega_s(r)\nu_1}{\nu_1\sqrt{1-\omega_s^2(r)}}\right) - T\left(\nu_2, \frac{\nu_1 - \omega_s(r)\nu_2}{\nu_2\sqrt{1-\omega_s^2(r)}}\right) , \end{aligned} \tag{S.1–2}$$

where $\nu_i = \nu(\boldsymbol{x}_i)$ is the local reduced threshold, $\omega_s(r) = \langle \delta(\boldsymbol{x}_1)\delta(\boldsymbol{x}_2)\rangle / \langle \delta(0)^2 \rangle$ the normalized two-point correlation function of the short-wavelength density field, $\mathrm{erfc}(x) \equiv \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} \mathrm{d}t$, and $T(z,u)$ the Owen T-function, given by

$$T(z,u) = \frac{1}{2\pi} \int_0^u \frac{e^{-\frac{(1+t^2)z^2}{2}}}{1+t^2}\, \mathrm{d}t \, . \tag{S.1–3}$$

For this simple modulation model, the N-point correlation function of PBHs can be computed analytically [22]. In particular, the two-point correlation function $\xi_{\mathrm{pbh}}(r)$ of PBHs is given by

$$1 + \xi_{\mathrm{pbh}}(r) = \frac{P_2(r)}{P_1^2} \overset{\nu \gg 1}{\longrightarrow} \frac{\left(1+\bar{\omega}(r)\right)^{3/2}}{\left(1-\bar{\omega}(r)\right)^{1/2}} \exp\left(\nu_g^2 \frac{\bar{\omega}(r)}{1+\bar{\omega}(r)}\right) , \tag{S.1–4}$$

where $\bar{\omega}(r) \equiv \frac{\omega_s(r) + \alpha^2 \omega_l}{1+\alpha^2}$ with $\alpha(\eta,\kappa) \equiv \delta_c \eta \kappa$ and $\kappa \equiv \sigma_l/\bar{\sigma}_s$ denoting the ratio of long- to short-mode variances. For simplicity, we adopt a model for the primordial power spectrum composed of two delta functions:

$$\begin{cases} \mathcal{P}_{s,\delta}(k) = \sigma_s^2 k_s \delta(k - k_s) \, , \\ \mathcal{P}_{l,\Phi}(k) = \sigma_l^2 k_l \delta(k - k_l) \, , \end{cases} \tag{S.1–5}$$

where $k_{s,l}$ represent the wavenumbers of the short and long modes, and $\sigma_{s,l}$ are the corresponding variances.

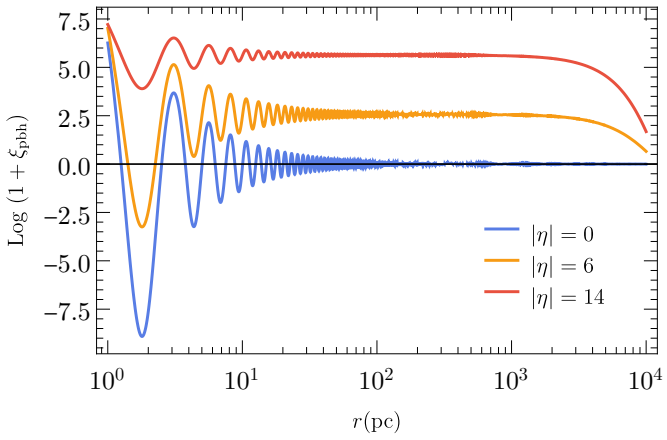

FIG. 4: The PBH two-point correlation function for $|\eta| = 0$, 6 and 14 given $\nu_g = 8.5$, $\kappa = 0.1313$.

Fig. 4 displays the PBH two-point correlation function for different coupling strengths $|\eta| = 0$, 6 and 14, given $\nu_g = 8.5$, $\kappa = 0.1313$, and $k_s \simeq 2.5\,\mathrm{pc}^{-1}$ corresponding to $m_{\rm pbh} = 30\,\mathrm{M}_\odot$. We set the ratio of long to short wavelengths as $k_l/k_s = 10^{-4}$. Two distinct features are observed in the correlation function. The first drop occurs at the characteristic scale of an individual PBH, $r_{\rm pbh} \sim 1/k_s$, while the second drop appears at the typical clustering scale set by the long mode, $r_{\rm cl} \sim 1/k_l$. When $|\eta| \neq 0$, a plateau emerges in the intermediate regime $r_{\rm pbh} < r < r_{\rm cl}$, where $\xi_{\rm pbh}(r)$ remains nearly constant at ξ_α. This plateau reflects the enhancement of PBH clustering induced by the long-mode modulation.

Appendix S.2: The total number of PBHs and the number density in a cluster

The total number of PBHs in a cluster can be estimated using the two-point correlation function [36, 37]:

$$N_{\rm cl} \approx \bar{n}_{\rm pbh} \int \mathrm{d}^3x\, \xi_{\rm pbh}(r) \approx \frac{4\pi}{3} \bar{n}_{\rm pbh} \xi_\alpha r_{\rm cl}^3 \tag{S.2–1}$$

where $\bar{n}_{\rm pbh}$ is the average comoving number density of PBHs:

$$\bar{n}_{\rm pbh} = \frac{\Omega_{\rm dm}}{\Omega_{\rm M}} f_{\rm pbh} \frac{\rho_{\rm eq}}{m_{\rm pbh}} a_{\rm eq}^3 \approx f_{\rm pbh} \left(\frac{30\,\mathrm{M}_\odot}{m_{\rm pbh}}\right) \mathrm{kpc}^{-3}\,. \tag{S.2–2}$$

Here, we have used the ratio of dark matter to total matter abundance $\Omega_{\rm dm}/\Omega_{\rm M} \approx 0.85$, the PBH fraction of dark matter $f_{\rm pbh} \equiv \Omega_{\rm pbh}/\Omega_{\rm dm}$, and the matter energy density at matter-radiation equality $\rho_{\rm eq} \equiv \Omega_{\rm M}\,\rho_{\rm crit,0}/a_{\rm eq}^3$, with the scale factor $a_{\rm eq} \simeq 1/3401$. The present critical density is taken to be $\rho_{\rm crit,0} \simeq 1.27 \times 10^{-7}\,\mathrm{M}_\odot\,\mathrm{pc}^{-3}$ and $\Omega_{\rm M} \simeq 0.3$.

We consider the PBH cluster to evolve as normal matter, diluting with the cosmic expansion until it decouples from the Hubble flow. This decoupling occurs when the energy density of the cluster, $\rho_{\rm cl} = m_{\rm pbh} n_{\rm cl}$, becomes comparable to the radiation energy density, $\rho_{\rm rad}$. After this point, the number density $n_{\rm cl}$ becomes constant in comoving coordinates. The cluster's energy density at decoupling, with scale factor $a_{\rm dec}$, is given by:

$$\rho_{\rm cl}(a_{\rm dec}) = \xi_\alpha \bar{\rho}_{\rm pbh}(a_{\rm dec}) = \xi_\alpha \bar{\rho}_{\rm pbh}(a_{\rm eq}) \left(\frac{a_{\rm eq}}{a_{\rm dec}}\right)^3 = \xi_\alpha\,(0.85 f_{\rm pbh} \rho_{\rm eq}) \left(\frac{a_{\rm eq}}{a_{\rm dec}}\right)^3\,, \tag{S.2–3}$$

where $\bar{\rho}_{\rm pbh} = m_{\rm pbh}\bar{n}_{\rm pbh}$. On the other hand, the radiation energy density redshifts as:

$$\rho_{\rm rad}(a_{\rm dec}) = \rho_{\rm eq} \left(\frac{a_{\rm eq}}{a_{\rm dec}}\right)^4\,. \tag{S.2–4}$$

Equating $\rho_{\rm cl}(a_{\rm dec}) = \rho_{\rm rad}(a_{\rm dec})$, we find the scale factor at decoupling:

$$a_{\rm dec} = \frac{a_{\rm eq}}{0.85 f_{\rm pbh} \xi_\alpha}\,, \tag{S.2–5}$$

which is valid for $0.85 f_{\rm pbh}\xi_\alpha > 1$. After decoupling and virialization, the cluster undergoes further contraction. We account for this by introducing the compactness parameter $C \gg 1$, which enhances the final cluster number density:

$$n_{\rm cl} = C(0.85 f_{\rm pbh}\xi_\alpha)^4 \frac{\rho_{\rm eq}}{m_{\rm pbh}} \simeq 2 \times 10^8\,\mathrm{pc}^{-3} \left(\frac{f_{\rm pbh}\xi_\alpha}{25}\right)^4 \left(\frac{C}{20}\right) \left(\frac{30\,\mathrm{M}_\odot}{m_{\rm pbh}}\right)\,. \tag{S.2–6}$$

Appendix S.3: The spin of clusters from tidal interactions

We adopt a model in which PBH clusters are approximated as homogeneous ellipsoids with semi-axes $r_1 \geq r_2 \geq r_3$ during their hierarchical merging phase leading to SMBH formation. The tidal torque yields

$$\tau_i = \epsilon_{ijk} Q_{jl} T_{lk}\,, \tag{S.3–1}$$

where Q_{ij} is the quadrupole moment of the cluster, $T_{ij} = GM_{\rm cl}\left(\delta_{ij} - 3\hat{x}_i\hat{x}_j\right)/r^3$ the tidal tensor of a source at (r, θ, ϕ) with the unit vector $\hat{x}_i = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. The ensemble-averaged squared tidal torque is computed by integrating over all external cluster configurations

$$\langle \tau^2 \rangle = \int_{r_{\rm vir}}^{\infty} P(r) \left(\sum_{i=1}^{3} \tau_i^2\right) \mathrm{d}^3x\,, \tag{S.3–2}$$

where $P(r)$ is the two-point spatial distribution function of PBH clusters, describing their correlation at separation r, and is given as follows.

We assume that all PBH clusters have the same mass and are homogeneously distributed within a maximum comoving distance $R^c_{\rm max}$, which defines the local clustering environment. The comoving mean inter-cluster distance is given by

$$\bar{R}^c_{\rm cl} \simeq \left(\frac{4\pi\bar{n}_{\rm pbh}}{3N_{\rm cl}}\right)^{-1/3} = \xi_\alpha^{1/3} r_{\rm cl}\,, \tag{S.3–3}$$

where $\xi_\alpha \gg 1$ ensures a high clustering factor, and hence $\bar{R}^c_{\rm cl} \gg r_{\rm cl}$. We set $R^c_{\rm max} \simeq (4/3)\bar{R}^c_{\rm cl}$ using the definition of the average distance $\bar{R}^c_{\rm cl} = \int_{|\vec{r}|<R^c_{\rm max}} |\vec{r}-\vec{r}_0|\,{\rm d}^3x / \int_{|\vec{r}|<R^c_{\rm max}} {\rm d}^3x$ for a reference cluster fixed at the origin $\vec{r}_0 = 0$. Evaluating the physical separation at the decoupling epoch $a_{\rm dec} \simeq a_{\rm eq}/(0.85 f_{\rm pbh}\xi_\alpha)$, we find:

$$R^p_{\rm max}(a_{\rm dec}) \simeq a_{\rm dec}\bar{R}^c_{\rm cl} = a_{\rm eq}/(0.85 f_{\rm pbh}\xi_\alpha)\cdot\frac{4}{3}\xi_\alpha^{1/3} r_{\rm cl} = \frac{4}{3}\left(C\xi_\alpha\right)^{1/3} r_{\rm vir}\,, \tag{S.3–4}$$

where we have used the virial radius expression:

$$r_{\rm vir} = \left(\frac{3M_{\rm cl}}{4\pi\rho_{\rm cl}}\right)^{1/3} = \left(\frac{3N_{\rm cl}}{4\pi n_{\rm cl}}\right)^{1/3} = (0.85 f_{\rm pbh}\xi_\alpha)^{-1} C^{-1/3} a_{\rm eq}\, r_{\rm cl}\,. \tag{S.3–5}$$

Thus, the probability distribution for finding another cluster at position $\vec{r}_0 + \vec{r}$, given a cluster at $\vec{r}_0 = 0$, is approximately:

$$P(r) \simeq \begin{cases} \frac{3}{4\pi}(R^p_{\rm max})^{-3}\,, & r_{\rm vir} < r \lesssim R^p_{\rm max} \\ 0\,, & \text{otherwise} \end{cases}\,. \tag{S.3–6}$$

Consequently, the ensemble-averaged tidal torque acting on a cluster is given by:

$$\langle\tau^2\rangle^{1/2} = \frac{27}{40\sqrt{5}}\left(\frac{GM^2_{\rm cl}}{r_{\rm vir}}\right)\left(C\xi_\alpha\right)^{-1/2}\left(\frac{a_{\rm dec}}{a}\right)^{3/2}\left[\left(1-\frac{r_2^2}{r_1^2}\right)^2 + \left(1-\frac{r_3^2}{r_1^2}\right)^2 + \left(\frac{r_2^2}{r_1^2}-\frac{r_3^2}{r_1^2}\right)^2\right]^{1/2}\,, \tag{S.3–7}$$

where we have used $R^p_{\rm max}(a_{\rm dec}) \simeq (4/3)\left(C\xi_\alpha\right)^{1/3} r_{\rm vir}$ and identified $r_{\rm vir} = r_1$ as the major axis of the ellipsoid. Since $\langle\tau^2\rangle^{1/2} \propto a^{-3/2}$, the tidal torque weakens with cosmic expansion, indicating that most of the spin is accumulated shortly after the clusters decouple from the Hubble flow.

The final spin of a cluster can be estimated by integrating the ensemble-averaged tidal torque over the time interval during which the cluster remains dynamically responsive to the tidal field of neighboring clusters. The upper limit of this interval is approximated by the typical time before the cluster fully merges into SMBH:

$$J_{\rm cl} = \int_{t_{\rm dec}}^{t_{\rm merg}} \langle\tau^2\rangle^{1/2}\,{\rm d}t \tag{S.3–8}$$

where $t_{\rm dec}$ is the decoupling time and $t_{\rm merg} \lesssim t_{\rm rm}$ denotes the typical merging time. After performing the time-dependent integral

$$\int_{t_{\rm dec}}^{t_{\rm merg}} \frac{a_{\rm dec}^{3/2}}{a^{3/2}}\,{\rm d}t = \frac{2}{\sqrt{8\pi G\rho_{\rm eq}/3}}\left(0.85 f_{\rm pbh}\xi_\alpha\right)^{-3/2}\ln\left(\frac{\sqrt{1+a_{\rm merg}/a_{\rm eq}}+\sqrt{a_{\rm merg}/a_{\rm eq}}}{\sqrt{1+1/(0.85 f_{\rm pbh}\xi_\alpha)}+\sqrt{1/(0.85 f_{\rm pbh}\xi_\alpha)}}\right) \tag{S.3–9}$$

with $a_{\rm eq} \simeq 1/3401$, and $a_{\rm merg} \sim 1/10$ around redshift $z \sim 9$, we estimate the dimensionless spin parameter $a_s \equiv cJ_{\rm cl}/GM^2_{\rm cl}$ of the cluster as:

$$a_s \simeq 5.6\left(\frac{f_{\rm pbh}}{10^{-5}}\right)^{-1/2}\left(\frac{\xi_\alpha}{10^6}\right)^{-1}\left(\frac{C}{20}\right)^{-1/6}\left(\frac{r_{\rm cl}}{10\,{\rm kpc}}\right)^{-1}\left[\left(1-\frac{r_2^2}{r_1^2}\right)^2 + \left(1-\frac{r_3^2}{r_1^2}\right)^2 + \left(\frac{r_2^2}{r_1^2}-\frac{r_3^2}{r_1^2}\right)^2\right]^{1/2}\,. \tag{S.3–10}$$

While clusters are also subject to tidal torques induced by background density fluctuations, we neglect these contributions as they are subdominant compared to the torques from neighboring clusters for the parameter range considered.

According to the peak theory, the non-sphericity of the long-wavelength curvature perturbation can be characterized by its ellipticity e and prolateness p, which follow a statistical distribution determined by the power spectrum of the initial perturbations and the peak value [17, 78, 80–82]. We assume a monochromatic mass distribution and calculate the corresponding peak value using the two-point correlation function. The axis ratios are then given by

$$\frac{r_2}{r_1} = \sqrt{\frac{1-3\bar{e}+\bar{p}}{1+3\bar{e}+\bar{p}}}\,, \quad \text{and} \quad \frac{r_3}{r_1} = \sqrt{\frac{1-3\bar{e}+\bar{p}}{1-2\bar{p}}} \tag{S.3–11}$$

with the average values $\bar{e} \simeq 0.09$ and $\bar{p} \simeq 0$ as our inputs.

Appendix S.4: Redshift evolution of the most massive black hole with various cluster densities

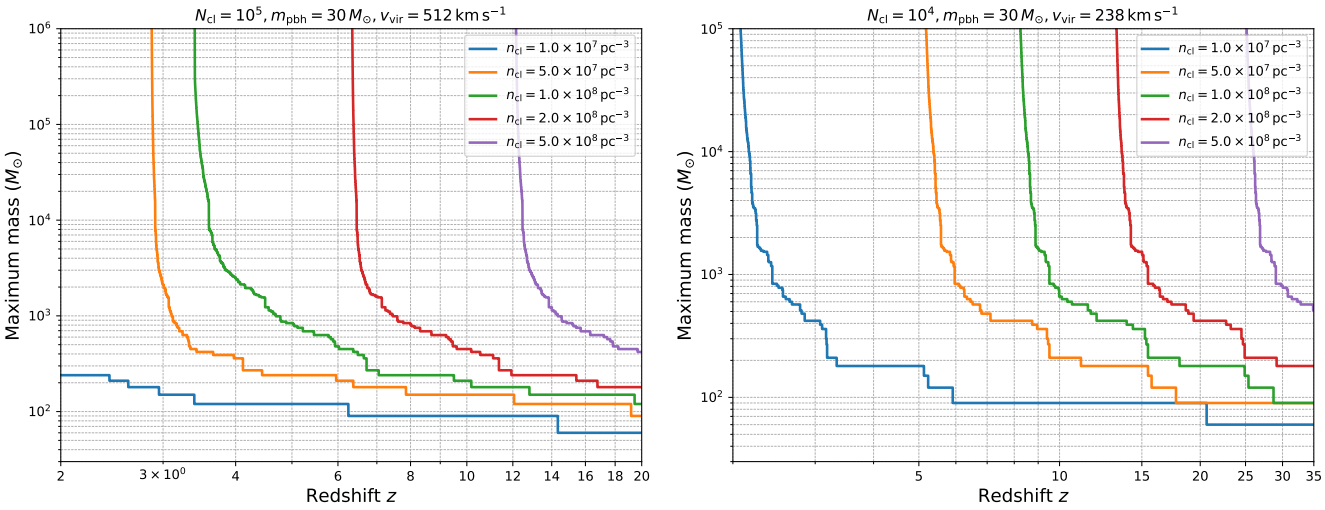

FIG. 5: Redshift evolution of the most massive black hole in a PBH cluster of $N_{\rm cl} = 10^5\,(10^4)$ PBHs of $m_{\rm pbh} = 30\,{\rm M}_\odot$ and virial velocity $v_{\rm vir} = 512\,(238)\,{\rm km\,s^{-1}}$ with various initial cluster densities $n_{\rm cl} = (1,\,5)\times10^7$ and $(1,\,2,\,5)\times10^8\,{\rm pc^{-3}}$.

Appendix S.5: Various timescales associated with a PBH cluster

Once two PBHs form a binary, they will eventually coalesce due to the emission of GWs. However, this process can be interrupted by encounters with a third PBH within the same cluster. To safely neglect such three-body interactions in our Monte Carlo simulation, we require that the inspiral timescale be shorter than the typical disruption timescale due to a third-body interaction. The coalescence timescale due to GW emission,

$$t_{\rm gw} \approx \frac{3c^5 r_{\rm p}^4 (1+\varepsilon)^{7/2}/(1-\varepsilon)^{1/2}}{85\,G^3(m_i+m_j)m_i m_j}\,, \tag{S.5–1}$$

must be shorter than the disruption timescale by a third PBH,

$$t_{\rm disrup} \approx \frac{1}{n_{\rm cl}\sigma_{\rm p} v_{\rm vir}}\,, \tag{S.5–2}$$

where $\varepsilon \lesssim 1$ is the initial orbital eccentricity and $\sigma_{\rm p} \equiv \pi r_{\rm p}^2 (1-\varepsilon)^{-2}$ the effective cross section of the newly formed binary with periapsis [24]

$$r_{\rm p} \approx \left(\frac{85\pi}{6\sqrt{2}}\frac{G^{7/2}(m_i+m_j)^{3/2}m_i m_j}{c^5 v_{\rm vir}^2}\right)^{2/7}\,. \tag{S.5–3}$$

Without mergers and losing energy through emitting GWs, an isolated PBH cluster will gradually shrink and evaporate when some PBHs undergo multiple gravitational scatterings during multiple (two-body) relaxation timescales to exceed the escape velocity, and this defines the evaporation timescale of a cluster:

$$t_{\rm evap} \approx \Gamma\, t_{\rm relx} \approx 3.4\,{\rm Myr}\left(\frac{\Gamma}{140}\right)\left(\frac{10}{\ln N_{\rm cl}}\right)\left(\frac{N_{\rm cl}}{10^5}\right)^{1/2}\left(\frac{m_{\rm pbh}}{30\,{\rm M}_\odot}\right)^{-1/2}\left(\frac{r_{\rm vir}}{0.02\,{\rm pc}}\right)^{3/2}\,, \tag{S.5–4}$$

where the relaxation time

$$t_{\rm relx} \approx \frac{0.1\,N_{\rm cl}}{\ln N_{\rm cl}}\frac{r_{\rm vir}}{v_{\rm vir}}\,, \tag{S.5–5}$$

and $1/\Gamma$ is the fractional loss of PBHs in the cluster, for which the energy is above the escape velocity in the Maxwellian distribution, per relaxation time $t_{\rm relx}$. The evaporation time is roughly the e-folding timescale of number reduction according to ${\rm d}N_{\rm cl}/{\rm d}t = -N_{\rm cl}/\Gamma\,t_{\rm relx} = -N_{\rm cl}/t_{\rm evap}$, and $\Gamma \approx 140$ for an isolated cluster. In reality, the seed effect [103, 104] of a PBH cluster will accumulate extra dark matter or baryons and increase the escape velocity such that $\Gamma \sim 3000\,(10^6)$ if the accumulated mass is comparable to (two times of) the cluster. Thus evaporation is not concerned.