# Shaping space-time wave packets beyond the paraxial limit using a dispersion magnifier

Dongha Kim, Cheng Guo, Peter B. Catrysse, and Shanhui Fan

*E. L. Ginzton Laboratory, Stanford University, 348 Via Pueblo, California 94305, USA*

Space-time wave packets (STWPs) have received significant attention since they can propagate in free space at arbitrary group velocity without dispersion and diffraction. However, at present, the generation of STWPs has been limited to the paraxial regime. Here we show that conventional optical elements can be used to extend STWPs beyond the paraxial regime. A dispersion magnifier, consisting of two lenses and a beam expander, applies spatiotemporal shaping to paraxial STWPs to create nonparaxial STWPs. The control of the magnification ratio results in versatile engineering capabilities on group velocity, beam diameter, and propagation distance. As an example, we numerically demonstrate long-distance propagation or slow group velocity of the output wave packet with subwavelength cross sections.

## I. INTRODUCTION

Space-time wave packets (STWPs) are a distinctive class of structured optical pulses possessing shape-preserving properties in space and time with engineered group velocities ($v_g$) [1]. These wave packets can be formed in either free space or dispersive media without the use of nonlinearity. Their formation involves exciting only those plane-wave components that have their frequencies ($\omega$) and longitudinal wave vectors ($k_z$) satisfying a dispersion relation $\omega = \omega(k_z)$. This dispersion relation is also referred to as space-time coupling. The preparation of space-time coupling has been proposed or demonstrated using various devices, such as multimode optical fibers [2,3], spatial light modulators [4,5], and photonic crystal slabs [6]. In particular, the high programmability of spatial light modulators has enabled the demonstration of various properties of STWPs, including arbitrary group velocity engineering [7], anomalous refraction and reflection [8], and complete localization of wave packets in all dimensions [9].

However, the demonstration of STWPs has not been achieved in nonparaxial regimes so far [1]. Here, the paraxial regime refers to the scenario where the diffraction angle $\theta$, i.e., the angle between the wave vector of the plane-wave components of the wave packet and the propagation axis, fulfills $\theta \approx \sin\theta$ with less than 1.5% error [10], which corresponds to $\theta < 10°$ or numerical aperture NA $< 0.17$. By generating STWPs in the nonparaxial regime, long-distance propagation of pulses with wavelength-scale cross section can be demonstrated, which may enable high-density information processing in free space [11,12] and transverse spatial resolution enhancement in light sheet fluorescence microscopy [13,14].

In this work, we theoretically propose a spatiotemporal pulse-shaping method, which we call a dispersion magnifier, that can generate STWPs in a nonparaxial regime and engineer their properties with conventional optical elements [Fig. 1(a)]. For an incident STWP in the paraxial regime, the dispersion magnifier can modify its dispersion relation by a stretching or shrinking transformation along the transverse wave-vector ($k_t$) axis, allowing access to both the paraxial and

nonparaxial regimes. As a result, various properties of STWPs can be manipulated by the dispersion magnifier, including transverse wave vector, group velocity, beam diameter, and propagation distance.

## II. CONCEPT

The dispersion magnifier consists of a collimating lens, an imaging lens, and a beam expander in between [Fig. 1(a)]. For the incident STWP, at each frequency $\omega$, the plane-wave components have the same longitudinal wave vector $k_z^{(i)}$. Their transverse wave vectors have the same magnitude $k_t^{(i)}$. The magnitude of wave vector $k_0$ is defined as $k_0 = \sqrt{(k_z^{(i)})^2 + (k_t^{(i)})^2}$. After the input STWP passes through the collimating lens, it forms a ring-shaped beam in real space, with each frequency component located at a ring with a radius $r_1$ of

$$r_1 = f_C \tan(\theta^{(i)}) = f_C k_t^{(i)} / k_z^{(i)}, \tag{1}$$

where $f_C$ is the focal length of the collimating lens and $\theta^{(i)}$ is the propagating angle of this frequency component in the input STWP. The ring-shaped beam then passes through the variable beam expander. The light at the ring with the radius $r_1$ is mapped to light at a ring with a radius $r_2$ as

$$r_2 = M r_1, \tag{2}$$

where $M$ is the magnification ratio of the beam expander. The output from the beam expander then passes through the imaging lens. The magnitude of the transverse wave vector of the output STWP ($k_t^{(f)}$) is related to the ring radius $r_2$ as

$$k_t^{(f)} = k_0 \sin(\theta^{(f)}) = \frac{r_2}{\sqrt{r_2^2 + f_O^2}}, \tag{3}$$

where $f_O$ is the focal length of the imaging lens and $\theta^{(f)}$ is the propagating angle of this frequency component in the output STWP. As a result, the relations between the input and
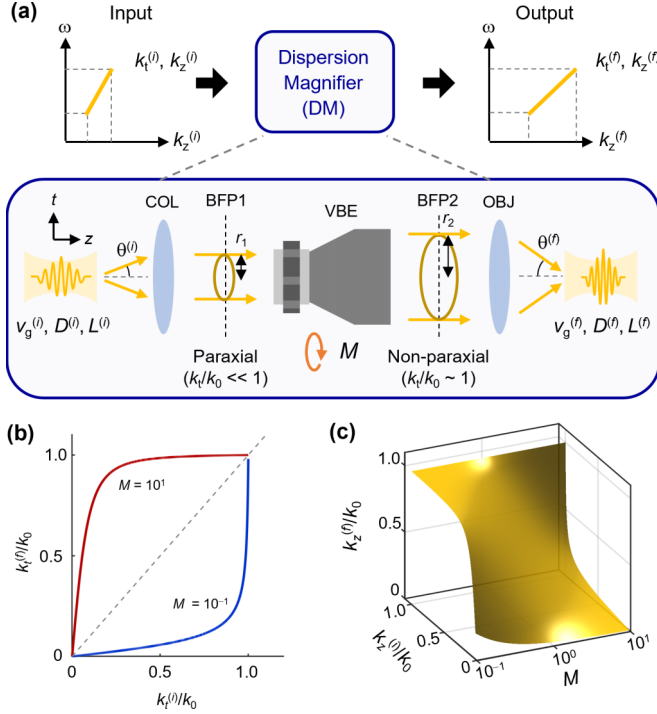
FIG. 1. (a) Schematic illustration of dispersion magnifier and its operation mechanism. COL (OBJ) is the collimating (imaging) lens, respectively. VBE is the variable beam expander. BFP1 (BFP2) is the back focal plane of COL (OBJ), respectively. $M$ is the magnification ratio. (b) Relation between the input and output transverse wave vectors $k_t^{(i)}/k_0 - k_t^{(f)}/k_0$ for the magnification ratios $M = 10^1$ (red) and $10^{-1}$ (blue), respectively. The gray dashed line indicates the condition of $k_t^{(i)} = k_t^{(f)}$. (c) Relation between the input and output longitudinal wave vectors $k_z^{(i)}/k_0 - k_z^{(f)}/k_0$ depending on $M = 10^{-1}$–$10^1$.

output wave-vector components can be derived by combining Eqs. (1)–(3) as

$$k_t^{(f)}/k_0 = \frac{M(k_t^{(i)}/k_0)}{\sqrt{1 + (M^2 - 1)(k_t^{(i)}/k_0)^2}}, \quad (4)$$

$$k_z^{(f)}/k_0 = \frac{k_z^{(i)}/k_0}{\sqrt{M^2 + (1 - M^2)(k_z^{(i)}/k_0)^2}}, \quad (5)$$

where we assumed $f_C = f_O$ for simplicity.

The transformation in Eq. (4) allows one to transform a STWP between the paraxial and nonparaxial regimes. As an illustration, in Fig. 1(b), we plot the relation of Eq. (4) for two different magnification ratios. The paraxial regime corresponds to $k_t/k_0 < 0.17$. Here and throughout the paper, we consider the magnification ratio $M$ to be in the range of $0.1 < M < 10$ for a dispersion magnifier, a range that is accessible in bulk lens systems [15]. For $M = 10$, an incident STWP with $k_t^{(i)}/k_0$ in the range of [0.017, 0.17], which is in the paraxial regime, will be transformed into a wave packet outside the paraxial regime. Conversely, for $M = 0.1$, an incident STWP with $k_t^{(i)}/k_0$ in the range of [0.17, 0.865], which is outside the paraxial regime, will be transformed into a wave packet within the paraxial regime. In Fig. 1(c), we plot the relation
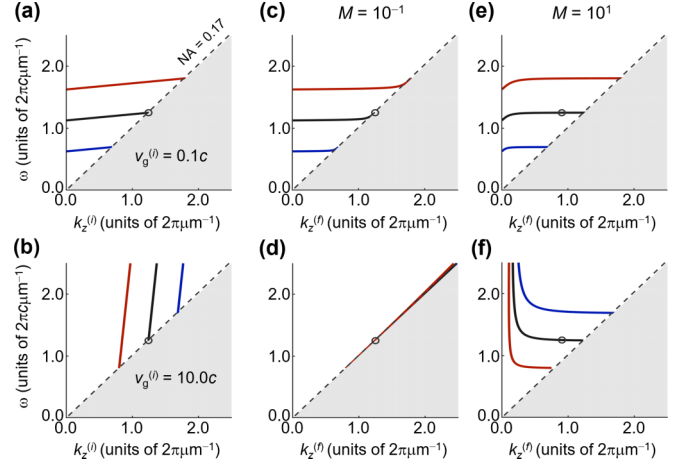


FIG. 2. (a) Dispersion of input STWPs ($\omega = v_g^{(i)} k_z + \omega_0$) with subluminal group velocity $v_g^{(i)} = 0.1$ and $\omega_0 = 1.5$ (red line), 1.0 (black line), and 0.5 (blue line). (b) Dispersion of input STWPs with superluminal group velocity $v_g^{(i)} = 10.0$ and $\omega_0 = -2$ (red line), $-6$ (black line), and $-10$ (blue line). [(c), (d)] Dispersion of output STWPs at magnification $M = 10^{-1}$ for (c) subluminal and (d) superluminal $v_g^{(i)}$. [(e), (f)] Dispersion of output STWPs at magnification $M = 10^1$ for (e) subluminal and (f) superluminal $v_g^{(i)}$. The black dashed line indicates the paraxial limit (NA = 0.17). The black circle indicates the wavelength of 800 nm. The white (gray-shaded) region corresponds to the region above (below) the light line. The units of $v_g$ and $\omega_0$ are $c$ and $2\pi c \ \mu m^{-1}$, respectively.

of Eq. (5) for different magnification ratios. When $M \neq 1$, the longitudinal wave vectors transform, resulting in the change of space-time coupling and hence the group velocity of the STWP. Related to our work, the manipulation of space-time coupling using conventional optical elements has been considered in Refs. [16,17]. References [16,17], however, are restricted to the paraxial regime. The implications of these elements for the nonparaxial regime have not been previously analyzed.

## III. STRETCHING AND SHRINKING OF STWP DISPERSION

In Fig. 2, we show how the space-time couplings of the STWPs are transformed by the dispersion magnifier. The space-time coupling of any of these STWPs lies above the light line of $\omega = ck_z$. With the scale of Fig. 2, the paraxial regime corresponds to a very narrow slice in the immediate vicinity of the light line. Thus, most of the plots here are concerned with the nonparaxial regime. We consider the input wave packets with a space-time coupling of $\omega = v_g^{(i)} k_z + \omega_0$. $\omega_0 = \omega_c [1 - \frac{v_g}{c} \cos(\theta_c)]$, where $\omega_c$ is the central frequency of the spectrum and $\theta_c$ is the propagating angle of the plane-wave component at $\omega = \omega_c$. The $v_g^{(i)}$ value is chosen to be subluminal at $0.1c$ and superluminal at $10.0c$ in Figs. 2(a) and 2(b), respectively. The space-time coupling for the corresponding output wave packets, for two magnification ratios of $M = 10^{-1}$ and $10^1$, are shown in Figs. 2(c)–2(f). In general, the space-time coupling for the output STWP is no longer linear. For $M = 10^{-1}$, a significant part of the space-time coupling is pushed towards the light line [Figs. 2(c) and 2(d)]. For
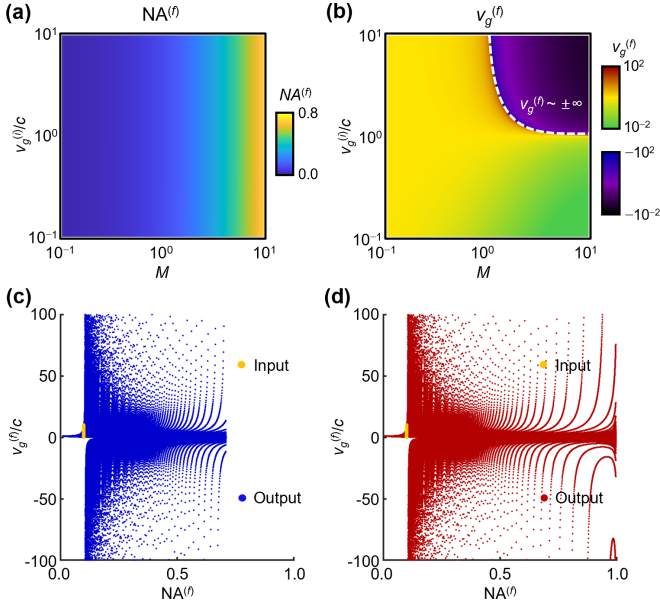
FIG. 3. (a) Calculated maximum numerical aperture ($NA^{(f)}$) and (b) group velocity $v_g^{(f)}$ of output STWPs for incident paraxial STWP ($NA^{(i)} = 0.1$) with group velocities of $v_g = 10^{-1}c$–$10^1 c$. [(c), (d)] Achievable combinations of $v_g^{(f)}$ and $NA^{(f)}$ in (c) a single dispersion magnifier and (d) two cascaded dispersion magnifiers.

$M = 10^1$, the space-time coupling, in general, is pushed away from the light line [Figs. 2(e) and 2(f)]. With the superluminal input wave packet [Fig. 2(b)], at $M = 10^1$, the output may exhibit a negative group velocity [blue curve, Fig. 2(f)]. The results here indicate that a rich set of space-time coupling behavior can be generated, especially in the nonparaxial regime.

## IV. NUMERICAL APERTURE AND GROUP VELOCITY

We investigate the engineering of the numerical aperture and group velocity of STWPs in a dispersion magnifier (Fig. 3). We assume the same kind of input wave packets as described in Fig. 2, with a center wavelength of $\lambda = 0.8$ μm, the group velocities in the range of $v_g = 0.1c$–$10c$, and a maximum numerical aperture of $NA = 0.1$. We further assume that these wave packets are sufficiently narrow-banded, such that the space-time coupling of the output wave packet can be approximated as linear. This narrow-band assumption is, in practice, applicable to most existing experiments on STWP [2,9]. For $v_g = 0.1c$ and $10c$, the input wave packets are represented by the black circles in Figs. 2(a) and 2(b), respectively. The calculated maximum numerical aperture $NA^{(f)} = k_t^{(f)}/k_0$ of output STWPs varies from 0.01 to 0.71, covering both paraxial and nonparaxial regimes depending on $M$ [Fig. 3(a)]. The NA engineering is independent of the incident group velocities. The calculated group velocity $v_g^{(f)} = \partial\omega/\partial k_z^{(f)}$ of the output STWPs can be dramatically changed from subluminal to superluminal ($3.1 \times 10^{-3}c$–$5.9 \times 10^3 c$), even including the negative group velocities from $-1.5 \times 10^4 c$ to $-3.2 \times 10^{-2} c$ [Fig. 3(b)].

We find all possible combinations of $NA^{(f)}$-$v_g^{(f)}$ that can be achieved in the dispersion magnifier. In Figs. 3(c) and 3(d), the input wave packets are represented by yellow dots.

Figure 3(c) considers the use of a single dispersion magnifier with $0.1 < M < 10$. For $M < 1$, the numerical aperture of the output wave packet $NA^{(f)} < NA^{(i)}$. The output wave packet remains in the paraxial regime, and the group velocity of the output wave packet resides in a narrower range as compared with that of the input wave packets. For $M > 1$, $NA^{(f)} > NA^{(i)}$. The output wave packet can be in the nonparaxial regime, and the group velocity of the output wave packet takes a much wider range spanning from $-100c$ to $100c$. In Fig. 3(c), with the magnification ratio $M$ restricted to less than 10, the maximum numerical aperture that can be achieved for the output wave packet is restricted to $NA^{(f)} < 0.71$. To achieve a higher numerical aperture, one can cascade two dispersion magnifers, resulting in the total magnification ratio of $0.01 < M < 100$. The results for two cascaded dispersion magnifiers are shown in Fig. 3(d) with the same input wave packets. We see that the numerical aperture for the output wave packet can reach unity in this case. Such capability for reaching high numerical aperture is important for creating wavelength-scale STWPs, as we discuss explicitly below.

## V. BEAM DIAMETER AND PROPAGATION DISTANCE

In the discussion above, we assume that the spectrum of a light bullet is along a line with zero thickness in the $\omega$-$k_z$ space, as shown in Fig. 2. Under such an assumption, a light bullet can propagate indefinitely without any distortion, but such a light bullet has infinite energy. In practice, the spectrum of a light bullet has a nonzero spread along the $k_z$ direction for each $\omega$. Such a nonzero spread is necessary for the light bullet to have finite energy. For such a practical light bullet, the propagation distance is no longer infinite. Assuming axial symmetry around the propagation axis $z$, we consider an incident STWP at $z = 0$ with its complex amplitudes of individual plane-wave components $\tilde{A}$ described as

$$\tilde{A}(k_x, k_y, z = 0, \omega)$$
$$= \exp\left(-\frac{[k_z - k_{z,c}(\omega)]^2}{2\delta k_z^2(\omega)}\right)\exp\left(-\frac{(\omega - \omega_c)^2}{2\Delta\omega^2}\right),$$

where $k_z = \sqrt{\omega^2/c^2 - k_x^2 - k_y^2}$, and $k_{z,c}(\omega)$ is the average space-time correlation for the wave packet. $\omega_c$ is the central frequency of the spectrum. $\Delta\omega \ll \omega_c$ is the frequency bandwidth of the spectrum and is related to the wavelength bandwidth $\Delta\lambda$ of the spectrum. $\delta k_z(\omega)$ is the linewidth of the longitudinal wave vector $k_z$, which is related to the frequency linewidth $\delta\omega$ of the space-time coupling as

$$\delta k_z(\omega) = \frac{\delta\omega}{c\cos(\theta_c)}\left(1 - \frac{c}{v_g}\cos(\theta_c)\right) \qquad (6)$$

as described in Appendix A. Here we assume that $\delta\omega \ll \omega$, which implies that $\delta k_z(\omega) \ll k_z(\omega)$. The pulse $A(x, y, z, t)$ of STWPs after propagation over a distance $z$ is

$$A(x, y, z, t) = \iiint d\omega\, dk_x\, dk_y$$
$$\times \tilde{A}(k_x, k_y, 0, \omega)e^{[i(k_x x + k_y y + k_z z - \omega t)]}. \qquad (7)$$

We define the beam diameter $D$ as the full width at half maximum of the intensity distribution along the transverse
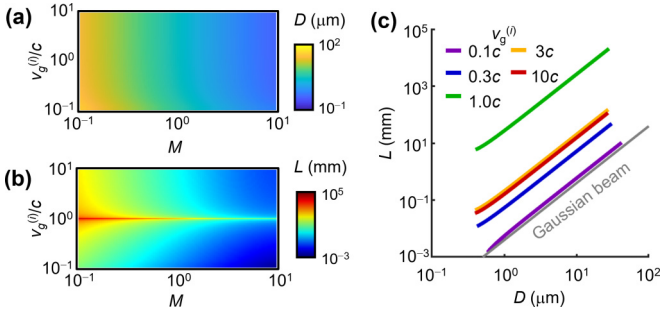
FIG. 4. (a) Calculated beam diameter ($D$) and (b) propagation distance ($L$) of output STWPs. The upper insets illustrate the definition of $D$ and $L$. (c) Combinations of $D$ and $L$ of output STWPs depending on the group velocity of input STWPs ($v_g^{(i)}$). The gray line presents the $D$-$L$ relation for a Gaussian beam with an 800-nm wavelength.

spatial axis at $z = 0$ and $t = 0$. The propagation distance $L$ is defined as the distance at which the maximum intensity attenuated to half of its value at $z = 0$. At any given $z$, the maximum intensity of the pulse is always located at $x = y = 0$ and reached at the time $t = z/v_g$.

We calculate the beam diameter $D$ and the propagation distance $L$ of output STWPs [Figs. 4(a) and 4(b)]. We use the same setup as in Fig. 3 and consider the output resulting from an input STWP passing through the dispersion magnifier. We assume the finite bandwidth of frequencies and wave vectors, which are defined by a wavelength bandwidth of $\Delta\lambda = 0.5$ nm and a frequency linewidth of $\delta\omega = 10^{-4}\omega$. Figure 4(a) presents the calculated beam diameter $D$ for different $v_g^{(i)}$ and $M$. The beam diameter can vary from the sub-millimeter-scale ($10^2$ µm) to the diffraction-limit of the center wavelength of the wave packet ($10^{-1}$ µm). The distribution of the beam diameter is mainly determined by the magnification ratio $M$, which controls the numerical aperture $\text{NA}^{(f)}$ of the output pulse. Figure 4(b) presents the calculated propagation distance $L$ for different $v_g^{(i)}$ and $M$. Unlike the beam diameter, $L$ depends significantly on both $M$ and $v_g^{(i)}$, varying from $10^{-3}$ to $10^5$ mm.

The distribution of $L$ in Fig. 4(b) does not match with the formula $L = c/\delta\omega/|1 - c/v_g|$ as derived in Refs. [5,6] since we are mostly outside the paraxial regime. Instead, the propagation distance $L$ can be expressed as (see Appendix B for derivation),

$$L \approx \frac{\sqrt{\ln 2}}{\delta k_z(\omega_c)} = \frac{c}{\delta\omega} \frac{\sqrt{\ln 2}\cos(\theta_c)}{|1 - c/v_g \cos(\theta_c)|}. \quad (8)$$

Equation (8) agrees well with numerical results, as shown in Appendix B.

We analyze the relation between the beam diameter $D$ and the propagation distance $L$ for the output STWPs. We consider the same set of input STWP as used in Fig. 4(a), with $v_g^{(i)}$ varying from $0.1c$ to $10c$. For each input STWP, we vary the magnification ratio $M$ to vary the output beam diameter $D$. The resulting propagation distance for the output wave packet is shown in Fig. 4(c). For every $v_g^{(i)}$, the propagation distance $L$ increases with the beam diameter $D$. For the same beam diameter, the propagation distance $L$ for all output wave
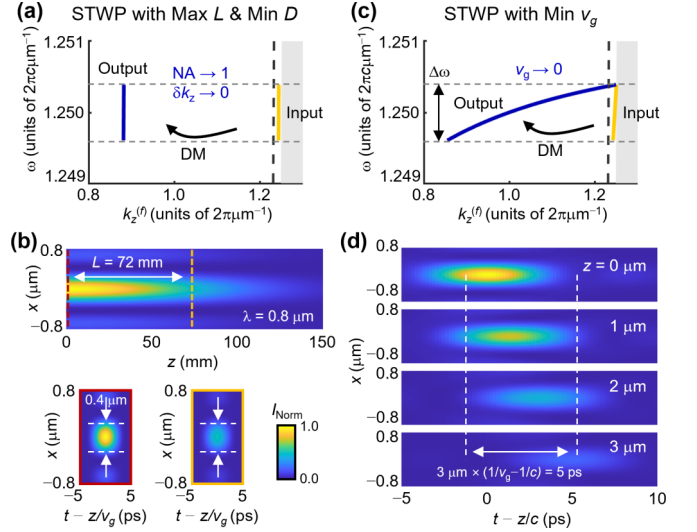


FIG. 5. (a) Dispersion of input (yellow line) and output STWP (blue line) for long-distance propagation of optical pulses with a subwavelength-scale cross section. (b) Numerical demonstration of the output STWP in panel (a). (Upper) Calculated transverse intensity profile along the $z$ axis at the temporal locations of maximum peak intensity $|A(x, 0, z, z/v_g)|^2$. The red (orange) dashed line indicates the beam waist (half-attenuated distance), respectively. (Lower) Calculated beam profile $|A(x, 0, z, t - z/v_g)|^2$ at the beam waist (red box) and the half-attenuated distance (orange box), respectively. (c) Dispersion of input (yellow line) and output STWP (blue line) for optical pulses with slow group velocity. (d) Numerical demonstration of output STWP in panel (c). Calculated beam profiles $|A(x, 0, z, t - z/c)|^2$ at $z = 0, 1, 2,$ and 3 µm, respectively. In panels (a) and (c), the black dashed line indicates the paraxial limit (NA = 0.17). The white (gray-shaded) region corresponds to the region above (below) the light line.

packets is significantly larger than that of a Gaussian wave packet. For the same beam diameter, Eq. (8) indicates that the propagation distance should depend strongly on the final group velocity $v_g^{(f)}$. Here we see a strong dependency on the input group velocity $v_g^{(i)}$ since these two group velocities are directly related as shown in Fig. 3(b).

## VI. APPLICATION OF DISPERSION MAGNIFIER

Finally, we show the capability of the dispersion magnifier to create STWPs with unusual characteristics. In Figs. 5(a) and 5(b), we numerically demonstrate long-distance propagation of optical pulses with subwavelength cross sections. For this purpose, we take an input wave packet used in Fig. 4 in the paraxial limit with a NA of 0.1, with $v_g^{(i)} = 0.9954c$, which has a space-time coupling very close to the light line as indicated by the yellow curve in Fig. 5(a). We pass the input wave packet through a dispersion magnifier with $M = 10$. The output space-time coupling is significantly away from the light line and has a much larger numerical aperture [blue curve, Fig. 5(a)]. The output wave packet has a subwavelength cross section, yet the subwavelength confinement persists even after propagation of approximately 100 mm.

In Figs. 5(c) and 5(d), we numerically demonstrate the creation of optical pulses with slow group velocity. In this

case, we use an input wave packet with a NA of 0.1, and with a $v_g^{(i)} = 0.1c$. Its space-time coupling is in the vicinity of the light line but has a slope that is different from that of the light line as shown by the yellow curve in Fig. 5(c). We pass such an input wave packet through a dispersion magnifier with $M = 10$. The space-time coupling of the output wave packet is shown as the blue curve in Fig. 5(c). It features a group velocity of $v_g^{(f)} = 2.970 \times 10^{-3}c$. [In general, for a given $M$, there is a complex and nonmonotomic dependency of $v_g^{(f)}$ on $v_g^{(i)}$ as noted in Fig. 3(b).] The propagation of the output wave packet is shown in Fig. 5(d). In this plot, with the choice of the horizontal coordinate, a pulse moving at the speed of light will have a constant center location. The shift in the center location as a function of $z$ indicates a slow light effect. From the plot, one can infer a group velocity of $v_g = 1.997 \times 10^{-3}c$, in agreement with the space-time coupling shown in Fig. 5(c).

## VII. CONCLUSION

In conclusion, we theoretically investigate the use of the dispersion magnifier for the spatiotemporal shaping of space-time wave packets. For an input wave packet in the paraxial regime, the output wave packets can be tuned to be in either paraxial or nonparaxial regimes and exhibit a large range of group velocity, beam diameter, and propagation length. Furthermore, we numerically demonstrate long-distance propagation of output wave packets with subwavelength cross sections, as well as the creation of an output wave packet with ultraslow group velocity. Our study provides a general approach for spatiotemporal pulse shaping of space-time wave packets, with implications for classical and quantum information processing [18,19] and in controlling nonlinear light-matter interactions [20–22].

## APPENDIX A: LINEWIDTH OF THE LONGITUDINAL WAVE VECTOR OF STWPs

We derive the linewidth of the longitudinal wave vector of STWPs. First, we define the frequency linewidth $\delta\omega$ of space-time coupling in the $\omega$-$k_t$ space, which is described in Fig. 6(a). The average space-time coupling $k_{t,c}(\omega)$ is plotted as a black line. The $k_{t,+}(\omega)$ and $k_{t,-}(\omega)$ are the space-time coupling with the frequency shift of $\delta\omega/2$ from the average space-time coupling, which are plotted as blue lines. The $k_{t,c}(\omega)$ and $k_{t,\pm}(\omega)$ can be expressed as

$$k_{t,c}^2(\omega) = \left(\frac{\omega}{c}\right)^2 - \left(\frac{\omega - \omega_0}{v_g}\right)^2,$$

$$k_{t,\pm}^2(\omega) = \left(\frac{\omega \pm \delta\omega}{c}\right)^2 - \left(\frac{\omega - \omega_0 \pm \delta\omega}{v_g}\right)^2,$$

where $\omega_0 = \omega_c[1 - \frac{v_g}{c}\cos(\theta_c)]$ and $\theta_c$ is the propagating angle of the plane-wave component with the frequency $\omega_c$. The
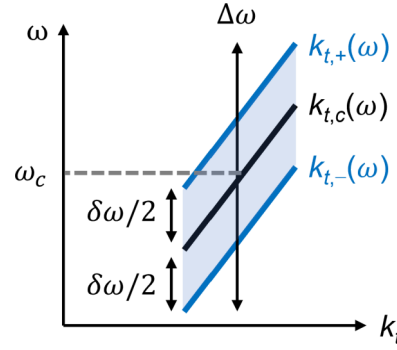


FIG. 6. Schematic illustration of space-time coupling with the finite frequency linewidth $\delta\omega$ in the $\omega$-$k_t$ space. The black curve [$k_{t,c}(\omega)$] indicates the average space-time coupling. The blue curves [$k_{t,+}(\omega)$ and $k_{t,-}(\omega)$] indicate the space-time coupling with the frequency shift of $\pm\delta\omega/2$, respectively. $\omega_c$ is the central frequency of the spectrum. $\Delta\omega$ is the frequency bandwidth of the pulse.

space-time coupling can be described in the $\omega$-$k_z$ space by coordinate transformation as

$$k_{z,c}^2(\omega) = \left(\frac{\omega - \omega_0}{v_g}\right)^2, \tag{A1}$$

$$k_{z,\pm}^2(\omega) = \left(\frac{\omega}{c}\right)^2 - \left(\frac{\omega \pm \delta\omega}{c}\right)^2 + \left(\frac{\omega - \omega_0 \pm \delta\omega}{v_g}\right)^2. \tag{A2}$$

At $\omega = \omega_c$, Eq. (A1) becomes

$$k_{z,c}(\omega_c) = \frac{\omega_c}{c}\cos(\theta_c), \tag{A3}$$

and to the first order of $\delta\omega$,

$$k_{z,\pm}(\omega_c) \approx \frac{\omega_c}{c}\cos(\theta_c)\left[1 \pm \frac{\delta\omega}{\omega_c\cos(\theta_c)^2}\left(\frac{c}{v_g}\cos(\theta_c) - 1\right)\right]. \tag{A4}$$

By combining Eqs. (A3) and (A4), the linewidth of the longitudinal wave vector at the central frequency $\omega_c$ can be obtained as

$$\delta k_z(\omega_c) = 2|k_{z,\pm}(\omega_c) - k_{z,c}(\omega_c)|$$

$$\approx \frac{\delta\omega}{c\cos(\theta_c)}\left|1 - \frac{c}{v_g}\cos(\theta_c)\right|. \tag{A5}$$

## APPENDIX B: DERIVATION OF PROPAGATION DISTANCE OF STWPs

We provide a derivation of the propagation distance of STWPs without making the paraxial approximation. As the maximum intensity of the pulse is always located at $x = y = 0$ and reached at the time $t = z/v_g$, the longitudinal distribution

of the maximum amplitude $A(0, 0, z, z/v_g)$ can be expressed as

$$A(0, 0, z, z/v_g) = \int_0^\infty d\omega \int_{-\infty}^\infty \int_{-\infty}^\infty dk_x dk_y \exp\left(-\frac{(\omega - \omega_c)^2}{2\Delta\omega^2}\right) \exp\left(-\frac{[k_z - k_{z,c}(\omega)]^2}{2\delta k_z^2(\omega)}\right) \exp\left[i\left(zk_z - \frac{z}{v_g}\omega\right)\right]. \qquad \text{(B1)}$$

The coordinate of wave-vector space can be transformed from $(k_x, k_y)$ to $(k_z, \phi)$ as $(k_x, k_y) = [\sqrt{(\omega/c)^2 - k_z^2}\cos\phi, \sqrt{(\omega/c)^2 - k_z^2}\sin\phi]$ and $dk_x dk_y = -k_z dk_z d\phi$. Equation (B1) becomes

$$A(0, 0, z, z/v_g) = \int_0^\infty d\omega \exp\left(-\frac{(\omega - \omega_c)^2}{2\Delta\omega^2} - i\frac{z}{v_g}\omega\right)\left[\int_0^\infty \int_0^{2\pi} k_z dk_z d\phi \exp\left(-\frac{[k_z - k_{z,c}(\omega)]^2}{2\delta k_z^2(\omega)} + izk_z\right)\right]. \qquad \text{(B2)}$$

The integration in Eq. (B2) can be carried out to yield

$$\int_0^\infty \int_0^{2\pi} k_z dk_z d\phi \exp\left(-\frac{[k_z - k_{z,c}(\omega)]^2}{2\delta k_z^2(\omega)} + izk_z\right) = 2\pi \exp\left(ik_{z,c}(\omega)z - \frac{1}{2}\delta k_z^2(\omega)z^2\right)\left[2\sqrt{\pi}\delta k_z^2(\omega)C(\omega)\right], \qquad \text{(B3)}$$

where $C(\omega) = [k_{z,c}(\omega) + iz\delta k_z^2(\omega)]/\sqrt{2}\delta k_z(\omega)$. By combining Eqs. (B2) and (B3) and $\omega - \omega_0 = v_g k_z$, Eq. (B2) becomes

$$A(0, 0, z, z/v_g) = \int_0^\infty d\omega \exp\left(-\frac{(\omega - \omega_c)^2}{2\Delta\omega^2} - i\frac{z}{v_g}\omega_0\right)\left[4\pi^{3/2}\exp\left(-\frac{1}{2}\delta k_z^2(\omega)z^2\right)\delta k_z^2(\omega)C(\omega)\right]. \qquad \text{(B4)}$$

Based on the narrow-band assumption, the dispersion of linewidths of the longitudinal wave vector can be approximated as constant, as $\delta k_z(\omega) \approx \delta k_z(\omega_c)$. Then, Eq. (B4) becomes

$$A(0, 0, z, z/v_g) = 4\sqrt{2}\pi^2 \delta k_z^2(\omega_c)\left(\frac{\omega_c}{c}\cos(\theta_c) + iz\delta k_z^2(\omega_c)\right)\exp\left(-\frac{1}{2}\delta k_z^2(\omega_c)z^2\right), \qquad \text{(B5)}$$

and the intensity distribution $|A(0, 0, z, z/v_g)|^2$ is

$$|A(0, 0, z, z/v_g)|^2 = 32\pi^4 \delta k_z^4(\omega_c)\left(\frac{\omega_c^2}{c^2}\cos^2(\theta_c) + \delta k_z^4(\omega_c)z^2\right)\exp\left[-\delta k_z^2(\omega_c)z^2\right]. \qquad \text{(B6)}$$

As $\delta k_z(\omega_c) \ll (\omega_c/c)\cos(\theta_c)$, Eq. (B6) can be approximated as

$$|A(0, 0, z, z/v_g)|^2 \approx 32\pi^4 \delta k_z^4(\omega_c)\frac{\omega_c^2}{c^2}\cos^2(\theta_c)\exp\left[-\delta k_z^2(\omega_c)z^2\right]. \qquad \text{(B7)}$$

Based on Eq. (B7), the propagation distance of STWPs is defined as the distance $L$ where the intensity attenuates into half compared to $z = 0$, which is expressed as

$$\frac{|A(0, 0, L, L/v_g)|^2}{|A(0, 0, 0, 0)|^2} = \exp\left\{-[\delta k_z(\omega_c)]^2 L^2\right\} = 1/2.$$

Therefore, the propagation distance $L$ is

$$L = \frac{\sqrt{\ln 2}}{\delta k_z(\omega_c)} \approx \frac{c}{\delta\omega}\frac{\sqrt{\ln 2}\cos(\theta_c)}{|1 - c/v_g\cos(\theta_c)|}, \qquad \text{(B8)}$$

where $\delta k_z(\omega_c) = \frac{\delta\omega}{c\cos(\theta_c)}(1 - \frac{c}{v_g}\cos(\theta_c))$, which is derived in Appendix A. In the paraxial regime ($\theta_c \ll 1$), Eq. (B8) becomes

$$L \approx \frac{c}{\delta\omega}\frac{\sqrt{\ln 2}}{|1 - c/v_g|}. \qquad \text{(B9)}$$

Equation (B9) can be compared to the expression $L = c/\delta\omega|1 - c/v_g|$ derived in Ref. [5] for STWPs in the paraxial regime. The factor of $\sqrt{\ln 2}$ difference results from the different definitions of propagation distance in our paper and the reference. Reference [5] defines the propagation distance in terms of a walk-off between the wave packet and the pilot

wave. Here we define the propagation distance in terms of the energy attenuation.

As a validation of Eq. (B8), we plot the relation between $L$ and $\sqrt{\ln 2}/\delta k_z(\omega_c)$ in Fig. 7 as blue dots for each $v_g^{(i)}/c$ and $M$ used in Fig. 4(b) in the main text. The numerical results agree excellently with Eq. (B8). [We note that to apply in Eq. (B8) for our case the $v_g$ in Eq. (B8) is the $v_g^{(f)}$ for each wave packet.]
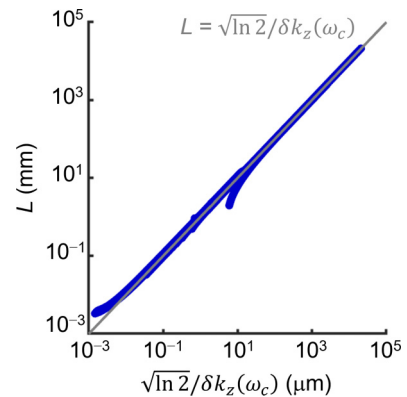


FIG. 7. Comparison between the propagation distance $L$ in Fig. 4(b) and the analytic expression $\sqrt{\ln 2}/\delta k_z(\omega_c)$. The gray line indicates the condition of $L = \sqrt{\ln 2}/\delta k_z(\omega_c)$.

[1] M. Yessenov, L. A. Hall, K. L. Schepler, and A. F. Abouraddy, Adv. Opt. Photon. **14**, 455 (2022).

[2] K. Stefańska, P. Béjot, K. Tarnowski, and B. Kibler, ACS Photon. **10**, 727 (2023).

[3] C. Guo and S. Fan, Phys. Rev. Res. **3**, 033161 (2021).

[4] H. E. Kondakci and A. F. Abouraddy, Nat. Photon. **11**, 733 (2017).

[5] M. Yessenov, L. Mach, B. Bhaduri, D. Mardani, H. E. Kondakci, G. K. Atia, M. A. Alonso, and A. F. Abouraddy, Opt. Express **27**, 12443 (2019).

[6] C. Guo, M. Xiao, M. Orenstein, and S. Fan, Light: Sci. Appl. **10**, 160 (2021).

[7] H. E. Kondakci and A. F. Abouraddy, Nat. Commun. **10**, 929 (2019).

[8] B. Bhaduri, M. Yessenov, and A. F. Abouraddy, Nat. Photon. **14**, 416 (2020).

[9] M. Yessenov, J. Free, Z. Chen, E. G. Johnson, M. P. J. Lavery, M. A. Alonso, and A. F. Abouraddy, Nat. Commun. **13**, 4573 (2022).

[10] F. L. Pedrotti, L. S. Pedrotti, and L. M. Pedrott, in *Introduction To Optics* (Cambridge University Press, Cambridge, England, 2017), Chap. 2, p. 28.

[11] S.-K. Liao, H.-L. Yong, C. Liu, G.-L. Shentu, D.-D. Li, J. Lin, H. Dai, S.-Q. Zhao, B. Li, J.-Y. Guan *et al.*, Nat. Photon. **11**, 509 (2017).

[12] H. J. Kang, J. Yang, B. J. Chun, H. Jang, B. S. Kim, Y.-J. Kim, and S.-W. Kim, Nat. Commun. **10**, 4438 (2019).

[13] F. O. Fahrbach and A. Rohrbach, Nat. Commun. **3**, 632 (2012).

[14] S. Takanezawa, T. Saitou, and T. Imamura, Nat. Commun. **12**, 2979 (2021).

[15] B. E. A. Saleh and M. C. Teich, in *Fundamentals of Photonics* (Wiley & Sons, New York, 1991), Chap. 3.2, pp. 90–91.

[16] M. A. Porras, Phys. Rev. A **97**, 063803 (2018).

[17] Z. Li, Y. Liu, Y. Leng, and R. Li, Sci. Rep. **12**, 16102 (2022).

[18] S. M. Walsh, S. F. E. Karpathakis, A. S. McCann, B. P. Dix-Matthews, A. M. Frost, D. R. Gozzard, C. T. Gravestock, and S. W. Schediwy, Sci. Rep. **12**, 18345 (2022).

[19] M. McLaren, M. Agnew, J. Leach, F. S. Roux, M. J. Padgett, R. W. Boyd, and A. Forbes, Opt. Express **20**, 23589 (2012).

[20] M. Soljačić, S. G. Johnson, S. Fan, M. Ibanescu, E. Ippen, and J. D. Joannopoulos, J. Opt. Soc. Am. B **19**, 2052 (2002).

[21] B. Corcoran, C. Monat, C. Grillet, D. J. Moss, B. J. Eggleton, T. P. White, L. O'Faolain, and T. F. Krauss, Nat. Photon. **3**, 206 (2009).

[22] P. Ornelas, I. Nape, R. de Mello Koch, and A. Forbes, Nat. Photon. **18**, 258 (2024).