Probing the Growth of Massive Black Holes with Black Hole–Host Galaxy Spin Correlations

Zhen Pan1 and Huan Yang1,2

1 Perimeter Institute for Theoretical Physics, Waterloo, ON N2L 2Y5, Canada; zpan@perimeterinstitute.ca, hyang@perimeterinstitute.ca
2 University of Guelph, Guelph, ON N1G 3G1, Canada

Received 2020 July 7; revised 2020 August 6; accepted 2020 August 21; published 2020 October 6

Abstract

Supermassive black holes are commonly found at the centers of their host galaxies, but their formation still remains an open question. In light of the tight correlation between the black hole (BH) mass and the velocity dispersions of the bulge component of the host galaxy, a BH–host galaxy coevolution scenario has been established. Such a description, however, still contains many theoretical uncertainties, including puzzles about the formation of BH seeds at high redshifts and the growth channel fueling these seeds. In this work, we systematically analyze the signatures of different growth channels on massive BH (MBH) spins. We show that different growth channels can be partially distinguished with the magnitudes of MBH spins inferred from extreme-mass-ratio inspirals detected by the Laser Interferometer Space Antenna. In addition, we propose measuring the correlation between the directions of MBH spins and their host galaxy spins, which is possible for extreme-mass-ratio inspirals happening in low-redshift galaxies (z \leq 0.3). With the inclusion of spin direction correlation different formation channels shall be significantly better constrained.

Unified Astronomy Thesaurus concepts: Supermassive black holes (1663); Galaxy mergers (608); Gravitational wave sources (677); Gravitational waves (678)

1. Introduction

Supermassive black holes (SMBHs) are commonly found at the centers of their host galaxies. Empirical correlations have also been extensively explored (Ferrarese & Merritt 2000; Kormendy & Ho 2013; McConnell & Ma 2013) between the black hole (BH) masses, \( M \), and different properties of their host galaxies, including the velocity dispersion \( \sigma_v \) of bulge stars in the host galaxies. The tight \( M-\sigma_v \) relation in combination with other correlations has inspired an interpretation that BHs and their host galaxies coevolve by regulating each other’s growth (Marconi & Hunt 2003; Kormendy & Ho 2013). The coevolution scenario provides a framework that connects the galaxy evolution with BH activities. However, there are still many important questions that this coevolution scenario provides no definitive answers for, especially the ones related to the SMBH formation, including the formation of BH seeds and the growth channels fueling these seeds.

BH seeds can generally be classified as light seeds with masses in the range of \( \sim (10^5, 10^6)M_\odot \) and heavy seeds within the range of \( \sim (10^4, 10^5)M_\odot \) (see, e.g., Rees 1984; Latif & Ferrara 2016; Haemmerlé et al. 2020 for reviews). The light seeds are thought to be results of the collapse of metal-free Population III stars (Madau & Rees 2001; Omukai 2001; Abel et al. 2002; Heger & Woosley 2002) and the heavy seeds are proposed to come from the direct collapse of a massive protogalactic gas cloud (Begelman et al. 2006; Mayer et al. 2010; di Matteo et al. 2012) or efficient merging stellar-mass compact objects in a gas-rich environment (Boco et al. 2020). Recently, Fragione & Silk (2020) proposed that repeated mergers of stellar-mass BHs in nuclear star clusters can produce both light and heavy seeds depending on the cluster masses and densities. Starting from the seeds at high redshifts, BHs further accumulate masses by either merging with other BHs or accreting gas, both of which seem to be compatible with available observations, including the \( M-\sigma_v \) relation (King 2003; Volonteri et al. 2003; Marconi et al. 2004; Murray et al. 2005; Volonteri & Rees 2005; King & Pringle 2006; Peng 2007). As shown by Sesana et al. (2008) and Sesana et al. (2011), both the seed formation and the BH growth history leave imprints on the mass function of massive BH (MBH) binaries, which can be probed from the LISA detection of MBH coalescence from redshift \( z \sim 10^{-15} \) to the local universe (Hughes 2002; Barausse et al. 2015; Amaro-Seoane et al. 2017).

Besides mass distributions, it is also natural to expect that different models of seed formation and BH growth will also lead to different signatures on MBH spins, as discussed in Berti & Volonteri (2008). MBH spins (both magnitudes and directions) can be accurately measured by LISA from extreme-mass-ratio-inspiral (EMRI) events (Haerta & Gair 2009; Babak et al. 2017; Gair et al. 2017), with the spin magnitudes expected to be measured with fractional uncertainty \( \sim (10^{-4}, 10^{-6}) \) and the spin directions expected to be constrained within \( \sim (10^{-2}) \) degree, which enables an accurate spin distribution reconstruction. In this work, we propose the inclusion of another observable—the correlation between the spin directions of MBHs and their host galaxies—to further sharpen our ability to distinguish different formation models. Such an observable relies on the host galaxy identification, which is possible for several percent of the EMRI events, and galaxy spectroscopic surveys (Bundy et al. 2015) to determine the galaxy spin orientation. We systematically analyze the spin signatures of different growth channels assuming the natural light seeds scenario, and explore how likely it is to distinguish various channels with the spin

3 We call BHs with masses \( \sim (10^5, 10^6)M_\odot \) MBHs.
4 Galaxy spin orientations inferred from spectroscopic surveys have been used to probe the initial conditions in the early universe (Motloch et al. 2020).
information of MBHs using a Bayes method. In particular, for
the first time we include the MBH–host galaxy spin
correlations in the analysis, which turns out to be a powerful
probe of these growth channels.

The structure of the paper is organized as follows. In
Section 2, we outline the growth channels and model their
signatures on MBH spins and on the BH–host galaxy spin
direction correlations. In Section 3, we introduce how well the
MBH spins can be extracted from the EMRI waveforms and
how many host galaxies of EMRIs can be identified in the
LISA mission. In Section 4, we show how to distinguish
different growth models from the spin information using
the method of Bayesian model selection. Some final remarks
are given in Section 5.

In this paper, we assume a flat ΛCDM cosmology with
Ω_m = 0.3 and H_0 = 70 km s^{-1} Mpc^{-1}, and we use geometrical
units G = c = 1.

2. Growth Mechanisms of MBHs and Their Impact on BH
Spins

In this section, we review different growth channels of
MBHs and obtain their corresponding implication on MBH
spins, including the magnitude distribution of MBH spins and
the MBH spin–galaxy spin correlation.

2.1. Accretion

Coherent Accretion. Accretion can be an efficient channel
for MBHs to gain their masses (Kawaguchi 2003; Kawaguchi
et al. 2004; King & Pringle 2006; Li 2012). If a central BH is
spun up with the large-scale gas fueling in a disk-like
configuration, the accretion is coherent. In the standard thin-
disk accretion, the BH could be spun up to a maximum value
a = 0.998 limited by the preferential accretion of low-angular-
momentum photons (Thorne 1974). In a magnetized disk, the
equilibrium spin is a ~ 0.95 (Gammie et al. 2004; Shapiro
2005). In this work, we do not intend to distinguish the subtle
differences arising from various assumptions on accretion physics, and we choose to describe the spin magnitudes of MBHs with a coarse probability distribution |a| ~ N(1, 0.05), where N(μ, σ) is a normal distribution with a mean value μ and a standard deviation σ (here we use σ = 0.05 to take account of typical variation of equilibrium BH spin magnitudes assuming different accretion physics). Following coherent accretion, the BH spins up to a nearly extremal state, with its spin direction nearly aligned with the rotation direction of the accretion disk. The initial spin magnitude and direction are no longer relevant after the BH mass increases by one or more e-folds. As a result, the BH spin direction should be aligned with the rotation direction \( \hat{L} \) of the large-scale gas disk. In reality, the gas disk may be turbulent and sometimes clumsy with its local rotation direction slightly off its mean value (see, e.g., Souza Lima et al. 2017). Though this variation is hard to calculate from the first principle, we may perform a rough estimate based on two quantities: the aspect ratio of the accretion disk \( h := H/r \) and the inclination angle \( \iota_{\text{gas,star}} \) between the gas disk and the stellar disk in the same galaxy. In the classical Mestel model of the gas disk (Mestel 1963), \( h \) is in the range of \( \sim 0.05, 0.1 \). The inclination angles \( \iota_{\text{gas,star}} \) are measured from galaxy spectroscopic surveys to be \( \sim 10^\circ \) (Krolik et al. 2019). Therefore, we take \( \delta L \approx 10^\circ \) as a reference and use an ansatz that \( \cos^{-1}(\hat{a} \cdot \hat{L}) \sim N(0, 10^\circ) \) in the following discussion.

Chaotic Accretion. The accretion is “chaotic” if it consists of many short chaotic episodes with a different accretion direction in each one. In this case, the distribution of BH spins is mainly determined by \( \Delta M/M \), the fractional BH mass increase in each episode. If \( \Delta M/M \gtrsim 1 \), the disk angular momentum in each episode is large enough to drive the BH to high spin no matter what the initial spin is; while in episodes with \( \Delta M/M \ll 1 \) BH tends to spin down, because the BH mass increases linearly with the number of accretion episodes \( N_{\text{acc}} \) while the angular momentum gain is proportional \( \sqrt{N_{\text{acc}}} \) due to the random-walk cancellation (King & Pringle 2006; King et al. 2008; Wang et al. 2009; Dotti et al. 2013; Volonteri et al. 2013; Li & Cao 2019; Zhang & Lu 2019).

In the beginning of each accretion episode, the BH spin is in general misaligned with the disk angular momentum, and the inner part of the disk will be wrapped in a viscous timescale (known as the Bardeen–Petterson effect (Bardeen & Petterson 1975)). The wrapped disk will exert a torque onto the central BH and align or antialign the BH spin with the angular momentum of the outer disk in a timescale (Scheuer & Feiler 1996; Natarajan & Pringle 1998; King et al. 2005; Lodato & Gerosa 2013; Gerosa et al. 2020) \( t_{\text{align}} \sim (M/M) \alpha^{5/3} a^{2/3} \approx h^{2/3} \) assuming a standard \( \alpha \)-disk description (Shakura & Sunyaev 1973), where \( M \) is the accretion rate and \( h = H/r \) is the aspect ratio of the disk. After the alignment, the central BH will be spun up as the gas is accreted. Together with the accretion timescale \( t_{\text{acc}} = \Delta M/M \), we obtain

\[
\frac{t_{\text{align}}}{t_{\text{acc}}} \approx 5 \times 10^{-3} |a|^{2/3} \alpha^{5/3} \left( \frac{\Delta M/M}{0.1} \right)^{2/3} \left( \frac{h}{0.1} \right)^{2/3}.
\]

Therefore, we can safely ignore the short alignment period in calculating the final BH spin of each accretion episode as long as \( \Delta M/M \gg 5 \times 10^{-3} |a|^{2/3} / \alpha^{5/3} / \left( \frac{\Delta M/M}{0.1} \right)^{2/3} \).

In each accretion episode, the BH angular momentum \( J \) changes with its mass \( M \) as

\[
dJ = l(a) M \Delta M_{\text{gas}} = \frac{l(a)}{\varepsilon(a)} M \Delta M,
\]

where \( J = aM^2 \), with \( a \) being the dimensionless spin, which is negative if the BH spin is antialigned with the angular momentum of the accretion disk; \( l(a) \) and \( \varepsilon(a) \) are the specific angular momentum and specific energy of particles on the innermost stable circular orbit (ISCO) \( r_{\text{ISCO}}(a) \) with the
is the mass element of accreted gas, a fraction \( \sim L_0 \) in each accretion episode, where solid lines are the simulation results and the dashed lines are the corresponding Gaussian fits. Here we take \( \sigma = 0.1 \mu \) as an example, and larger \( \sigma \), say, \( \sigma = 0.2 \mu \) only slightly broadens \( P_{\mu}, \mu = 1 \) and makes little change to remaining three distributions.

Following explicit forms (Bardeen et al. 1972)

\[
e(a) = \frac{r_{\text{isco}}^3/2 - 2 r_{\text{isco}}^{1/2} + a}{r_{\text{isco}}^{3/2} (r_{\text{isco}}^{1/2} - 3 r_{\text{ISCO}}^{1/2} + 2a)};
\]

\[
l(a) = \frac{r_{\text{ISCO}}^2 - 2 a r_{\text{ISCO}}^{1/2} + a^2}{r_{\text{ISCO}}^3 (r_{\text{ISCO}}^{1/2} - 3 r_{\text{ISCO}}^{1/2} + 2a)};
\]

\( dM_{\text{gas}} \) is the mass element of accreted gas, a fraction \( 1 - e(a) \) of which is converted to radiation escaping to infinity and the remaining fraction \( e(a) \) is absorbed by the BH. As a result, we obtain the following simple evolution equation

\[
da = \left[ \frac{l(a)}{e(a)} - 2a \right] d \ln M,
\]

where the specific angular momentum in retrograde accretion is larger than in direct accretion \( l(a < 0) > l(a > 0) \), leading to a larger spin magnitude change in retrograde accretion than in direct accretion for the same mass increase \( \Delta (\ln M) \).

We consider a simple chaotic accretion model: in the beginning of each episode (after the short alignment process), the BH spin is assumed to be aligned or antialigned with the angular momentum of the accretion disk with equal chance and the BH increases by \( \Delta M/M \sim N(\mu, 0.1/\mu) \). To take account of the equilibrium BH spin as in the coherent accretion case, we enforce a spin distribution \( P(a) \propto N(1) \) for \( a > 0.9 \). In Figure 1, we show four example models of chaotic accretion (ChA1, ChA2, ChA3, ChA4 with \( \mu = 4, 1, 0.2, 0.02 \) respectively), where the spin distribution \( P(|a|) \) is the same as that of coherent accretion for \( \mu \geq 1 \) because the angular momentum gain is large enough to drive the BH to high spin whatever the initial spin is, and \( P(|a|) \) peaks at zero for \( \mu < 1 \). In the case of \( \mu \ll 1 \), the BH will be spun up \(|a| \) in direct accretion \( (a > 0) \) and will be spun down \(|a| \) in retrograde accretion \( (a < 0) \). As the spinning down is more efficient than the spinning up (see the explanation following Equation (4)), the net result is an equilibrium distribution \( P(|a|) \) that peaks at \( a = 0 \) and decreases with \( |a| \). In the case of chaotic accretion, we expect no BH-host galaxy spin direction correlation, i.e., \( \hat{a} \cdot \hat{L} \sim \mathcal{U}(-1, 1) \), where \( \mathcal{U} \) is a uniform distribution.

2.2. Mergers

Dry Mergers. MBHs may merge following the merger of their host galaxies. This is considered as a natural growth channel for MBHs considering galaxies commonly harbor MBHs. These mergers can be further classified as wet mergers (mergers in a gas-rich environment) and dry mergers (mergers in a gas-poor environment) (see, e.g., Colpi 2014 for a review).

For dry mergers there are three main phases along the path to the final coalescence (Begelman et al. 1980): (i) an early phase of pairing when MBHs migrate inwards driven by the dynamical friction with background stars, until the two MBHs form a Keplerian binary (Chandrasekhar 1943; Begelman et al. 1980); (ii) a binary hardening phase when the binary separation decreases by ejecting stars of close encounters (Yu 2002; Milosavljević & Merritt 2003); (iii) a gravitational inspiral phase when the binary orbital decay is driven by the emission of GWs until the final coalescence. For mergers of nearly equal mass BHs, the spins of remnant BHs peak around \( a \approx 0.69 \), while mergers of small mass ratio BHs tend to produce remnant BHs with larger spin dispersion (Barausse & Rezzolla 2009). In the dry mergers of binary BHs, there is no apparent mechanism that aligns the BH spin directions \( \hat{a}_1, \hat{a}_2 \) with their orbital direction \( \hat{L}_{12} \) or aligns the orbital direction \( \hat{L}_{12} \) with the host galaxy spin direction \( \hat{L} \), i.e., \( \hat{a}_1, \hat{a}_2, \hat{L}_{12} \) are randomly oriented. Berti & Volonteri (2008) investigated the cosmological spin evolution of BHs driven by dry mergers (or isotropy mergers in their language) and they found the distribution of MBH spins peaks around \( \sim 0.7 \) with a long tail extending to small spins (we fit their histogram of BH spins with a skewed Gaussian distribution and plot in Figure 2). The spin direction of the remnant BH should also be randomly oriented with respect to the spin direction of the host galaxy, i.e., \( \hat{a} \cdot \hat{L} \sim \mathcal{U}(-1, 1) \).

Wet Mergers. If the two merging galaxies are gas rich, a stellar disk and a gas disk form in the remnant galaxy, so that the two MBHs undergo roughly four different phases before the final coalescence (Mayer et al. 2007): (i) an early pairing phase when MBHs wander beyond the scope of the gas disk and migrate inwards due to the dynamical friction with background stars; (ii) a pairing phase when the motion of MBHs is influenced by the gravity of the gas disk (called circumnuclear disk) and migrate inwards driven by the torque from the density-wave excitations in the disk (similar to the Type I planet migration) (Dotti et al. 2006, 2007; Mayer et al. 2007; Colpi et al. 2009; Mayer 2013); (iii) a hardening phase when the two MBHs form a Keplerian binary surrounded by a circumbinary disk and migrate under the binary-disk coupling arising from two opposing actions that the binary tidal fields open gaps in the disk whereas viscous torque fills the gaps (similar to the Type II planet migration) (Armitage & Natarajan 2002; Armitage 2013); (iv) an inspiral phase dominated by GW emission.

As an MBH with mass \( M \) rams into the disk with an inclination angle \( \iota \), the inclination angle will be damped by the BH–disk interaction. Now we calculate the timescale of this process. Assuming the disk volume density, surface density, circular velocity, and sound speed at radius \( r \) are \( \rho(r), \Sigma(r), V_c(r), \) and \( c_s(r) \), respectively. The disk aspect ratio \( \delta \) is approximately \( \delta \approx c_s(r)/V_c(r) \) and we take \( \delta = 0.1 \) as a fiducial value. Gas bounded by the BH within radius \( r_s = GM/V_c^2 \) will be shocked and accelerated to roughly the same velocity of the BH. In a self-gravitating circumnuclear disk, the BH speed is
In each accretion episode, there are misaligned. It is still not clear whether the interaction peaks on $pi$ (as shown above. In combination with the intrinsic is not easy to measure, whereas its 2D projection are also derived respectively. However, produced in the GM $R$ galaxy spin direction correlations \( \hat{L} \cdot \hat{L} \). Therefore, the disk-BH interacting force is roughly \( F = \hat{\nu}_{\text{gas}} V d V \) with \( \hat{\nu}_{\text{gas}} \) being the amount of shocked and accelerated gas per unit time and the inclination damping timescale is

\[
I_{\text{damp}} = \frac{M V_d \sin \iota}{F} \frac{\pi \sin \iota}{h} = \frac{M}{(G M_d^2 \Sigma(r)) V_d} r \sin \iota, \tag{5}
\]

where \( F = F \sin \iota \) is the component perpendicular to the disk and factor \( \pi \sin \iota / h \) takes account of the fact that only a fraction of the BH orbit is inside the disk for \( \iota > h \). We consider a Mestel model of the circumnuclear disk (Mestel 1963; Escala et al. 2005). The disk is self-gravitating and axisymmetric, with a constant rotational velocity \( V_d \), which is related to the total disk mass \( M_d \) and the disk size \( R_d \) by \( V_d = \sqrt{G M_d / R_d} \). The gas mass within radius \( r \) is \( \hat{M}_{\text{gas}}(r) = M_d r / R_d \) and disk surface density at each radius is \( \Sigma(r) = M_d / (2 \pi R_d r) \). Using these relations, we find

\[
I_{\text{damp}} = \frac{M_{\text{gas}}(r) 2 \pi r \sin \iota}{M V_d} \quad (r = r_f),
\]

where \( r_f \) is where \( M_{\text{gas}}(r) = M \), i.e., the transition radius from the circumnuclear disk phase to the circumbinary disk phase. The timescale of inclination damping is much shorter than the typical migration timescale, so we expect that the gas disk and the binary BH orbit are coplanar at the end of the circumnuclear disk phase, \( \hat{L}_2 \cdot \hat{L} \approx 1 \).

In the circumbinary disk phase, the inner part of the disk will be warped if the BH spin \( \hat{a}_{1,2} \) and the rotation direction of the disk \( \hat{L} \) are misaligned. It is still not clear whether the interaction between the BH binary and the warped disk can efficiently align them, though a number of studies have been performed previously (see, e.g., Dotti et al. 2010; Maio et al. 2013; Lodato & Gera 2013; Gerosa et al. 2020). However, as shown by Barausse & Rezzolla (2009) and Hofmann et al. (2016), the spin direction of the remnant BH \( \hat{a} \) produced in the final coalescence is roughly aligned with the direction of the total angular momentum of the binary BH system at the beginning of the gravitational inspiral phase to high precision with \( \cos^{-1}(\hat{a} \cdot \hat{J}_0) \approx N(0, 5') \). Here \( J_0 \) is the initial total angular momentum dominated by the orbital angular momentum \( \hat{L}_2 \) at large separations, and \( L_2 \) aligns with the spin direction of the gas disk \( L \) as shown above. In combination with the intrinsic \( \sim 10^2 \) variation in the gas disk direction \( \hat{L} \) (see Section 2.1), we obtain \( \cos^{-1}(\hat{a} \cdot \hat{L}) \approx N(0, \sqrt{(5')^2 + (10')^2}) \) . Berti & Volonteri (2008) also simulated the spin magnitudes of MBHs from aligned mergers and we again fit the histogram they obtained with a skewed Gaussian distribution (Figure 2).

At this point, we have outlined the basic pictures of the different growth channels of MBHs. The corresponding BH spin distribution \( P(\hat{a}) \) and the BH–host galaxy spin direction correlations \( \hat{a} \cdot \hat{L} \) are also derived respectively. However, \( \hat{a} \cdot \hat{L} \) is not an ideal observable, because the 3D galaxy spin direction \( \hat{L} \) is not easy to measure, whereas its 2D projection \( \hat{L}_2 \) onto the plane perpendicular to the line of sight (LoS) can be accurately measured to \( \approx 1^\circ \) precision in galaxy spectroscopic surveys (e.g., MaNGA, Bundy et al. 2015; Krolewski et al. 2019). Therefore, we also need to calculate the probability distributions of 2D spin direction correlation \( \hat{a} \cdot \hat{L}_2 \). For a given distribution \( P(\hat{a} \cdot \hat{L}) \), we sample 32 data \( \hat{a} \cdot \hat{L} \) points in concordance with the probability distribution; and for each pair of \( \hat{a}, \hat{L} \), we uniformly sample 1024 directions of LoS, project the 3D directions \( \hat{a}, \hat{L} \) onto the plane perpendicular to the LoS and calculate the 2D correlation \( \hat{a} \cdot \hat{L}_2 \). We finally obtain a histogram of all \( \hat{a} \cdot \hat{L}_2 \) data points and fit it with a smooth distribution function \( P(\hat{a} \cdot \hat{L}_2) \) (Figure 2). We find \( P(\hat{a} \cdot \hat{L}_2) \approx P(\hat{a} \cdot \hat{L}) \) for Wet Merger and Coherent Accretion, and \( P(\hat{a} \cdot \hat{L}_2) \) peaks on \( \pm 1 \) for Dry Merger and Chaotic Accretion with uniform \( P(\hat{a} \cdot \hat{L}) \) due to projection distortion, i.e., there is a large chance of projecting \( \hat{a} \cdot \hat{L} \approx 0 \) to
\[ \hat{a}_i \cdot \hat{L}_{1i} \approx \pm 1, \text{ while the chance of projecting } \hat{a}_i \cdot \hat{L}_{1i} \approx 0 \text{ is small.} \]

### 3. LISA Detection of EMRIs and Host Galaxies Identification

The expected EMRI rate depends on the mass function of the MBH population at different redshifts, the fraction of MBHs living in dense stellar cusps where stellar-mass BHs are produced, EMRI rate per MBH and properties of stellar-mass BHs in the cusps. Babak et al. (2017) quantified each of these astrophysical ingredients with semianalytical models and calculated the corresponding expected EMRI rates. They found that tens to thousands of EMRIs per year should be detectable by LISA taking into account astrophysical uncertainties. In particular, \( \sim 6 \) to \( \sim 180 \) low-redshift \( (z \leq 0.5) \) EMRIs are expected to be detected by LISA per year for the majority of the models considered (Gair et al. 2017). For all detectable EMRIs, the typical fractional errors of intrinsic parameters, e.g., redshifted masses and MBH spins, are found in the range of \( (10^{-6}, 10^{-5}) \) (Babak et al. 2017). Luminosity distance can be constrained with precision \( \sigma (\ln D_L) \approx \rho^{-1} \) and the median sky resolution is approximately \( \sigma (\Omega_L) \approx 0.05 (\rho/100)^{-5/2} \) degree\(^2\), where \( \rho \) is the signal-to-noise ratio (S/N) of the EMRI event (McGee et al. 2020).

Based on these results, we expect there is a fraction of low-redshift EMRIs that can be localized to small 3D error boxes containing a single galaxy. Therefore the host galaxies of these EMRIs can be identified from the corresponding LISA observation. Following the approach in Babak et al. (2017), we consider a power-law mass function of the EMRI population with a redshift-independent EMRI rate \( R_0 \) at redshift \( z \leq 0.5 \),

\[
\frac{dR_0}{d\ln M} \propto M^{\alpha} \quad \text{for} \quad 10^5 M_\odot < M < 10^6 M_\odot \quad (7)
\]

with power index \( \alpha = -0.5 \) or 0. At each redshift \( z \), we sample 128 EMRIs with the MBH mass sampled from Equation (7), the MBH spin chosen as \( \alpha = 0.98 \), the companion BH mass set to be \( m = 10 M_\odot \), the binary orbital eccentricity at plunge being \( e_p = 0.1 \), the luminosity distance \( D_L(z) \), and eight randomly sampled angles (including the source sky localization angles \( (\theta_0, \phi_0) \) and the MBH spin direction angles \( (\theta_0, \phi_0) \) that uniquely determine the binary configuration at coalescence (see Chua & Gair 2015; Chua et al. 2017 for details). For each EMRI, we model its GWs with the Augment Analytic Kludge (AAK) waveform (Chua et al. 2017) and record the time-domain waveforms \( h_1(t) \) and \( h_{\Pi}(t) \) in the last two years before coalescence (I and II mark the two orthogonal LISA channels), where (Barack & Cutler 2004; Rubbo et al. 2004)

\[
\begin{align*}
    h_1(t) &= h_{+}(t)F_{11}^+(t) + h_{\times}(t)F_{11}^\times(t), \\
    h_{\Pi}(t) &= h_{+}(t)F_{\Pi}^+(t) + h_{\times}(t)F_{\Pi}^\times(t),
\end{align*}
\]

(8) with \( h_{+\times} \) being the waveforms of two polarizations and \( F_{11\times} \) are the corresponding detector antenna patterns of the two channels.

The \( S/N \) of EMRI GWs is calculated as

\[
\rho = \sqrt{\langle h^*_1 h_1 \rangle + \langle h^*_\Pi h_\Pi \rangle} \quad \text{with the inner product defined as}
\]

\[
\langle u|v \rangle := 4 \int_0^{\infty} \mathcal{R}[u(f)v^*(f)] \frac{df}{P_n(f)},
\]

where \( \mathcal{R} \) denotes the real part and \( P_n(f) \) is the combination of one-side spectral density of the LISA detector noise and the residual foreground of unresolvable galactic binary white dwarfs (Amaro-Seoane et al. 2017; Robson et al. 2019). To keep consistent with the criteria used in Babak et al. (2017), we choose \( \rho = 20 \) as the threshold of EMRI detections. For EMRIs with \( \rho > 20 \), we forecast the model parameter constraints with Fisher

\[
F_{ij} = \left( \frac{\partial h_1}{\partial \lambda_i} \frac{\partial h_1}{\partial \lambda_j} \right) + \left( \frac{\partial h_\Pi}{\partial \lambda_i} \frac{\partial h_\Pi}{\partial \lambda_j} \right),
\]

(10) where \( \lambda_{ij} \) (\( i, j = 1, \ldots, 13 \)) are the EMRI model parameters briefed in the previous paragraph. The 1\( \sigma \) uncertainty of parameter \( \lambda_i \) is \( \sigma (\lambda_i) = \sqrt{C_{ii}} \) with \( C_{ij} := F_{ij}^{-1} \) being the covariance matrix.

The volume of the 3D error box is

\[
\delta V = r^3(z) \times \pi \sin \theta_f \sqrt{C_{\theta_\phi \phi_\phi} - C_{\theta_\phi \phi_\phi} C_{\phi_\phi \phi_\phi} \sigma (\ln D_L)},
\]

(11) where \( r(z) = \int_0^{\infty} \frac{dz}{\mathcal{P}(f)} \) is the comoving distance from the EMRI source to the Earth, with \( H(z) \) being the Hubble expansion rate. Inside the error box, we expect to see \( \delta N_{\text{gal}} \) galaxies, with \( \delta N_{\text{gal}} = D_{\text{gal}} \delta V \). Here the average number of galaxies per comoving volume is chosen as \( D_{\text{gal}} = 10^{-2} \text{ Mpc}^{-3} \) (consistent with Kuns et al. 2020). In Figure 3, we show the \( \delta N_{\text{gal}} \) distribution at each redshift, where we find that 13% of the detectable EMRIs with \( z < 0.5 \) can be traced back to their host galaxies, i.e., \( \sim 8 \) to \( \sim 234 \) EMRIs and their host galaxies can be identified by LISA in the maximum mission duration of \( \approx 10 \) years (Amaro-Seoane et al. 2017).

In Figure 4, we also show the sky resolution \( \delta \Omega \) and spin direction \( \hat{\alpha} \) resolution \( \delta \Omega \) of detectable EMRIs at each redshift, where \( \delta \Omega_{\text{a}} := 2 \pi \sin \theta_f \sqrt{C_{\theta_\phi \phi_\phi} - C_{\theta_\phi \phi_\phi} C_{\phi_\phi \phi_\phi} \sigma (\ln D_L)} \) is defined in a similar way. They will be used in estimating the data error bars of MBH spins and MBH–host galaxy spin direction correlations.

### 4. Bayesian Model Selection

As explained in the previous two sections, different growth channels will leave different imprints on MBH spins and on MBH–host galaxy spin direction correlations. The former can be measured by LISA from EMRIs, and the latter can be measured by LISA in combination with galaxy spectroscopic surveys. In this section, we will quantitatively explore how likely these channels can be distinguished given data of MBH spins and MBH–host galaxy spin correlations.

According to Bayes theorem, the posterior probability \( P(\lambda|D, m) \) for the parameters \( \lambda \) of a model \( m \) given data \( D \) is related to the likelihood \( P(D|\lambda, m) \) of seeing data \( D \) under model \( m \) with model parameter \( \lambda \) by

\[
P(\lambda|D, m) = \frac{P(D|\lambda, m)P(\lambda|m)}{\mathcal{E}(D|m)},
\]

(11) where \( P(\lambda|m) \) is the prior and \( \mathcal{E}(D|m) = \int P(D|\lambda, m)P(\lambda|m)d\lambda \) is the evidence of model
\( m \) given data \( D \). To determine the (de)preference of models \( m_1 \) over \( m_2 \) based on data \( D \), we calculate the Bayes factor

\[
B_{m_2}^{m_1}(D) = \frac{\mathcal{E}(D|m_1)}{\mathcal{E}(D|m_2)}. \tag{12}
\]

If \( B_{m_2}^{m_1} > 1 \), \( m_1 \) is a better model than \( m_2 \); and if \( B_{m_2}^{m_1} < 1 \), \( m_2 \) is better. According to Jeffreys' evidence scale, \( \ln B_{m_2}^{m_1} > 5 \) is interpreted as \( m_1 \) being overwhelmingly better than \( m_2 \) and equivalently \( \ln B_{m_2}^{m_1} < -5 \) as \( m_2 \) being overwhelmingly better.

In the context of this work, there is no free parameter in the considered models (Figure 2), thus the evidence calculation is simplified as \( \mathcal{E}(D|m) = \mathcal{P}(D|m) \). In order to calculate the likelihood \( \mathcal{P}(D|m) \), we first divide data into \( N_D \) bins \( B_i \) \((i = 1, \ldots, N_D)\), and count the number of events in each bin \( n_i \). Each model will predict an average probability of events occurring in each bin \( P_i(m) = \int_{B_i} \mathcal{P}(d|m) \, dd \) (see Figure 2 for the probability distribution functions \( \mathcal{P}(d|m) \)). Events in different bins are independent and events in each bin should satisfy Poisson distribution with an average probability \( \rho_i \); therefore, the likelihood is written as (Gair et al. 2010, 2011)

\[
\mathcal{P}(D|m) = \prod_{i=1}^{N_D} \frac{[N_D(m)P_i(m)^{n_i} e^{-N_D(m)}P_i(m)]}{n_i!},
\]

\[
= [N_D(m)]^{n_D} e^{-N_D(m)} \prod_{i=1}^{N_D} \frac{[P_i(m)]^{n_i}}{n_i!}, \tag{13}
\]

where \( N_D(m) \) is the number of events predicted by model \( m \), and \( N_D(m)P_i(m) \) is the expected number of events in bin \( B_i \), and we have used the normalization condition \( \sum_i P_i(m) = 1 \) in the second equal sign. That is to say, different models predict not only different distributions \( P_i(m) \) of data but also different occurrences \( N_D(m) \) of data, both of which contribute to model selections. In fact, the total number \( N_D(m) \) is commonly more uncertain than the distribution \( P_i(m) \). In this paper, we will only use the distribution \( P_i(m) \) information to distinguish different channels, i.e., we set \( N_D(m) \) to be the same for different models, and we obtain

\[
\mathcal{P}(D|m) = \text{const} \times \prod_{i=1}^{N_D} \frac{[P_i(m)]^{n_i}}{n_i!}. \tag{14}
\]

In reality, any data point is subject to some measurement uncertainty, \( D = \{d^j \pm \Delta d^j \} \). We model the true value of event \( j \) with a probability distribution \( N(d - d^j, \Delta d^j) \) and assign a fractional occurrence \( \int_B N(d - d^j, \Delta d^j) \, dd \) into each bin \( B \).

In the continuum limit (small bins limit), Equation (14) simplifies as

\[
\mathcal{P}(D|m) = \text{const} \times \prod_{j=1}^{N_D} P_j(m), \tag{15}
\]

where \( P_j(m) = \int_B P(d|m)N(d - d^j, \Delta d^j) \, dd \).

In our case, data \( D \) includes both the MBH spins \( D_1 = \{\alpha_1^j \pm \delta_\alpha^j \} \) \((j = 1, \ldots, N_j)\) of the EMRIs detected by LISA and the MBH–host galaxy spin correlations \( D_2 = \{\hat{a}_1 \cdot \hat{L}_d \pm \delta_{aL} \} \) \((j = 1, \ldots, N_j)\), where the error bar \( \delta_{aL} \) is obtained from Fisher forecasts explained in the previous section, while \( \delta_{aL} \) depends on the spin direction uncertainty \( \delta\Omega_a \), the sky location uncertainty \( \delta\Omega_d \), the angle between the spin direction and the LoS \( \theta_{a,\text{LoS}} \), and the uncertainty \( \delta_{aL} \) of \( \hat{L}_d \). As shown in Bundy et al. (2015) and Krolewski et al. (2019), the galaxy spin direction \( \hat{L}_d \) can be measured with uncertainty \( \delta_{aL} \approx 1^\circ \). From Figure 4, the sky resolution \( \delta\Omega_d \) of LISA turns out to be \( \sim 30 \) times better than that of the MBH spin direction \( \delta\Omega_a \). As a result, we find \( \delta_{aL} \) is dominated by the uncertainty of the MBH spin direction \( \delta\Omega_a \). In terms of azimuthal angles, \( \hat{a}_1 \cdot \hat{L}_d \pm \delta_{aL} \) can be expressed as \( \cos^{-1}(\phi_{aL} \pm \delta\phi_{aL}) \), where \( \phi_{aL} \) is the angle between \( \hat{a}_1 \) and \( \hat{L}_d \), and \( \delta\phi_{aL} \) is its uncertainty.

To illustrate how different growth channels can be distinguished from each other based on the MBH spin magnitudes data \( D_1 \) and MBH–host galaxy spin correlation data \( D_2 \), we take a model \( m_2 \) as the true underlying model and generate 256 realizations of \( D_1 \) and \( D_2 \) sampled from the corresponding distributions shown in Figures 2 and 3. We conservatively take the data sizes of \( N_1 = 60 \) and \( N_2 = 8 \). For each realization of data, we can calculate the Bayes factors \( B_{m_2}^{m_1} \) of model \( m_1 \) relative to \( m_2 \) given data \( D_1 \) or given both data \( D_{1,2} \) using Equations (12, 15). In Table 1, we list the results of \( \ln B_{m_2}^{m_1} \), with \( m_2 \) being Wet Merger or Coherent Accretion, and \( m_1 \) being Dry Merger (DM), Wet Merger (WM), Coherent Accretion (CoA), or Chaotic Accretion (ChA). If WM is the true underlying model, we find it can be distinguished from DM/CoA/ChA1 with overwhelming evidence \( (\ln B_{m_2}^{m_1} > 5) \) at 12/2.1/2.1 (\( \sigma \)) confidence level, while it is indistinguishable from ChA2 with data \( D_1 \) only. Taking data \( D_2 \) into consideration, the Bayes factor contrasts \( \ln B_{m_2}^{m_1} \) increase by 15 for \( m_1 = \text{DM}/\text{ChA1}/\text{ChA2} \), and WM can be distinguished from DM/CoA/ChA1/ChA2 with overwhelming evidence at

![Figure 3](image-url)
In each column, we take \((\text{WM})\) inside are simply because they predict distinct spin distributions that can easily be found if the underlying model is CoA. The Astrophysical Journal, 195.

The formation of BH seeds and the growth history of MBHs leave imprints on the mass function of MBHs, on the distribution of MBH spins driven by coherent accretion. However, the realistic distributions in each channel may be considerably different. The bias or theoretical uncertainty in assessing the model distribution will inevitably affect the results of model selection. It is difficult to nail down all the theoretical uncertainties in this study, as it depends on many details of accretion that are hard to model from first principle: the disk thickness, the magnetic field strength, the configuration of magnetic fields lines, and the matter emission properties of the inner disk. More theoretical efforts along this direction are needed, otherwise in the coming epoch of LISA some of our understanding of MBH formation may be limited by the accuracy of modeling given all the astrophysical processes involved, instead of the data uncertainty.

In reality, more than one dominant channel may play important roles, so that the detected data may imply a mixed distribution from various channels. Then the question becomes how to determine the mixing ratios of various channels based on LISA observations and corresponding electromagnetic counterpart measurements. There have been some efforts toward more accurately modeling the MBH growth taking account of mixed channels and detailed astrophysics (see, e.g., Barausse 2012; Sesana et al. 2014; Kulier et al. 2015; Bhattacharyya & Mangalam 2020; Sayeb et al. 2020; Zhang et al. 2020). We expect that a similar discussion can be applied taking the full spin magnitude and MBH spin–galaxy spin correlation into account.

In this paper, we have shown the huge potential of probing the MBH growth via the MBH–galaxy spin direction correlation into account.
correlations, in addition to the spin magnitude distribution. As shown by Kuns et al. (2020), host galaxies ofstellar-mass binary BH mergers can be identified from combined observations of a deci-hertz GW detector and a ground-based detector to redshift $z \approx 0.3$, i.e., $\sim400$ pairs of binary BH–host galaxies would be identified per year (assuming a constant merger rate $60$ Gpc$^{-3}$ yr$^{-1}$). If there is an nonnegligible correlation between the orbital angular momentum of field-borne binaries and the rotation direction of their host galaxies, it will be interesting to explore what we can learn from these stellar-mass systems.

Z.P. and H.Y. are supported by the Natural Sciences and Engineering Research Council of Canada and in part by Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported in part by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Colleges and Universities.

ORCID iDs

Zhen Pan @ https://orcid.org/0000-0001-9608-009X
Huan Yang @ https://orcid.org/0000-0002-9965-3030

References