Weak lensing of gravitational waves in wave optics: Beyond the Born approximation

Morifumi Mizuno◉ and Teruaki Suyama

Department of Physics, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8551, Japan

The Universe's matter inhomogeneity gravitationally affects the propagation of gravitational waves (GWs), causing the lensing effect. Particularly, the weak lensing of GWs has been studied within the range of the Born approximation to constrain the small-scale power spectrum. In this work, the validity of the Born approximation is investigated by accounting for the higher-order terms in the gravitational potential Φ. To do so, we formulate the post-Born approximation and derive the magnification K and the phase modulation S up to third order in Φ. We find that the average of S and K is nonzero and that the average of S depends on the size of the point mass. Due to this size dependency, the signal is enhanced, and the number of GW events required for detecting the average of S decreases. We find that this number can become comparable to or even smaller than the number required for detecting the variance of S in certain scenarios. In addition, it is verified that, for lensing by dark low-mass halos, the post-Born corrections are a few orders of magnitude smaller than the Born approximation at $f \geq 0.01$ Hz. However, in the presence of the point mass, there is a condition under which the Born approximation fails. We derive the correction terms to the Born approximation and identify the condition under which the Born approximation no longer holds. For the magnification, the Born approximation is valid as long as the wavelength of GWs is larger than the Schwarzschild radius of lenses, while for the phase modulation, this condition is modified due to the physical size of the point mass.

I. INTRODUCTION

When light travels across the Universe, its trajectory is bent by the gravitational potential of intervening massive objects. This phenomenon called gravitational lensing (GL) is quite useful in astrophysics and cosmology (e.g., [1–3]). For instance, it can be used to measure the cosmological parameters. It can probe the abundance of dark compact objects.

According to general relativity, GL also occurs for the gravitational waves (GWs) [4]. One notable feature of the GL of GWs is that geometrical optics, which is a perfect approximation in most cases for light, no longer holds for GWs in some cases since the wavelength of GWs is typically much larger than that of light and the diffraction effect becomes important [5–7]. In such cases, wave optics must be used to deal with the GL. In wave optics, contrary to the geometrical optics where the starting point is the lens equation, the lensing signal is represented by the so-called amplification factor defined as a ratio of the lensed waveform to the unlensed one (e.g., [8]). This quantity is a complex number and all the information of the lensing is encoded in it. Its absolute value and argument represent the amplification and phase modulation of the lensed wave, respectively.

In [9], GL of GWs caused by dark matter fluctuations was studied. It was shown that there is a length scale of the matter power spectrum below which the contribution to the lensing signal is suppressed due to its wavy nature. This scale, the "Fresnel scale," depends on the GW frequency. Thus, by measuring the lensing signal at multiple frequencies and its frequency dependence, we can probe the matter power spectrum at the Fresnel scale. This idea has been investigated in more detail by updating the matter power spectrum as well as adding the compact objects in [10]. In [11], it was shown that the lensing signal of the dark matter fluctuations is hugely amplified by a massive object located on the line of sight which itself causes strong lensing.

When the lensing signal is weak, it is natural to keep only the terms first order in the gravitational potential Φ (i.e., Born approximation). The contributions of the higher-order terms are expected to be suppressed compared to the leading-order contribution. In geometrical optics, this has

been explicitly demonstrated in [12–16]. One naively expects that a similar conclusion can be drawn for the case of wave optics.

In [9], the amplification factor sourced by the dark matter fluctuations was obtained under the Born approximation (thus the variance of the lensing signal is second order in Φ). Typical magnitude of the amplitude and the phase fluctuations was found to be $\mathcal{O}(10^{-2} - 10^{-3})$. Thus, the lensing signal is weak and this would naturally justify the validity of the Born approximation. However, there are two issues that need to be investigated regarding the Born approximation. First, although the post-Born corrections are expected to be small, it is not known how much they are suppressed actually. When the measurements of the lensing signal become available in the future, quantitative computation of the magnitude of the post-Born corrections is indispensable to correctly extract the matter power spectrum as well as to understand the level of the precision under consideration. Second, the Born approximation used in [9] apparently breaks down at large wave frequency since the gravitational potential in the wave equation is associated with the frequency. Notice that this issue does not appear in the geometrical optics since the lens equation is independent of the frequency of light. While it is known how the lens equation emerges in the wave optics, it is not obvious how the breakdown of the Born approximation for the large frequency in the wave optics is reconciled with the Born approximation in the geometric optics. In this paper, we take a first step toward addressing these issues by extending the previous studies to next higher orders in the gravitational potential. We first reformulate the wave equation to make its structure more tractable. We then derive the expression of the lensing signal up to third order in the gravitational potential. Expansion to this order is necessary to evaluate the variance of the post-Born corrections. As we will demonstrate, the post-Born corrections are suppressed by a few orders of magnitude compared to the leading-order signal except in a high-frequency region. Interestingly, when the post-Born corrections are included, the average of the lensing signal does not vanish. This average depends on the frequency in a nontrivial manner and thus cannot be absorbed into the change of the parameters characterizing the unlensed waveform. Our analysis suggests an interesting possibility to make use of the average of the lensing signal as an additional observable to probe the matter power spectrum.

II. FORMULATION

A. Lensing signal beyond the Born approximation

In this section, we reformulate the wave equation and show how the post-Born corrections are derived. Throughout this paper, we assume that the gravitational potential is small ($\Phi \ll 1$) and the Universe is flat. We also ignore the polarization of GWs since the polarization tensor

in the geometrical optics is parallel transported along the null geodesics [17] and hence the change of the polarization tensor would be suppressed by a factor of $\mathcal{O}(\Phi)$ and observationally irrelevant.

The presence of mass fluctuation creates the distortion on spacetime, causing the deviation from the Friedmann-Lemaître-Robertson-Walker (FLRW) metric. This effect is small in most of the astrophysical situations and it is a good approximation to write the metric as [18]

$$
\begin{aligned}
ds^2 &= g^B_{\mu\nu} dx^\mu dx^\nu \\
&= a^2(\eta)[-(1 + 2\Phi)d\eta^2 + (1 - 2\Phi)dx^2],
\end{aligned} \tag{2.1}
$$

where η and x are a conformal time and a comoving coordinate, and $a(\eta)$ is a scale factor. If the wavelength of GWs is much smaller than the typical radius of the curvature of the background metric, the propagation of GWs becomes the same as the wave equation of the massless scalar field ϕ: $\partial_\mu(\sqrt{-g^B}g^{\mu\nu}_B \partial_\nu \phi) = 0$. The expansion of the Universe causes attenuation of ϕ as $\phi \propto 1/a$. We extract this effect by redefining the GW amplitude ϕ as $\phi \to \phi/a$. Then, the wave equation becomes [8]

$$
(\nabla^2 + \omega^2)\tilde{\phi} = 4\omega^2\Phi\tilde{\phi}, \tag{2.2}
$$

in the frequency space. $\tilde{\phi}(\omega, x)$ is the Fourier transform of $\phi(\eta, x)$[1] and the higher-order terms in Φ have been ignored. It is common to represent the lensed waveform in terms of the amplification factor, which is the ratio of the lensed and unlensed waveform, namely, $F = \tilde{\phi}/\tilde{\phi}_0$ [8], where the unlensed waveform is given by $\tilde{\phi}_0 = e^{i\omega\chi}/\chi$ in terms of χ, which is the (comoving) distance from the source. Using the amplification factor F, Eq. (2.2) is rewritten as

$$
2i\omega \frac{\partial F}{\partial \chi} + \frac{1}{\chi^2}\nabla^2_\theta F = 4\omega^2\Phi F, \tag{2.3}
$$

where the polar coordinate (χ, θ, ϕ) is used and $\nabla^2_\theta = \partial^2/\partial\theta^2 + \sin\theta^{-1}\partial/\partial\theta + \sin\theta^{-2}\partial/\partial\phi^2$ is the two-dimensional Laplace operator on two-sphere. In Eq. (2.3), we have assumed GWs propagate along the line of sight and are confined in the region $\theta \ll 1$. Therefore, ∇_θ can be interpreted as the operator on a two-dimensional flat surface perpendicular to the line of sight.

In order to evaluate the effects of the post-Born approximation, we find it convenient to deal with a new variable J defined as $F = e^{i\omega J}$. Using this new variable J, Eq. (2.3) becomes

$$
\left(\frac{\partial}{\partial\chi} - \frac{i}{2\omega\chi^2}\nabla^2_\theta\right)J = -2\Phi - \frac{1}{2\chi^2}(\nabla_\theta J)^2. \tag{2.4}
$$

---

[1] It is defined by $\phi(\eta, x) = \int \frac{d\omega}{2\pi} e^{-i\omega\eta}\tilde{\phi}(\omega, x)$. Thus, ω is the comoving (angular) frequency.

This differential equation can be written as an integral equation by Green's function of the linear operator acting on the right-hand side,

$$J(\chi_s, \boldsymbol{\theta}) = \int_0^{\chi_s} d\chi \exp\left[i\frac{W(\chi,\chi_s)\nabla_\theta^2}{2\omega}\right]$$
$$\times \left(-2\Phi(\chi,\boldsymbol{\theta}) - \frac{1}{2\chi^2}(\nabla_\theta J)^2\right), \quad (2.5)$$

where $W(\chi,\chi_s) = \frac{1}{\chi} - \frac{1}{\chi_s}$. In geometric optics, $W(\chi,\chi_s)$ is sometimes called the lensing efficiency function [15]. The change of variable to J allows us to partially take into account the higher-order terms in the gravitational potential which are not included in the previous studies [9,10,19]. Defining $J^{(n)}$ as the term proportional to nth order of the gravitational potential, $J^{(n)}$ can be calculated iteratively order by order as

$$J^{(1)}(\chi_s, \boldsymbol{\theta}) = \int_0^{\chi_s} d\chi \exp\left[i\frac{W(\chi,\chi_s)\nabla_\theta^2}{2\omega}\right](-2\Phi(\chi,\boldsymbol{\theta})), \quad (2.6)$$

$$J^{(2)}(\chi_s, \boldsymbol{\theta}) = -\int_0^{\chi_s} d\chi \exp\left[i\frac{W(\chi,\chi_s)\nabla_\theta^2}{2\omega}\right]\frac{(\nabla_\theta J^{(1)}(\chi,\boldsymbol{\theta}))^2}{2\chi^2}, \quad (2.7)$$

$$J^{(3)}(\chi_s, \boldsymbol{\theta}) = -\int_0^{\chi_s} d\chi \exp\left[i\frac{W(\chi,\chi_s)\nabla_\theta^2}{2\omega}\right]$$
$$\times \frac{\nabla_\theta J^{(1)}(\chi,\boldsymbol{\theta}) \cdot \nabla_\theta J^{(2)}(\chi,\boldsymbol{\theta})}{\chi^2},$$
$$\vdots \quad (2.8)$$

In the geometrical optics limit (i.e., large ω), $\lim_{\omega\to\infty} J^{(n)}$ of any n becomes real and the correction term at $\mathcal{O}(1/\omega)$ becomes imaginary. Thus, $\lim_{\omega\to\infty} J$ is nothing but the difference between the arrival time of the geodesic under the influence of Φ and the one without Φ. At the leading order, this reduces to the standard expression of the Shapiro time delay. At $\mathcal{O}(1/\omega)$, J gives the magnification in geometrical optics. In particular, $i\omega J^{(1)}$ reduces to the standard formula of the convergence (e.g., [1,2]).

In the literature, the Born approximation refers to the approximation to truncate the expansion of F up to first order in Φ. Meanwhile, since our expansion is performed for J, even the truncation at $J^{(1)}$ partially captures the higher-order terms not included in the previous studies (see also footnote 2). In spite of such a difference at the conceptual level, there is practically no difference as to whether the Born approximation refers to the first-order truncation for F or J since the variation of F in the former case is nothing but $J^{(1)}$.

Our aim is to investigate the leading correction to the Born approximation of the lensing signal caused by the dark matter fluctuations. To this end, we compute the average and the variance of J and investigate how they are affected by the post-Born approximation by treating Φ as a random variable. The average trivially vanishes in the Born approximation. Since the average of $J^{(2)}$ does not vanish in general, we truncate the evaluation of the average at this order. The leading post-Born correction to the variance comes from the cross term $J^{(1)}J^{(2)}$ and thus it is $\mathcal{O}(\Phi^3)$. This is nonvanishing only when Φ is non-Gaussian. The next leading correction, which is $\mathcal{O}(\Phi^4)$, remains finite even when Φ is Gaussian. Thus, the correction at $\mathcal{O}(\Phi^4)$ may dominate over the one at $\mathcal{O}(\Phi^3)$ in some cases, especially when Φ is nearly Gaussian. Because of this reason, we compute the variance up to $\mathcal{O}(\Phi^4)$. To make our calculation consistent up to this order, we need to keep the expansion up to $J^{(3)}$ since the cross term $J^{(1)}J^{(3)}$ is $\mathcal{O}(\Phi^4)$.

For clarity, we define new differential operators $(W\nabla)^{(2)}$ and $(W\nabla)^{(3)}$ as

$$(W\nabla)^{(2)} = W(\chi,\chi_s)\nabla_{\theta 12}^2 + W(\chi_1,\chi)\nabla_{\theta 1}^2 + W(\chi_2,\chi)\nabla_{\theta 2}^2, \quad (2.9)$$

$$(W\nabla)^{(3)} = W(\chi,\chi_s)\nabla_{\theta 123}^2 + W(\chi_3,\chi)\nabla_{\theta 3}^2 + W(\chi',\chi)\nabla_{\theta 12}^2$$
$$+ W(\chi_1,\chi')\nabla_{\theta 1}^2 + W(\chi_2,\chi')\nabla_{\theta 2}^2. \quad (2.10)$$

Using these notations, we obtain the following expressions of J up to third order:

$$J^{(1)}(\chi_s, \boldsymbol{\theta}) = -2\int_0^{\chi_s} d\chi \exp\left[i\frac{W(\chi,\chi_s)\nabla_\theta^2}{2\omega}\right]\Phi(\chi,\boldsymbol{\theta}), \quad (2.11)$$

$$J^{(2)}(\chi_s, \boldsymbol{\theta}) = -2\int_0^{\chi_s} \frac{d\chi}{\chi^2}\int_0^\chi d\chi_1 \int_0^\chi d\chi_2 \exp\left[i\frac{(W\nabla)^{(2)}}{2\omega}\right]$$
$$\times \nabla_{\theta 1}\Phi_1 \cdot \nabla_{\theta 2}\Phi_2, \quad (2.12)$$

$$J^{(3)}(\chi_s, \boldsymbol{\theta}) = -4\int_0^{\chi_s} \frac{d\chi}{\chi^2}\int_0^\chi d\chi_3 \int_0^\chi \frac{d\chi'}{\chi'^2}\int_0^{\chi'} d\chi_1 \int_0^{\chi'} d\chi_2$$
$$\times \exp\left[i\frac{(W\nabla)^{(3)}}{2\omega}\right]\nabla_{\theta 12}(\nabla_{\theta 1}\Phi_1 \cdot \nabla_{\theta 2}\Phi_2)$$
$$\cdot \nabla_{\theta 3}\Phi_3. \quad (2.13)$$

Note that $\Phi_i = \Phi(\chi_i, \boldsymbol{\theta})$ and $\nabla_{\theta i}$ only acts on Φ_i (when there are more than two subscript numbers at the corner of ∇_θ, it means the operator acts on the gravitational potentials that have the corresponding subscripts). $J^{(1)}$ corresponds to the Born approximation and subsequent terms ($J^{(2)}$ and $J^{(3)}$) are the post-Born corrections.

The information about the phase and the magnification of GWs is encoded in the real and imaginary part of J,

respectively. Conventionally, the phase modulation and the magnification are denoted as S and K [9], and we follow the same notation in this paper. S and K are related to J as

$$S(\omega) = \omega \mathrm{Re}(J), \qquad (2.14)$$

$$K(\omega) = -\omega \mathrm{Im}(J). \qquad (2.15)$$

In this definition, the amplification factor is written as $F(\omega) = e^{K(\omega)} e^{iS(\omega)}.$[2] As we have already pointed out, the Shapiro time delay describes the time lag caused by the gravitational potential. In the observation of GWs, the Shapiro time delay is not measurable. Therefore, this degree of freedom needs to be removed from the phase modulation. We redefine the physical phase modulation as

$$\frac{S_{\mathrm{ph}}(\omega)}{\omega} = \frac{S(\omega)}{\omega} - \lim_{\omega \to \infty} \frac{S(\omega)}{\omega}. \qquad (2.16)$$

From now on, the term phase modulation always means this physical quantity even if it is not explicitly mentioned. With these in mind, the phase modulation and the magnification are then explicitly given by

$$S^{(1)} = -2\omega \int_0^{\chi_s} d\chi \left[\cos\left[\frac{W(\chi,\chi_s)\nabla_\theta^2}{2\omega} \right] - 1 \right] \Phi, \quad (2.17)$$

$$S^{(2)} = -2\omega \int_0^{\chi_s} \frac{d\chi}{\chi^2} \int_0^\chi d\chi_1 \int_0^\chi d\chi_2 \left[\cos\left[\frac{(W\nabla)^{(2)}}{2\omega} \right] - 1 \right] \\ \times \nabla_{\theta 1}\Phi_1 \cdot \nabla_{\theta 2}\Phi_2, \qquad (2.18)$$

$$S^{(3)} = -4\omega \int_0^{\chi_s} \frac{d\chi}{\chi^2} \int_0^\chi d\chi_3 \int_0^\chi \frac{d\chi'}{\chi'^2} \int_0^{\chi'} d\chi_1 \int_0^{\chi'} d\chi_2 \\ \times \left[\cos\left[\frac{(W\nabla)^{(3)}}{2\omega} \right] - 1 \right] \nabla_{\theta 12}(\nabla_{\theta 1}\Phi_1 \cdot \nabla_{\theta 2}\Phi_2) \cdot \nabla_{\theta 3}\Phi_3, \qquad (2.19)$$

$$K^{(1)} = 2\omega \int_0^{\chi_s} d\chi \sin\left[\frac{W(\chi,\chi_s)\nabla_\theta^2}{2\omega} \right] \Phi, \qquad (2.20)$$

$$K^{(2)} = 2\omega \int_0^{\chi_s} \frac{d\chi}{\chi^2} \int_0^\chi d\chi_1 \int_0^\chi d\chi_2 \sin\left[\frac{(W\nabla)^{(2)}}{2\omega} \right] \\ \times \nabla_{\theta 1}\Phi_1 \cdot \nabla_{\theta 2}\Phi_2. \qquad (2.21)$$

$$K^{(3)} = 4\omega \int_0^{\chi_s} \frac{d\chi}{\chi^2} \int_0^\chi d\chi_3 \int_0^\chi \frac{d\chi'}{\chi'^2} \int_0^{\chi'} d\chi_1 \int_0^{\chi'} d\chi_2 \\ \times \sin\left[\frac{(W\nabla)^{(3)}}{2\omega} \right] \nabla_{\theta 12}(\nabla_{\theta 1}\Phi_1 \cdot \nabla_{\theta 2}\Phi_2) \cdot \nabla_{\theta 3}\Phi_3. \qquad (2.22)$$

$S^{(1)}$ and $K^{(1)}$ have been derived in the previous works [9,10] and ours reproduce their results. To the best of our knowledge, higher-order terms $S^{(2)}, S^{(3)}, K^{(2)}, K^{(3)}$ are the new results. In the high-frequency limit of Eqs. (2.20) and (2.21), the magnification computed from the above expressions reproduces the result derived in [12–16] under the post-Born approximation in geometric optics, which is demonstrated in Appendix A.

B. Statistics of K and S

The situation we have in mind is the lensing caused by the dark matter inhomogeneities randomly distributed in the whole Universe. This means that the K and S behave in a stochastic manner for individual GW events. Thus, the comparison between the theoretical prediction and observation is possible only for the statistical quantities. This motivates us to compute the average and the variance of the lensing signal.

To this end, we first notice that, for the ensemble average of the functions of the gravitational potential, the following equations hold under the Limber approximation. For arbitrary functions $F(x)$, $G(x)$, $H(x)$, $I(x)$ of differential operator x, we have

$$\langle F(\nabla_{\theta 1})\Phi_1 G(\nabla_{\theta 2})\Phi_2 \rangle \\ = \delta^D(\chi_1 - \chi_2) \int \frac{d^2 k_\perp}{(2\pi)^2} F(i\chi_1 k_\perp) G(-i\chi_1 k_\perp) P_\Phi(k_\perp, \chi_1), \qquad (2.23)$$

$$\langle F(\nabla_{\theta 1})\Phi_1 G(\nabla_{\theta 2})\Phi_2 H(\nabla_{\theta 3})\Phi_3 \rangle_c \\ = \delta^D(\chi_1 - \chi_3)\delta^D(\chi_2 - \chi_3) \int \frac{d^2 k_{1\perp}}{(2\pi)^2} \int \frac{d^2 k_{2\perp}}{(2\pi)^2} \\ \times F(i\chi_3 k_{1\perp}) G(i\chi_3 k_{2\perp}) H(-i\chi_3 k_{1\perp} - i\chi_3 k_{2\perp}) \\ \times B_\Phi(k_{1\perp}, k_{2\perp}, |k_{1\perp} + k_{2\perp}|, \chi_1), \qquad (2.24)$$

$$\langle F(\nabla_{\theta 1})\Phi_1 G(\nabla_{\theta 2})\Phi_2 H(\nabla_{\theta 3})\Phi_3 I(\nabla_{\theta 4})\Phi_4 \rangle_c \\ = \delta^D(\chi_1 - \chi_4)\delta^D(\chi_2 - \chi_4)\delta^D(\chi_3 - \chi_4) \int \frac{d^2 k_{1\perp}}{(2\pi)^2} \int \frac{d^2 k_{2\perp}}{(2\pi)^2} \\ \times \int \frac{d^2 k_{3\perp}}{(2\pi)^2} F(i\chi_4 k_{1\perp}) G(i\chi_4 k_{2\perp}) H(i\chi_4 k_{3\perp}) \\ \times I(-i\chi_4 k_{1\perp} - i\chi_4 k_{2\perp} - i\chi_4 k_{3\perp}) \\ \times T_\Phi(k_{1\perp}, k_{2\perp}, k_{3\perp}, -k_{1\perp} - k_{2\perp} - k_{3\perp}, \chi_1). \qquad (2.25)$$

---

[2]In [9,10,19], F was written as $F = 1 + K + iS$, then K and S were obtained up to first order in Φ, and finally, exponentiation $F \approx (1 + K)e^{iS}$ was done. In our approach, the exponentiation procedure is naturally incorporated from the outset by using the variable J.

Here P_Φ, B_Φ, T_Φ are the power spectrum, bispectrum, and trispectrum of Φ, and $\langle \cdots \rangle_c$ indicates the connected term. They are characterized by

$$\langle \tilde{\Phi}(\boldsymbol{k}_1,\chi)\tilde{\Phi}(\boldsymbol{k}_2,\chi)\rangle = (2\pi)^3\delta^D(\boldsymbol{k}_1+\boldsymbol{k}_2)P_\Phi(k_1,\chi), \quad (2.26)$$

$$\langle \tilde{\Phi}(\boldsymbol{k}_1,\chi)\tilde{\Phi}(\boldsymbol{k}_2,\chi)\tilde{\Phi}(\boldsymbol{k}_3,\chi)\rangle_c$$
$$= (2\pi)^3\delta^D(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)B_\Phi(k_1,k_2,k_3,\chi), \quad (2.27)$$

$$\langle \tilde{\Phi}(\boldsymbol{k}_1,\chi)\tilde{\Phi}(\boldsymbol{k}_2,\chi)\tilde{\Phi}(\boldsymbol{k}_3,\chi)\tilde{\Phi}(\boldsymbol{k}_4,\chi)\rangle_c$$
$$= (2\pi)^3\delta^D(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3+\boldsymbol{k}_4)T_\Phi(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4,\chi),$$
$$(2.28)$$

where $\tilde{\Phi}(\boldsymbol{k})$ is the Fourier transform of Φ. With these definitions, we are ready to derive the average and the variance of the post-Born corrections, which we will address in the following.

1. Average

At the level of the Born approximation, the average of K and S is zero. This does not happen beyond the Born approximation. Thus, the average of K and S fully represents the effects of the post-Born corrections. The leading-order correction is $\mathcal{O}(\Phi^2)$, and we evaluate $\langle K \rangle, \langle S \rangle$ at this order. From Eqs (2.18), (2.21), and (2.23), we obtain the following expressions:

$$\langle S \rangle = 2\omega \int_0^{\chi_s} \frac{d\chi}{\chi^2} \int_0^\chi d\chi_1 \chi_1^2 \int \frac{d^2\boldsymbol{k}_\perp}{(2\pi)^2} k_\perp^2$$
$$\times \left(1 - \cos\left[\frac{(\chi-\chi_1)\chi_1}{\chi\omega}k_\perp^2\right]\right)P_\Phi(k_\perp,\chi_1), \quad (2.29)$$

$$\langle K \rangle = -2\omega \int_0^{\chi_s} \frac{d\chi}{\chi^2} \int_0^\chi d\chi_1 \chi_1^2 \int \frac{d^2\boldsymbol{k}_\perp}{(2\pi)^2} k_\perp^2$$
$$\times \sin\left[\frac{(\chi-\chi_1)\chi_1}{\chi\omega}k_\perp^2\right]P_\Phi(k_\perp,\chi_1). \quad (2.30)$$

At this stage, there are three things worth mentioning. First, it is suggestive to rewrite the above relations in terms of the filter functions $F_S^{(2)}, F_K^{(2)}$ as

$$\langle S \rangle = 2\int_0^{\chi_s}\frac{d\chi}{\chi^3}\int_0^\chi d\chi'\chi'^3(\chi-\chi')\int\frac{dk}{2\pi}F_S^{(2)}k^5P_\Phi(k,\chi'), \quad (2.31)$$

$$\langle K \rangle = -2\int_0^{\chi_s}\frac{d\chi}{\chi^3}\int_0^\chi d\chi'\chi'^3(\chi-\chi')\int\frac{dk}{2\pi}F_K^{(2)}k^5P_\Phi(k,\chi'), \quad (2.32)$$

where

$$F_S^{(2)} = \frac{1-\cos k^2r_F^2}{k^2r_F^2}, \qquad F_K^{(2)} = \frac{\sin k^2r_F^2}{k^2r_F^2}, \quad (2.33)$$

and r_F defined by $r_F^2 = \chi'(\chi-\chi')/(\omega\chi)$ is the Fresnel scale [9]. By writing in this way, it is manifest that the frequency dependence of $\langle S \rangle$ and $\langle K \rangle$ is solely encoded in the filter functions. These filter functions are suppressed below the Fresnel scale $k^{-1} < r_F$. Physically, the filter functions describe the diffraction effect that lowers the lensing signal when the size of matter fluctuations is below this scale. In [9], it was argued that $\langle S^2 \rangle$ and $\langle K^2 \rangle$ (within the Born approximation) are insensitive to the matter fluctuations below the Fresnel scale. Our result demonstrates that a similar conclusion holds for $\langle S \rangle, \langle K \rangle$. Second, since, unlike in the case of geometric optics, both $\langle S \rangle$ and $\langle K \rangle$ depend on the GW frequency due to the frequency dependence of the Fresnel scale, we can extract the matter power spectrum at the Fresnel scale by measuring $\langle S \rangle$ and $\langle K \rangle$ at multiple frequencies and how they vary as the frequency is changed. This suggests a possibility that, in addition to $\langle S^2 \rangle$ and $\langle K^2 \rangle$, $\langle S \rangle$ and $\langle K \rangle$ can be used as new observables to probe the matter power spectrum at the Fresnel scale. Notice that, contrary to the case of the cosmological perturbations where the average of the perturbations is absorbed into the FLRW background, $\langle S \rangle$ and $\langle K \rangle$ cannot be absorbed into the unlensed waveform since (i) the frequency dependence of the average is different from that of the unlensed waveform and (ii) each merger event has a different unlensed waveform. Third, $\langle S \rangle$ is a positive definite for any ω. Thus, if the measurement of $\langle S \rangle$ gives a negative value, we can robustly conclude that it is not due to the lensing by the matter fluctuations but due to something else.

We also derive the expressions of the next leading-order contributions to the average coming from the higher-order statistical quantities (bispectrum),

$$\langle S^{(3)} \rangle = -4\omega \int_0^{\chi_s}\frac{d\chi}{\chi^2}\int_0^\chi\frac{d\chi'}{\chi'^2}\int_0^{\chi'}\chi_3^4 d\chi_3 \int\frac{d\boldsymbol{k}_{1\perp}}{(2\pi)^2}$$
$$\times \int\frac{d\boldsymbol{k}_{2\perp}}{(2\pi)^2}\frac{(\boldsymbol{k}_{\perp1}\cdot\boldsymbol{k}_{\perp2})B_\Phi(k_1,k_2,k_3,\chi)}{k_{\perp1}^2k_{\perp2}^2}$$
$$\times \left(1-\cos\left[\frac{\chi_3(\chi-\chi_3)}{2\omega\chi}(k_{\perp1}^2+k_{\perp2}^2+(\boldsymbol{k}_{\perp1}+\boldsymbol{k}_{\perp2})^2)\right.\right.$$
$$\left.\left.+\frac{\chi_3^2(\chi'-\chi)}{\omega\chi\chi'}\boldsymbol{k}_{\perp1}\cdot\boldsymbol{k}_{\perp2}\right]\right), \quad (2.34)$$

$$\langle K^{(3)} \rangle = -4\omega \int_0^{\chi_s}\frac{d\chi}{\chi^2}\int_0^\chi\frac{d\chi'}{\chi'^2}\int_0^{\chi'}\chi_3^4 d\chi_3 \int\frac{d\boldsymbol{k}_{1\perp}}{(2\pi)^2}$$
$$\times \int\frac{d\boldsymbol{k}_{2\perp}}{(2\pi)^2}\frac{(\boldsymbol{k}_{\perp1}\cdot\boldsymbol{k}_{\perp2})B_\Phi(k_1,k_2,k_3,\chi)}{k_{\perp1}^2k_{\perp2}^2},$$
$$\times \sin\left[\frac{\chi_3(\chi-\chi_3)}{2\omega\chi}(k_{\perp1}^2+k_{\perp2}^2+(\boldsymbol{k}_{\perp1}+\boldsymbol{k}_{\perp2})^2)\right.$$
$$\left.+\frac{\chi_3^2(\chi'-\chi)}{\omega\chi\chi'}\boldsymbol{k}_{\perp1}\cdot\boldsymbol{k}_{\perp2}\right]. \quad (2.35)$$

By incorporating the bispectrum contributions to the average of S and K, we are able to assess whether the effects of the non-Gaussianity terms on $\langle S \rangle$ and $\langle K \rangle$ are significant.

2. Variance

In the same way, the rms of the magnification up to fourth order in Φ is given by

$$\langle K^2 \rangle = \langle (K^{(1)})^2 \rangle + 2\langle K^{(1)}K^{(2)} \rangle + 2\langle K^{(1)}K^{(3)} \rangle + \langle (K^{(2)})^2 \rangle. \tag{2.36}$$

The variance of the magnification up to the same order is then written as $\Delta_K^2 = \langle K^2 \rangle - \langle K^{(2)} \rangle^2$. We define the post-Born corrections to the variance as $\Delta_K^2 = \langle K_{\text{Born}}^2 \rangle + \delta_{K^2}$. In this definition, it is possible that $\delta_{K^2} < 0$. As we mentioned earlier, the third-order term in Φ is necessary because it couples with the first-order term. At this order, the result will depend on whether Φ is Gaussian or non-Gaussian. In the diagrammatic language, the variance contains both disconnected ($\delta_{K^2,\text{dc}}$) and connected ($\delta_{K^2,\text{c}}$) parts,

$$\delta_{K^2} = \delta_{K^2,\text{dc}} + \delta_{K^2,\text{c}}. \tag{2.37}$$

As for the disconnected part, we find that it consists of three distinct terms,

$$\delta_{K^2,\text{dc}} = \langle (K^{(2)})^2 \rangle_{\text{dc}} + 2\langle K^{(1)}K^{(3)} \rangle_{\text{dc}} - \langle K^{(2)} \rangle_{\text{dc}}^2, \tag{2.38}$$

where the subscript dc should be understood that the corresponding quantity is obtained by treating Φ as a Gaussian variable. The connected part also consists of three terms,

$$\delta_{K^2,\text{c}} = 2\langle K^{(1)}K^{(2)} \rangle_{\text{c}} + 2\langle K^{(1)}K^{(3)} \rangle_{\text{c}} + \langle (K^{(2)})^2 \rangle_{\text{c}}. \tag{2.39}$$

The first term in Eq. (2.36) is nothing but the variance in the Born approximation and has been already derived in the literature [9]. For completeness, we will provide its expression below. For the Gaussian variable, the n-point correlation function is completely specified by the two-point function, i.e., the matter power spectrum. Using this fact, we find that each term can be written as

$$\langle (K^{(1)})^2 \rangle = 4\omega^2 \int_0^{\chi_s} d\chi \int \frac{d^2 \boldsymbol{k}_\perp}{(2\pi)^2} \sin^2\left[\frac{(\chi_s - \chi)\chi}{2\chi_s \omega} k_\perp^2 \right] P_\Phi(k_\perp), \tag{2.40}$$

$$\begin{aligned}
\langle (K^{(2)})^2 \rangle_{\text{dc}} = \langle K^{(2)} \rangle_{\text{dc}}^2 + 16\omega^2 &\int_0^{\chi_s} \frac{d\chi}{\chi^2} \int_0^\chi \frac{d\chi'}{\chi'^2} \int_0^{\chi'} d\chi_1 \int_0^{\chi'} d\chi_2 \int \frac{d^2 \boldsymbol{k}_{1\perp}}{(2\pi)^2} \int \frac{d^2 \boldsymbol{k}_{2\perp}}{(2\pi)^2} \\
&\times \sin\left[\frac{(\chi_s - \chi_1)\chi_1}{2\chi_s \omega} k_{1\perp}^2 + \frac{(\chi_s - \chi_2)\chi_2}{2\chi_s \omega} k_{2\perp}^2 + \frac{(\chi_s - \chi)\chi_1 \chi_2}{\omega \chi_s \chi} \boldsymbol{k}_{1\perp} \cdot \boldsymbol{k}_{2\perp} \right] \\
&\times \sin\left[\frac{(\chi_s - \chi_1)\chi_1}{2\chi_s \omega} k_{1\perp}^2 + \frac{(\chi_s - \chi_2)\chi_2}{2\chi_s \omega} k_{2\perp}^2 + \frac{(\chi_s - \chi')\chi_1 \chi_2}{\omega \chi_s \chi'} \boldsymbol{k}_{1\perp} \cdot \boldsymbol{k}_{2\perp} \right] \\
&\times \chi_1^2 \chi_2^2 (\boldsymbol{k}_{1\perp} \cdot \boldsymbol{k}_{2\perp})^2 P_\Phi(k_{1\perp}) P_\Phi(k_{2\perp}),
\end{aligned} \tag{2.41}$$

$$\begin{aligned}
\langle K^{(1)}K^{(3)} \rangle_{\text{dc}} = -16\omega^2 &\int_0^{\chi_s} \frac{d\chi}{\chi^2} \int_0^\chi \frac{d\chi'}{\chi'^2} \int_0^{\chi'} d\chi_1 \int_0^{\chi'} d\chi_2 \int \frac{d^2 \boldsymbol{k}_{1\perp}}{(2\pi)^2} \int \frac{d^2 \boldsymbol{k}_{2\perp}}{(2\pi)^2} \\
&\times \sin\left[\frac{(\chi_s - \chi_2)\chi_2}{2\chi_s \omega} k_{2\perp}^2 \right] \sin\left[\frac{(\chi_s - \chi_2)\chi_2}{2\chi_s \omega} k_{2\perp}^2 + \frac{(\chi - \chi')\chi_1 \chi_2}{\omega \chi \chi'} \boldsymbol{k}_{1\perp} \cdot \boldsymbol{k}_{2\perp} + \frac{(\chi - \chi_1)\chi_1}{\chi \omega} k_{1\perp}^2 \right] \\
&\times (\chi_1^2 \chi_2^2 (\boldsymbol{k}_{1\perp} \cdot \boldsymbol{k}_{2\perp})^2 + \chi_1^3 \chi_2 k_{1\perp}^2 (\boldsymbol{k}_{1\perp} \cdot \boldsymbol{k}_{2\perp})) P_\Phi(k_{1\perp}) P_\Phi(k_{2\perp}).
\end{aligned} \tag{2.42}$$

As these expressions show, the computation of $\delta_{K^2,\text{dc}}$ requires multiple integrations in eight variables. Among these eight variables, the integral with respect to the angle between $\boldsymbol{k}_{1\perp}$ and $\boldsymbol{k}_{2\perp}$ can be analytically performed and the result is written in terms of the Bessel functions. Thus, practically, the number of variables in the integration is six. The concrete expression of $\delta_{K^2,\text{dc}}$ which we will

evaluate numerically in the next section is given in Appendix B.

The connected part $\delta_{K^2,\text{c}}$ given by Eq. (2.39) comes from the non-Gaussianity of the matter fluctuations: the matter bispectrum, trispectrum, and so forth. The first term in Eq. (2.39), which is $\mathcal{O}(\Phi^3)$, is written in terms of the bispectrum as

$$\langle K^{(1)} K^{(2)} \rangle_c = -4\omega^2 \int_0^{\chi_s} \frac{d\chi}{\chi^2} \int_0^{\chi} d\chi_3 \chi_3^2 \int \frac{d^2 \mathbf{k}_{1\perp}}{(2\pi)^2} \int \frac{d^2 \mathbf{k}_{2\perp}}{(2\pi)^2} \sin \left[\frac{(\chi_s - \chi_3)\chi_3}{2\omega\chi_s} |\mathbf{k}_{1\perp} + \mathbf{k}_{2\perp}|^2 \right]$$

$$\times \sin \left[\frac{(\chi_s - \chi_3)\chi_3}{2\omega\chi_s} (k_{1\perp}^2 + k_{2\perp}^2) + \frac{(\chi_s - \chi)\chi_3^2}{\omega\chi_s\chi} \mathbf{k}_{1\perp} \cdot \mathbf{k}_{2\perp} \right]$$

$$\times (\mathbf{k}_{1\perp} \cdot \mathbf{k}_{2\perp}) B_\Phi(k_{1\perp}, k_{2\perp}, |\mathbf{k}_{1\perp} + \mathbf{k}_{2\perp}|). \tag{2.43}$$

The other two terms, which are $\mathcal{O}(\Phi^4)$, are written in terms of the trispectrum as

$$\langle (K^{(2)})^2 \rangle_c = -8\omega^2 \int_0^{\chi_s} \frac{d\chi}{\chi^2} \int_0^{\chi} \frac{d\chi'}{\chi'^2} \int_0^{\chi'} d\chi_4 \chi_4^4 \int \frac{d^2 \mathbf{k}_{1\perp}}{(2\pi)^2} \int \frac{d^2 \mathbf{k}_{2\perp}}{(2\pi)^2} \int \frac{d^2 \mathbf{k}_{3\perp}}{(2\pi)^2}$$

$$\times \sin \left[\frac{(\chi_s - \chi)\chi_4^2}{2\omega\chi_s\chi} |\mathbf{k}_{1\perp} + \mathbf{k}_{2\perp}|^2 + \frac{(\chi - \chi_4)\chi_4}{2\omega\chi} (k_{1\perp}^2 + k_{2\perp}^2) \right]$$

$$\times \sin \left[\frac{(\chi_s - \chi')\chi_4^2}{2\omega\chi_s\chi'} |\mathbf{k}_{1\perp} + \mathbf{k}_{2\perp}|^2 + \frac{(\chi' - \chi_4)\chi_4}{2\omega\chi'} (k_{3\perp}^2 + |\mathbf{k}_{1\perp} + \mathbf{k}_{2\perp} + \mathbf{k}_{3\perp}|^2) \right]$$

$$\times (\mathbf{k}_{1\perp} \cdot \mathbf{k}_{2\perp})[\mathbf{k}_{3\perp} \cdot (\mathbf{k}_{1\perp} + \mathbf{k}_{2\perp} + \mathbf{k}_{3\perp})] T_\Phi(\mathbf{k}_{1\perp}, \mathbf{k}_{2\perp}, \mathbf{k}_{3\perp}, -\mathbf{k}_{1\perp} - \mathbf{k}_{2\perp} - \mathbf{k}_{3\perp}), \tag{2.44}$$

$$\langle K^{(1)} K^{(3)} \rangle_c = 8\omega^2 \int_0^{\chi_s} \frac{d\chi}{\chi^2} \int_0^{\chi} \frac{d\chi'}{\chi'^2} \int_0^{\chi'} d\chi_4 \chi_4^4 \int \frac{d^2 \mathbf{k}_{1\perp}}{(2\pi)^2} \int \frac{d^2 \mathbf{k}_{2\perp}}{(2\pi)^2} \int \frac{d^2 \mathbf{k}_{3\perp}}{(2\pi)^2} \sin \left[\frac{(\chi_s - \chi_4)\chi_4}{2\omega\chi_s} |\mathbf{k}_{1\perp} + \mathbf{k}_{2\perp} + \mathbf{k}_{3\perp}|^2 \right]$$

$$\times \sin \left[\frac{(\chi_s - \chi)\chi_4^2}{2\omega\chi_s\chi} |\mathbf{k}_{1\perp} + \mathbf{k}_{2\perp} + \mathbf{k}_{3\perp}|^2 + \frac{(\chi - \chi_4)\chi_4}{2\omega\chi} k_{3\perp}^3 + \frac{(\chi - \chi')\chi_4^2}{2\omega\chi\chi'} |\mathbf{k}_{1\perp} + \mathbf{k}_{2\perp}|^2 + \frac{(\chi' - \chi_4)\chi_4}{2\omega\chi'} (k_{1\perp}^2 + k_{2\perp}^2) \right]$$

$$\times (\mathbf{k}_{1\perp} \cdot \mathbf{k}_{2\perp})[(\mathbf{k}_{1\perp} + \mathbf{k}_{2\perp}) \cdot \mathbf{k}_{3\perp}] T_\Phi(\mathbf{k}_{1\perp}, \mathbf{k}_{2\perp}, \mathbf{k}_{3\perp}, -\mathbf{k}_{1\perp} - \mathbf{k}_{2\perp} - \mathbf{k}_{3\perp}). \tag{2.45}$$

In order to evaluate the connected part, we need to determine the bispectrum and the trispectrum of the matter fluctuations. In this paper, we only focus on the lowest-order term in the non-Gaussianity part, which is the contribution from the bispectrum.

The above formulation is for the variance of K. The variance of S can be formulated in exactly the same manner. The quantities of S corresponding to Eqs. (2.40)–(2.45) are obtained by replacing all the sine functions as $\sin(\cdots) \to 1 - \cos(\cdots)$.

III. HALO MODEL

Our results in the previous section are described by the power spectrum and bispectrum of the potential $P_\Phi(k, \chi)$ and $B_\Phi(k_1, k_2, k_3, \chi)$. As mentioned, the bispectrum is the only term considered in this paper to capture the non-Gaussianity. In the actual computations of the average and the variance of S and K, we need the spectra of matter instead of the gravitational potential and they are obtained through the Poisson equation,

$$P_\Phi(k, \chi) = \left(\frac{3H_0^2 \Omega_m}{2} \right)^2 \frac{1}{a^2(\chi)k^4} P_\delta(k, \chi), \tag{3.1}$$

$$B_\phi(k_1, k_2, k_3, \chi)$$
$$= -\left(\frac{3H_0^2 \Omega_m}{2} \right)^3 \frac{1}{a^3(\chi)k_1^2 k_2^2 k_3^2} B_\delta(k_1, k_2, k_3, \chi). \tag{3.2}$$

The precise dependence of the matter power spectrum at a small scale, which is important for the frequency range of our interest, is very difficult to compute from the first principle due to complex physical processes such as baryonic physics. To circumvent this issue, we adopt the formulation of the halo model described in [10] which provides a useful phenomenological approach with reasonable computational costs.

While it is advisable to refer to [10] for more detail, we would like to give a brief summary of their findings. In [10], the power spectrum was computed using their halo model, which incorporates subhalos and baryonic matter such as galaxies and stars. Their model's power spectrum shows a close match to the one computed through hydrodynamical simulations within the range covered by the simulation ($k < 30h$ Mpc^{-1}), confirming the reliability of the model at least within this range. Then, they compute the power spectrum at an even smaller scale (up to as small as $k \sim 10^8 h$ Mpc^{-1}) with their halo model assuming that the halo model approach provides a reasonable estimation of

the power spectrum. Under this assumption, they found that the power spectrum at $k \sim 10^6 h \, \mathrm{Mpc}^{-1}$ is predominantly determined by the two components: dark low-mass halos ($1 h^{-1} M_\odot \lesssim M \lesssim 10^4 h^{-1} M_\odot$, no stars within them, thus dark) and the point mass (ordinary stars, neutron stars, and black holes which causes the shot noise in the signal). Considering that the weak lensing effect on GWs is highly sensitive to the matter power spectrum at the corresponding scale [9], they show that, within a certain frequency range of GWs ($10 \lesssim f \lesssim 100$ Hz), the lensing dispersion is also sensitive to the abundance of dark low-mass halos and primordial black holes (point mass) and can be used to probe them.

Following their result that dark low-mass halos and the point mass are important, we only focus on their contributions to the matter power spectrum and bispectrum, and we ignore the effect from other sources such as the density distribution within a galaxy. In addition, we separately compute the lensing signal from the halos and the point mass (shot noise) for the following two reasons. First, the power spectrum and bispectrum for these two components exhibit different characteristics (the power spectrum with subhalos included is strongly suppressed at small scales, while the shot noise contributes to all scales of the spectra equally). By analyzing their effects separately, we can gain a clearer understanding of the individual contributions. Second, there is a technical aspect to consider. While analytical computations can be partially carried out for the shot noise contribution, full numerical computation poses challenges due to rapid oscillation in the integration processes. On the other hand, the subhalo contribution cannot be evaluated analytically due to the lack of a simple analytical form for the power spectrum and bispectrum, leaving numerical calculations the only option. Consequently, separating the calculation of these two components allows us to effectively address the computational challenges associated with each contribution (the contribution from the shot noise is separately studied in Sec. IV E).

Following [10], we estimate the effect of subhalos by computing the analytic subhalo mass function based on the extended Press-Schechter theory with the tidal stripping and the dynamical friction effects included.

Figure 1 shows the matter power spectrum used in this paper evaluated at $z = 0, 1$ (solid line). The dotted lines are the power spectrum without the subhalo contributions, illustrating the enhancement of the small-scale power spectrum due to the subhalos. We also show with the red dashed lines the power spectrum computed by the halofit model [20]. Halofit is a fitting formula whose functional form is motivated by the halo model and is calibrated against the N-body simulation at $k \lesssim 30 h \, \mathrm{Mpc}^{-1}$. The halo model is indeed consistent with the fitting formula within the calibration range. Note that, in this figure, the vertical axis is scaled $k^2 P_\delta(k)$ as this is the contribution to gravitational lensing per $\log k$. At $k \sim 10^{-2} h \, \mathrm{Mpc}^{-1}$, a slight hump can be observed. This scale represents the peak of the linear power
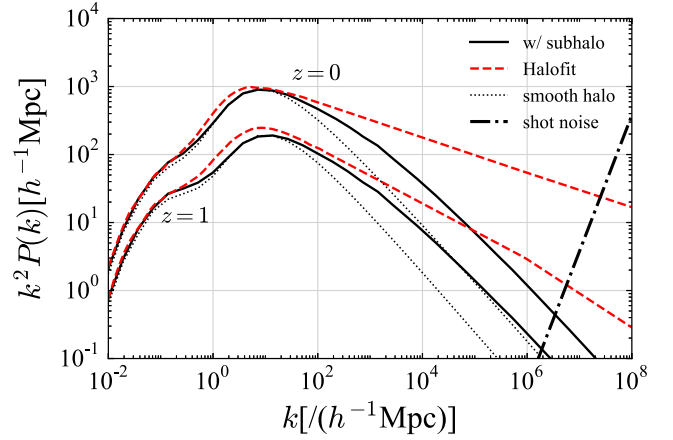

FIG. 1. The power spectrum for the halo model with subhalos included (solid) is significantly enhanced by the presence of subhalos at small scales compared to the power spectrum of the halo model without subhalos (dotted). At large scales, the halo model is fairly accurate with the halofit (red dashed). The halofit is a fitting formula given in [20] whose parameters are calibrated on the N-body simulation results at $k < 30 h \, \mathrm{Mpc}^{-1}$. The shot noise effect (compact object with mass $0.5 M_\odot$ and mass fraction $f_p = 0.01$) on the power spectrum is illustrated by the dash-dotted line emerging at $k \sim 10^6 h \, \mathrm{Mpc}^{-1}$. The hump seen at $k = 10^{-2}$–$10^{-1} h \, \mathrm{Mpc}^{-1}$ represents the peak of the linear power spectrum, while the peak at $k \sim 10^1 h \, \mathrm{Mpc}^{-1}$ represents the scale of the biggest halos considered in the model.

spectrum, while the peak observed at $k \sim 10 h \, \mathrm{Mpc}^{-1}$ indicates the scale of the largest halos. As the scale moves toward the lower side (high k), the power spectrum decreases as the contribution from the larger halos (halos whose radius is greater than the scale of interest) becomes less and less significant. The dash-dotted straight line emerging at $k = 10^6 h \, \mathrm{Mpc}^{-1}$ represents the shot noise effect due to the point mass. The shot noise is evaluated using a simple model, where all point masses are the same type of object and are randomly distributed throughout the Universe. This model does not take into account the time evolution of the point mass either. The power spectrum based on this model is simply given by a constant $P_{\mathrm{shot}} = f_p^2 / \bar{n}$, where f_p and $\bar{n}$ are the mass fraction of the point mass to the total matter and the number density of the point mass, respectively. In order to compute $\bar{n}$, we use the relation $\bar{\rho} \Omega_m f_p = m_p \bar{n}$, and $H_0^2 = \frac{8\pi G}{3} \bar{\rho}$ with mass $m = 0.5 M_\odot$ and mass fraction $f_p = 0.01$. The fiducial value of f_p is consistent with the measured abundance of stars [21].

Next, we would like to provide an approximation formula for the power spectrum,

$$P_\delta(k, \chi) = P_\delta(k_0, \chi) \left(\frac{k}{k_0} \right)^{\frac{k_0}{P_\delta(k_0, \chi)} \frac{dP_\delta(k_0, \chi)}{dk_0}} = B(\chi) k^{-b}. \qquad (3.3)$$

This is equivalent to the Taylor expansion of $\log P_\delta$ with respect to $\log k$ and is useful to approximately evaluate the

asymptotic behavior of the average and the variance of S and K (particularly, $\delta_{S^2,\mathrm{dc}}$, the correction term to the variance of the phase modulation). k_0 is an arbitrary scale around which $P(k, \chi)$ is expanded, thus we can take $k_0 = \sqrt{H_0 \omega} \sim 1/(\mathrm{Fresnel\, scale})$ so that the variation of the power spectrum around the Fresnel scale is properly evaluated.

In addition to the power spectrum, we also need to evaluate the bispectrum for calculating the lensing dispersion. In the halo model formulation, the total bispectrum is the sum of so-called 1-halo, 2-halo, and 3-halo terms,

$$B_\delta = B_{1\mathrm{H}} + B_{2\mathrm{H}} + B_{3\mathrm{H}}, \qquad (3.4)$$

where each term is calculated by following the formalism presented in [22]. Since we are interested in the small-scale bispectrum, we need to incorporate the subhalo contribution to this expression. To do so, we follow the formulation in [23], where the 1-halo term is separated into seven terms as

$$B_{1\mathrm{H}} = B_{\mathrm{sss}} + B_{\mathrm{ssc}} + B_{\mathrm{s1c}} + B_{\mathrm{s2c}} + B_{1\mathrm{c}} + B_{2\mathrm{c}} + B_{3\mathrm{c}}. \qquad (3.5)$$

The notation used here is the same as the one in [23] (s and c mean smooth and clump, respectively). We evaluate each term using the same mass function and density profile of the subhalos in [10].

Note that we have ignored the contributions from subhalos to the 2-halo term and 3-halo term.

The rationale for this assumption is as follows: The 2-halo term refers to the three-point correlation involving two points from the same halo and the third point from a different halo. In the case of an equilateral or flattened configuration of a triangle, this term becomes subdominant compared to the 1-halo term. For instance, when k_1, k_2, and k_3 are equal (equilateral), and their corresponding scales are smaller than the size of a main halo, the 2-halo term is significantly suppressed due to a very small correlation between these two halos. However, the 1-halo term remains relevant because of the matter density fluctuation and the presence of subhalos in a main halo. The only triangle configuration where the 2-halo term becomes relevant is when the scales corresponding to k_1 and k_2 are of the order of the halo size, while the scale corresponding to k_3 is much larger than the size of a main halo. In such cases, the contribution of subhalos can be ignored because, at scales much larger than the size of a main halo, even the main halos can be treated as point mass objects. Similar considerations can be made for the 3-halo term. The 3-halo term only becomes relevant when the scale corresponding to k_1, k_2, and k_3 is significantly larger than the size of a main halo. In such cases, the structure of individual main halos becomes irrelevant.

In Fig. 2, we show the bispectrum evaluated at three different configurations (equilateral, flattened, and squeezed)

and two different redshifts ($z = 0, 1$). The squeezed configuration is set to $k_1 = k_2 = k, k_3 = 0.01k$. The solid (dotted) lines show the bispectrum with (without) subhalos. The red dashed lines are the BiHalofit model given in [24] (a fitting formula for the matter bispectrum calibrated on the simulations at $k < 30h\ \mathrm{Mpc}^{-1}$). The trend observed in the bispectrum is the same as the one in the power spectrum: The inclusion of subhalos enhances the small-scale bispectrum, while the large-scale behavior (the small hump representing the peak linear bispectrum and the peak representing the largest scale of halos) is compatible with the fitting formula. However, there is a notable difference from the power spectrum which is intrinsic to the bispectrum. In one squeezed configuration ($k1 = k2 = k, k_3 = 0.01k$), it is observed that the subhalo enhancement to the bispectrum has the peak at $k_1 = k_2 \sim 10^3$ and $k_3 \sim 10h\ \mathrm{Mpc}^{-1}$. This represents the correlation within the main halo where the larger scale corresponds to the scale of the largest halos, while the smaller scale involves the subhalo scale. The straight dash-dotted lines are the shot noise effect, which is calculated by assuming that it is given by a constant $B_{\mathrm{shot}} = f_p^3/\bar{n}^2$ with $m = 0.5M_\odot$ and $f_p = 0.01$. Precisely speaking, the shot noise effect on the bispectrum contains not just the constant term, instead, it is given by [25]

$$B_{\mathrm{shot}} = \frac{f_p}{\bar{n}} \left(P_\delta(k_1) + P_\delta(k_2) + P_\delta(k_3) \right) + \frac{f_p^3}{\bar{n}^2}, \qquad (3.6)$$

where $P\delta(k)$ is the power spectrum without the shot noise effect. In fact, the first three terms in Eq. (3.6) come from the coupling of the non-shot-noise effect and the shot noise effect. However, as mentioned earlier, we treat the shot noise separately, and the isolation of the shot noise from the rest of the terms results in ignoring the coupling terms.

Up to this point, we have assumed that the halo model can be a reasonable estimation for such a small scale as $k \sim 10^6 h\ \mathrm{Mpc}^{-1}$. However, it is not clear whether the uncertainty of the spectra based on the halo model is reasonably suppressed even at scales much smaller than the ones covered by the simulations. Since our aim in this paper is to provide formalism to compute the post-Born corrections given that the matter power spectrum and bispectrum are properly evaluated, we simply take it for granted that our matter power spectrum and bispectrum models are valid at all scales.

Thus, the numerical values of the post-Born corrections that will be given later should not be understood as the robust quantitative prediction of the post-Born effects.[3] Once the

---

[3]In addition to the uncertainty of the matter power spectrum caused by the use of the halo model, there are other types of uncertainties coming from our ignorance about the microscopic properties of dark matter and primordial power spectrum on small scales, both of which affect the shape of the matter power spectrum at small scales.
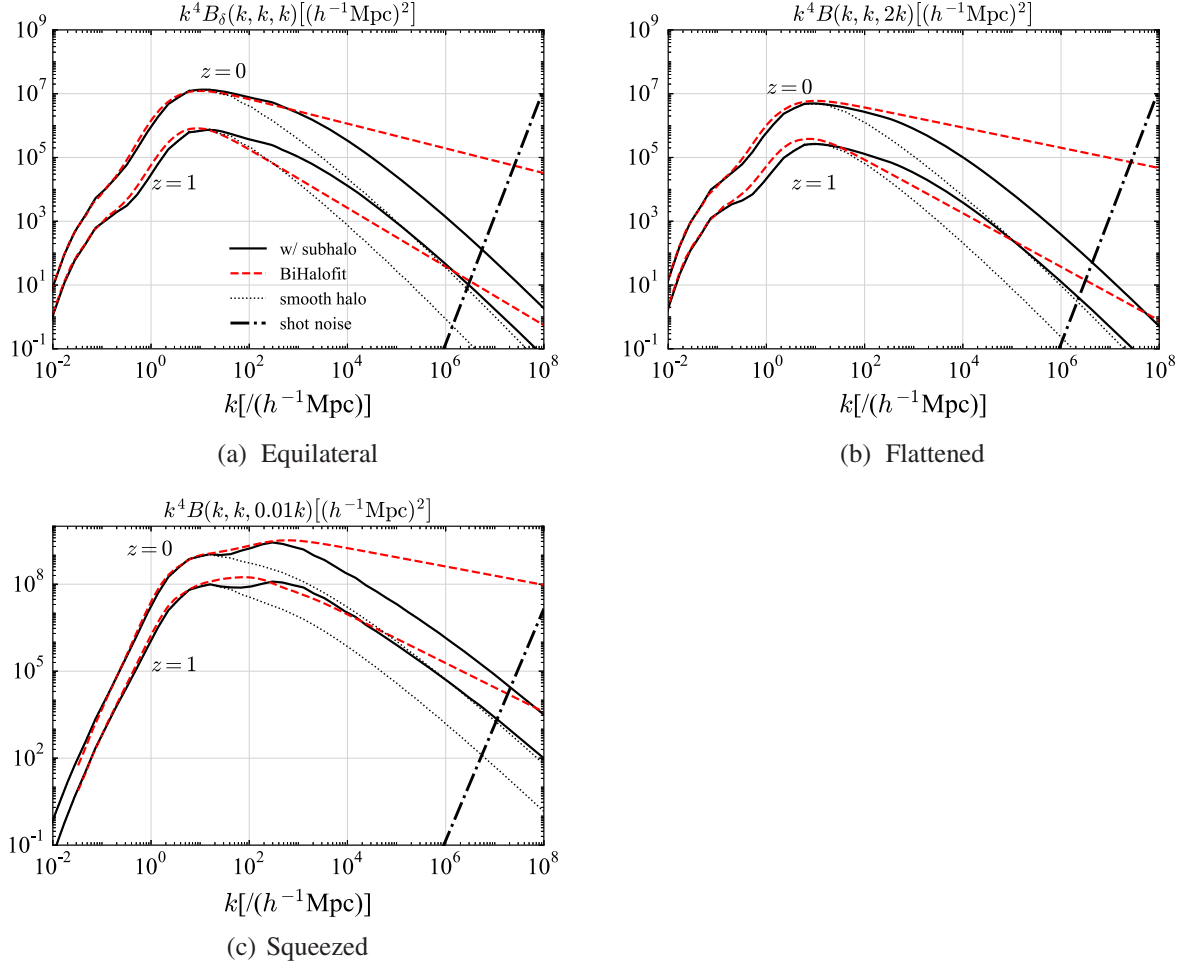
$k^4 B_\delta(k,k,k)[(h^{-1}\mathrm{Mpc})^2]$

(a) Equilateral

$k^4 B(k,k,2k)[(h^{-1}\mathrm{Mpc})^2]$

(b) Flattened

$k^4 B(k,k,0.01k)[(h^{-1}\mathrm{Mpc})^2]$

(c) Squeezed

FIG. 2. In all configurations of a triangle [(a) equilateral, (b) flattened, (c) squeezed)], the halo model with subhalos (solid) is significantly enhanced at small scales compared to the smooth (without subhalos) model (dotted). At large scales, the halo model is fairly consistent with the BiHalofit model [24] (red dashed) in all configurations at different redshifts ($z = 0, 1$). The shot noise (dash-dotted) becomes relevant above $k \sim 10^7 h\ \mathrm{Mpc}^{-1}$. In this squeezed bispectrum ($k_1 = k_2 = k, k_3 = 0.01k$), there is a peak at $k \sim 10^3 h\ \mathrm{Mpc}^{-1}$ due to the subhalo enhancement.

correct matter power spectrum and bispectrum are obtained, it is straightforward to evaluate the post-Born corrections by using the formulation given in this paper. We emphasize that our general conclusion that the post-Born corrections are subdominant compared to the Born approximation remains unaffected by the refinement of the matter power spectrum in the absence of the shot noise.

The cosmological parameters we use to compute the halo model, as well as the lensing dispersion, are $h = 0.6739$, $\Omega_m = 0.3147$, $\Omega_\Lambda = 0.6888$, $n_s = 0.9665$, and $\sigma_8 = 0.8102$.

IV. POST-BORN EFFECT ON S AND K

A. Numerical computation

In order to evaluate the lensing signal, it is essential to compute the multivariable integral over a considerably wide range for highly oscillatory functions. To achieve this, we use the Monte Carlo algorithm to compute the lensing dispersion. We ensure that the error of our calculation is smaller than

10% by progressively increasing the number of sample points until the fluctuation in the results is less than 10%. In the case of $\langle S^{(3)} \rangle$ and $\langle S_{\mathrm{Born}} S^{(2)} \rangle$, which exhibit a slow convergence rate, the computation error is ensured to be less than 30%. Additionally, we assess the impact of the integration range on the final result by varying its width and we confirm that the results remain unchanged (the default integration range is set to $10^{-4} h^{-1} < k < 10^{12} h\ \mathrm{Mpc}^{-1}$). By undertaking these checks, we can guarantee that the results are indeed converged. We have to keep in mind that the error in our calculation, which we estimate to be 10%, arises from the slow convergent rate of these integrals. Thus, this number cannot be interpreted as the error of the actual observed S and K.

B. Born approximation

For completeness, we first evaluate the variance of S and K under the Born approximation. Formally, $\langle K_{\mathrm{Born}}^2 \rangle$ is

given by Eq. (2.40) and $\langle S^2_{\text{Born}} \rangle$ is given by the same equation with the sin function being replaced with $1 - \cos$. In Fig. 5, we show $\langle S^2_{\text{Born}} \rangle$ as a function of f with the corresponding Fresnel scale $1/\bar{r}_F = \sqrt{H_0\omega}$ represented by the second horizontal axis at the top. Irrespective of the source redshift, $\langle S^2_{\text{Born}} \rangle$ exhibits a general behavior pertaining to the power spectrum. As the frequency increases from $f = 10^{-18}$ Hz, it initially rises and forms a hump around $f = 10^{-15}$ Hz ($k \sim 10^{-2}h$ Mpc^{-1}). It then reaches a peak at $f = 10^{-11}$ Hz ($k \sim 10^1 h$ Mpc^{-1}) and gradually decreases as the frequency moves toward higher values. As seen in Fig. 1, these scales correspond to the peak of the linear power spectrum and the largest scale of the halos, respectively.

In order to understand this feature, let us write $\langle S^2_{\text{Born}} \rangle$ in terms of the matter power spectrum as

$$\langle S^2_{\text{Born}} \rangle = 4\omega^2 \left(\frac{3H_0^2 \Omega_m}{2} \right)^2 \int_0^{\chi_s} \frac{d\chi}{a^2(\chi)} \int \frac{dk}{2\pi k^3}$$
$$\times \left(1 - \cos \left[\frac{(\chi_s - \chi)\chi}{2\chi_s \omega} k^2 \right] \right)^2 P_\delta(k, \chi). \quad (4.1)$$

The integration over k is dominated by the integrand around the scale where the argument of the cosine function becomes $\mathcal{O}(1)$. Thus, approximating P_δ at that scale as a single power

law $P_\delta \propto k^{-b}$, we find the scaling $\langle S^2_{\text{Born}} \rangle \propto \omega^{1-\frac{b}{2}}$, which explains the behavior of the purple curve in Fig. 3 and the reason why it reflects the matter power spectrum at the corresponding Fresnel scale. Notice that the Taylor expansion of the cosine in $1/\omega$ in the high-frequency regime, which naively gives the scaling $\langle S^2_{\text{Born}} \rangle \propto \omega^{-2}$, does not make sense due to the divergence of k integration stemming from the coefficient of ω^{-2}.

The fact that the scaling of $\langle S^2_{\text{Born}} \rangle$ depends on b is one of the advantages of the weak lensing of GWs as the measurement of the frequency dependence of $\langle S^2_{\text{Born}} \rangle$ at a detectable frequency range of GWs provides a direct probe of the slope of the matter spectrum at small scales (as small as or smaller than $k \sim 10^5$–$10^6 h$ Mpc^{-1}).

The orange solid line in Fig. 6 shows $\langle K^2_{\text{Born}} \rangle$ as a function of f with the corresponding Fresnel scale at the top. Contrary to the phase modulation, the variance of K approaches a constant value in the high-frequency limit, which is nothing but the variance of the convergence κ in geometrical optics. Also, we can observe a hump at $f \sim 10^{-15}$ Hz in the same way as the phase modulation. At frequencies below $f = 10^{-11}$ Hz, $\langle K^2_{\text{Born}} \rangle$ decreases as the frequency is lowered. In order to understand this feature, let us write $\langle K^2_{\text{Born}} \rangle$ in terms of the matter power spectrum as
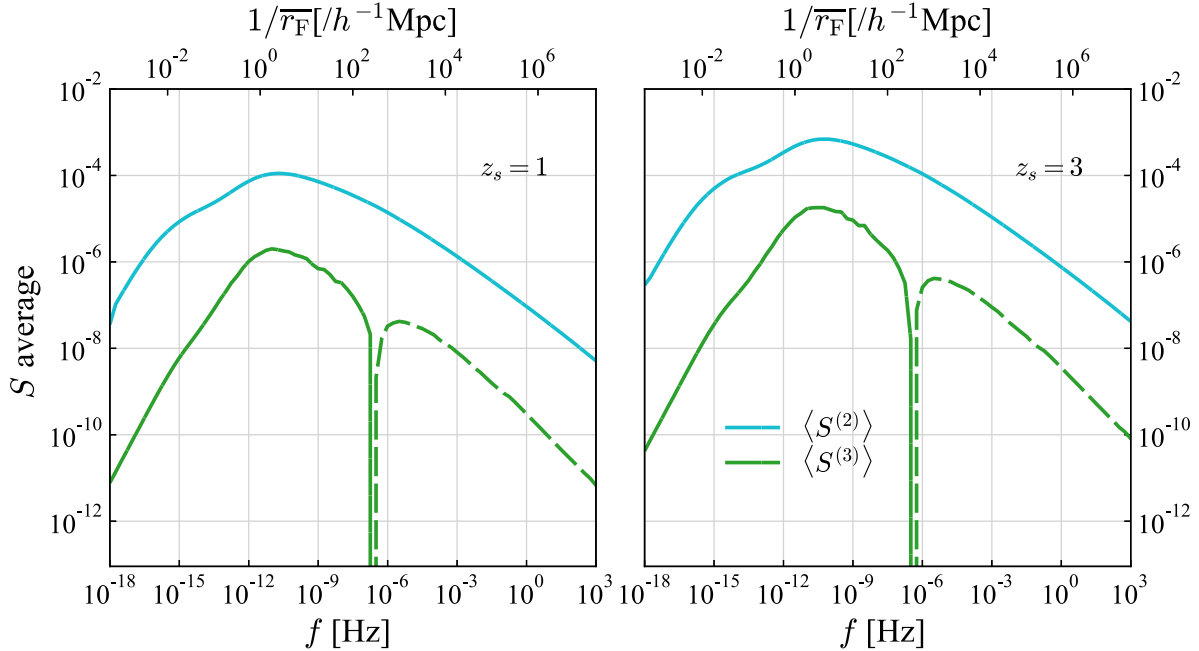

FIG. 3. It is clear that $\langle S^{(3)} \rangle$ (green) is a few orders of magnitude smaller than $\langle S^{(2)} \rangle$ (cyan), confirming the validity of using $\langle S^{(2)} \rangle$ as an approximation of $\langle S \rangle$. In both figures, $\langle S^{(2)} \rangle$ experiences a hump at $f \sim 10^{-15}$ Hz ($k \sim 10^{-2}h$ Mpc^{-1}) and reaches the peak at $f \sim 10^{-11}$ Hz ($k \sim 10^1 h$ Mpc^{-1}), tracing the behavior of the power spectrum. $\langle S^{(3)} \rangle$ also exhibits a similar behavior (the hump at $f \sim 10^{-15}$ Hz and the peak at $f \sim 10^{-11}$ Hz). However, the change of signature at $f \sim 10^{-7}$ Hz seems to be a reflection of enhancement on the squeezed bispectrum caused by subhalos. The solid (dashed) line indicates the $+(-)$ value.

$$\langle K_{\text{Born}}^2 \rangle = 4\omega^2 \left(\frac{3H_0^2 \Omega_m}{2} \right)^2 \int_0^{\chi_s} \frac{d\chi}{a^2(\chi)} \int \frac{dk}{2\pi k^3}$$

$$\times \sin^2 \left[\frac{(\chi_s - \chi)\chi}{2\chi_s \omega} k^2 \right] P_\delta(k, \chi). \qquad (4.2)$$

The difference between the magnification and the phase arises from the difference in the filter function. As mentioned above, $\langle S_{\text{Born}}^2 \rangle$ is predominantly determined by the power spectrum at the Fresnel scale due to the filter function being relevant only when k is comparable to the Fresnel scale. On the other hand, the filter function for $\langle K_{\text{Born}}^2 \rangle$ indicates that it provides relevant contributions whenever the argument of $\sin(\cdots)$ is comparable to or smaller than unity. Thus, $\langle K_{\text{Born}}^2 \rangle$ computes the weighted sum of the power spectrum at all scales above the Fresnel scale. For this reason, the magnification in wave optics can be viewed as the convergence in geometric optics with the smoothing radius being replaced by the Fresnel scale. Since the scale above the Fresnel scale contributes to the magnification, the strong lensing due to galaxies might cause relevant effects, even if the Fresnel scale for typical GWs ($k \sim 10^6 h \, \text{Mpc}^2$) is much smaller than the size of a galaxy ($k \sim 10^3 h \, \text{Mpc}^{-3}$). However, as partially mentioned in [10], the removal of strong lensing by galaxies allows us to extract the underlying lensing signal primarily due to the dark matter halos and the point mass.

C. Average

Having understood the behavior of the magnification and phase modulations in the Born approximation, let us proceed to the post-Born corrections. The cyan curve in Fig. 3 shows the nontrivial leading-order term of $\langle S \rangle (= \langle S^{(2)} \rangle)$ and the green line shows their non-Gaussian correction $\langle S^{(3)} \rangle$. They are both shown as a function of GW frequency f and are numerically computed based on Eqs. (2.31) and (B15) combined with Eqs. (3.1) and (3.2). The source redshift is taken to be $z_s = 1$ for the left and $z = 3$ for the right. As mentioned in the previous section, $\langle S^{(2)} \rangle$ should be always positive and our numerical result also confirms this property.

Figure 3 implies that $\langle S \rangle$ shows close similarities to $\langle S_{\text{Born}}^2 \rangle$: it has a hump around $f = 10^{-15}$ Hz and a peak at $f = 10^{-11}$ Hz, representing the hump and peak of the power spectrum as $\langle S_{\text{Born}}^2 \rangle$ does. Based on the same analysis, we found that $\langle S^{(2)} \rangle$ scales as $\omega^{(2-b)/4}$, thus, the scaling behavior is the same as $\langle S_{\text{Born}}^2 \rangle$ as well as $\langle S^{(2)} \rangle$ and $\langle S_{\text{Born}} \rangle$ having the same order of magnitude $\langle S \rangle \sim \langle S_{\text{Born}}^2 \rangle$. We also evaluate $\langle S^{(3)} \rangle$ (the correction term to $\langle S \rangle$) that originates due to the bispectrum contribution. $\langle S^{(3)} \rangle$ is shown to be about 2 orders of magnitude smaller than $\langle S^{(2)} \rangle$, which guarantees that the correction to $\langle S \rangle$ due to the non-Gaussian effect is negligible and the estimation of $\langle S \rangle$ solely by the power spectrum term is still reliable.

Based on the result of $\langle S_{\text{Born}}^2 \rangle$ and $\langle S^{(2)} \rangle$, it is expected that the frequency dependence of $\langle S^{(3)} \rangle$ also reflects the behavior of the bispectrum. Indeed, $\langle S^{(3)} \rangle$ traces the general behavior of the bispectrum: it initially increases with a small hump at $f = 10^{-15}$ Hz and reaches a peak at $f = 10^{-11}$ Hz, after which it decreases. However, there is a frequency scale at $f \sim 10^{-7}$ Hz (corresponding Fresnel scale is 10^2–$10^3 h \, \text{Mpc}^{-1}$) where the signature of $\langle S^{(3)} \rangle$ flips. Our numerical computation reveals that this change of signature $\langle S^{(3)} \rangle$ is attributed to the squeezed bispectrum. Nevertheless, the precise influence of different squeezed bispectrum configurations, such as the ratio of the larger and smaller scales, on the change of signature has not been elucidated yet. However, it is worth noting that this scale significantly exceeds the scale of our interest ($f \sim 0.01$–1000 Hz) and would also be predominantly influenced by galactic structures. For this reason, the specific behavior of $\langle S^{(3)} \rangle$ around this frequency is not particularly relevant in our analysis compared to the scale around $f \sim 1$ Hz. Also, as we will see later, the hump in $\langle S^{(3)} \rangle$ at $f \sim 10^{-15}$ Hz is relatively smooth compared to the hump seen in $\delta_{S^2,\text{c}}$, though its specific physical meaning is unclear to us yet.

The olive curve in Fig. 4 shows the nontrivial leading-order correction to $\langle K \rangle$, while the brown dashed line is the next leading-order contribution. As observed in $\langle K_{\text{Born}}^2 \rangle$, $\langle K \rangle$ initially increases and approaches a constant.

We find that $\langle K \rangle$ is always negative within the frequency range of our calculation, which is not obvious from the formulation of $\langle K \rangle$. Although the exact reason for this negative sign is still unclear, a similar result appears in geometric optics. In geometric optics, the mean convergence $\langle \kappa \rangle$ is shown to follow the simple relation $\langle \kappa \rangle = -2 \langle \kappa^2 \rangle < 0$,[4] hence negative, as demonstrated analytically [26] and in N-body simulation [27]. Since K is reduced to the convergence κ in weak lensing when taking the geometric optics limit ($\omega \to \infty$), $\langle K \rangle$ should be negative, at least when $\omega \to \infty$ is taken.

While $\langle K^{(3)} \rangle$ (the correction to $\langle K \rangle$) also shows similar behavior to $\langle K^{(2)} \rangle$, it is smaller than $\langle K^{(2)} \rangle$ by about 2 orders of magnitude. Similar to the case of $\langle S \rangle$, the correction to $\langle K \rangle$ from the non-Gaussian effect is small, and therefore the estimation of $\langle K \rangle$ solely by the power spectrum term is reliable.

D. Variance

Before discussing the significance of the post-Born corrections to the variance solely due to the presence of the dark matter halos, it is important to clarify that we have

---

[4]By performing integration by parts for the integration over χ_1 in Eq. (2.30), we can verify that $\langle K \rangle$ and $\langle K_{\text{Born}}^2 \rangle$ derived in this paper reproduce the relation $\langle K \rangle = -\langle K_{\text{Born}}^2 \rangle$ in the high-frequency limit.
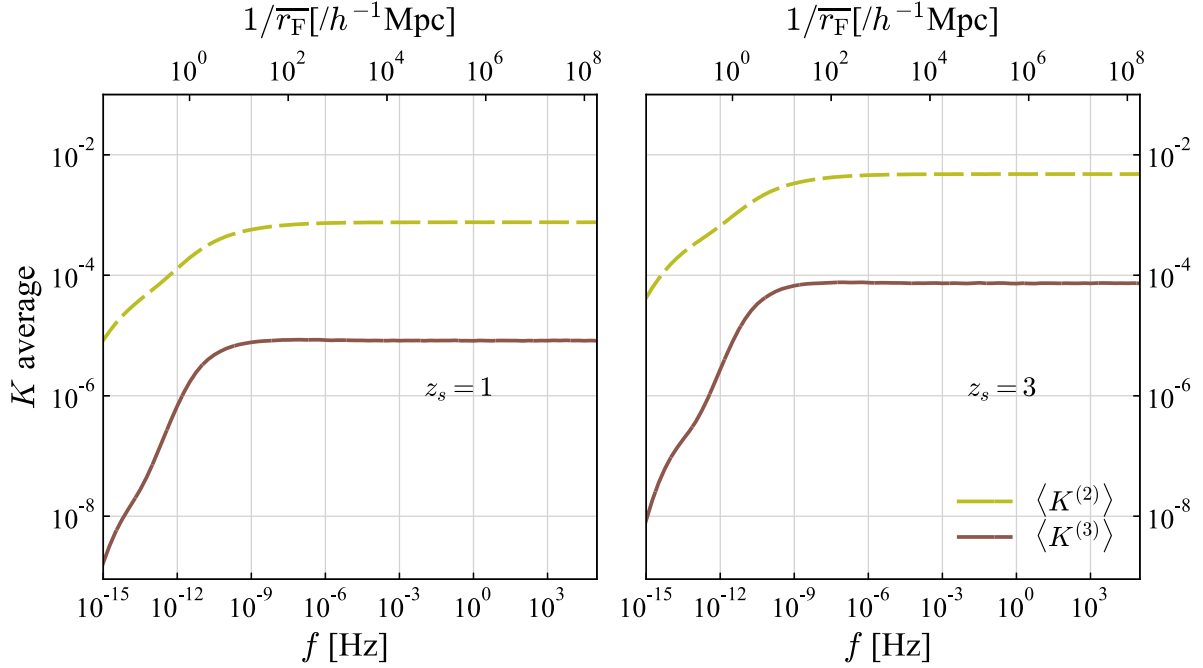
FIG. 4. As expected, both $\langle K^{(2)} \rangle$ (olive) and $\langle K^{(3)} \rangle$ (brown) approach a constant value as the frequency increases. Since $\langle K^{(3)} \rangle$ is more than an order of magnitude smaller than $\langle K^{(2)} \rangle$, it is reasonable to use $\langle K^{(2)} \rangle$ as an approximation of $\langle K \rangle$. At the low end of these figures, $K^{(3)}$ decreases more rapidly than $\langle K^{(2)} \rangle$, indicating the small non-Gaussian effect at large scales. We can see that the hump also exists ($f \sim 10^{-15}$ Hz) for the average of K. The solid (dashed) line indicates the $+(-)$ value.

successfully calculated the post-Born corrections numerically for almost all cases using the expressions for $\delta_{S^2,\mathrm{dc}}$, $\delta_{S^2,\mathrm{c}}$, $\delta_{K^2,\mathrm{dc}}$, $\delta_{K^2,\mathrm{c}}$ presented in Appendix B. However, there is one case that poses computational challenges when computing the Gaussian correction to the phase $\delta_{S^2,\mathrm{dc}}$, particularly above frequency $f = 10^{-7}$ Hz. The reason for this difficulty arises from the cancellation of significant digits. Specifically, $\langle (S^{(2)})^2 \rangle_{\mathrm{dc}}$ and $2\langle S^{(1)} S^{(3)} \rangle_{\mathrm{dc}}$ in $\delta_{S^2,\mathrm{dc}}$ have very similar values but differ in their signature at high frequency, causing the cancellation of almost all contributions from each term and leaving very small differences. It is known that, in geometric optics, the translation invariance of the correlation functions plays a key role in this type of cancellation for K [12–16]. To overcome this issue, we employ an alternative approximation method to obtain $\delta_{S^2,\mathrm{dc}}$ at frequencies above $f \sim 10^{-7}$ Hz. In Appendix C, we develop a method to approximately obtain $\delta_{S^2,\mathrm{dc}}$ given the power spectrum following a singe-power law at high wave number. On the other hand, the power spectrum can be approximated by a single power law around an arbitrary reference scale k_0 using Eq. (3.3). Due to the expectation that the power spectrum at the Fresnel scale dominates the contribution to $\delta_{S^2,\mathrm{dc}}$, we chose the reference scale to be the approximated value of the corresponding Fresnel scale, namely, $k_0 = \sqrt{H_0 \omega}$. By utilizing the approximated power spectrum, which takes the form of a power law around the Fresnel scale and our developed computation method, we

can effectively compute $\delta_{S^2,\mathrm{dc}}$ with reasonable accuracy using the following expression:

$$\delta_{S^2,\mathrm{dc}} = \frac{3}{4} \left(\frac{3H_0^2 \Omega_m}{2} \right)^4 \int_0^{\chi_s} d\chi_1 \int_0^{\chi_1} d\chi_2 \frac{W^4(\chi_1, \chi_s) \chi_1^4 \chi_2^4}{a^2(\chi_1) a^2(\chi_2)} \frac{1}{(2\pi)^2}$$
$$\times \left\{ \frac{B(\chi_2)}{\omega^{\frac{b_2}{2}-1}} \int_0^\infty dk_1 k_1 P_\delta(k_1, \chi_1) \left[\frac{\chi_2^2 W(\chi_2, \chi_s)}{2} \right]^{\frac{b_2}{2}-1} \right.$$
$$\times \left. \left(1 - \frac{2^{\frac{b_2}{2}}}{2} \right) \Gamma \left(1 - \frac{b_2}{2} \right) \sin \left[\frac{b_2}{2} \right] + (1 \leftrightarrow 2) \right\}, \quad (4.3)$$

where $b_2 = b(\chi_2)$ and $B(\chi_2)$ are calculated using Eq. (3.3). Note that, in analyzing the frequency dependency of $\delta_{S^2,\mathrm{dc}}$, it can be regarded as a constant with respect to the variations in χ due to the relatively small variation of $b(\chi)$ with changing χ.

Figure 5 shows the variance of S including the Born approximation $\langle S_{\mathrm{Born}}^2 \rangle$ and two types of post-Born corrections (Gaussian $\delta_{S^2,\mathrm{dc}}$ and non-Gaussian $\delta_{S^2,\mathrm{c}}$ with the non-Gaussian correction only containing the bispectrum contribution). The source redshift is taken to be $z_s = 1, 3$ for the left and right graphs, respectively. This result suggests that the post-Born corrections to the variance of S are subdominant at all frequency ranges considered in the paper. Also, it implies that the smallness of the post-Born corrections remains valid regardless of whether the matter density is Gaussian or non-Gaussian, as long as the
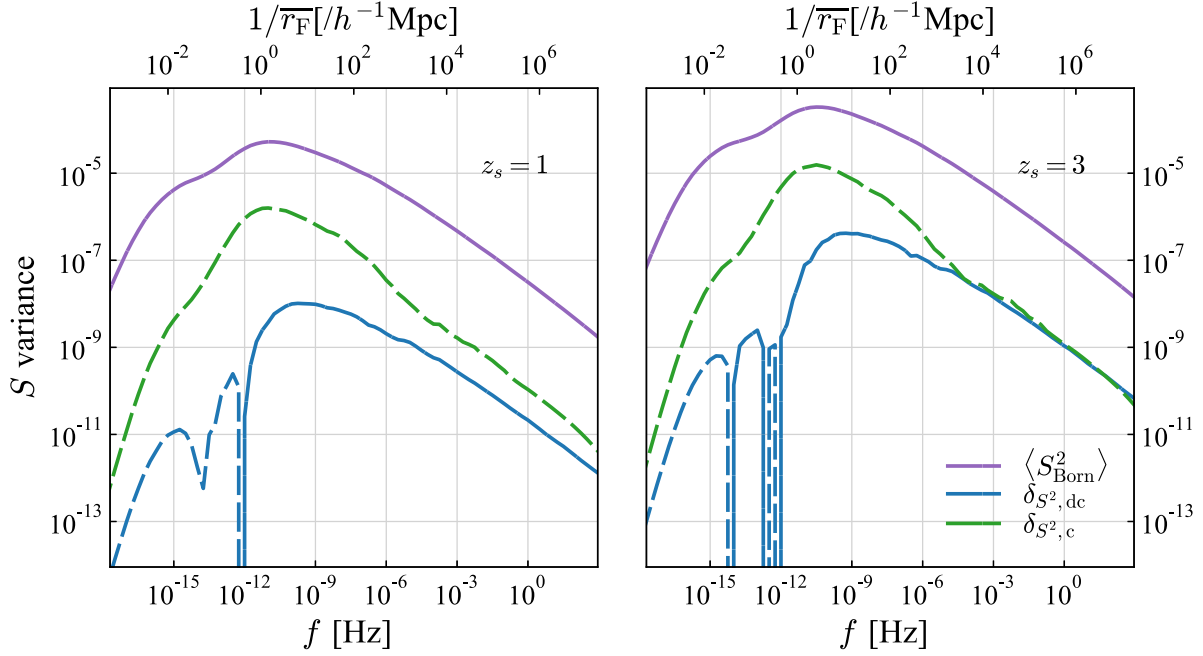
FIG. 5. In the frequency range of our interest $f > 0.01$ Hz ($k \sim 10^5 h$ Mpc^{-1}), $\langle S^2_{\mathrm{Born}} \rangle$ (purple) is more than 2 orders of magnitude larger than both $\delta_{S^2,\mathrm{dc}}$ (blue) and $\delta_{S^2,\mathrm{c}}$ (green), showing that the Born approximation remains valid under the halo model. In this frequency range, the post-Born corrections $\delta_{S^2,\mathrm{dc}}$, and $\delta_{S^2,\mathrm{c}}$ mainly enhance the amplitude of $\langle S^2_{\mathrm{Born}} \rangle$, which is interpreted as the effect of the halos being distributed unevenly. While $\langle S^2_{\mathrm{Born}} \rangle$ and $\delta_{S^2,\mathrm{c}}$ show the hump and the peak at $f \sim 10^{-15}$ Hz ($k \sim 10^5 h$ Mpc^{-1}, the scale corresponding to the peak of linear spectra) and at $f \sim 10^{11}$ Hz ($k \sim 10^5 h$ Mpc^{-1}, the scale corresponding to the size of the largest halos), $\delta_{S^2,\mathrm{dc}}$ changes its signature multiple times in this range. Also, at $f \sim 10^{-5}$ Hz, there is a sudden change of the slope of $\delta_{S^2,\mathrm{c}}$, respecting the effect of the enhanced squeezed bispectrum by subhalos. The solid (dashed) line indicates the $+(-)$ value.

contribution to the variance is primarily attributed to the dark matter halos. Note that the computation of $\delta_{S^2,\mathrm{dc}}$ using Eq. (4.3) is reliable as it nicely coincides with the result of the numerical computation of $\delta_{S^2,\mathrm{dc}}$ at $f \sim 10^{-7}$ Hz.

The behavior of $\delta_{S^2 dc}$ is drastically different from $\langle S^2_{\mathrm{Born}} \rangle$ below $f = 10^{-9}$ Hz. While $\langle S^2_{\mathrm{Born}} \rangle$ exhibits a hump and peak at $f = 10^{-15}$ and $f = 10^{-11}$ Hz, the Gaussian correction $\delta_{S^2,\mathrm{dc}}$ changes its signature multiple times in this region. On the other hand, the non-Gaussian correction $\delta_{S^2,\mathrm{c}}$ shows generally the same behavior as $\langle S^2_{\mathrm{Born}} \rangle$: it increases with a small hump at $f \sim 10^{-15}$ Hz and reaches the peak at $f = 10^{-11}$ Hz, after which it gradually decreases. However, there is a frequency scale at $f = 10^{-5}$ Hz corresponding to $k \sim 10^4 h$ Mpc^{-1} where the behavior of $\delta_{S^2,\mathrm{c}}$ changes. Similar to $\langle S^{(3)} \rangle$, it appears that this change of behavior represents the effect of subhalos through the dependence of $\langle S^{(3)} \rangle$ on the squeezed bispectrum. Moreover, by comparing Figs. 3 and 5 at $f \sim 10^{-15}$ Hz, it is clear that the hump observed in Fig. 3 is relatively less prominent than the one observed in Fig. 5.

Since the frequencies $f \sim 10^{-15}$ and $f \sim 10^{-11}$ Hz correspond to the peak that appears in the linear spectrum and the scale of the largest halos in the halo model ($b \sim 0$ and $b \sim 2$ for each scale), it is suggested that $\delta_{S^2,\mathrm{dc}}$ encodes information about this scale. In fact, Eq. (4.3) partially captures this behavior. For example, the sine function in Eq. (4.3) enforces this term to be suppressed when $b \sim 0$. Note that this estimation should not be taken too seriously as Eq. (4.3) cannot be used when $b < 2$. In addition, identifying the exact scale is also challenging since the error of the numerical calculation inevitably increases at this specific frequency. Thus, further investigation is needed for understanding the more precise property of $\delta_{S^2,\mathrm{dc}}$ at around this frequency.

On the other hand, the frequency above $f = 10^{-9}$ Hz provides important physical insight. By examining the frequency dependence of $\delta_{S^2,\mathrm{dc}}$ in Eq. (4.3), it is clear that $\delta_{S^2,\mathrm{dc}}$ scales as $\omega^{1-b/2}$, the same scaling property as $\langle S^2_{\mathrm{Born}} \rangle$. The underlying reason for this scaling can be understood as follows. In the case of $\langle S^2_{\mathrm{Born}} \rangle$, the main contribution arises solely from the power spectrum (thus two-point correlation function) evaluated at the Fresnel scale. However, as Eq. (4.3) implies, the main contribution to $\delta_{S^2,\mathrm{dc}}$ comes from the product of the power spectrum evaluated at two scales: the Fresnel scale and the scale that primarily contributes to $\int dk k P(k)$. This leads to the power spectrum evaluated at the Fresnel scale producing $\omega^{1-b/2}$, while the other one contributes mainly to the amplitude of $\delta_{S^2,\mathrm{dc}}$.

The significance of this cross term contribution is that information about large-scale density fluctuation is encoded

in the variance of S through higher-order terms. The above discussion about the behavior of the Gaussian correction $\delta_{S^2,\mathrm{dc}}$ immediately implies that the correlations between density fluctuations at the Fresnel scale and much larger scales are the main contributing factor to $\delta_{S^2,\mathrm{dc}}$. Physically, this correlation arises from the fact that small-scale density fluctuations are more likely to grow in regions with a high matter density (i.e., large-scale density fluctuations are significant). In fact, the main contribution to the non-Gaussian correction $\delta_{S^2,\mathrm{c}}$ at high frequencies also comes from this correlation, since $\delta_{S^2,\mathrm{c}}$ is particularly affected by the squeezed bispectrum, which is by definition the correlation between large- and small-scale density fluctuation.

In order to understand this effect more intuitively, let us consider a universe in which lensing only occurs by dark matter halos. In this case, $\langle S^2_{\mathrm{Born}}\rangle$ fails to capture the uneven distribution of the halos. The rationale for this is that $\langle S^2_{\mathrm{Born}}\rangle$ is solely determined by the abundance of halos at the Fresnel scale, and information about unevenness (such as bispectrum) is absent. However, it is expected that the true $\langle S^2\rangle$ would deviate from $\langle S^2_{\mathrm{Born}}\rangle$ simply because the lensing effect is enhanced (suppressed) in regions with a high (low) halo number density compared to regions with average number density. Our analysis indicates that the post-Born approximation can capture the uneven distribution of lens objects by accounting for the correlation between the density fluctuation at the Fresnel scale and the density fluctuations at much larger scales.

In terms of the frequency range of our interest, which is above $f = 0.01$ Hz and corresponds to the scale dominated by the dark low-mass halos, our result suggests that the post-Born corrections to the variance are estimated to be $\delta_{S^2}/\langle S^2_{\mathrm{Born}}\rangle \lesssim \mathcal{O}(10^{-2})$. Also, it is important to mention that the Gaussian correction term $\delta_{S^2,\mathrm{dc}}$ and the non-Gaussian correction term $\delta_{S^2,\mathrm{c}}$ are relatively the same order of magnitude in this frequency range.

Let us next investigate the variance of K. Figure 6 shows the correction to $\langle K^2_{\mathrm{Born}}\rangle$ in the case where the source redshift is $z_s = 1$ and $z_s = 3$, respectively. Similar to $\langle K^2_{\mathrm{Born}}\rangle$ and $\langle K\rangle$, the correction to the variance is suppressed at low frequency and approaches a constant value in the high-frequency limit. One difference between the post-Born correction to the magnification and the phase modulation is the relative magnitude of the non-Gaussianity term to the Gaussianity term. In the case of the magnification, the non-Gaussian term $\delta_{K^2,\mathrm{c}}$ is almost 2 orders of magnitude larger than the Gaussian term $\delta_{K^2,\mathrm{dc}}$ in the frequency range $f = 0.01\text{–}1000$ Hz, while those terms are relatively the same order of magnitude for the phase modulation. Due to this effect, the post-Born correction to the variance of K in the high-frequency range is mainly caused by the non-Gaussian term $\delta_{K^2} \approx \delta_{K^2,\mathrm{c}}$. The relative magnitude of the post-Born correction is given by $\delta_{K^2}/\langle K^2_{\mathrm{Born}}\rangle \lesssim \mathcal{O}(10^{-1})$ in the frequency range where the magnification can be treated as a constant. On the other hand, when the frequency is smaller than $f = 10^{-9}$ Hz, the
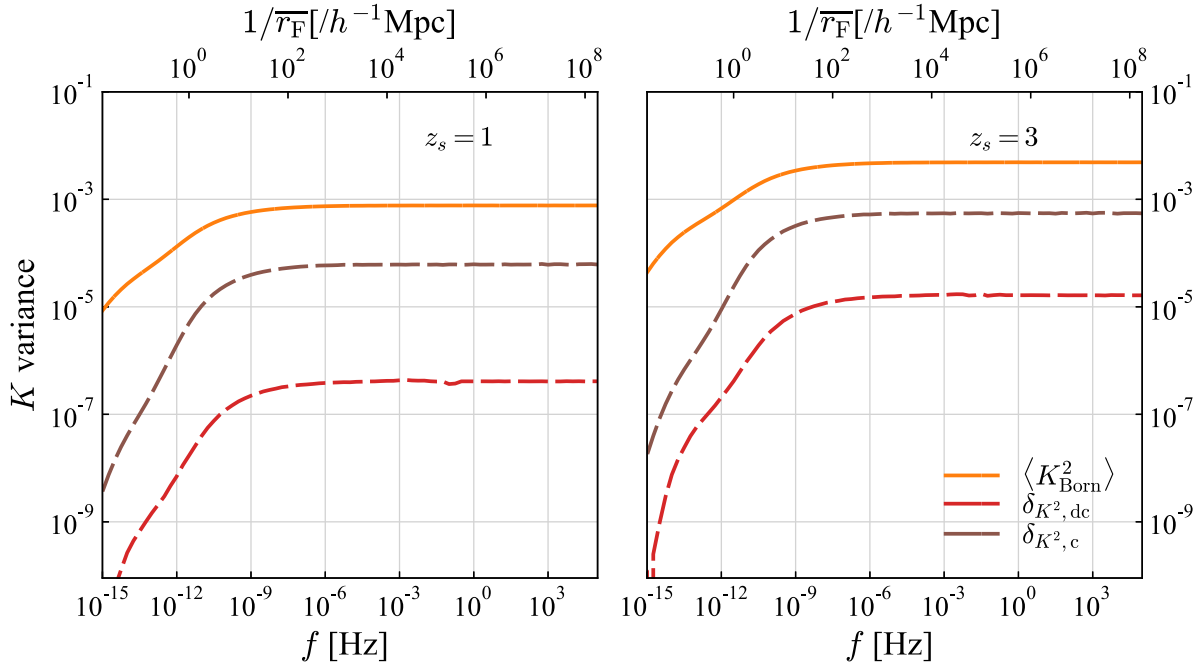

FIG. 6. The Born approximation $\langle K^2_{\mathrm{Born}}\rangle$ (orange) is the dominant effect on the K, with $\delta_{K^2,\mathrm{c}}$ (brown) being the second. Gaussian contribution $\delta_{K^2,\mathrm{c}}$ (red) is more than an order of magnitude smaller than the non-Gaussian term. The solid (dashed) line indicates the $+(-)$ value.

non-Gaussian term exhibits a faster decrease compared to the Gaussian correction term, which can be seen by the reduction of the difference between $\delta_{K^2,\text{c}}$ and $\delta_{K^2,\text{dc}}$ in Fig. 6. This behavior reflects the fact that the non-Gaussianity is less significant at larger scales.

E. Shot noise contribution

Up to this point, we have not considered the effect of the shot noise coming from the Poisson distributed stars (or any other dark compact objects) because (i) the low-frequency GWs are not strongly affected by the point masses and (ii) numerical integrations face technical challenges associated with the highly oscillatory behavior of the integrand. However, the shot noise may become important at high frequencies, as suggested in [10], and the effect of the post-Born correction needs to be considered. Thus we discuss the contribution of the shot noise in this subsection. At the level of the post-Born approximation, the lensing signals have a dependency on not only the terms purely representing the shot noise but also the cross term between the shot noise terms and the smooth halo terms. This is due to the fact that the higher-order corrections to S and K contain the bispectrum and the product of the power spectrum. The evaluation of the cross terms is beyond the scope of this paper and we will only focus on the pure contribution of the shot noise.

Formally, the shot noise is given by adding a constant to the matter power spectrum and the bispectrum, namely,

$$P_\delta = \frac{f_p^2}{\bar{n}}, \qquad (4.4)$$

$$B_\delta = \frac{f_p^3}{\bar{n}^2}, \qquad (4.5)$$

where $\bar{n}$ is the average number density of an individual star and f_p is the mass fraction of the point mass to total matter density. For simplicity, we ignore the time variation of $\bar{n}$ due to stellar evolution. As mentioned earlier, we chose $f_p = 0.01$ and $m = 0.5M_\odot$ as a fiducial value, which is consistent with observations [21]. The shot noise effect on the lensing signal is caused by the point mass, which, in reality, possesses a finite physical size. To account for this finite size, it is reasonable to introduce a cutoff scale k_c, which characterizes the regime where the point mass approximation holds, since the cutoff scale represents the size of stars in this case, and we chose $k_c = 4 \times 10^{13}h$ Mpc^{-1} ($\sim 10^{-6}$ km^{-1}).

We first evaluate the shot noise contribution to $\langle S_{\text{Born}}^2 \rangle$ and $\langle K_{\text{Born}}^2 \rangle$. As long as it is smaller than the Fresnel scale, the size of the stars k_c does not have a relevant contribution to the results, allowing us to effectively take $k_c \to \infty$. Substituting Eq. (4.4) for the power spectrum that appears in Eq. (4.1), we can analytically perform the integral over k by using the formula $\int_0^\infty \frac{dx}{x^3}(1 - \cos x^2)^2 = 1/(4\pi)$. Then, we obtain

$$\langle S_{\text{Born}}^2 \rangle_{\text{shot}} = \frac{1}{4}\left(\frac{3H_0\Omega_m}{2}\right)^2 \int_0^{\chi_s} \frac{d\chi}{a^2(\chi)}\chi^2 W(\chi,\chi_s) \times \omega\frac{f_p^2}{\bar{n}}. \qquad (4.6)$$

Similarly, using the formula $\int_0^\infty \frac{dx}{x^3}\sin^2 x^2 = 1/(4\pi)$, we find that $\langle K_{\text{Born}}^2 \rangle_{\text{shot}} = \langle S_{\text{Born}}^2 \rangle_{\text{shot}}$. Thus, both the variance of S and K under the Born approximation diverge in the high-frequency limit. Note that this divergence comes from neglecting the size of stars. In reality, the point mass approximation breaks down when the Fresnel scale becomes smaller than the size of stars. By properly incorporating the size of stars, $\langle S_{\text{Born}}^2 \rangle_{\text{shot}}$ and $\langle K_{\text{Born}}^2 \rangle_{\text{shot}}$ can be shown to remain finite in the high-frequency limit.

1. Shot noise contribution to average

Let us next investigate the average $\langle S \rangle$ and $\langle K \rangle$. Plugging Eqs. (3.1) and (4.4) into the expression $\langle S \rangle$ given by Eq. (2.29), we find that the integration over k diverges logarithmically at large k. Hence, we need the cutoff wave number k_c which physically represents the inverse of the size of the stars. With this cutoff, we can perform the integration over k and the result is given by

$$\langle S \rangle_{\text{shot}} = 2\left(\frac{3H_0\Omega_m}{2}\right)^2 \int_0^{\chi_s} d\chi \int_0^\chi d\chi' \frac{\chi'^2}{\chi^2}\frac{1}{a^2(\chi')}\frac{1}{4\pi}$$
$$\times \text{Cin}\left(k_c^2\frac{\chi'^2 W(\chi',\chi)}{\omega}\right) \times \omega\frac{f_p^2}{\bar{n}}, \qquad (4.7)$$

where $\text{Cin}(x)$ is the cosine integral defined as $\text{Cin}(x) = \int_0^x dt\frac{1-\cos t}{t}$. When x is sufficiently large, the cosine integral is approximated as $\text{Cin}(x) \approx \log(e^\gamma x)$, where γ is Euler's constant. We usually consider the case where χ_s takes the cosmological distance ($\chi_s \approx 1/H_0$). In this specific case, we can further approximate this expression by

$$\langle S \rangle_{\text{shot}} \sim \frac{1}{2\pi}\left(\frac{3H_0\Omega_m}{2}\right)^2 \int_0^{\chi_s} d\chi \int_0^\chi d\chi' \frac{\chi'^2}{\chi^2}\frac{1}{a^2(\chi')}$$
$$\times \omega\frac{f_p^2}{\bar{n}}\log\left(\frac{k_c^2}{H_0\omega}\right). \qquad (4.8)$$

Since the cutoff scale k_c only appears in $\log(\cdots)$, it can be eliminated by the weighted subtraction of S evaluated at different frequencies. For example, when $\langle S(\omega_1)/\omega_1 - S(\omega_2)/\omega_2 \rangle$ is computed using the shot noise power spectrum, $k_c^2/H_0\omega$ that appeared as an argument of $\log(\cdots)$ is replaced by ω_2/ω_1.

As for $\langle K \rangle$, the integral over k does not diverge and we can practically take $k_c \to \infty$. Then, the result is given by

$$\langle K \rangle_{\text{shot}} = -\frac{1}{4} \left(\frac{3H_0\Omega_m}{2} \right)^2 \int_0^{\chi_s} d\chi \int_0^{\chi} d\chi' \frac{\chi'^2}{\chi^2} \frac{1}{a^2(\chi')} \times \omega \frac{f_p^2}{\bar{n}}.$$

$$(4.9)$$

$$\langle S^{(3)} \rangle_{\text{shot}} = 0, \tag{4.10}$$

$$\langle K^{(3)} \rangle_{\text{shot}} = 8\omega^2 \left(\frac{3H_0^2\Omega_m}{2} \right)^3 \int_0^{\chi_s} \frac{d\chi}{a(\chi)^3} \chi^2 W(\chi,\chi_s) J \times \frac{f_p^3}{\bar{n}^2},$$

$$(4.11)$$

The third-order terms $\langle S^{(3)} \rangle$ and $\langle K^{(3)} \rangle$ can be calculated in a similar way. Using Eqs. (3.2) and (4.5), we obtain the following expressions:

where J is just a number given by the following integral:

$$J = \frac{1}{12} \int_0^\infty \frac{d\xi_1}{2\pi} \int_0^\infty \frac{d\xi_2}{2\pi} \int_0^{2\pi} \frac{d\phi}{2\pi} \frac{\sin(\xi_1^2) + \sin(\xi_2^2) + \sin(\xi_1^2 + \xi_2^2 + 2\xi_1\xi_2\cos\phi) - 2\sin(\xi_1^2 + \xi_2^2 + \xi_1\xi_2\cos\phi)}{\xi_1\xi_2(\xi_1^2 + \xi_2^2 + 2\xi_1\xi_2\cos\phi)}, \tag{4.12}$$

which is found to be $J \simeq 0.0021$. Note that $\langle S^{(3)} \rangle_{\text{shot}}$ becomes exactly zero. In fact, $\langle S^{(3)} \rangle_{\text{shot}}$ is formally written in the same way as $\langle K^{(3)} \rangle_{\text{shot}}$ is given,

$$\langle S^{(3)} \rangle_{\text{shot}} = 8\omega^2 \left(\frac{3H_0^2\Omega_m}{2} \right)^3 \int_0^{\chi_s} \frac{d\chi}{a(\chi)^3} \chi^2 W(\chi,\chi_s) L \times \frac{f_p^3}{\bar{n}^2}, \tag{4.13}$$

where L is given by an integral

$$L = \frac{1}{6} \int_0^\infty \frac{d\xi_1}{2\pi} \int_0^\infty \frac{d\xi_2}{2\pi} \int_0^{2\pi} \frac{d\phi}{2\pi} \frac{2\sin^2(\xi_1^2 + \xi_2^2 + \xi_1\xi_2\cos\phi) - \sin^2(\xi_1^2) - \sin^2(\xi_2^2) - \sin^2(\xi_1^2 + \xi_2^2 + 2\xi_1\xi_2\cos\phi)}{\xi_1\xi_2(\xi_1^2 + \xi_2^2 + 2\xi_1\xi_2\cos\phi)}. \tag{4.14}$$

However, this integral is found to be exactly zero. This implies that the corrections to $\langle S \rangle$ come from much higher-order terms, but it is not clear the rationale for $\langle S^{(3)} \rangle$ being zero.

Figure 7 shows the shot noise contribution to the average of S. For both cases (the source redshift $z_s = 1$, 3), the shot noise effect becomes dominant at slightly above $f = 0.1$ Hz. This frequency that the shot noise takes over
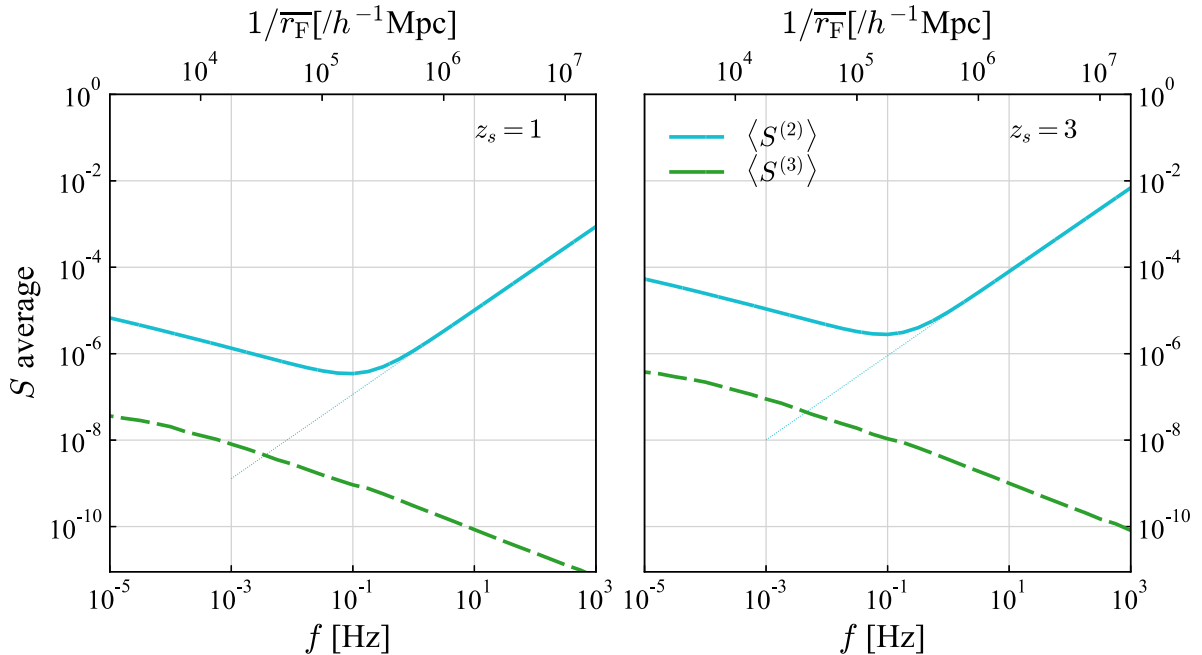

FIG. 7. The cyan line shows the value of $\langle S \rangle$ including the halo and the shot noise terms. The shot noise is represented by the thin cyan straight line ascending in the upper right direction. The shot noise term overcomes the halo term at $f \sim 10^{-1}$ Hz. Compared to the shot noise effect on $\langle S_{\text{Born}}^2 \rangle$ in Fig. 9, the onset of the shot noise effect for $\langle S \rangle$ occurs at a lower frequency than $\langle S_{\text{Born}}^2 \rangle$ due to $\langle S^{(2)} \rangle$ depending on the size of the point mass k_c. The shot noise effect on $\langle S^{(3)} \rangle$ (green) vanishes, indicating that the higher-order terms are necessary to provide the corrections to $\langle S^{(2)} \rangle$. The solid (dashed) line indicates the $+(-)$ value.

is lower than the one for the variance. This is due to the enhancement of the shot noise effect on $\langle S \rangle$, which arises from the logarithmic factor in Eq. (4.8).

Figure 8 shows the shot noise contribution to the average of K. The shot noise effect is hidden until the frequency becomes $f = 10^4$ Hz. Thus, the frequency lower than this is not affected by the shot noise effect. When the frequency is above $f \sim 10^4$, the shot noise effect becomes dominant. However, soon after the shot noise effect dominates $\langle K \rangle$, the higher-order contribution $\langle K^{(3)} \rangle$ overcomes the lower-order term $\langle K^{(2)} \rangle$. When the higher-order terms dominate the lower-order terms, it is an indication that the perturbative approach fails. Taking the ratio of Eq. (4.9) to Eq. (4.11) and assuming the redshift is not too large (truncating the second- or higher-order terms in z_s), we find the following condition for the validity of the perturbative approach for the computation of $\langle K \rangle$:

$$\left| \frac{\langle K^{(3)} \rangle_{\text{shot}}}{\langle K^{(2)} \rangle_{\text{shot}}} \right| = 1.7 G m_p \omega \left(1 + \frac{z_s}{2} \right) < 1. \qquad (4.15)$$

As we will show later, this condition is analogous to the condition under which the Born approximation for the magnification can be reliably applied.

2. Shot noise contribution to variance

Finally, we evaluate the post-Born correction to the variance. First, we consider the corrections to the phase modulation. For the same reason as $\langle S \rangle_{\text{shot}}$, the integral for $\delta_{S^2,\text{dc,shot}}$ and $\delta_{S^2,\text{c,shot}}$ requires the cutoff scale k_c to avoid divergence.

We can compute the approximation of the Gaussian correction $\delta_{S^2,\text{dc}}$, using the property of Eq. (B10). When k_1 and k_2 are large enough, the main contribution to $\delta_{S^2,\text{dc,shot}}$ comes from the factor $\frac{1}{k_1 k_2}$ due to the cancellation by oscillations. In other words, it is possible to make the following approximation:

$$\left[\frac{1}{k_1 k_2} \mathcal{F}_{S,12} - \frac{1}{k_1^2} \mathcal{F}_{S,1} - \frac{1}{k_2^2} \mathcal{F}_{S,2} \right]$$
$$\sim -\frac{1}{2 k_1 k_2} \left(1 - \cos \left(\frac{\chi_1^2 W(\chi_1, \chi_s)}{2\omega} k_1^2 \right) \right)$$
$$\times \left(1 - \cos \left(\frac{\chi_2^2 W(\chi_2, \chi_s)}{2\omega} k_2^2 \right) \right), \qquad (4.16)$$

where k_1 and k_2 are taken to be sufficiently large due to the dominant contribution to $\delta_{S^2,\text{dc,shot}}$ coming from such a region. Using this approximation as well as other simplifications used in the computation of $\langle S \rangle_{\text{shot}}$, we obtain
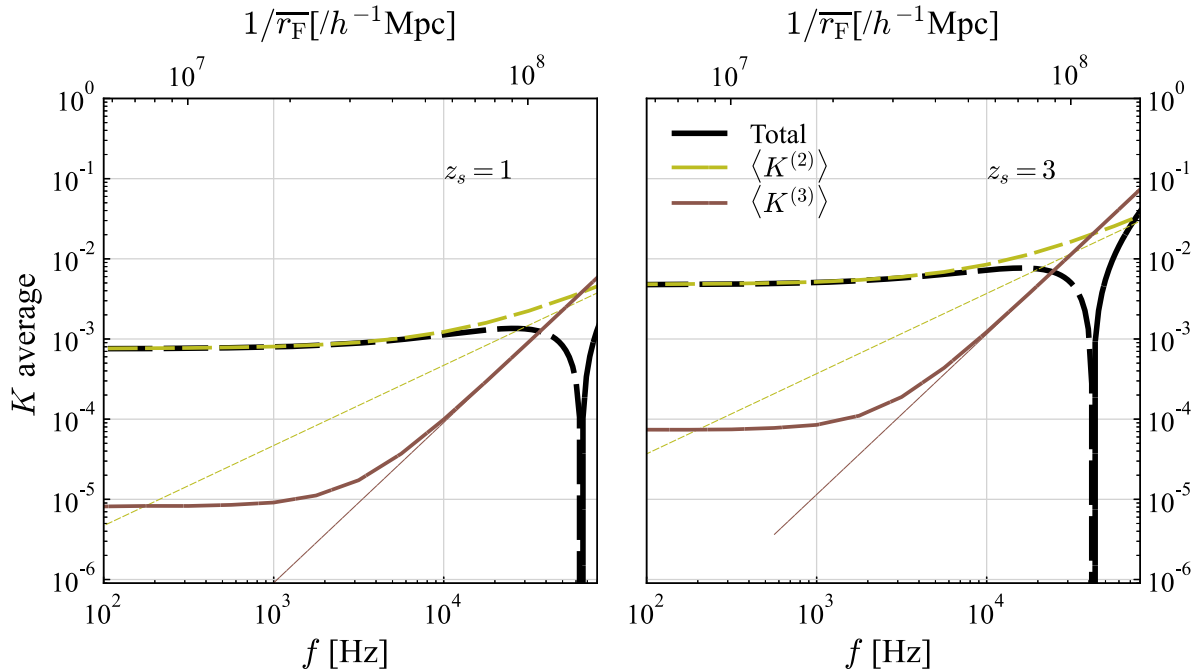

FIG. 8.   The back lines show the total $\langle K \rangle$. $\langle K^{(2)} \rangle$ (olive) is dominated by the halo term until $f \sim 10^4$ Hz. However, $\langle K^{(3)} \rangle$ (brown) exceeds $\langle K^{(2)} \rangle$ around $f \sim 5 \times 10^4$ Hz as well, indicating that the perturbative approach fails around this frequency. Equation (4.15) provides a condition under which $\langle K^{(2)} \rangle$ is a reliable approximation of $\langle K \rangle$, which is equivalent to the weak lensing condition presented in [10] up to a constant prefactor. Note that the precise behavior of $\langle K \rangle$ around frequencies where the halo and the shot noise are of similar magnitude (transition frequency) is not captured in these figures due to the exclusion of cross terms. The solid (dashed) line indicates the $+(-)$ value.

$$\delta_{S^2,\text{dc,shot}} \sim -\frac{1}{2\pi^2}\left(\frac{3H_0^2\Omega_m}{2}\right)^4 \int_0^{\chi_s}\frac{d\chi}{\chi^2}\int_0^{\chi}\frac{d\chi'}{\chi'^2}\int_0^{\chi'}d\chi_1\int_0^{\chi'}d\chi_2$$

$$\times\frac{1}{a^2(\chi_1)}\frac{1}{a^2(\chi_2)}\chi_1^2\chi_2^2\omega^2\left(\frac{f_p^2}{\bar{n}}\right)^2\left(\log\left[\frac{k_c^2}{H_0\omega}\right]\right)^2.$$

$$(4.17)$$

The correction from the bispectrum term can be calculated using Eq. (4.18) and the expression for the bispectrum Eq. (4.5). Since we consider the bispectrum term as the main contribution to the non-Gaussian correction $\delta_{S^2,\text{c,shot}} = 2\langle S_{\text{Born}}S^{(2)}\rangle$ holds. Adopting the same

approximation we used to calculate $\langle S\rangle_{\text{shot}}$ and $\delta_{S^2,\text{dc,shot}}$, we have the following result:

$$\delta_{S^2,\text{c,shot}} \sim \frac{1}{4\pi^2}\left(\frac{3H_0^2\Omega_m}{2}\right)^3 \int_0^{\chi_s}\frac{d\chi}{a^3}\chi^2 W(\chi,\chi_s)$$

$$\times \omega^2\frac{f_p^3}{\bar{n}^2}\left(\log\left[\frac{k_c^2}{H_0\omega}\right]\right)^2.$$

$$(4.18)$$

Similarly, the logarithmic factor arises due to the presence of the cutoff.

As for the Gaussian correction $\delta_{K^2,\text{dc}}$, it does not diverge even if we take $k_c \to \infty$. Taking this limit, the formal expression is given by

$$\delta_{K^2,\text{dc,shot}} = 16\left(\frac{3H_0^2\Omega_m}{2}\right)^4 \int_0^{\chi_s}\frac{d\chi}{\chi^2}\int_0^{\chi}\frac{d\chi'}{\chi'^2}\int_0^{\chi'}d\chi_1\int_0^{\chi'}d\chi_2\frac{1}{a^2(\chi_1)}\frac{1}{a^2(\chi_2)}\frac{1}{(2\pi)^2}$$

$$\times \int_0^{\infty}dk_1\int_0^{\infty}dk_2\left[\frac{1}{k_1k_2}\mathcal{F}_{K,12} - \frac{1}{k_1^2}\mathcal{F}_{K,1} - \frac{1}{k_2^2}\mathcal{F}_{K,2}\right]\times\omega^2\frac{f_p^2}{\bar{n}}.$$

$$(4.19)$$

The definition of $\mathcal{F}_{K,12}, \mathcal{F}_{K,1}, \mathcal{F}_{K,2}$ is given in Appendix B. We have not been able to find an analytic method to approximately compute the integration over k_1, k_2 because it requires careful analytical treatment. Although it is, in principle, possible to compute the integral numerically, it turned out to be quite complicated to achieve it. This arises from the fact that the expression is highly oscillatory at large k. However, it is expected that the non-Gaussian correction $\delta_{K^2,\text{dc}}$ is sufficiently smaller than the Gaussian correction $\delta_{K^2,\text{c,shot}}$ due to the reason we discuss below.

On the other hand, $\delta_{K^2,\text{c,shot}} = 2\langle K_{\text{Born}}K^{(2)}\rangle$ is computed using Eqs. (2.43) and (4.5) as

$$\delta_{K^2,\text{c,shot}} = 16\left(\frac{3H_0^2\Omega_m}{2}\right)^3 \int_0^{\chi_s}\frac{d\chi}{a^3}\chi^2 W(\chi,\chi_s)I\times\omega^2\frac{f_p^3}{\bar{n}^2},$$

$$(4.20)$$

where I is just a number given by the following integral:

$$I = \frac{1}{2}\int_0^{\infty}\frac{d\xi_1}{2\pi}\int_0^{\infty}\frac{d\xi_2}{2\pi}\int_0^{2\pi}\frac{d\phi}{2\pi}\frac{\sin(\xi_1^2+\xi_2^2+2\xi_1\xi_2\cos\phi)\sin(\xi_1^2+\xi_2^2+\xi_1\xi_2\cos\phi)\sin(\xi_1\xi_2\cos\phi)}{\xi_1\xi_2(\xi_1^2+\xi_2^2+2\xi_1\xi_2\cos\phi)},$$

$$(4.21)$$

which is found to be $I \sim -0.0038$.

Figure 9 shows the shot noise effect on the variance of the phase modulation, indicating that the shot noise is subdominant until f becomes greater than $f \sim 1$ Hz. As for the post-Born corrections, it can be seen that the non-Gaussian contribution $\delta_{S^2,\text{c}}$ is the dominant contribution compared to the Gaussian contribution $\delta_{S^2,\text{dc}}$. This can be understood by considering that $\delta_{S^2\text{c,shot}}$ is proportional to B_{shot} while $\delta_{S^2,\text{dc,shot}}$ is proportional to P_{shot}^2. Since P_{shot}^2 is smaller than B_{shot} by a factor of $f_p(= 0.01)$, as we can see in Eqs. (4.4) and (4.5), the effect from the non-Gaussian term $\delta_{S^2,\text{c,shot}}$ is dominant compared to the Gaussian term $\delta_{S^2,\text{dc,shot}}$. Thus, it can be concluded that the post-Born correction is primarily determined by the non-Gaussian term ($\delta_{S^2,\text{shot}} \approx \delta_{S^2,\text{c,shot}}$).

An important observation is that, in this point mass scenario ($m_p = 0.5M_\odot, f_p = 0.01, k_c = 4\times10^{13}h$ Mpc^{-1}), the post-Born term $\delta_{S^2,\text{shot}}$ surpasses the Born approximation $\langle S_{\text{Born}}^2\rangle_{\text{shot}}$ at around $f \sim 20$ Hz. This indicates the breakdown of the Born approximation around this frequency.

A similar trend can be observed for the variance of K in Fig. 10. In this case, the post-Born term $\delta_{K^2,\text{c,shot}}$ exceeds the Born result $\langle K_{\text{Born}}^2\rangle$ at around $f = 10000$ Hz. It is expected that $\delta_{K^2,\text{dc,shot}}$ is smaller than $\delta_{K^2,\text{c,shot}}$ for similar reasons as $\delta_{S^2,\text{dc,shot}}$ being smaller than $\delta_{S^2,\text{c,shot}}$ (bispectrum is much larger than the square of the power spectrum). Therefore, we can consider the correction to the variance of K to be dominated by the non-Gaussian term $\delta_{K^2,\text{shot}} \approx \delta_{K^2,\text{c,shot}}$.
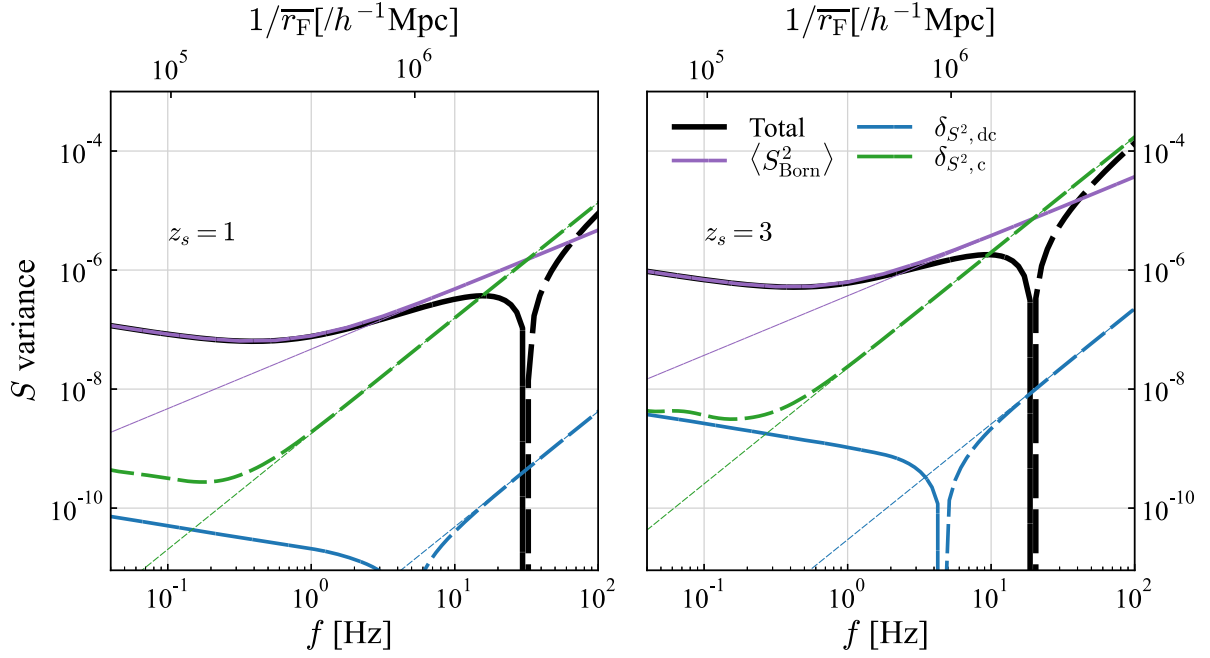
FIG. 9.   The black lines show the total $\langle S^2 \rangle$. The shot noise effect dominates $\langle S^2_{\text{Born}} \rangle$ (purple) above $f \sim 1$ Hz, while for $\delta_{S^2,\text{c}}$ (green) the shot noise takes over the halo term at $f \sim 0.3$ Hz. The shot noise from the Gaussian correction $\delta_{S^2,\text{dc}}$ (blue) is subdominant compared to $\delta_{S^2,\text{c}}$ due to $B_{\text{shot}} \gg P^2_{\text{shot}}$. At $f \sim 20$ Hz, $\delta_{S^2} (= \delta_{S^2,\text{c}} + \delta_{S^2,\text{dc}})$ exceeds $\langle S^2_{\text{Born}} \rangle$, indicating the breakdown of the Born approximation. Since δ_{S^2} is enhanced by the $\log(\cdots)$ factor which reflects the physical size of the point mass, the breakdown frequency for $\langle S^2 \rangle$ is lower than that for $\langle K^2 \rangle$ under the same point mass scenario. Note that the precise behavior of $\delta_{S^2,\text{dc}}$ and $\delta_{S^2,\text{c}}$ evaluated at the transition frequency are imprecise due to the exclusion of the cross term. The solid (dashed) line indicates the $+(-)$ value.
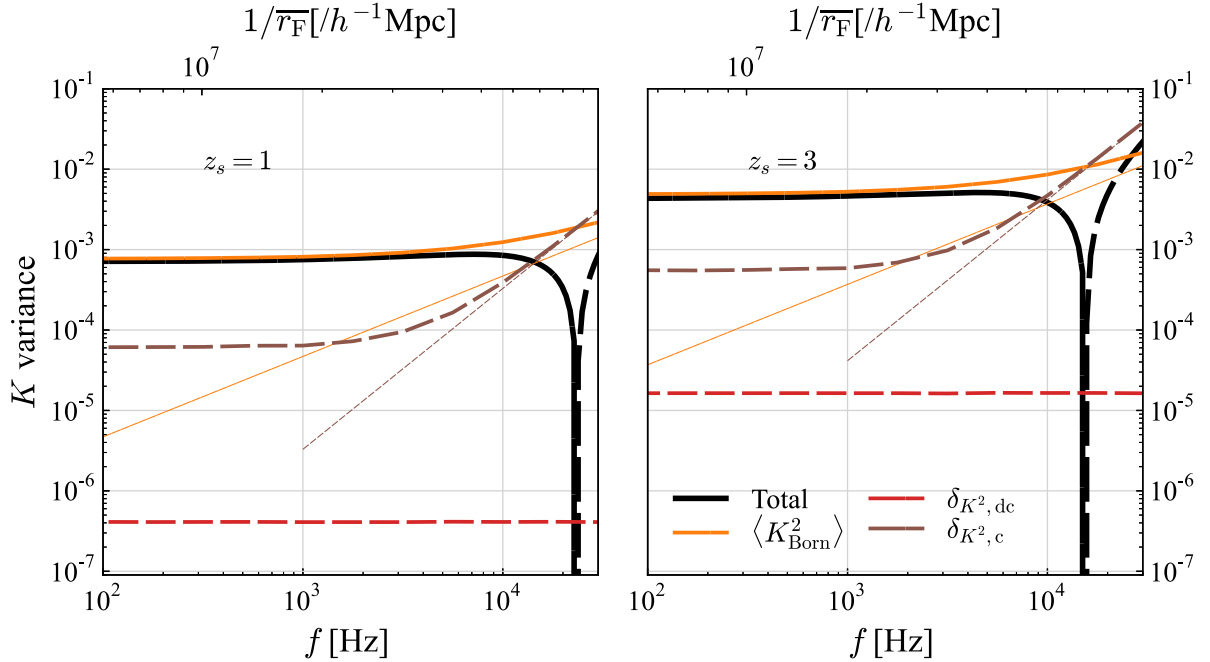

FIG. 10.   The black lines show the total $\langle K^2 \rangle$. The shot noise effect on $\langle K^2_{\text{Born}} \rangle$ (orange) is subdominant until $f \sim 10^4$ Hz, while the non-Gaussian correction $\delta_{K^2,\text{c}}$ (brown) takes over $\langle K^2_{\text{Born}} \rangle$ at a similar frequency, implying the breakdown of the Born approximation. We do not show the shot noise effect on the Gaussian correction $\delta_{K^2,\text{dc}}$ due to computational challenges but it is expected to be smaller than the non-Gaussian correction $\delta_{K^2,\text{c}}$ because $B_{\text{shot}} \gg P^2_{\text{shot}}$. For the magnification, the condition under which the Born approximation holds is given by Eq. (4.23), which ensures that the point mass does not cause strong lensing. Thus, the breakdown of the Born approximation can be attributed to the variance of K being dominated by rare strong lensing events to which the weak lens approximation cannot be applied. Note that the precise behavior of $\delta_{K^2,\text{dc}}$ and $\delta_{K^2,\text{c}}$ evaluated at the transition frequency is not accurate due to the exclusion of the cross terms. The solid (dashed) line indicates the $+(-)$ value.

Based on these considerations, it is possible to provide the general condition under which the Born approximation holds for the shot noise. This condition can be derived by computing the relative magnitude of the post-Born corrections $\delta_{S^2,\text{shot}}$, $\delta_{K^2,\text{shot}}$ to the Born approximation $\langle S^2_{\text{Born}}\rangle_{\text{shot}}$, $\langle K^2_{\text{Born}}\rangle_{\text{shot}}$. By assuming that the source redshift is not exceedingly large and truncating the second- or higher-order terms in z_s is justified, we obtain the following conditions:

$$\left|\frac{\delta_{S^2,\text{shot}}}{\langle S^2_{\text{Born}}\rangle_{\text{shot}}}\right| = c_S G m_p \omega \left(1+\frac{z_s}{2}\right)\left(\log\left[\frac{k_c^2}{H_0\omega}\right]\right)^2 < 1,$$

$$(4.22)$$

$$\left|\frac{\delta_{K^2,\text{shot}}}{\langle K^2_{\text{Born}}\rangle_{\text{shot}}}\right| = c_K G m_p \omega \left(1+\frac{z_s}{2}\right) < 1, \qquad (4.23)$$

where a factor of order unity c_S and c_K are found to be approximately $c_S = 4/\pi$ and $c_K = 3.1$ in this study. Note that the presence of the Hubble parameter H_0 in Eq. (4.22) arises from the assumption that the source and the lens redshifts are of cosmological order $\chi(z_s) \sim 1/H_0$. The general trend observed in Eqs. (4.22) and (4.23) is that the Born approximation provides a reliable estimation when $G m_p \omega$ is small. Physically, there are two ways to interpret the factor $G m_p \omega$. One is to consider this as a ratio of the Schwarzschild radius of the point mass to the wavelength of GWs, while the other views it as a square of the ratio of the Einstein radius of the point mass to the Fresnel scale of GWs. In the second interpretation, the distances to the source and lens from the observer are assumed to be the same order of magnitude. Also, Eq. (4.23) is the same as the one derived in [10] up to constant, which is based on the requirement that strong lensing by the point mass does not occur.

V. DISCUSSION

In this section, we summarize our main findings and discuss the possibility of the applications and the detectability of the post-Born effect.

A. Validity of the Born approximation

The weak lensing of gravitational waves offers the advantage of probing the scale corresponding to the Fresnel scale. In the case of typical GWs seen by the ground-based detectors ($f = 10 \sim 1000$ Hz), the Fresnel scale can reach values as small as a few parsecs. At such a small scale, a high degree of non-Gaussianity is expected. However, strong non-Gaussianity does not automatically indicate that the post-Born corrections are large. To gain a better understanding of this, let us begin by examining the general case of statics before delving into our specific cases.

Suppose $X[\delta]$ is a physical quantity such as K and S evaluated by a random variable δ (in this case, it is the matter density fluctuation δ). Here, the Born approximation $X_{\text{Born}}[\delta]$ is usually interpreted as an approximation of $X[\delta]$ by the leading-order terms of its Taylor series, thus $X[\delta] = X_{\text{Born}}[\delta] + \delta X[\delta]$, where $\delta X[\delta]$ is the post-Born correction.

For non-Gaussianity, it can be characterized by comparing the skewness of X with its variance while ensuring that they have the same dimension. In other words, the dimensionless parameter $\langle X^3\rangle^{2/3}/\langle X^2\rangle$ can be used to quantify the degree of non-Gaussianity in X. If this quantity is sufficiently small compared to 1, it suggests that X is almost Gaussian. When this is comparable to 1, it indicates a strong deviation from Gaussianity.

On the other hand, the post-Born corrections to $\langle X^2_{\text{Born}}\rangle$ are given by the higher-order terms, which in this case are $2\langle X_{\text{Born}}\delta X\rangle + \langle \delta X^2\rangle$. Since non-Gaussianity and the post-Born corrections are intrinsically different, even if X_{Born} exhibits strong non-Gaussianity ($\langle X^3_{\text{Born}}\rangle^{2/3} \sim \langle X^2_{\text{Born}}\rangle$), the post-Born corrections can be still small ($\langle X^2_{\text{Born}}\rangle \gg 2\langle X_{\text{Born}}\delta X\rangle + \langle \delta X^2\rangle$).

Now, let us turn to more specific cases, particularly those involving the presence of dark matter halos and the shot noise.

1. Without the shot noise

When considering only the dark matter halos, we found that the corrections to the Born approximation are significantly small, especially for the phase modulation. For frequencies higher than $f = 0.01$ Hz, the ratio of the correction terms to the leading-order term is $\delta_{S^2}/\langle S^2_{\text{Born}}\rangle \lesssim \mathcal{O}(10^{-2})$. Similarly, the corrections to the magnification are small, although not excessively so: $\delta_{K^2}/\langle K^2_{\text{Born}}\rangle \lesssim \mathcal{O}(10^{-1})$. Note that the frequencies below $f = 0.01$ Hz correspond to scales larger than $k \sim 10^5 h$ Mpc^{-1} and may be strongly influenced by baryonic matter, specifically in the form of galaxies [10].

Our findings indicate that, in the absence of the shot noise, the Born approximation remains still valid across the frequency ranges where the primary contributions to $\langle S^2\rangle$ and $\langle K^2\rangle$ are attributed to dark matter halos. Since the post-Born corrections account for the effect of the halos being unevenly distributed, the suppression of the post-Born corrections implies that halos can be treated as though they are uniformly distributed when computing the lensing signal.

This result does not contradict the expectation that the matter distribution is highly non-Gaussian at small scales. In fact, our analysis revealed that S and K show significant non-Gaussian behavior. Figure 11 shows the degree of non-Gaussianity in S and K. As these figures indicate, $\langle S^3_{\text{Born}}\rangle^{3/2}/\langle S^2_{\text{Born}}\rangle$ and $\langle K^3_{\text{Born}}\rangle^{3/2}/\langle K^2_{\text{Born}}\rangle$ exceed unity at high frequencies, exhibiting a strong deviation
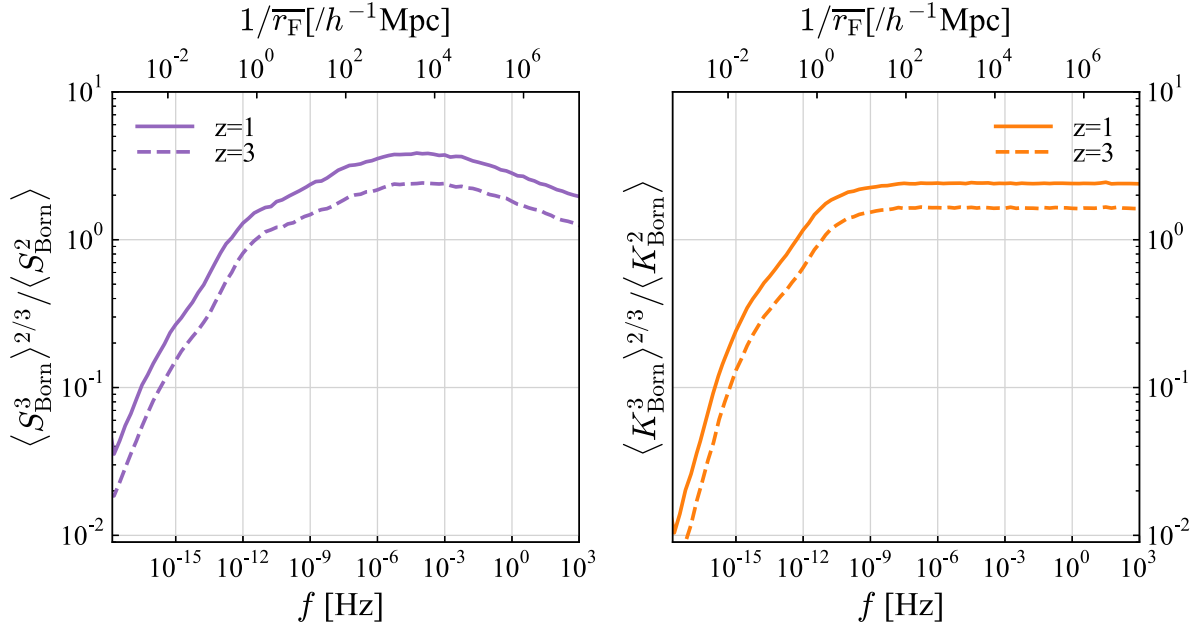
FIG. 11.   For the phase modulation (left), the degree of non-Gaussianity exhibits a rapid increase, followed by a deceleration at $f \sim 10^{-12}$ Hz, and reaching a peak at $f \sim 10^{-4}$ Hz. After that, it gradually decreases. A similar trend can be observed for the non-Gaussianity of K (right). However, $\langle K_{\mathrm{Born}}^3 \rangle^{3/2} / \langle K_{\mathrm{Born}}^2 \rangle$ reaches a constant value once the frequency exceeds $f \sim 10^{-11}$ Hz. In both cases, the degree of non-Gaussianity is large at $f > 10^{-10}$ Hz ($k \sim 10^1 h$ Mpc^{-1}), which is intuitively true. Also, it is observed that the non-Gaussianity in S and K decreases as the source redshift becomes larger. This is because S and K become more Gaussian when they are able to pass many dark matter halos on average due to the central limit theorem. This behavior is also consistent with the discussion about the non-Gaussianity of S and K found in [10].

from Gaussian behavior. Note that, in Fig. 11, it can be seen that the degree of non-Gaussianity decreases as the source of GWs moves further away (at higher redshift). This behavior is consistent with [10]. Essentially, when the source of GWs is distant, they are more likely to traverse multiple halos. When they pass through many halos, the overall lensing effect is described by the sum of each individual lensing event that occurs during their propagation. Due to the central limit theorem, this summation process leads to the reduction of non-Gaussianity.

2. With the shot noise

In the presence of the shot noise, we found that the applicability of the Born approximation depends on two factors: the satisfaction of the weak lensing condition [Eq. (4.22) for S and Eq. (4.23) for K] and the dominance of the shot noise contribution to the variance. For simplicity, let us focus on the validity of the Born approximation for the magnification. The same argument can be applied to the phase modulation as well.

For the magnification, the Born approximation remains valid if the shot noise effect is subdominant or Eq. (4.23) is satisfied. However, if Eq. (4.23) is not satisfied and the shot noise is the dominant contribution to the variance, the Born approximation breaks down. In this context, the breakdown of the Born approximation specifically refers to the

situation where the variance of S and K computed using the leading-order terms of S and K in Φ no longer provide a reliable estimate of the true variance. This also implies that the perturbative approach fails since adding up a finite number of higher-order terms does not necessarily improve the accuracy of variance estimation. At this stage, a full-order analysis or simulation is required to obtain the true distribution of S and K. The simulation approach has been taken in geometric optics [27], but further investigation is needed within the framework of wave optics.

A similar condition to Eq. (4.23) is also derived in [10] (the only difference is a unity-order prefactor) under the requirement for the absence of strong lensing. Thus, the violation of Eq. (4.23) implies the existence of a specific configuration of lenses and a GW source that can lead to strong lensing such as the point mass being very close to the line of sight. It is important to note that the violation of Eq. (4.23) does not by itself imply the breakdown of the use of the Born approximation for the variance. As mentioned above, the breakdown requires not only the violation of Eq. (4.23) but also the dominance of the shot noise effects on the variance over other effects. Therefore, even if there are objects that can potentially cause strong lensing, the Born approximation is still valid as long as the contribution of these lenses to the variance is subdominant compared to the contribution of other lenses that do not cause strong lensing (such as dark low-mass halos).

Due to this property, the shot noise signal can help constrain the nature of point masses. For example, if Eq. (4.23) is satisfied and the shot noise effect dominates the variance (this is true for the phase modulation when the parameters for the shot noise is $f_p = 0.01$, $m = 0.5 M_\odot$ and the frequency of GWs is around $f = 1$ Hz), it is possible to estimate the parameters of the point masses such as m and f_p [10]. In addition to this, our analysis provides a method to include the correction to the Born approximation. By using the post-Born terms calculated in this study, the shot noise contribution to the variance is modified as

$$\langle S^2 \rangle_{\text{shot}} = \langle S^2_{\text{Born}} \rangle_{\text{shot}}$$
$$\times \left\{ 1 - \frac{4}{\pi} G m_p \omega \left(1 + \frac{z_s}{2}\right) \left(\log\left[\frac{k_c^2}{H_0 \omega}\right]\right)^2 \right\}, \tag{5.1}$$

$$\langle K^2 \rangle_{\text{shot}} = \langle K^2_{\text{Born}} \rangle_{\text{shot}} \left\{ 1 - 3.1 G m_p \omega \left(1 + \frac{z_s}{2}\right) \right\}. \tag{5.2}$$

In the moderately high-frequency region, where the perturbative approach is still useful but the accuracy of the Born approximation is uncertain, this modification will enable us to more accurately estimate the variance of S and K produced by the point masses with specific parameters m_p, k_c, and f_p.

On the other hand, the scarcity of strongly lensed signals can place constraints on the abundance of lens objects capable of causing strong lensing. For example, if we consider a scenario where $m = 50 M_\odot$ instead of $m = 0.5 M_\odot$ while maintaining $f_p = 0.01$, Eq. (4.23) indicates that the frequency at which the breakdown of the Born approximation for the magnification shifts is from $f \sim 10000$ to $f \sim 100$ Hz. This scenario ($m = 50 M_\odot, f_p = 0.01$) corresponds to the universe in which the 50 solar mass black holes as part of dark matter are as prevalent as the stellar components.

As this frequency range falls within the sensitivity of current ground-based detectors, there is a possibility of detecting the strong lensing signal caused by such black holes. If the number of strong lensing events is small enough so that their impact on the variance is subdominant, the abundance of such black holes can be constrained by this information.

Such a scenario ($m = 50 M_\odot, f_p = 0.01$) has been attracting great interest recently after the observations of such massive black holes by the GW experiments. It is under active investigation whether the abundance of primordial black holes comparable to $f_p \simeq 0.01$ is consistent with the existing observations [28]. GL of GWs studied in this paper provides an alternative path to test this possibility (see also [10]).

B. Average as an additional probe

We found that the ensemble average of K and S is no longer zero at the level of the post-Born approximation. This provides the possibility to detect the average of S and K and use them as an additional probe for matter abundance. However, it is crucial to assess the validity of the approximation of the average by only considering the power spectrum term. Therefore, we will now examine the reliability of this approximation.

In the absence of the shot noise, the main contribution to the average of S and K comes from the power spectrum and the contribution from the higher-order terms containing the bispectrum is subdominant. This is consistent with the behavior observed in the variance, where the correction terms to the Born approximation are found to be subdominant. In this case, the average of S and K is roughly of the same order as their variance ($\langle S \rangle \sim \langle S^2 \rangle$, $\langle K \rangle \sim -\langle K^2 \rangle$).

In the presence of the shot noise, we found that there are cases where the computation of the average of K by accounting only for the matter power spectrum breaks down, which is the same condition as the breakdown of the Born approximation for $\langle K^2 \rangle$ up to a constant prefactor. Therefore, if the Born approximation for the variance of K is valid, then the computation of the average of K by accounting only for the power spectrum contribution remains valid.

Regarding the phase modulation, the contribution from the bispectrum term due to the shot noise becomes exactly zero $\langle S^{(3)} \rangle = 0$. This suggests that including higher-order terms such as trispectrum would be necessary to capture the corrections to $\langle S \rangle$ in the presence of the shot noise. However, as long as the Born approximation for $\langle S^2 \rangle$ holds, it is expected that the approximation of $\langle S \rangle$ using the power spectrum contribution alone is valid. This presumption is reasonable because if the Born approximation for the variance holds, it implies that the lensing signal is weak and the first term in the perturbative approach offers a reliable approximation.

Based on these considerations, the average calculation is valid as long as the Born approximation for the variance also holds. Now, let us shift our focus to the average of S, as it can play a significant role in probing the properties of the point masses. Specifically, by combining $\langle S \rangle$ and $\langle S^2 \rangle$, we can probe the size of the shot noise constituent, as well as its mass and abundance, since the shot noise has different effects on $\langle S \rangle$ and $\langle S^2 \rangle$. This analysis cannot be performed by considering the variance alone because the size dependency does not appear in the variance. The obtained properties of the shot noise can be compared with the properties of stars inferred by other astronomical observations. This provides the test of whether the sources causing the shot noise in the gravitational lensing of GWs are stars or other types of compact objects that have not been detected by non-GW observations.

It is also important to mention that, according to our formulation, $\langle S \rangle$ is always positive. If the negative value of $\langle S \rangle$ is detected, it means an indication of the presence of something outside the lensing effect. It could mean the presence of new matter that interacts with gravity in an unusual way or the violation of general relativity, which could lead to new physics.

C. Detectability of the post-Born effect

In [29], it is suggested that the amplitude and phase fluctuation of GWs can be measured with an accuracy of 1/SNR, where SNR is the signal-to-noise ratio. According to [10], the accuracy of measurement is improved by combining many gravitational events. In [10], it is argued that the required accuracy is written as $\sim (2/N_{\text{event}})^{1/4}(1/\text{SNR})$, which yields $N_{\text{event}} \sim 3 \times 10^5$ as the number of GW events (with SNR = 50) required for detecting the lensing signal. We would like to perform a similar estimation of N_{event} required to detect the post-Born effect.

First, let us consider the number of GW events required for detecting the average of S and K. Since our purpose is to estimate N_{event} by the back-of-the-envelope calculations, we will consider the following toy model which simplifies the situation without losing the essential point. Suppose we have succeeded in inferring the source parameters and hence the unlensed waveform from the GW measurement. Then, the residual signal, which we denote by s, remains after subtracting the unlensed waveform from the measured waveform and consists of the uncertainties n of the unlensed waveform and the lensing signal X, namely,

$$s_{X,i} = n_i + X_i, \qquad (5.3)$$

where i labels the GW events, while X takes either S or K. For simplicity, we assume that both n_i and X_i are Gaussian random variables and each GW event is independent of the others. In this case, the ensemble average of the quantities computed from n_i and X_i is given by

$$\langle n_i n_j \rangle = \left(\frac{1}{\text{SNR}}\right)^2 \delta_{ij}, \qquad (5.4)$$

$$\langle X_i \rangle = \mu_X, \qquad (5.5)$$

$$\langle X_i X_j \rangle = \sigma_X^2 \delta_{ij} + \mu_X^2, \qquad (5.6)$$

$$\langle n_i X_j \rangle = 0. \qquad (5.7)$$

Here, μ_X and σ_X are the values of both the average and the standard deviation of the phase modulation and the magnification. All the other quantities can be computed from the combination of these relations. The first relation $\langle n_i n_j \rangle = \delta_{ij}/\text{SNR}^2$ is about the accuracy of detecting the

phase and magnification fluctuation mainly discussed in [29].

In reality, we are only able to detect a finite number of GW events. Thus, it is convenient to introduce the estimator of the average μ_X defined as

$$\mathcal{E}_\mu = \frac{1}{N_{\text{event}}} \sum_{i=1}^{N_{\text{event}}} s_{X,i}. \qquad (5.8)$$

This quantity is an approximated version of the ensemble average, thus taking $N_{\text{event}} \to \infty$ reproduces μ_X. Indeed, computing the average and the variance of $\mathcal{E}_\mu$, we obtain

$$\langle \mathcal{E} \rangle_\mu = \mu_X, \qquad (5.9)$$

$$\langle \mathcal{E}_\mu^2 \rangle - \langle \mathcal{E}_\mu \rangle^2 \sim \left(\frac{1}{\text{SNR}}\right)^2 \frac{1}{N_{\text{event}}}. \qquad (5.10)$$

It is important to mention that we have used the assumption $1/\text{SNR} \gg \mu_X, \sigma_X$ to derive the second equation. This result shows that $\mathcal{E}_\mu$ fluctuates around μ_X with a fluctuation width of about $\left(\frac{1}{\text{SNR}}\right)\frac{1}{\sqrt{N_{\text{event}}}}$. In order to confidently conclude that the average is nonzero, $\mu_X > \left(\frac{1}{\text{SNR}}\right)\frac{1}{\sqrt{N_{\text{event}}}}$ needs to be satisfied. Using this restriction, we can estimate that $N_{\text{event},\mu_X} \sim \left(\frac{1}{\text{SNR}}\right)^2 \frac{1}{\mu_X^2}$ is at least necessary to detect the average of K and S.

Next, we consider the number of events for detecting the variance. In this case, we need at least two independent measurements of the residual for the same GW event if it is difficult to distinguish the lensing signal from the uncertainty associated with the unlensed waveform by using one measurement alone. In the following, we assume measurements by two detectors. For this purpose, we denote the signals from two different measurements (1 and 2) to be $s_{X,1,i} = n_{1,i} + X_i$, $s_{X,2,i} = n_{2,i} + X_i$ and assume that one measurement noise is independent of the other's $\langle n_{1,i} n_{2,j} \rangle = 0$. The detectability is calculated in the same way above by introducing the estimator of the variance,

$$\mathcal{E}_{\sigma_X^2} = \frac{1}{N_{\text{event}}} \sum_{i=1}^{N_{\text{event}}} s_{X,1,i} s_{X,2,i} - \frac{1}{N_{\text{event}}^2} \sum_{i=1}^{N_{\text{event}}} s_{X,1,i} \sum_{j=1}^{N_{\text{event}}} s_{X,2,j}. \qquad (5.11)$$

From this expression, we obtain the expressions of the ensemble average of $\mathcal{E}_{\sigma_X^2}$,

$$\langle \mathcal{E}_{\sigma_X^2} \rangle = \sigma_X^2, \qquad (5.12)$$

$$\langle \mathcal{E}_{\sigma_X^2}^2 \rangle - \langle \mathcal{E}_{\sigma_X^2} \rangle^2 \sim \left(\frac{1}{\text{SNR}}\right)^4 \frac{1}{N_{\text{event}}}. \qquad (5.13)$$

The interpretation of this result is exactly the same as $\mathcal{E}_{\mu_X}$ that $\mathcal{E}_{\sigma_X^2}$ fluctuates around σ_X^2 with a width of about

$(\frac{1}{\mathrm{SNR}})^2 \frac{1}{\sqrt{N_{\mathrm{event}}}}$. Therefore, the number of gravitational wave events required to detect the variance is given by $N_{\mathrm{event},\sigma_X^2} \sim (\frac{1}{\mathrm{SNR}})^4 \frac{1}{\sigma_X^4}$.

Now, let us examine the detectability of the phase modulation and the magnification. Table I presents the order of magnitude for the average and variance of S and K, along with the post-Born corrections to the variance. We consider the scenario where the source redshift is $z_s = 3$, and the signal-to-noise ratio is SNR $= 50$. The shot noise effect we include corresponds to lensing by point masses with $m = 0.5 M_\odot$, $f_p = 0.01$, and $k_c = 4 \times 10^{13} h$ Mpc^{-1}.

For the phase modulation, we focus on the frequency range of $f = 0.01$–10 Hz, which falls within the range where the Born approximation is valid. Within this frequency range, the lensing signal is dominated by both the dark low-mass halos and the point masses, with specific dominance depending on the frequency. At the lower end of this range ($f = 0.01$ Hz), the signal is primarily attributed to the dark matter halos. However, as the frequency of GWs increases, the shot noise effect becomes more significant, taking over the halo contribution at around $f = 1$ Hz for the Born variance and $f = 0.3$ Hz for the average.

In this frequency range, the variance remains relatively stable, while the average increases moderately. The typical order of the average is around $\mathcal{O}(10^{-5})$, but at the higher end of this range $f = 10$ Hz, it can be enhanced by up to $\mathcal{O}(10^{-4})$. On the other hand, the order of the Born variance remains $\mathcal{O}(10^{-6})$, even at the higher end of the range. This difference arises from the dependency of the average on the size of the point mass. Using the formalism we developed above, the number of GW events required to detect $\langle S \rangle$ is estimated to be $\mathcal{O}(10^6)$ in the middle of this frequency range. However, at the higher-frequency end, the required number reduces to $\mathcal{O}(10^4)$. On the other hand, the number of events

required to detect $\langle S^2 \rangle$ is $\mathcal{O}(10^5)$ in the middle-frequency range and $\mathcal{O}(10^4)$ at the high end. As a result, the detection cost for the average is comparable to the detection cost for the variance at high frequencies in which the shot noise dominates. In a slightly different scenario, the detection of the average might be easier than the detection of the variance. For instance, if the point masses we considered here are not ordinary stars but black holes with the same mass and mass fraction (thus, k_c becomes much bigger), the required number for detecting the average decreases while the number for the variance remains the same.

Note that, if the signal-to-noise ratio is much larger than SNR $= 50$, the number of GW events required to detect the variance becomes significantly smaller compared to the number required for the detection of the average. This is because the number of required events for the variance scales as $1/\mathrm{SNR}^2$, while the number for the average scales as $1/\mathrm{SNR}^2$. Hence, the situation where the average might be easier to detect is when the SNR is not excessively high.

In the case of the post-Born corrections to the variance, their relative magnitude compared to the Born variance prior to the onset of the shot noise is $\mathcal{O}(10^{-3})$. However, once the shot noise effect becomes dominant, their relative magnitude is described by $f/20$ Hz. In this case, the ratio of the number of GW events required to resolve this correction to the number required to detect the Born variance scales as $(20\ \mathrm{Hz}/f)^2$. This means that, even if the corrections to the Born variance are 10%, resolving it would require 100 times more GW events than those needed to detect the Born variance. If the SNR is 100, which is expected to be achieved in the future [30], the number of events to resolve the Born variance reduces to $\mathcal{O}(10^{-3})$. Assuming a total of 10^5 GW events are observed, it would be possible to resolve the post-Born corrections that exceed 1% of the Born variance. This corresponds to frequencies around $f \sim 2$ Hz, which is already close to the breakdown frequency ($f \sim 20$ Hz).

These considerations indicate that the post-Born corrections are challenging to detect except in the vicinity of the breakdown frequency. In the frequency range where the perturbative approach holds but the accuracy of the Born approximation becomes less reliable, including the post-Born correction [Eqs. (5.1) and (5.2)] can yield a more accurate estimate of the variance caused by the point mass lens.

Next, let us consider the magnification. In Table I, we consider the frequency range of $f = 0.01$–1000 Hz. This frequency range is chosen based on the validity of the Born approximation, which holds until Eq. (4.23) breaks down, which occurs at around $f = 10000$ Hz. As shown above, the magnification has a broader frequency range within which the Born approximation is valid compared to the phase modulation.

The order of magnitude of the magnification is much larger than that of the phase, making SNR $= 50$ sufficient

TABLE I. In the scenario where the shot noise consists of point masses with $m_p = M_\odot$, $f_p = 0.01$, and $k_c = 4 \times 10^{13} h$ Mpc^{-1}, with $z_s = 3$ and SNR $= 50$. In this case, the number of GW events required to detect $\langle S^2 \rangle$ and $\langle S \rangle$ can be comparable to $\mathcal{O}(10^4)$ at $f \sim 10$ Hz due to the enhancement of $\langle S \rangle$ by its dependence on the physical size of the point mass. On the other hand, the magnification can be much more easily observed. However, a similar number of GW evens is expected to be required to extract the wave-dependent part from K.

	$10^{-2} \leq f \leq 10^1$ Hz	$N_{\mathrm{events}}(\mathrm{SNR} = 50)$
$\langle S \rangle$	$\mathcal{O}(10^{-5}\text{–}10^{-4})$	$\mathcal{O}(10^6 \sim 10^4)$
$\langle S_{\mathrm{Born}}^2 \rangle$	$\mathcal{O}(10^{-6})$	$\mathcal{O}(10^5 \sim 10^4)$
$\delta_{S^2}/\langle S_{\mathrm{Born}}^2 \rangle$	$\sim \frac{f}{20\ \mathrm{Hz}}$	$\gtrsim \mathcal{O}(10^5) \times (\frac{20\ \mathrm{Hz}}{f})^2$
	$10^{-2} \leq f \leq 10^3$ Hz	$N_{\mathrm{events}}(\mathrm{SNR} = 50)$
$\langle K \rangle$	-5×10^{-3}	$\mathcal{O}(10^2)$
$\langle K_{\mathrm{Born}}^2 \rangle$	5×10^{-3}	$\mathcal{O}(1)$
$\delta_{K^2}/\langle K_{\mathrm{Born}}^2 \rangle$	5×10^{-2}	$\mathcal{O}(10)$

to resolve $\langle K^2 \rangle$, while around 100 GW events are required to resolve $\langle K \rangle$. Even the post-Born correction to the variance can be resolved with just $\mathcal{O}(10)$ GW events.

However, there is an important consideration to make. The magnification approaches a constant value as the frequency increases, representing the geometric optics limit. Since the geometric optics limit lacks frequency dependence, it cannot be used to probe the matter abundance at the Fresnel scale. In order to extract the pure wave effect, which can be used to probe the scale corresponding to the Fresnel scale, the constant term in the magnification needs to be subtracted. However, as calculated in [10], this pure frequency-dependent part is of the same order as the phase. This can be also understood by considering the consistency relation for the variance of the phase modulation and the magnification, namely, $\langle K_{\mathrm{Born}}^2(2f) \rangle - \langle K_{\mathrm{Born}}^2(f) \rangle = \langle S_{\mathrm{Born}}^2(f) \rangle$ [19]. Therefore, the magnification needs to be determined at the same level of accuracy as the phase to extract the wave-dependent part that is superimposed on the constant part. Consequently, a similar number of GW events is required to make the magnification as useful as the phase in extracting information about the matter abundance at the Fresnel scale.

VI. CONCLUSION

In this paper, we have investigated the weak lensing of GWs beyond the Born approximation by including the higher-order terms in the gravitational potential Φ. To do this, we adopted a new formulation for the equation governing the GL of GWs. Instead of using the amplification factor F defined as the ratio of the lensed to unlensed waveform, we introduced a new variable J defined as $F = e^{i\omega J}$. This process allows us to partially include the nonlinear effect of Φ and reduces the complexity of calculating the higher-order terms. We then derived the expression of the phase modulation S and the magnification K up to third order in Φ and calculated the post-Born corrections to the average and variance. In computing the post-Born corrections, we considered both Gaussian (product of the bispectrum) and non-Gaussian (bispectrum) terms up to the lowest nontrivial order in Φ. To evaluate the validity of the Born approximation, we numerically computed $\langle S \rangle, \langle K \rangle, \delta_{K^2}, \delta_{S^2}$ by using the matter power spectrum and bispectrum obtained by the phenomenological halo model including subhalos.

We found that, at the level of the post-Born approximation, $\langle S \rangle$ and $\langle K \rangle$ are no longer zero. We also confirmed, by computing the contribution to $\langle S \rangle$ and $\langle K \rangle$ from the bispectrum terms, that evaluating $\langle S \rangle$ and $\langle K \rangle$ by solely using the power spectrum still provides a reliable estimation. While $\langle S \rangle$ and $\langle K \rangle$ typically have the same order as $\langle S^2 \rangle$ and $\langle K^2 \rangle$, the presence of the point masses (shot noise) can particularly enhance $\langle S \rangle$, due to the dependency of $\langle S \rangle$ on the physical size of the point masses. We then estimate the number of GW events required to observe $\langle S \rangle$ and $\langle K \rangle$

and found that, while detecting the average generally requires a larger number of events than the variance, the number required to observe $\langle S \rangle$ can be of the order of $\mathcal{O}(10^4)$ at $f \sim 10$ Hz with SNR $= 50$. This number is comparable to, or potentially even smaller than, the number required to detect $\langle S^2 \rangle$, depending on the nature of the point masses.

As for the post-Born corrections to the variance, we found that their primary contribution comes from uneven distributions of the target halos with the corresponding Fresnel scale. Our findings show that the corrections to $\langle S^2 \rangle$ in the absence of the shot noise are 2 orders of magnitude smaller than the Born approximation at $f > 0.01$ Hz and $z_s \leq 3$. This also indicates that the halos can be treated as if they are uniformly distributed when computing $\langle S^2 \rangle$ and $\langle K^2 \rangle$. In addition, the post-Born corrections do not pose relevant issues in the absence of the shot noise unless SNR for GWs is excessively high.

In the presence of the shot noise, we determined the conditions under which the Born approximation fails. The validity of the Born approximation is guaranteed when the point mass does not dominate $\langle S^2 \rangle$ and $\langle K^2 \rangle$ or when strong lensing by the point mass does not occur. However, when these conditions are violated simultaneously, the variance is predominantly determined by rare strong lensing events, and the Born approximation no longer predicts the true variance. Furthermore, the breakdown frequency for $\langle S^2 \rangle$ is lower compared to the one for $\langle K^2 \rangle$ due to the enhancing factor pertaining to the physical size of the point masses. Since the breakdown frequency may fall within the sensitivity range of current detectors in certain scenarios (such as $f \sim 20$ Hz for $\langle S^2 \rangle$ with $m_p = 0.5 M_\odot, f_p = 0.01$, and $k_c = 4 \times 10^{13} h$ Mpc^{-1}), careful analysis of the lensing signal is required. For example, when the frequency of GWs approaches the breakdown frequency from below and the accuracy of the Born approximation becomes less trustable, the modification to the Born approximation given in Eqs. (5.1) and (5.2) can be used to provide a more accurate estimation of $\langle S^2 \rangle$ and $\langle K^2 \rangle$. When the frequency is above the breakdown frequency, a perturbative approach fails to provide a reliable result. Thus, in this case, a separate study involving a full-order analysis is needed to effectively constrain the property of the point masses.

ACKNOWLEDGMENTS

APPENDIX A: GEOMETRIC OPTICS LIMIT

In [12–16], the post-Born approximation is discussed under geometric optics. Although geometric optics has

been widely used in the gravitational lensing, a fundamentally more accurate description for the GL of GWs is wave optics. In this sense, wave optics should be able to encompass everything that could be derived in geometric optics.

In geometric optics, we take the large frequency limit ($\omega \to \infty$) from the outset and start from the geodesic equation which does not contain ω. In this appendix, we demonstrate explicitly that the magnification in the high-frequency limit under the post-Born approximation in wave optics coincides with the one derived based on geometric optics. In order to calculate the magnification under geometric optics, we need the convergence κ and shear γ up to second order and first order in Φ, respectively. According to [12–16], they are given by

$$\kappa^{(1)}(\boldsymbol{\theta}_0, \chi_s) = \int_0^{\chi_s} d\chi \chi^2 W(\chi, \chi_s) \Phi_{ii}(\chi), \tag{A1}$$

$$\kappa^{(2)}(\boldsymbol{\theta}_0, \chi_s) = -2 \int_0^{\chi_s} d\chi \int_0^{\chi} d\chi' \chi^2 \chi'^2 W(\chi, \chi_s) W(\chi', \chi) \Phi_{ij}(\chi) \Phi_{ij}(\chi')$$
$$- 2 \int_0^{\chi_s} d\chi \int_0^{\chi} d\chi' \chi^3 \chi' W(\chi, \chi_s) W(\chi', \chi) \Phi_{iik}(\chi) \Phi_k(\chi'), \tag{A2}$$

$$\gamma_1^{(1)}(\boldsymbol{\theta}_0, \chi_s) = \int_0^{\chi_s} d\chi \chi^2 W(\chi, \chi_s) (\Phi_{11}(\chi) - \Phi_{22}(\chi)), \tag{A3}$$

$$\gamma_2^{(1)}(\boldsymbol{\theta}_0, \chi_s) = 2 \int_0^{\chi_s} d\chi \chi^2 W(\chi, \chi_s) \Phi_{12}(\chi). \tag{A4}$$

The gravitational potential is evaluated at the straight line along which the unlensed ray propagates, namely, $\Phi(\chi) = \Phi(\boldsymbol{\theta}_0, \chi)$. The magnification $\mu_{\text{geo}}(\boldsymbol{\theta}_0, \chi_s)$ is the inverse of the determinant of the Jacobian matrix $A(\boldsymbol{\theta}_0, \chi) = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2-\Omega \\ -\gamma_2+\Omega & 1-\kappa+\gamma_1 \end{pmatrix}$ and, up to second order in Φ, μ_{geo} is given by

$$\mu_{\text{geo}}(\boldsymbol{\theta}_0, \chi_s) = 1 + 2\kappa^{(1)} + 2\kappa^{(2)} + 3(\kappa^{(1)})^2 + (\gamma_1^{(1)})^2 + (\gamma_2^{(1)})^2$$
$$= 1 + 2\kappa^{(1)}(\boldsymbol{\theta}_0, \chi_s) + 2(\kappa^{(1)}(\boldsymbol{\theta}_0, \chi_s))^2 - 4 \int_0^{\chi_s} d\chi \int_0^{\chi} d\chi' \chi^3 \chi' W(\chi, \chi_s) W(\chi', \chi) \Phi_{iik}(\chi) \Phi_k(\chi')$$
$$+ 4 \int_0^{\chi_s} d\chi \int_0^{\chi} d\chi' \chi^2 \chi'^2 W(\chi, \chi_s)^2 \Phi_{ij}(\chi) \Phi_{ij}(\chi'). \tag{A5}$$

Up to this order, Ω does not appear in the magnification as Ω itself is already second order in Φ.

We now show the magnification computed in wave optics based on the formulation given in this paper reduces to Eq. (A5) in the high-frequency limit. Prior to that, we write the approximated solution of the lens equation up to first order in Φ,

$$\delta\boldsymbol{\Theta}(\boldsymbol{\theta}_0, \chi) = -2 \int_0^{\chi} d\chi' W(\chi', \chi) \nabla_\theta \Phi(\boldsymbol{\theta}_0, \chi'). \tag{A6}$$

In wave optics, the magnification effect is encoded in K as $\mu_{\text{wave}}(\boldsymbol{\theta}, \omega) = e^{2K}$, where $\boldsymbol{\theta}$ is the position of the source on the source plane $\boldsymbol{\theta} = \boldsymbol{\theta}_0 + \delta\boldsymbol{\Theta}(\boldsymbol{\theta}_0, \chi_s)$. Taking $\omega \to \infty$ of Eqs. (2.20) and (2.21) yields

$$K^{(1)}(\boldsymbol{\theta}, \chi_s, \omega \to \infty) = -2 \int_0^{\chi_s} d\chi \chi^2 W(\chi, \chi_s) \Phi_{ii}(\boldsymbol{\theta}, \chi)$$
$$= \kappa^{(1)}(\boldsymbol{\theta}, \chi_s), \tag{A7}$$

$$
\begin{aligned}
K^{(2)}(\boldsymbol{\theta},\chi_s,\omega\to\infty) &= \int_0^{\chi_s}\frac{d\chi}{\chi^2}\int_0^{\chi}d\chi_1\int_0^{\chi}d\chi_2[W(\chi,\chi_s)\nabla_{\theta12}^2 + W(\chi_1,\chi)\nabla_{\theta1}^2 + W(\chi_2,\chi)\nabla_{\theta2}^2]\nabla_{\theta1}\Phi_1\cdot\nabla_{\theta2}\Phi_2 \\
&= -\int_0^{\chi_s}d\chi_1 W(\chi_1,\chi_s)\nabla_\theta(\nabla_\theta^2\Phi(\chi_1))\cdot(-2)\int_0^{\chi_s}d\chi_2 W(\chi_2,\chi_s)\nabla_\theta\Phi(\chi_2) \\
&\quad -2\int_0^{\chi_s}d\chi\int_0^{\chi}d\chi'\chi^3\chi' W(\chi,\chi_s)W(\chi',\chi)\Phi_{iik}(\chi)\Phi_k(\chi') + 2\int_0^{\chi_s}d\chi\int_0^{\chi}d\chi'\chi^2\chi'^2 W(\chi,\chi_s)^2\Phi_{ij}(\chi)\Phi_{ij}(\chi') \\
&= -\nabla_\theta K^{(1)}\cdot\delta\boldsymbol{\Theta} - 2\int_0^{\chi_s}d\chi\int_0^{\chi}d\chi'\chi^3\chi' W(\chi,\chi_s)W(\chi',\chi)\Phi_{iik}(\chi)\Phi_k(\chi') \\
&\quad + 2\int_0^{\chi_s}d\chi\int_0^{\chi}d\chi'\chi^2\chi'^2 W(\chi,\chi_s)^2\Phi_{ij}(\chi)\Phi_{ij}(\chi').
\end{aligned} \tag{A8}
$$

The magnification in wave optics up to second order is then given by

$$
\begin{aligned}
\mu_{\mathrm{wave}}(\boldsymbol{\theta},\chi_s,\omega\to\infty) &= 1 + 2K^{(1)} + 2(K^{(1)})^2 + 2K^{(2)} \\
&= 1 + 2\kappa^{(1)}(\boldsymbol{\theta}-\delta\boldsymbol{\Theta},\chi_s) + 2(\kappa^{(1)}(\boldsymbol{\theta}_0,\chi_s))^2 - 4\int_0^{\chi_s}d\chi\int_0^{\chi}d\chi'\chi^3\chi' W(\chi,\chi_s)W(\chi',\chi)\Phi_{iik}(\chi)\Phi_k(\chi') \\
&\quad + 4\int_0^{\chi_s}d\chi\int_0^{\chi}d\chi'\chi^2\chi'^2 W(\chi,\chi_s)^2\Phi_{ij}(\chi)\Phi_{ij}(\chi') \\
&= \mu_{\mathrm{geo}}(\boldsymbol{\theta}_0,\chi_s).
\end{aligned} \tag{A9}
$$

Therefore, the result of geometric optics is indeed derived by taking the high-frequency limit of wave optics. It is important to mention again that the lens plane $\boldsymbol{\theta}_0$ is used in geometric optics whereas, in wave optics, the source plane $\boldsymbol{\theta}$ is the fundamental variable. This difference manifests itself in the argument of both μ_{geo} and μ_{wave}. We have shown that, at least up to second order in Φ, our formulation reduces to the well-known result in geometric optics. This consistency strongly supports the validity of

the discussion about the post-Born approximation of the lensing in wave optics.

APPENDIX B: POST-BORN VARIANCE OF S AND K

The correction of the variance of K to the Born approximation is described by Eq. (2.38), and a similar relation holds for S. This equation is rewritten by using the matter power spectrum given in Eq. (3.1),

$$
\begin{aligned}
\delta_{X^2,\mathrm{dc}} &= 16\left(\frac{3H_0^2\Omega_m}{2}\right)^4\int_0^{\chi_s}\frac{d\chi}{\chi^2}\int_0^{\chi}\frac{d\chi'}{\chi'^2}\int_0^{\chi'}d\chi_1\int_0^{\chi'}d\chi_2\frac{1}{a^2(\chi_1)}\frac{1}{a^2(\chi_2)}\frac{1}{(2\pi)^2} \\
&\quad \times \omega^2\int_0^\infty dk_1\int_0^\infty dk_2 P_\delta(k_1,\chi_1)P_\delta(k_2,\chi_2)\left[\frac{1}{k_1 k_2}\mathcal{F}_{12} - \frac{1}{k_1^2}\mathcal{F}_1 - \frac{1}{k_2^2}\mathcal{F}_2\right],
\end{aligned} \tag{B1}
$$

where

$$
\begin{aligned}
\mathcal{F}_{12} &= \chi_1^2\chi_2^2\int_0^{2\pi}\frac{d\phi}{2\pi}\cos^2\phi\Bigg\{ F\left(\frac{\chi_1^2 W(\chi_1,\chi_s)}{2\omega}k_1^2 + \frac{\chi_2^2 W(\chi_2,\chi_s)}{2\omega}k_2^2 + \frac{\chi_1\chi_2 W(\chi,\chi_s)}{\omega}k_1 k_2\cos\phi\right) \\
&\quad \times F\left(\frac{\chi_1^2 W(\chi_1,\chi_s)}{2\omega}k_1^2 + \frac{\chi_2^2 W(\chi_2,\chi_s)}{2\omega}k_2^2 + \frac{\chi_1\chi_2 W(\chi',\chi_s)}{\omega}k_1 k_2\cos\phi\right) \\
&\quad - F\left(\frac{\chi_1^2 W(\chi_1,\chi_s)}{2\omega}k_1^2\right)F\left(\frac{\chi_1^2 W(\chi_1,\chi_s)}{2\omega}k_1^2 + \frac{\chi_2^2 W(\chi_2,\chi)}{\omega}k_2^2 + \frac{\chi_1\chi_2 W(\chi',\chi)}{\omega}k_1 k_2\cos\phi\right) \\
&\quad - F\left(\frac{\chi_2^2 W(\chi_2,\chi_s)}{2\omega}k_2^2\right)F\left(\frac{\chi_2^2 W(\chi_2,\chi_s)}{2\omega}k_2^2 + \frac{\chi_1^2 W(\chi_1,\chi)}{\omega}k_1^2 + \frac{\chi_1\chi_2 W(\chi',\chi)}{\omega}k_1 k_2\cos\phi\right)\Bigg\},
\end{aligned} \tag{B2}
$$

$$\mathcal{F}_1 = \chi_1\chi_2^3 \int_0^{2\pi} \frac{d\phi}{2\pi} \cos\phi F\left(\frac{\chi_1^2 W(\chi_1,\chi_s)}{2\omega}k_1^2\right) F\left(\frac{\chi_1^2 W(\chi_1,\chi_s)}{2\omega}k_1^2 + \frac{\chi_2^2 W(\chi_2,\chi)}{\omega}k_2^2 + \frac{\chi_1\chi_2 W(\chi',\chi)}{\omega}k_1 k_2 \cos\phi\right), \quad \text{(B3)}$$

$$\mathcal{F}_2 = \chi_2\chi_1^3 \int_0^{2\pi} \frac{d\phi}{2\pi} \cos\phi F\left(\frac{\chi_2^2 W(\chi_2,\chi_s)}{2\omega}k_2^2\right) F\left(\frac{\chi_2^2 W(\chi_2,\chi_s)}{2\omega}k_2^2 + \frac{\chi_1^2 W(\chi_1,\chi)}{\omega}k_1^2 + \frac{\chi_1\chi_2 W(\chi',\chi)}{\omega}k_1 k_2 \cos\phi\right). \quad \text{(B4)}$$

Here, X is either K or S. $F(x)$ is defined as $F(x) = \sin x$ for the magnification K and $F(x) = 1 - \cos x$ for the phase modulation S. Even though χ_1 and χ_2 are symmetrical and can be expressed by either one of two terms, we explicitly write both terms so that the symmetry can be captured easily. The integral with respect to ϕ can be performed analytically by using the identities regarding Bessel functions,

$$\int_0^{2\pi} \frac{d\phi}{2\pi} \cos\phi \sin(x\cos\phi) = J_1(x), \quad \text{(B5)}$$

$$\int_0^{2\pi} \frac{d\phi}{2\pi} \cos^2\phi \cos(x\cos\phi) = \frac{1}{2}[J_0(x) - J_2(x)]. \quad \text{(B6)}$$

In addition to this, $\int_0^{2\pi} \frac{d\phi}{2\pi} \cos\phi \cos(x\cos\phi) = \int_0^{2\pi} \frac{d\phi}{2\pi} \cos^2\phi \sin(x\cos\phi) = 0$ holds by virtue of the antisymmetric nature of the integrand.

For the magnification, $\mathcal{F}_{K,12}, \mathcal{F}_{K,1}, \mathcal{F}_{K,2}$ are given by

$$\mathcal{F}_{K,12} = \frac{\chi_1^2\chi_2^2}{4}\left\{\left(\left(1 - \cos\left(\frac{\chi_1^2 W(\chi_1,\chi)}{\omega}k_1^2\right) - \cos\left(\frac{\chi_2^2 W(\chi_2,\chi)}{\omega}k_2^2\right) + \cos\left(\frac{\chi_1^2 W(\chi_1,\chi_s)}{\omega}k_1^2 + \frac{\chi_2^2 W(\chi_2,\chi)}{\omega}k_2^2\right)\right.\right.\right.$$
$$\left.+ \cos\left(\frac{\chi_2^2 W(\chi_2,\chi_s)}{\omega}k_2^2 + \frac{\chi_1^2 W(\chi_1,\chi)}{\omega}k_1^2\right)\right)\left[J_0\left(\frac{\chi_1\chi_2 W(\chi',\chi)}{\omega}k_1 k_2\right) - J_2\left(\frac{\chi_1\chi_2 W(\chi',\chi)}{\omega}k_1 k_2\right)\right]$$
$$- \cos\left(\frac{\chi_1^2 W(\chi_1,\chi_s)}{\omega}k_1^2 + \frac{\chi_2^2 W(\chi_2,\chi_s)}{\omega}k_2^2\right)\left[J_0\left(\frac{\chi_1\chi_2 W(\chi,\chi_s) + \chi_1\chi_2 W(\chi',\chi_s)}{\omega}k_1 k_2\right)\right.$$
$$\left.\left.- J_2\left(\frac{\chi_1\chi_2 W(\chi,\chi_s) + \chi_1\chi_2 W(\chi',\chi_s)}{\omega}k_1 k_2\right)\right]\right\}, \quad \text{(B7)}$$

$$\mathcal{F}_{K,1} = \frac{\chi_1\chi_2^3}{2}\left\{\sin\left(\frac{\chi_1^2 W(\chi_1,\chi_s)}{\omega}k_1^2 + \frac{\chi_2^2 W(\chi_2,\chi)}{\omega}k_2^2\right) - \sin\left(\frac{\chi_2^2 W(\chi_2,\chi)}{\omega}k_2^2\right)\right\} J_1\left(\frac{\chi_1\chi_2 W(\chi',\chi)}{\omega}k_1 k_2\right), \quad \text{(B8)}$$

$$\mathcal{F}_{K,2} = \frac{\chi_2\chi_1^3}{2}\left\{\sin\left(\frac{\chi_2^2 W(\chi_2,\chi_s)}{\omega}k_2^2 + \frac{\chi_1^2 W(\chi_1,\chi)}{\omega}k_1^2\right) - \sin\left(\frac{\chi_1^2 W(\chi_1,\chi)}{\omega}k_1^2\right)\right\} J_1\left(\frac{\chi_1\chi_2 W(\chi',\chi)}{\omega}k_1 k_2\right). \quad \text{(B9)}$$

In exactly the same way, the similar equations are derived for the phase modulation,

$$\mathcal{F}_{S,12} = \chi_1^2\chi_2^2\left\{-\frac{1}{2} + \frac{1}{2}\cos\left(\frac{\chi_1^2 W(\chi_1,\chi_s)}{2\omega}k_1^2\right) + \frac{1}{2}\cos\left(\frac{\chi_2^2 W(\chi_2,\chi_s)}{2\omega}k_2^2\right) - \frac{1}{2}\cos\left(\frac{\chi_1^2 W(\chi_1,\chi_s)}{2\omega}k_1^2 + \frac{\chi_2^2 W(\chi_2,\chi_s)}{2\omega}k_2^2\right)\right.$$
$$\times\left[J_0\left(\frac{\chi_1\chi_2 W(\chi,\chi_s)}{\omega}k_1 k_2\right) + J_0\left(\frac{\chi_1\chi_2 W(\chi',\chi_s)}{\omega}k_1 k_2\right) - J_2\left(\frac{\chi_1\chi_2 W(\chi,\chi_s)}{\omega}k_1 k_2\right) - J_2\left(\frac{\chi_1\chi_2 W(\chi',\chi_s)}{\omega}k_1 k_2\right)\right]$$
$$+ \frac{1}{4}\cos\left(\frac{\chi_1^2 W(\chi_1,\chi_s)}{\omega}k_1^2 + \frac{\chi_2^2 W(\chi_2,\chi_s)}{\omega}k_2^2\right)\left[J_0\left(\frac{\chi_1\chi_2 W(\chi,\chi_s) + \chi_1\chi_2 W(\chi',\chi_s)}{\omega}k_1 k_2\right)\right.$$
$$\left.- J_2\left(\frac{\chi_1\chi_2 W(\chi,\chi_s) + \chi_1\chi_2 W(\chi',\chi_s)}{\omega}k_1 k_2\right)\right] + \left(\left(\frac{1}{4} + \sin^2\left(\frac{\chi_1^2 W(\chi_1,\chi_s)}{4\omega}k_1^2\right)\right)\cos\left(\frac{\chi_1^2 W(\chi_1,\chi_s)}{2\omega}k_1^2 + \frac{\chi_2^2 W(\chi_2,\chi)}{\omega}k_2^2\right)\right.$$
$$\left.+ \sin^2\left(\frac{\chi_2^2 W(\chi_2,\chi_s)}{4\omega}k_2^2\right)\cos\left(\frac{\chi_2^2 W(\chi_2,\chi_s)}{2\omega}k_2^2 + \frac{\chi_1^2 W(\chi_1,\chi)}{\omega}k_1^2\right)\right)$$
$$\times\left[J_0\left(\frac{\chi_1\chi_2 W(\chi',\chi)}{\omega}k_1 k_2\right) - J_2\left(\frac{\chi_1\chi_2 W(\chi',\chi)}{\omega}k_1 k_2\right)\right]\right\}, \quad \text{(B10)}$$

$$\mathcal{F}_{S,1} = 2\chi_1^2\chi_2^3\sin^2\left(\frac{\chi_1^2 W(\chi_1,\chi_s)}{4\omega}k_1^2\right)\sin\left(\frac{\chi_1^2 W(\chi_1,\chi_s)}{2\omega}k_1^2 + \frac{\chi_2^2 W(\chi_2,\chi)}{\omega}k_2^2\right)J_1\left(\frac{\chi_1\chi_2 W(\chi',\chi)}{\omega}k_1 k_2\right), \quad \text{(B11)}$$

$$\mathcal{F}_{S,2} = 2\chi_2^2\chi_1^3\sin^2\left(\frac{\chi_2^2 W(\chi_2,\chi_s)}{4\omega}k_2^2\right)\sin\left(\frac{\chi_2^2 W(\chi_2,\chi_s)}{2\omega}k_2^2 + \frac{\chi_1^2 W(\chi_1,\chi)}{\omega}k_1^2\right)J_1\left(\frac{\chi_1\chi_2 W(\chi',\chi)}{\omega}k_1 k_2\right). \quad \text{(B12)}$$

For the non-Gaussian correction, we only consider the bispectrum term as the only relevant contribution, thus $\delta_{X^2 c} = 2\langle X_{\text{Born}}X^{(2)}\rangle_c$. Based on this assumption, Eq. (2.43) can be rewritten as

$$\delta_{S^2 c} = 8\omega^3\left(\frac{3H_0^2\Omega_m}{2}\right)^3\int_0^{\chi_s}\frac{d\chi}{a^3}\int_0^\infty\frac{dk_1}{2\pi}\int_0^\infty\frac{dk_2}{2\pi}\int_0^{2\pi}\frac{d\phi}{2\pi}\frac{B_\delta(k_1,k_2,k_3,\chi)}{k_1 k_2 k_3^2}$$
$$\times\left(1 - \cos\left[\frac{r_F^2}{2}k_3^2\right]\right)\left\{r_F^2\boldsymbol{k}_1\cdot\boldsymbol{k}_2 - 2\sin\left[\frac{r_F^2}{2}\boldsymbol{k}_1\cdot\boldsymbol{k}_2\right]\cos\left[\frac{r_F^2}{2}\frac{k_1^2 + k_2^2 + k_3^2}{2}\right]\right\}, \quad \text{(B13)}$$

$$\delta_{K^2,c} = 16\omega^3\left(\frac{3H_0^2\Omega_m}{2}\right)^3\int_0^{\chi_s}\frac{d\chi}{a^3}\int_0^\infty\frac{dk_1}{2\pi}\int_0^\infty\frac{dk_2}{2\pi}\int_0^{2\pi}\frac{d\phi}{2\pi}\frac{B_\delta(k_1,k_2,k_3,\chi)}{k_1 k_2 k_3^2}$$
$$\times\sin\left[\frac{r_F^2}{2}k_3^2\right]\sin\left[\frac{r_F^2}{2}\frac{k_1^2 + k_2^2 + k_3^2}{2}\right]\sin\left[\frac{r_F^2}{2}\boldsymbol{k}_1\cdot\boldsymbol{k}_2\right], \quad \text{(B14)}$$

where $k_3^2 = k_1^2 + k_2^2 + 2k_1 k_2\cos\phi$ and $\boldsymbol{k}_1\cdot\boldsymbol{k}_2 = k_1 k_2\cos\phi$. The bispectrum contribution to the average of S and K, Eqs. (2.34) and (2.35), can be rewritten in a similar way as

$$\langle S^{(3)}\rangle = 4\omega^3\left(\frac{3H_0^2\Omega_m}{2}\right)^3\int_0^{\chi_s}\frac{d\chi}{a^3}\int_0^\infty\frac{dk_1}{2\pi}\int_0^\infty\frac{dk_2}{2\pi}\int_0^{2\pi}\frac{d\phi}{2\pi}\frac{B_\delta(k_1,k_2,k_3,\chi)}{k_1 k_2 k_3^2}$$
$$\times\left\{\frac{4k_3^2\sin^2(\frac{r_F^2}{4}(k_1^2 + k_2^2 + k_3^2))}{(k_1^2 + k_2^2 + k_3^2)} - 2\sin^2\frac{r_F^2 k_3^2}{2} - \frac{r_F^2}{2}(\boldsymbol{k}_1\cdot\boldsymbol{k}_2)k_3^2\right\}, \quad \text{(B15)}$$

$$\langle K^{(3)}\rangle_c = 4\omega^3\left(\frac{3H_0^2\Omega_m}{2}\right)^3\int_0^{\chi_s}\frac{d\chi}{a^3}\int_0^\infty\frac{dk_1}{2\pi}\int_0^\infty\frac{dk_2}{2\pi}\int_0^{2\pi}\frac{d\phi}{2\pi}\frac{B_\delta(k_1,k_2,k_3,\chi)}{k_1 k_2 k_3^2}$$
$$\times\left\{\sin[r_F^2 k_1^2] + \sin[r_F^2 k_2^2] + \sin[r_F^2 k_3^2] - 2\sin\left[\frac{r_F^2}{2}(k_1^2 + k_2^2 + k_3^2)\right]\right\}. \quad \text{(B16)}$$

Note that, in computing bispectrum contribution numerically, symmetrizing the wave number variables k_1, k_2, k_3 reduces the computational cost.

APPENDIX C: HIGH-FREQUENCY BEHAVIOR OF $\delta_{S^2,\text{dc}}$

In this appendix, we would like to approximately derive the high-frequency behavior of $\delta_{S^2,\text{dc}}$ in order to avoid the difficulty of numerical computation associated with the cancellation of significant digits. When the frequency of GWs is high, $\delta_{S^2,\text{dc}}$ is mainly affected by the large k region of the matter power spectrum. Given that the corresponding Fresnel scale mainly contributes to the lensing, we can expand the power spectrum around the approximated Fresnel scale $(1/\sqrt{H_0\omega})$ as

$$P_\delta(k,\chi) = P_\delta(k_0,\chi)\left(\frac{k}{k_0}\right)^{\frac{d\log P_\delta(k_0,\chi)}{d\log k_0}} = B(\chi)k^{-b(\chi)}, \quad \text{(C1)}$$

where $(k_0 = \sqrt{H_0\omega})$. In this way, $B(\chi)$ and b_χ are both functions of redshift and the frequency of GWs. We compute $B(\chi)$ and $b(\chi)$ numerically using our power spectrum at each frequency. Keeping this in mind, we only consider the case that is relevant to our discussion.

We usually deal with the GW sources whose distance from Earth is roughly given by $1/H_0$, so the corresponding Fresnel scale is $1/\sqrt{\omega H_0}$. The high-frequency behavior in this context is then interpreted as the satisfaction of the condition $k_L \ll \sqrt{H_0\omega}$. Defining $\int\cdots\int \equiv 16(\frac{3H_0^2\Omega_m}{2})^4\int_0^{\chi_s}\frac{d\chi}{\chi^2}\int_0^\chi\frac{d\chi'}{\chi'^2}\int_0^{\chi'}d\chi_1\int_0^{\chi'}d\chi_2\frac{1}{a^2(\chi_1)}\frac{1}{a^2(\chi_2)}\frac{1}{(2\pi)^2}$, the post-Born approximation of the variance of S is given by

$$\delta_{S^2,\mathrm{dc}} = \int \cdots \int \omega^2 \int_0^\infty dk_1 \int_0^\infty dk_2 P_\delta(k_1,\chi_1) P_\delta(k_2,\chi_2) \left[\frac{1}{k_1 k_2} \mathcal{F}_{S,12} - \frac{1}{k_1^2} \mathcal{F}_{S,1} - \frac{1}{k_2^2} \mathcal{F}_{S,2} \right]. \tag{C2}$$

The definition of $\mathcal{F}_{S,12}, \mathcal{F}_{S,1}, \mathcal{F}_{S,2}$ is the same as the ones in Appendix B. Change of variables $k_1 = \sqrt{\omega}\xi_1$ and $k_2 = \sqrt{\omega}\xi_2$ and separating the integral area at k_L yields

$$\begin{aligned}
\delta_{S^2,\mathrm{dc}} =& \int \cdots \int \omega^2 \int_0^{\frac{k_L}{\sqrt{\omega}}} d\xi_1 \int_0^{\frac{k_L}{\sqrt{\omega}}} d\xi_2 P_\delta(\sqrt{\omega}\xi_1,\chi_1) P_\delta(\sqrt{\omega}\xi_2,\chi_2) \left[\frac{1}{\xi_1 \xi_2} \mathcal{F}_{12} - \frac{1}{\xi_1^2} \mathcal{F}_1 - \frac{1}{\xi_2^2} \mathcal{F}_2 \right] \\
&+ \int \cdots \int \omega^{2-\frac{b}{2}} B(\chi_2) \int_0^{\frac{k_L}{\sqrt{\omega}}} d\xi_1 \int_{\frac{k_L}{\sqrt{\omega}}}^\infty d\xi_2 P_\delta(\sqrt{\omega}\xi_1,\chi_1) \xi_2^{-b} \left[\frac{1}{\xi_1 \xi_2} \mathcal{F}_{12} - \frac{1}{\xi_1^2} \mathcal{F}_1 - \frac{1}{\xi_2^2} \mathcal{F}_2 \right] \\
&+ \int \cdots \int \omega^{2-\frac{b}{2}} B(\chi_1) \int_{\frac{k_L}{\sqrt{\omega}}}^\infty d\xi_1 \int_0^{\frac{k_L}{\sqrt{\omega}}} d\xi_2 P_\delta(\sqrt{\omega}\xi_2,\chi_2) \xi_1^{-b} \left[\frac{1}{\xi_1 \xi_2} \mathcal{F}_{12} - \frac{1}{\xi_1^2} \mathcal{F}_1 - \frac{1}{\xi_2^2} \mathcal{F}_2 \right] \\
&+ \int \cdots \int \omega^{2-b} B(\chi_1) B(\chi_2) \int_{\frac{k_L}{\sqrt{\omega}}}^\infty d\xi_1 \int_{\frac{k_L}{\sqrt{\omega}}}^\infty d\xi_2 \xi_2^{-b} \xi_1^{-b} \left[\frac{1}{\xi_1 \xi_2} \mathcal{F}_{12} - \frac{1}{\xi_1^2} \mathcal{F}_1 - \frac{1}{\xi_2^2} \mathcal{F}_2 \right]. \tag{C3}
\end{aligned}$$

Since these four terms contribute to $\delta_{S^2,\mathrm{dc}}$ in a different way, we will compute the contribution from each term separately. To begin with, we consider the first term. In the high-frequency limit, the integral range $\int_0^{\frac{k_L}{\sqrt{\omega}}}$ is restricted in a very small area so the contribution from the first term in Eq. (C3) comes from the region where $\xi_1, \xi_2 \ll 1$ holds. Since ξ_1 and ξ_2 are both order $1/\sqrt{\omega}$ in this integral range, the expansion of $\mathcal{F}$ in $1/\sqrt{\omega}$ up to leading order yields

$\mathcal{F} = \mathcal{O}(1/\omega^4)$. Considering that ω^2 is multiplied in the expression, we can conclude that the first term is proportional to ω^{-2}.

The second and third terms in Eq. (C3) are symmetrical with respect to the subscript 1, 2, so they have the same contribution. In the second term, ξ_1 is still restricted in the area where $\xi_1 \ll 1$, whereas ξ_2 is no longer small. In this case, we can expand $\mathcal{F}$ only in terms of ξ_1 and keep the ξ_2 term untouched, then we have

$$\begin{aligned}
\mathcal{F} =& \xi_1 \left[\xi_2 \left(\cos\left(\frac{\chi_2^2 W(\chi_2,\chi_s)}{2} \xi_2^2 \right) - \cos^2\left(\frac{\chi_2^2 W(\chi_2,\chi_s)}{2} \xi_2^2 \right) \right) C \right. \\
&\left. + \left(\frac{\sin\left(\frac{\chi_2^2 W(\chi_2,\chi_s)}{2} \xi_2^2 \right)}{2\xi_2} - \frac{\sin\left(\chi_2^2 W(\chi_2,\chi_s)\xi_2^2 \right)}{4\xi_2} \right) D_1 + C\xi_2 \sin^2\left(\frac{\chi_2^2 W(\chi_2,\chi_s)}{2} \xi_2^2 \right) \right] + \mathcal{O}(\xi_1^3), \tag{C4}
\end{aligned}$$

where C, D_1 are

$$C = \frac{3}{8} \chi_1^4 \chi_2^4 W(\chi,\chi_s) W(\chi',\chi_s), \tag{C5}$$

$$D_1 = \chi_1^6 \chi_2^2 \{ -W(\chi',\chi) + W(\chi,\chi_s) \}. \tag{C6}$$

Combining these notations, the second term is calculated as

$$\begin{aligned}
(\mathrm{2nd}) =& \int \cdots \int \omega^{2-\frac{b}{2}} B(\chi_2) \int_0^{\frac{k_L}{\sqrt{\omega}}} d\xi_1 P_\delta(\sqrt{\omega}\xi_1,\chi_1) \int_{\frac{k_L}{\sqrt{\omega}}}^\infty d\xi_2 \xi_2^{-b} \left[\frac{1}{\xi_1 \xi_2} \mathcal{F}_{12} - \frac{1}{\xi_1^2} \mathcal{F}_1 - \frac{1}{\xi_2^2} \mathcal{F}_2 \right] \\
=& \int \cdots \int \frac{1}{\omega^{\frac{b}{2}-1}} B(\chi_2) \int_0^\infty dk_1 k_1 P_\delta(k_1,\chi_1) \int_0^\infty d\xi_2 \frac{1}{\xi_2^b} [(\cdots)C + (\cdots)D_1 + (\cdots)C] \\
=& \int \cdots \int \frac{1}{\omega^{\frac{b}{2}-1}} B(\chi_2) \int_0^\infty dk_1 k_1 P_\delta(k_1,\chi_1) \left(\frac{\chi_2^2 W(\chi_2,\chi_s)}{2} \right)^{\frac{b-2}{2}} \left[CI_c + \frac{2}{\chi_2^2 W(\chi_2,\chi_s)} D_1 I_d + CI_e \right]. \tag{C7}
\end{aligned}$$

From the first line to the second, we used Eq. (C4) and took the integral range from zero to infinity. We can safely perform this approximation due to the fact that the integral converges. Here, I_c, I_d, I_e are just numbers defined as

$$I_c = \int_0^\infty dx \frac{\cos x^2 - \cos^2 x^2}{x^{b-1}}, \tag{C8}$$

$$I_d = \int_0^\infty dx \frac{2\sin x^2 - \sin 2x^2}{4x^{b+1}}, \tag{C9}$$

$$I_e = \int_0^\infty dx \frac{\sin^2 x^2}{x^{b-1}}. \tag{C10}$$

The fourth term in Eq. (C3) is calculated in a similar way,

$$(4\text{th}) = \int \cdots \int \frac{1}{\omega^{\frac{b}{2}-1}} B(\chi_1)B(\chi_2) \frac{k_L^{-b+2}}{b-2} \left\{ \left(\frac{\chi_2^2 W(\chi_2, \chi_s)}{2} \right)^{\frac{b-2}{2}} \left[CI_c + \frac{2}{\chi_2^2 W(\chi_2, \chi_s)} D_1 I_d + CI_e \right] + (1 \Leftrightarrow 2) \right\}. \tag{C11}$$

Note that the fourth term is essentially determined by k_L, which is the arbitrary scale. However, the second and third term are determined by the scale at which P_δ changes from an increasing function to a decreasing function due to the dependence on $\int_0^\infty dk_1 k_1 P_\delta(k_1)$. Since we can take k_L to be sufficiently larger than this scale, it is justified to ignore the contribution from the fourth term, and we have the following expression for $\delta_{S^2,\text{dc}}$:

$$\delta_{S^2,\text{dc}} = 2 \times 16 \left(\frac{3H_0^2 \Omega_m}{2} \right)^4 \int_0^{\chi_s} \frac{d\chi}{\chi^2} \int_0^\chi \frac{d\chi'}{\chi'^2} \int_0^{\chi'} d\chi_1 \int_0^{\chi'} d\chi_2 \frac{1}{a^2(\chi_1)} \frac{1}{a^2(\chi_2)} \frac{1}{(2\pi)^2}$$
$$\times \frac{B(\chi_2)}{\omega^{\frac{b}{2}-1}} \int_0^\infty dk_1 k_1 P_\delta(k_1, \chi_1) \left[CI_c + \frac{2}{\chi_2^2 W(\chi_2, \chi_s)} D_1 I_d + CI_e \right] \left(\frac{\chi_2^2 W(\chi_2, \chi_s)}{2} \right)^{\frac{b_2-2}{2}}. \tag{C12}$$

This expression can be further simplified by changing the order of integral and using some formula for the gamma function, and finally, we have

$$\delta_{S^2,\text{dc}} = \frac{3}{4} \left(\frac{3H_0^2 \Omega_m}{2} \right)^4 \int_0^{\chi_s} d\chi_1 \int_0^{\chi_1} d\chi_2 \frac{W^4(\chi_1, \chi_s)\chi_1^4\chi_2^4}{a^2(\chi_1)a^2(\chi_2)} \frac{1}{(2\pi)^2}$$
$$\times \left\{ \frac{B(\chi_2)}{\omega^{\frac{b}{2}-1}} \int_0^\infty dk_1 k_1 P_\delta(k_1, \chi_1) \left[\frac{\chi_2^2 W(\chi_2, \chi_s)}{2} \right]^{\frac{b_2}{2}-1} \left(1 - 2^{\frac{b_2}{2}} \right) \Gamma\left(1 - \frac{b_2}{2} \right) \sin\left[\frac{b_2}{2} \right] + (1 \leftrightarrow 2) \right\}. \tag{C13}$$

It is clear from this expression that this term depends not only on the scale corresponding to the Fresnel scale but also on the factor $\int_0^\infty dk_1 k_1 P_\delta(k_1, \chi_1)$ that is mainly contributed by the large-scale matter fluctuation. This means that the information pertaining to the larger-scale fluctuation is encoded within the small scale through the higher-order terms. In the physics context, the correlation between the large- and small-scale matter fluctuations arises from the fact that the regions where the large-scale matter fluctuation is significant have higher matter density than areas with small fluctuation, and in this region, the small-scale matter fluctuation is more likely to grow and be amplified simply due to the abundance of matter available.

[1] M. Bartelmann, Gravitational lensing, Classical Quantum Gravity **27**, 233001 (2010).

[2] R. Mandelbaum, Weak lensing for precision cosmology, Annu. Rev. Astron. Astrophys. **56**, 393 (2018).

[3] M. Oguri, Strong gravitational lensing of explosive transients, Rep. Prog. Phys. **82**, 126901 (2019).

[4] C. Misner, K. Thorne, and J. Wheeler, *Gravitation* (W. H. Freeman and Company, San Francisco, 1973).

[5] H. C. Ohanian, On the focusing of gravitational radiation, Int. J. Theor. Phys. **9,** 425 (1974).

[6] R. Takahashi and T. Nakamura, Wave effects in gravitational lensing of gravitational waves from chirping binaries, Astrophys. J. **595,** 1039 (2003).

[7] T. T. Nakamura, Gravitational Lensing of Gravitational Waves from Inspiraling Binaries by a Point Mass Lens, Phys. Rev. Lett. **80,** 1138 (1998).

[8] T. T. Nakamura and S. Deguchi, Wave optics in gravitational lensing, Prog. Theor. Phys. Suppl. **133,** 137 (1999).

[9] R. Takahashi, Amplitude and phase fluctuations for gravitational waves propagating through inhomogeneous mass distribution in the Universe, Astrophys. J. **644,** 80 (2006).

[10] M. Oguri and R. Takahashi, Probing dark low-mass halos and primordial black holes with frequency-dependent gravitational lensing dispersions of gravitational waves, Astrophys. J. **901,** 58 (2020).

[11] M. Oguri and R. Takahashi, Amplitude and phase fluctuations of gravitational waves magnified by strong gravitational lensing, Phys. Rev. D **106,** 043532 (2022).

[12] C. Shapiro and A. Cooray, The Born and lens-lens corrections to weak gravitational lensing angular power spectra, J. Cosmol. Astropart. Phys. 03 (**2006**) 007.

[13] S. Hilbert, J. Hartlap, S. D. M. White, and P. Schneider, Ray-tracing through the millennium simulation: Born corrections and lens-lens coupling in cosmic shear and galaxy-galaxy lensing, Astron. Astrophys. **499,** 31 (2009).

[14] E. Krause and C. M. Hirata, Weak lensing power spectra for precision cosmology: Multiple-deflection, reduced shear and lensing bias corrections, Astron. Astrophys. **523,** A28 (2010).

[15] G. Pratten and A. Lewis, Impact of post-Born lensing on the CMB, J. Cosmol. Astropart. Phys. 08 (2016) 047.

[16] A. Petri, Z. Haiman, and M. May, Validity of the Born approximation for beyond Gaussian weak lensing observables, Phys. Rev. D **95,** 123503 (2017).

[17] P. Schneider, J. Ehlers, and E. E. Falco, *Gravitational Lenses* (Springer-Verlag, Berlin, 1992).

[18] S. Dodelson, *Modern Cosmology* (Academic Press, New York, 2003).

[19] M. Inamori and T. Suyama, Universal relation between the variances of distortions of gravitational waves owing to gravitational lensing, Astrophys. J. Lett. **918,** L30 (2021).

[20] R. Takahashi, M. Sato, T. Nishimichi, A. Taruya, and M. Oguri, Revising the halofit model for the nonlinear matter power spectrum, Astrophys. J. **761,** 152 (2012).

[21] M. Fukugita, C. J. Hogan, and P. J. E. Peebles, The cosmic baryon budget, Astrophys. J. **503,** 518 (1998).

[22] A. Cooray and R. K. Sheth, Halo models of large scale structure, Phys. Rep. **372,** 1 (2002).

[23] D. Dolney, B. Jain, and M. Takada, Effects of halo substructure on the power spectrum and bispectrum, Mon. Not. R. Astron. Soc. **352,** 1019 (2004).

[24] R. Takahashi, T. Nishimichi, T. Namikawa, A. Taruya, I. Kayo, K. Osato, Y. Kobayashi, and M. Shirasaki, BiHalofit: Fitting formula of non-linear matter bispectrum, Astrophysics Source Code Library, record ascl:2106.031, 2021.

[25] A. Oddo, E. Sefusatti, C. Porciani, P. Monaco, and A. G. Sánchez, Toward a robust inference method for the galaxy bispectrum: Likelihood function and model selection, J. Cosmol. Astropart. Phys. 03 (**2020**) 056.

[26] N. Kaiser and J. A. Peacock, On the bias of the distance-redshift relation from gravitational lensing, Mon. Not. R. Astron. Soc. **455,** 4518 (2016).

[27] R. Takahashi, M. Oguri, M. Sato, and T. Hamana, Probability distribution functions of cosmological lensing: Convergence, shear, and magnification, Astrophys. J. **742,** 15 (2011).

[28] P. Villanueva-Domingo, O. Mena, and S. Palomares-Ruiz, A brief review on primordial black holes as dark matter, Front. Astron. Space Sci. **8,** 87 (2021).

[29] L. Lindblom, B. J. Owen, and D. A. Brown, Model waveform accuracy standards for gravitational wave data analysis, Phys. Rev. D **78,** 124020 (2008).

[30] B. P. Abbott *et al.*, Exploring the sensitivity of next generation gravitational wave detectors, Classical Quantum Gravity **34,** 044001 (2017).