Radiation guiding with surface plasmon polaritons

Zhanghua Han and Sergey I Bozhevolnyi

Institute of Technology and Innovation, University of Southern Denmark, Odense M, DK-5230 Denmark
E-mail: seib@iti.sdu.dk

Received 4 July 2012, in final form 31 October 2012
Published 19 December 2012
Online at stacks.iop.org/RoPP/76/016402

Abstract
Surface plasmon polaritons (SPPs) are electromagnetic (EM) modes propagating along metal–dielectric interfaces, in which surface collective excitations of free electrons in the metal are coupled to evanescent EM fields in the dielectric. Various SPP modes can be supported by flat and curved, single and multiple surfaces, exhibiting remarkable properties, including the possibility of concentrating EM fields beyond the diffraction limit, i.e. on the nanoscale, while enhancing local field strengths by several orders of magnitude. This unique feature of SPP modes, along with the ever-increasing demands for miniaturization of photonic components and circuits, generates an exponentially growing interest in SPP-mediated radiation guiding and SPP-based waveguide components. Here we review the current status of this rapidly developing field, starting with a brief presentation of the main planar SPP modes along with the techniques employed for their excitation and manipulation by sets of nanoparticles. We then describe in detail various SPP-based waveguide configurations that ensure two-dimensional mode confinement in the plane perpendicular to the propagation direction and compare their characteristics. Excitation of SPP waveguide modes and recent progress in the development of SPP-based waveguide components are also discussed, concluding with our outlook on challenges and possible future developments in this field.

(Some figures may appear in colour only in the online journal)

This article was invited by Horst-Guenter Rubahn.

Contents

1. Introduction 2
2. Planar SPP modes 2
2.1. Surface plasmon polaritons 3
2.2. SPP modes of thin metal films 4
2.3. Gap SPP modes 6
3. Micro-optics of planar SPP waves 7
3.1. SPP excitation 7
3.2. SPP manipulation by sets of nanoparticles 9
4. SPP waveguides 10
4.1. SPP guiding by dielectric ridges 10
4.2. SPP guided by metal stripes 13
4.3. Slot and gap waveguides 16
4.4. Channel plasmon polaritons 18
4.5. SPP guiding by nanowires 20
4.6. SPP guiding along metal wedges 22
4.7. Comparison of SPP waveguides 23
5. Excitation of SPP waveguide modes 26
5.1. Prism and grating coupling 26
5.2. Excitation with strongly focused radiation 27
5.3. Fiber end-fire coupling 27
5.4. Hybrid integration (coupling via dielectric waveguides) 28
6. SPP waveguide components 29
6.1. Passive SPP components (bends and splitters) 29
6.2. Wavelength selective SPP components 30
6.3. Radiation modulation with SPP components 30
6.4. SPP generation and detection 31
6.5. Quantum plasmonics 31
7. Outlook 32
Acknowledgment 33
References 33
1. Introduction

Surface plasmon polaritons (SPPs), often shortened to surface plasmons (SPs), represent electromagnetic (EM) excitations, which are coupled to surface collective oscillations of free electrons in a metal, thereby forming two-dimensional (2D) bound waves propagating along metal–dielectric interfaces and exponentially decaying into neighboring media [1, 2]. The topology of the metal surface determines the characteristics of SP modes ranging from localized SPs (LSPs) of individual particles to various propagating SPPs existing at flat and curved, and single and multiple surfaces, including SPP modes of complex particle arrays and metal nanostructures [1, 2]. The field of plasmonics [3], which has been established and explosively expanded during the last ten years, taking advantage of the rapid progress in nanoscience and nanotechnology [4–6], comprises various physical phenomena associated with localized and propagating SP modes [7]. LSP modes bring about unique plasmonic phenomena of subwavelength light confinement and EM field enhancement, often resulting together in extreme light concentration, i.e. in strongly localized (within nm-sized volumes) giant EM fields [8, 9]. SPP modes, whose propagation length is limited by inevitable EM absorption in metals, can also be strongly localized in the cross section perpendicular to the propagation direction [10]. Recognition of this unique feature of SPP-based waveguides, implying that plasmonic waveguides could transport the same huge bandwidth of information as in conventional (dielectric-based) photonics and yet not be limited by diffraction to sub-micrometer cross sections, has attracted enormous attention to the field of plasmon-based nanophotonics [11–13].

Many different SPP-based waveguide configurations, each offering specific advantages and suffering from particular limitations, have been developed over a decade of intensive research into plasmon-based nanophotonics, a field whose exploration has been primarily driven by the tantalizing perspective of combining the benefits of the compactness of an electronic circuit and the bandwidth of a photonic network [10]. There appears, however, to be a common feature, namely the trade-off between the mode confinement and propagation loss, which is found in all plasmonic waveguides [12]. Since it is the metal boundaries which enforce squeezing the EM fields of SPP modes in the lateral cross section (perpendicular to the propagation direction), a decrease in the SPP mode cross section is accompanied by an increase of a fraction of the EM energy being concentrated inside a metal, where it gets absorbed due to a nonzero (and positive) imaginary part of the metal permittivity responsible for ohmic loss. However, the general trend of stronger SPP mode confinement leading to shorter propagation lengths turns out to be quite different for different SPP waveguide configurations, being also strongly influenced by the approach used to determine the mode confinement [14]. Another general trend observed in SPP waveguides is related to the properties of dielectrics being brought into contact with metals in the SPP waveguide configurations: dielectrics (assumed to be lossless) with higher refractive indices result in a larger propagation loss (i.e. in shorter propagation distances), due to the smaller field penetration depths in dielectrics [15] and thereby larger parts of the mode power residing in the metal. The selection of dielectrics to be employed can be governed by the intention to realize a particular functionality, making use of material effects such as thermo- and electro-optical, but the aforementioned trend should be kept in mind when arriving at the final choice of dielectrics.

This review aims to provide both an introduction to a variety of SP waveguide configurations, whose very extensive list (being constantly extended further) can hardly be covered within the limited space of this review, and a careful consideration of the advantages and limitations of most developed SPP waveguide configurations, that should facilitate the choice and design of a particular configuration for a specific application. We start with reviewing the main characteristics of SPP modes supported by the simplest planar multilayer configurations (section 2). SPPs supported by an individual metal–dielectric interface were historically the first SPP modes whose excitation, propagation and routing in the surface plane using metal nanostructures have been extensively investigated for the purposes of realizing SP circuitry [1, 11]. This research is briefly overviewed in section 3. Next we consider the SPP waveguide configurations that ensure 2D mode confinement in the lateral cross section by using 3D confined geometries of metals and/or dielectrics involved. Section 4 is concluded with a comparison of different waveguide configurations with respect to their characteristics (mode confinement, propagation length, cross talk and bend loss) that are most important for the realization of SPP waveguide components and circuits. Section 5 is devoted to a brief review of those SPP waveguide excitation techniques, with an emphasis on coupling from dielectric waveguides, which provides a possible solution to alleviate the propagation loss of SPP waveguides by integrating photonic and plasmonic waveguides on the same platform. Progress in the realization of various SPP components to provide both passive and active functionalities as well as devices to generate and detect SPP signals is discussed in section 6. The whole paper ends in section 7 with our outlook on the future challenges and possible research directions of plasmonics for radiation guiding and information transmission.

2. Planar SPP modes

Most metals are opaque at frequencies below ultraviolet because the real part of their dielectric susceptibilities is negative, making the propagation of transverse electromagnetic (TEM) waves impossible. Consequently, an EM wave incident on a metal–dielectric interface from the side of the dielectric is almost totally reflected, being partially absorbed due to a nonzero and positive imaginary part of the dielectric permittivity responsible for ohmic loss. One can, however, excite a bound (surface) EM mode propagating along a metal–dielectric interface with the propagation constant exceeding that of the EM waves propagating in the dielectric. This mode, called a SPP, features a transverse magnetic (TM) polarized EM field, i.e. polarized in the propagation plane, that decays
expansively into both neighboring media [7, 16]. In the general case of 2D multilayer configurations, SPP modes can be viewed as hybrid modes with primarily TEM waves in dielectrics being coupled (near metal–dielectric interfaces) to free collective electron (plasma) oscillations in metals, which are primarily longitudinal. It is then clear that the transverse electric (TE) waves, which are polarized perpendicular to the propagation plane and thereby purely transverse (in the sense that \( \nabla \cdot E = 0 \) everywhere), cannot be coupled to the longitudinal electron oscillations in metals and form SPP modes [17]. The same conclusion can be obtained by considering the boundary conditions that require the continuity of the tangential (parallel to an interface) components of EM fields across the interface, implying the continuity of the lateral derivative (normal to the interface) of the electric field of the TE wave, a condition which is impossible to satisfy for bound modes [7, 16]. In the following subsections, we consider only the TM (p-polarized) waves and the corresponding SPP modes in single- and double-interface configurations.

2.1. Surface plasmon polaritons

SPP modes supported by an individual metal–dielectric interface, usually called simply SPPs, can be viewed as consisting of two evanescent waves decaying exponentially into both neighboring media (figure 1). Since the tangential field components are continuous across the interface (implying thereby that the wave-vector component \( k_x \) parallel to the interface is conserved), the SPP electric field can be written as follows

\[
E^{\pm}(z) = (E_x^{\pm}(0), 0, E_z^{\pm}) = \begin{cases} E_x^0 \cos(k_{SPP}z), & z > 0 \\ E_x^m \cos(k_{SPP}z), & z < 0 \end{cases} \exp(-z\sqrt{k^2_{SPP} - \varepsilon_z}),
\]

where \( E_x^0 \) and \( E_x^m \) are the amplitudes of the corresponding electric-field components in the dielectric (metal) medium, \( \varepsilon_z \) and \( \varepsilon_z \) are the permittivities of the dielectric and metal, respectively, \( k_0 = 2\pi/\lambda \) is the light wavenumber, and \( k_{SPP} = k_x \) is the mode propagation (complex) constant yet to be determined. Since the electric-field components should satisfy Coulomb’s law \( \nabla \cdot E = 0 \) (both in the upper and lower half-spaces), the normal (perpendicular to the interface) field components can be related to the tangential one:

\[
E_z^0 = \frac{ik_{SPP}}{\sqrt{k^2_{SPP} - \varepsilon_z k_0^2}} E_x^0 \quad \text{and} \quad E_z^m = -\frac{ik_{SPP}}{\sqrt{k^2_{SPP} - \varepsilon_m k_0^2}} E_x^m.
\]

Finally, using the appropriate boundary condition for the normal electric-field components: \( \varepsilon_d E_z^d = \varepsilon_m E_z^m \), results in the SPP dispersion relation:

\[
k_{SPP}(\omega) = \frac{\omega}{c} \sqrt{\frac{\varepsilon_m(\omega)\varepsilon_d}{\varepsilon_m(\omega) + \varepsilon_d}}.
\]

The necessary conditions for an SPP (as a bound and propagating SP mode) to exist can be obtained from the requirement of the square root expressions in equations (1a), (1b) and (3) to have a positive real part, leading to

\[
\Re[\varepsilon_d - \varepsilon_m(\omega)] < 0 \quad \text{and} \quad \Re[\varepsilon_d + \varepsilon_m(\omega)] < 0,
\]

\[
\Rightarrow \Re[\varepsilon_m(\omega)] < -\varepsilon_d.
\]

For most metals and dielectrics, this condition is satisfied in the long-wavelength part of the visible and in the infrared. The real part of the propagation constant determines the SPP wavelength \( \lambda_{\text{SPP}} = 2\pi/|k_{\text{SPP}}| \), which is always smaller than the light wavelength in the dielectric: \( \lambda_{\text{SPP}} < \lambda = \lambda_0/\sqrt{\varepsilon_d} \), and the SPP (intensity) propagation length \( L_{\text{SPP}} = 1/\text{Im}(2k_{\text{SPP}}) \). Using the free-electron (Drude) approximation for the metal permittivity: \( \varepsilon_m(\omega) = 1 - (\omega_p/\omega)^2 \) with \( \omega_p \) being the plasma frequency, one can work out the frequency dependencies of the SPP characteristics in the lossless limit [7, 16, 17]. It is seen, for example, that the SPP wavelength approaches zero at the high-frequency (short-wavelength) limit of the SPP dispersion, i.e. when \( \omega \to \omega_p/\sqrt{1 + \varepsilon_d} \). The corresponding limiting frequency of SPPs, which satisfies the requirement: \( \varepsilon_m(\omega) + \varepsilon_d = 0 \) (equation (3)), is also the frequency of electrostatic surface waves, called SPs, whose dispersion relation can be obtained using Laplace’s equation for a scalar potential [1].

In practice, the SPP damping caused by ohmic losses (i.e. damping of the electron oscillations associated with the SPP) is a very important factor, limiting both the extent of the decrease of SPP wavelength (for short wavelengths) and the SPP propagation length (figure 2). The latter circumstance, i.e. a rather limited SPP propagation length, poses a serious problem when developing SPP-based waveguide components for practical applications. For example, the normalized SPP propagation length for the gold–air interface at the telecom wavelength of 1.55 \( \mu \text{m} \) can be estimated using the permittivity \( \varepsilon_m = -132 + 12.7i \) [18] as \( L_{\text{SPP}}/\lambda_{\text{SPP}} \approx 220 \), indicating thereby the size of playground available for realizing SPP-based photonic components. This ratio (figure 2) also pertains to the upper limit for the quality factor of the corresponding SPP-based resonators or interferometers, which is given by \( Q_{\text{SPP}} = 2\pi L_{\text{SPP}}/\lambda_{\text{SPP}} \).

The SPP dispersion relation (equation (3)) allows one to explicitly relate the amplitudes of normal and
tangential electric-field components in the dielectric and metal (equation (2)):
\[
E_{z}^{d} = i\sqrt{-\varepsilon_{m}/\varepsilon_{d}}E_{x}^{0} \quad \text{and} \quad E_{z}^{m} = -i\sqrt{-\varepsilon_{d}/\varepsilon_{m}}E_{x}^{0}.
\]
Typically, metals exhibit large magnitudes of the dielectric permittivity, i.e. $|\varepsilon_{m}| \gg \varepsilon_{d}$, resulting in the normal component being dominant in the dielectric and the tangential component being dominant in the metal. The SPP electric field is thereby primarily transverse in the dielectric and longitudinal in the metal, reflecting the hybrid nature of SPPs that combine the features of both the propagating EM waves in dielectrics and free-electron oscillations in metals [10, pp 1–31]. Since SPP damping occurs due to ohmic losses in the metal, it is the longitudinal electric-field component of the SPP in the metal that determines the SPP damping, a circumstance which is important for understanding the properties of SPP modes supported by thin metal films considered in the next subsection.

The penetration depths (or the decay lengths) of the SPP electric field in the dielectric $d_{d}$ and metal $d_{m}$ are also different as seen from the corresponding SPP field distributions (equations (1a) and (1b)) and can be explicitly expressed by making use of the dispersion relation (equation (3)) as follows (for the $1/e$ amplitude decrease):

\[
d_{d(m)} = |k_{x,d(m)}|^{-1} = \frac{\lambda}{2\pi} \sqrt{\frac{\varepsilon_{m} + \varepsilon_{d}}{\varepsilon_{d(m)}^{2}}}.
\]

so that $\varepsilon_{d}d_{d} = |\varepsilon_{m}|d_{m}$. The latter relation implies that the field penetration in the metal is typically much shorter than that in the dielectric. It is interesting that the penetration depth in metals depends rather weakly on the wavelength staying at the level of a few tens of nanometers, while that in dielectrics increases fast and nonlinearly with the wavelength. These trends can be accounted for by making use of the Drude expression for the dielectric permittivity $\varepsilon_{m}(\omega)$ together with equation (6), resulting in the following asymptotic expressions for the penetration depths in the long-wavelength limit [10, pp 1–31]:

\[
d_{m} \approx c/\omega_{p}, \quad \text{and} \quad d_{d} \approx (\lambda/2\pi)^{2} \cdot (\omega_{p}/c\varepsilon_{d}).
\]

Finally, using the above relations (equations (6) and (7)), one can approximate the fraction of the SPP electric-field energy in the dielectric as simply being proportional to $(|\varepsilon_{m}|/\varepsilon_{d})^{2}$, which is somewhat counterintuitive as, normally, light is drawn toward regions with a higher refractive index. The obtained expression signifies that the SPP electric field is progressively being forced out of the metal for longer wavelengths since, in the Drude approximation, $|\varepsilon_{m}| \propto \omega^{-2}$. One can also confirm here the trend, which we have mentioned in section 1, that dielectrics (assumed to be lossless) with higher refractive indices result in larger propagation losses because more mode power is located in the metal.

### 2.2. SPP modes of thin metal films

Many different SPP modes can be found in multiple-interface systems, when the SPPs associated with individual metal–dielectric interfaces start interacting with each other [17, 19–21]. Considering the SPP modes associated with two metal–dielectric interfaces, one finds that the SPP modes can be supported by either a thin metal film abutted with dielectrics or a thin dielectric layer surrounded by metals, often called the insulator–metal–insulator (IMI) or metal–insulator–metal (MIM) structures, correspondingly [22]. We commence by considering the SPP modes in the symmetric IMI configuration of a thin metal film with the thickness $t$ being embedded in the dielectric (extending indefinitely on both sides of the film). When two identical SPP modes start overlapping with each other for small thicknesses of the metal film, the propagation constants of the symmetric and anti-symmetric (individual) mode combinations become different [17, 23]. Note that the symmetric (with respect to the film mid-plane) configuration of the transverse component $E_{z}$ or $H_{z}$ corresponds to the anti-symmetric one for the longitudinal components $E_{x}$ and vice versa (see equation (2)). Hereafter the $E_{z}$-symmetric or $H_{z}$-symmetric configuration is called the symmetric SPP mode [19]. Since the SPP damping is mainly determined by the longitudinal SPP electric-field component, the symmetric SPP mode exhibiting the odd symmetry of $E_{z}$, which thereby crosses zero (changing its sign) at the mid-plane of the metal film, experiences considerably smaller attenuation than the anti-symmetric SPP mode, which is the $E_{x}$-symmetric configuration. The symmetric SPP mode is therefore conventionally called the long-range SPP (LR-SPP) and the anti-symmetric one—the short-range SPP (SR-SPP) [10, 19], albeit the opposite terminology (related to the symmetry of the longitudinal SPP component) can also be found in use [23, 24].

#### 2.2.1. Long-range SPPs

Using the appropriate boundary conditions for the normal and tangential electric-field components, as in the case of SPPs and the aforementioned symmetry of the corresponding field component distributions,
allows one to obtain the LR-SPP dispersion relation:

\[
\tanh (k_{z}^{(m)} t/2) = - (\varepsilon_m k_{z}^{(d)}/\varepsilon_d k_{z}^{(m)})
\]

with

\[
k_{z}^{(m,d)} = -\sqrt{k_{LRSSP}^{2} - \varepsilon_m d k_{0}^{2}}.
\]  

(8)

Note that, unlike the case of SPPs, the LR-SPP propagation constant \( k_{LRSSP} \) cannot be presented in an explicit form. However, for sufficiently thin metal films (\( t \to 0 \)), one can use the approximation \( \tanh x \approx x \) resulting in the following simple expression:

\[
k_{LRSSP} \approx k_0 \sqrt{\varepsilon_d + (tk_0 \varepsilon_d/2)^2} \cdot [1 - (\varepsilon_d/\varepsilon_m)]^2.
\]  

(9)

It is seen that, in the limit of very thin films, the LR-SPP propagation constant approaches the light line in the dielectric, signifying that the LR-SPP field being forced out of the metal (in metal the main field component \( E_z \) crosses zero) spreads progressively into the dielectric (equation (1a)). Note that both the real and imaginary part of the difference between the LR-SPP propagation constant and that of the light in the dielectric decreases quadratically with the decrease in the film thickness (equation (9)), implying a rapid evolution of the LR-SPP into a purely photonic mode [24]. At the same time and in the same manner, the LR-SPP propagation loss decreases to zero and the LR-SPP propagation length increases infinitely, a remarkable feature that allows one to employ the LR-SPPs for developing practical plasmonic components for integrated optics [25].

The trend of decreasing the LR-SPP propagation loss when employing progressively thinner metal films is illustrated (figure 3) by considering the LR-SPP propagation at the telecom wavelength of 1550 nm (equation (3)).

The propagation loss for thin metal films is similar to that of gold [18], the propagation loss below 0.1 dB mm\(^{-1}\) is feasible. Finally, it should be borne in mind that the requirement of a symmetric dielectric environment is essential to maintain the LR-SPP propagation loss at a minimum level. For example, for a 10 nm thick gold film embedded in BCB, the LR-SPP mode was found to have the propagation loss increasing from 0.17 dB mm\(^{-1}\) for the symmetrical structure to \(-0.4\) dB mm\(^{-1}\) for the refractive index difference between the top and bottom cladding layers of \(-0.006\) [25].

2.2.2. Short-range SPPs. The SR-SPP field distribution exhibits the symmetry opposite to that of LR-SPP, resulting (along with the appropriate boundary conditions) in the following dispersion relation:

\[
\tanh (k_{z}^{(m)} t/2) = -(\varepsilon_d k_{z}^{(m)}/\varepsilon_m k_{z}^{(d)})
\]

with

\[
k_{z}^{(m,d)} = \sqrt{k_{SRSSP}^{2} - \varepsilon_m d k_{0}^{2}}.
\]  

(10)

which can be simplified for thin metal films as follows:

\[
k_{SRSSP} \approx k_0 \sqrt{\varepsilon_d + (2\varepsilon_d/(tk_0 \varepsilon_d))^2}.
\]  

(11)

Contrary to the above case of LR-SPPs, the SR-SPP propagation constant \( k_{SRSSP} \) increases indefinitely with its real and imaginary parts:

\[
k_{SRSSP} \approx -(2\varepsilon_d/(tk_0 \varepsilon_d)) \to \infty
\]

when the metal film thickness \( t \) decreases to zero. This means that both the SR-SPP wavelength and propagation length decrease, approaching zero for infinitely thin films. It can also be shown that, in this limit, their ratio tends to a constant determined by the ratio between the real and imaginary parts of the metal susceptibility:

\[
L_{SRSSP}/\lambda_{SRSSP} \to [Re(\varepsilon_m)]/[4\pi Im(\varepsilon_m)],
\]

a ratio that can be rather large in the infrared. Furthermore, the SR-SPP field efficiently filling the metal (the main component \( E_z \) is nearly constant) is also tightly bound to the metal film (equation (1a)). Since, at the same time, the main SR-SPP component in the dielectric (\( E_z \)) is anti-symmetric, the SR-SPP should exhibit a high reflectivity at the film termination, due to an impedance mismatch with the propagating EM waves (due to asymmetry and a very short field penetration in the dielectric). This non-trivial feature can be successfully exploited when designing various plasmonic resonator structures [26].

Concluding this section, let us consider the properties of the LR- and SR-SPP modes supported by a thin gold film surrounded by air at the excitation wavelength of 775 nm. Using the corresponding gold permittivity \( \varepsilon_m = -23.6 + 1.69i \) [18] and the appropriate dispersion relations (equations (8) and (10)), the mode effective index (real part) and propagation length were calculated for both SPP modes (figure 4). Note that, in both cases, one has to solve a transcendental equation for a complex solution and that the approximate values obtained directly from equations (9) and (11) can only be used for very thin (<40 nm) films. It turned out, however, that a very good approximation can be found in the closed form by making use of the following simplification:

\[
\sqrt{k_{LRSSP}^{2} - \varepsilon_m k_0^{2}} \approx k_0 \sqrt{\varepsilon_d - \varepsilon_m}
\]

in both equations (8) and (10). The corresponding dispersion relations can then be
written down as follows [27]:

\[
k_{LSR-SPP} \approx k_0 \times \sqrt{\varepsilon_d + (\varepsilon_d - \varepsilon_m)(\varepsilon_d/\varepsilon_m)^2 \cdot \tanh^{-12}(0.5k_0t/\sqrt{\varepsilon_d - \varepsilon_m})}
\]

(12)

It is seen that the explicit solutions above result in accurate values of the mode effective index and propagation loss for both LR- and SR-SPPs in the whole range of film thicknesses, down to the thickness of 10 nm (figure 4).

2.3. Gap SPP modes

Let us now consider the SPP modes in the symmetric MIM configuration of a thin dielectric layer with the thickness \( t \) being sandwiched between two metal surfaces (with metal extending indefinitely on both sides of the dielectric). When two identical SPP modes start overlapping with each other for small layer thicknesses, the propagation constants of the symmetric and anti-symmetric (individual) mode combinations become different [17, 20]. Moreover, within a certain range of parameters, a symmetric MIM waveguide configuration can support simultaneously three guided SPP modes, so that in addition to the (usual) forward propagating symmetric and anti-symmetric SPP modes one can find a backward propagating anti-symmetric SPP mode [28]. Backward modes in SPP waveguides are somewhat exotic due to very large propagation losses, but might be of interest for realizing negative refraction at optical frequencies [29, 30]. However, the only SPP mode surviving for all values \( t \) of the gap between metal surfaces, the so-called gap SPP (G-SPP), is the mode exhibiting an odd symmetry of the longitudinal electric-field component \( E_z \) and, consequently, an even symmetry of the transverse field component \( E_t \) (figure 1). The corresponding MIM configurations, often also called plasmonic slot waveguides, have been extensively studied both theoretically [31] and experimentally [32] due to the strong mode confinement that can be achieved while ensuring fairly low attenuation.

Applying the appropriate boundary conditions for the normal and tangential electric-field components as in the case of SPPs and the aforementioned symmetry of the corresponding field component distributions, one obtains the following G-SPP dispersion relation [17, 22]:

\[
\tanh(k_{G-SPP}t/2) = -\frac{\varepsilon_d E_t(m)}{\varepsilon_m k_{G-SPP}}
\]

with

\[
k_{G-SPP}^{m,d} = \sqrt{k_{G-SPP}^2 - \varepsilon_m d_k^2}.
\]

(13)

where \( k_{G-SPP} \) denotes the propagation constant of the fundamental G-SPP mode with the transverse field component \( E_t \) having the same sign across the gap. The G-SPP properties determine the SPP guiding in different channel structures, including the SPP guiding in V grooves and trenches [33], which are considered in detail elsewhere [34]. The most important G-SPP feature is that, similarly to the above considered case of SR-SPP, the G-SPP propagation constant increases indefinitely with its real and imaginary parts: \( k_{G-SPP} \rightarrow \infty \) as \( t \rightarrow \infty \), while the gap width \( t \) decreases to zero [27]. This means that the G-SPP mode can be squeezed indefinitely in its lateral cross section, albeit at the price of decreasing the G-SPP propagation length, reflecting the general trade-off in plasmonics between the mode confinement and propagation loss.

General trends in the G-SPP characteristics are illustrated by numerically solving equation (13) for MIM waveguides formed by air gaps in gold at several light wavelengths in the interval between visible and telecom wavelengths (figure 5).

The following dielectric constants of gold were used in the simulations: \( n = 0.166 + 3.15i (\lambda = 653 \text{ nm}) \), \( 0.174 + 4.86i (775 \text{ nm}) \), \( 0.272 + 7.07i (1033 \text{ nm}) \) and \( 0.55 + 11.5i (1550 \text{ nm}) \) [18]. It is interesting that the G-SPP propagation length first increases when the gap width decreases from large values corresponding to uncoupled individual SPPs. This surprising fact indicates that G-SPP modes with better confinement can exhibit larger propagation lengths [35]. The longest G-SPP propagation length is \( \approx 8\% \) larger than the
SPP propagation length (for \( t \to \infty \)) independently on the wavelength, though it is achieved for different gaps that can be evaluated simply as \( l_{\text{eff}} \approx 0.5\lambda |\Re(\varepsilon_m)| + 1^{0.5} \). The explanation of this remarkable effect is related to the fact that, with the decrease of the gap width, the electric field of the G-SPP (composed of two exponents) approaches quickly to the electrostatic (capacitor) mode, which is constant across the gap (see the inset in figure 5). Consequently, the fractional electric-field energy concentrated in the gap first increases (and the damping in the metal decreases) when the gap width decreases, reaches its maximum and then starts to decrease with the field being squeezed from the gap into the metal. An exact analysis of the process of energy redistribution is quite complicated because the corresponding dispersion relation is implicit (equation (13)), but one can actually obtain a very good explicit expression for the fractional electric-field energy in the gap, resulting in practically the same formula for the optimum gap width [34].

The G-SPP characteristics discussed above indicate that the G-SPPs exploit in the most efficient way the available dielectric space (gap) between the metal walls, thereby minimizing the absorption loss [34, 35]. This remarkable feature stimulates investigations of various G-SPP-based geometries, from configurations ensuring 2D mode confinement in the lateral cross section by laterally varying the gap width, such as the SPP gap [36] and V-groove [37] waveguides, to configurations ensuring radiation nanofocusing [38] and plasmonic analogs of black holes [39]. These configurations advantageously exploit the fact that the G-SPP effective index is strongly dependent on the gap width, increasing drastically with its decrease. In order to facilitate the consideration and assessment of various G-SPP-based configurations, it is highly desirable to find an explicit expression for the G-SPP propagation constant that can be used instead of the G-SPP dispersion given by equation (13). It turned out that, for not too small gaps, i.e. when \( t > (\lambda \varepsilon_m)/(\pi |\varepsilon_m|) \), the G-SPP propagation constant can be approximated as follows [40]:

\[
k_{\text{G-SPP}} \approx k_0 \sqrt{\varepsilon_d + \frac{2\varepsilon_d\varepsilon_m - \varepsilon_m}{k_0(|\varepsilon_m|)}}.
\]  

The G-SPP characteristics calculated for moderately wide air gaps in gold at several wavelengths using the exact (implicit) dispersion relation (equation (13)) and the above explicit (analytic) formula (figure 6) indicate that the obtained approximation (equation (14)) gives fairly accurate values for both the G-SPP propagation length and effective index, in fact progressively more accurate for longer wavelengths because of larger values of \( |\varepsilon_m| \). The validity of this simple formula and its usefulness for the analysis of different G-SPP-based waveguide configurations has further been demonstrated in a number of publications [39, 41–43].

3. Micro-optics of planar SPP waves

The first plasmonic analogs of optical components were developed using the concept of micro-optics of planar SPP waves [44]. Planar SPP waves literally mean they are SPP waves on an infinite planar metal/dielectric interface discussed in section 2.1, as opposed to those supported by, e.g., metal stripes and surrounding dielectrics. These SPP waves are only confined in one dimension normal to the metal/dielectric interface, resulting from the intrinsic properties of SPP waves exponentially decaying in that direction. In this context, manipulation of planar SPP waves can be categorized to 2D optics [44, 45]. Although the confinement is different from that provided by the 3D plasmonic waveguides (featuring 2D mode confinement in the plane perpendicular to the propagation direction), which will mainly be discussed in the next section, planar SPP waves provide the transport of EM radiation along metal/dielectric interfaces; the micro-optics of planar SPP waves can be exploited for various applications. Plasmonic devices based on planar SPP waves provide functionalities including reflecting, refracting, focusing, routing and so on, are like those optical components in conventional free-space optics. Besides the applications in basic plasmonic components, the planar SPP waves are characterized by a strongly enhanced electric field and high sensitivity to the material properties on the metal surface; they will be of great interest in bio-sensors technologies especially for lab-on-chip applications.

3.1. SPP excitation

In order to investigate the micro-optics of plasmonic devices based on planar SPP waves, it is a prerequisite for one to develop appropriate approaches so as to enable local excitation of SPP waves propagating on the metal surface with a well-defined direction. Using a prism made of a high-index material is a conventional way of providing the wave-vector matching between free-space light and SPP waves; this method, which is widely used in optical sensors based on SP resonances, has already been well established and described in detail in the literature [16]. However, prisms are usually quite bulky; in addition, this approach is also limited with respect to the range of wave-vector matching by available prism materials.
and their refractive indices. In the well-known Kretschmann configuration, a metal film thickness should also be chosen with great care (i.e. neither too thin nor too thick) in order to optimize the SPP excitation efficiency without resulting in excessive SPP leakage into the substrate. One can therefore choose to use artificial surface defects on thick metal surfaces that are intentionally designed to circumvent the problem of wave-vector mismatch by scattering, i.e. by matching wave-vectors of scattered field components to those of SPP waves. Different types of defects on the metal surfaces have been explored and investigated to obtain a collimated SPP wave with a low divergence angle. Ditlbacher et al [46, 47] excited a SP wave using a strongly focused laser beam at the free-space wavelength of 515 nm onto the silver film with a single silver nanoparticle (diameter 200 nm and height 60 nm) fabricated by first generating the SiO2 nanostructures on the glass substrate with electron beam lithography (EBL), SiO2 deposition and lift-off, and then coating the whole sample with a silver film so that a nanoparticle in the form of a bump could be realized. However, it was found using fluorescence imaging that the obtained SPP wave has a cos²δ (where δ is the azimuthal angle) intensity distribution, shown in figure 7(a), not quite satisfying in terms of directionality. Using the same fabrication procedure, the same group also managed to realize SPP excitation using a silver nanowire (width 300 nm, height 60 nm and length 20 µm) illuminated by a focused laser beam with the polarization perpendicular to the nanowire axis and demonstrated a more directed SP propagation [47], as illustrated in figure 7(b). They also observed that the SP beam divergence angle could be tuned by changing the diameter of the exciting laser spot [46]. Radko et al also launched the SPP wave by illuminating a 180 nm wide straight gold ridge on a 50 nm thick gold film evaporated on a quartz substrate with a focused laser beam from the tunable Ti : Sapphire laser (wavelength 700–860 nm) [48]. The fabrication procedure was somewhat different: the gold ridges were realized using EBL and lift off on a flat gold film which was evaporated on the quartz substrate beforehand.

The approach of SPP excitation with individual nanowires was further extended by using multiple nanowires, allowing one to increase the excitation efficiency. For an incident focused laser beam normal to the nanowire array surface, the array period should be equal to the SPP wavelength $\lambda_{SPP}$ so that the generated SPP waves from different nanowires will interfere constructively with each other, resulting in an increase in the SPP excitation efficiency. An ensemble of three parallel oriented ridges with width 200 nm and height 60 nm was used for SPP excitation at the wavelength of 800 nm and the efficiency of local SP excitation was experimentally measured, showing that 15% efficiency could be achieved when the ridge period was in close correspondence with the SPP wavelength, significantly larger than the efficiency (4.5%) when only one such kind of ridge was used [49]. In that work, bidirectional SPP excitations were realized on both sides of the 3-ridge array. Using 11 gold ridges, Radko et al have achieved an efficiency as high as 45% at the excitation wavelength of 816 nm [50]. The high efficiency was attributed to the fact that when the focused laser beam was normally illuminating certain places of the 11-ridge area, the SPP excitation on the right side of the gold ridge was suppressed [50], shown in figure 8. The suppression was related to the second-order Bragg reflection from ridges, which redirected the SPP back to the left, as well as the out-coupling of the propagating SPP since the ridge number was large. Further more, the gold ridge parameters in [50] were optimized with a larger height compared to the structures in [49], which might be another reason for the dramatic increase of the efficiency.

Besides the bump or protrusion type of defect on the metal surface, one can also introduce some slits or nano-holes perforating a metal film to assist with the SPP excitation.
In addition to wave-vector matching, this approach has the advantage that the excitation laser beam can illuminate the optically thick metal film from the back side so that the detection of the SPP waves will have less background noise from the incident beam. When one single straight slit in a metal film is illuminated in such a way, two SPP beams will be excited propagating bi-directionally due to symmetry. It is reasonable to expect that the excitation efficiency can be influenced by directing excited SPP beams in one direction. Exploiting this idea, several different approaches have been proposed and investigated, including the use of an asymmetric slit composed of a conventional nanoslit with a nanogroove just on one side of it [51] or incorporating a periodic array of grooves carved on one side of the original slit [52]. The efficiency was also found to increase when a highly oblique incidence was used [53]. The oblique incidence approach also works for two identical one-dimensional subwavelength slits perforated in a metal screen with the distance between them comparable to the wavelength, achieving a highly tunable splitting ratio between the two SPP beam propagating in two opposite directions or even a unidirectional SPP excitation [54]. When two nano-slits are separated with a larger distance, under normal incidence an interference phenomenon will exhibit in the total intensity of the far-field double-slit pattern, arising as a consequence of the excitation of SPs propagating from one slit to the other [55]. This plasmonic analog of the famous Young’s double-slit experiment in physical optics has suggested its broad applications in the SP assisted nanolithography [56, 57].

A circular or elliptical slit milled into optically thick metallic films can excite SPPs propagating toward certain places within the circle or ellipse, thus they can work as plasmonic lenses and focus SPPs, concentrating the EM field at the focal points [58, 59]. It is also found that similar to conventional lenses, the focus position for such a circular plasmonic lens can be tuned by adjusting the light incidence angle [59] to the axis of the lens. Besides the use of nano-slits, an array of nanometric holes that are milled into a silver film and arranges on a quarter circle can also launch SPPs with constructive interference, allowing the SPP to be focused into an intense spot having a subwavelength width [60].

2D nano-hole arrays with a larger scale have also been utilized to launch the SPP waves and convert them back to free propagating light [61]. Also investigated is how the properties of the locally launched SP beams, such as divergence or uniformity, can be tuned and optimized by adjusting the shape of the micro nano-hole grating [62]. The authors also found therein that oblique illumination together with the choice of an appropriate period of the nano-hole array would allow the enhancement of a single SP beam at the expense of the other, resulting in a unidirectional SPP beam similar to the results shown in figure 8(c). With respect to SPP excitation with high unidirectional propagation, it can be concluded that oblique incidence is a good choice for nanoslit or nano-hole structures, since the reflection from the incident beam will have little influence on the SPP detection on the other side of the metal film.

3.2. SPP manipulation by sets of nanoparticles

As mentioned before, due to the intrinsic evanescent properties of SPP waves in the direction perpendicular to the metal surface, the SPP waves can only propagate freely in the other two dimensions; plasmonic components based on these planar SPP waves can find their counterparts as those conventional free-space optical components. In order to investigate the micro-optics of planar SPPs, one needs to realize the essential functionality of reflection, refraction and focusing as those in conventional free-space optics and there are many ways of achieving this goal.

Based on the fact that light tends to propagate in a high-index medium, by fabricating dielectric structures of defined geometry on top of a gold film, the SPP dispersion relation at the metal interface can be shifted and the SPPs will also travel to the region with a higher effective index. Guided by this principle, using a cylindrical and triangular shaped thin layer of SiO₂, some basic optical elements for SPs like prisms and lenses with the functionality of focusing, refraction and total internal reflection (TIR), have been experimentally demonstrated [63]. Instead of using a real dielectric, one can also use a periodic structure with period much smaller than the wavelength to work as an effective medium to change the SPP dispersion relation. Such kinds of gold nanoparticles with a 100 nm period square lattice on top of a gold film were found to have the SPP effective index increasing inside the array by a factor of 1.08 (for the wavelength 800 nm) with respect to the SPP index at a flat surface [48]. Constructing such a lattice into a triangular-outlined structure, one can observe that the SPP beam propagating across the border between the lattice structure and the smooth gold surface will experience slight refraction, according to which one can calculate the effective SPP index inside the array using Snell’s law. By shaping this gold nanoparticle array with a higher SPP effective index into a circular or a rectangular geometry, the SPP beam on the gold surface can be deflected and focused by the former or guided along the latter’s geometry, demonstrating their functions similar to a lens or a waveguide respectively.

Nanoparticles and their arrays can also be used to manipulate the SPP waves directly. Using their fabrication procedure described above (section 3.1), Ditlbacher et al excited the SPP wave on the silver/PMMA interface and demonstrated, employing fluorescence imaging, the reflection of the excited SPP beam by an ensemble of nanoparticles arranged in parallel lines to act as a Bragg reflector [45]. The geometrical parameters of the nanoparticles were designed with the incident angle of the SPP beam with respect to the nanoparticle lines to fulfil the first order Bragg condition. Introducing another Bragg reflector along with a single row of nanoparticles acting as a beam splitter, this group has realized a Mach–Zehnder interferometer (MZI) for SPP waves [45]. In addition to the use of a straight line row of nanoparticles as Bragg reflectors, one can also use nanoparticle arrays arranged into elliptical reflecting mirrors, which have two foci; the SPP beam excited at one focus would be steered toward the other focus [64].

Bragg reflection of SPP waves propagating in all directions along the surface plane can be achieved by using 2D periodic
arrays of nanoscaterees, manifesting thereby the occurrence of the SPP band-gap effect [65]. Similar to photonic crystal structures, SPP guidance can then be realized along line defects, i.e. channels free of scaterees, introduced in periodic arrays of nanoparticles [66]. One can also achieve inhibition of SPP propagation due to multiple SPP scattering by randomly arranged nanoparticles (leading to the strong SPP localization) and, conversely, SPP propagation along scattering-free channels cut through those arrays [67]. With respect to the usage of individual lines of nanoparticles for SPP manipulation, it was demonstrated that a parabolic chain of gold nanoparticles fabricated on a thin gold film focuses incident SPP waves inside the parabola [68] operating thereby just like a metallic parabolic mirror in conventional optics. The influence of excitation wavelength and geometrical system parameters has been investigated in detail with the help of leakage radiation microscopy (LRM) imaging [48].

The inverse structure of a nanoparticle array, i.e. the nano-hole array performed through a metal film, can also be utilized to steer the propagation of planar SPP waves on the same metal surface. By carefully designing a nano-array structure perforated through a silver film, an experimental realization of a plasmonic Airy beam formed from an in-plane propagating SPP wave on the silver surface has recently been reported [69], demonstrating the versatility of this approach in the application of SPP manipulation.

In principle, SPP optical elements allow us to realize a variety of 2D optical circuit functionalities. Many optical elements in conventional free-space optics may find their counterparts in SPP micro-optics. Such plasmonic devices could lead to the technologically easy and cost-effective realization of all-optical switches, integrated photonic devices and optical lab-on-a-chip applications based on the SPP micro-optics platform.

4. SPP waveguides

From this section onward, we will start focusing on 3D SPP waveguides with full 2D lateral confinement in the plane perpendicular to the propagation direction. The tight mode confinement provided by plasmonic waveguides leads to intensive research interest in the exploration of their application for high photonic integration. To date, various plasmonic waveguides have been proposed and investigated; then the sub-discipline of integrated plasmonics was established. Although the general trade-off between propagation length and mode confinement still exists for all those 3D plasmonic waveguides, different waveguides have their own special features and may find suitable applications under different circumstances. In this section, we will discuss the mode guiding mechanism for various plasmonic waveguides. Note that those elementary structures discussed in section 2 with single or multiple metal/dielectric interfaces, along with the basic SPP modes supported by them, although in the 2D case, provide the foundation for 3D plasmonic waveguiding. The mode guiding in 3D plasmonic waveguides may result from one of the basic SPP modes supported in 1D and the effective index contrast in the other dimension of the 2D cross section, or from the basic SPP modes supported in both dimensions, depending on the specific waveguide geometry. Due to scope limitation, we will only review some very popular SPP waveguide representatives that have attracted intensive research interest in recent years. Following the sequence in section 2, we start with plasmonic waveguides with only one metal/dielectric interface and then move on to plasmonic waveguides with multiple metal/dielectric interfaces. The tight mode confinement provided by these plasmonic waveguides lays the foundation for the realization of various plasmonic components, which will be discussed in section 6. Emphasis is put on the geometrical influence to the mode properties of various plasmonic waveguides, which is the basis of nanofocusing with different structured structures. At the end of this section, a comparison of five SPP waveguides in terms of some important characteristics will be briefly presented.

The calculations for all the plasmonic mode properties in this section are realized with the finite element method implemented in the commercially available software COMSOL Multiphysics, which will be abbreviated as COMSOL in the rest part of this paper. Unless otherwise stated, scattering boundary conditions are used for the numerical simulations.

4.1. SPP guiding by dielectric ridges

4.1.1. Dielectric-loaded SPP waveguides. SPPs on an infinite metal/dielectric interface can only provide confinement in the direction perpendicular to the interface. When a thin dielectric stripe (normally a polymer like polymethyl methacrylate (PMMA) is used due to the simplicity in the fabrication process) is present on the metal surface, the lateral confinement can also be provided by the contrast in the mode effective index of SPP at the metal/dielectric/air interface and metal/air interface. This waveguide geometry with the cross section schematically shown in figure 9(a), termed as a dielectric loaded surface plasmon-polariton waveguide (DLSPPW), is one of the most popular plasmonic waveguides in recent years and several fabrication approaches have been exploited to realize it [70–72]. Since polymers like PMMA can work as both the resist and the dielectric core for the DLSPPW, it is quite a convenient and simple process to fabricate plasmonic devices based on DLSPPW using deep ultraviolet (DUV) lithography, for those devices with larger dimensions working at telecom [73], or the standard EBL, for those working at near infrared [74].

Figure 9(b) shows a typical mode profile of the DLSPPW at 780 nm. The dielectric stripe, assumed to be PMMA with a dimension of 300 nm by 300 nm and an index of 1.49 at 780 nm, sits on top of a 70 nm gold layer (index 0.18 + 4.92i at 780 nm [18]), which is deposited on the silica substrate with an index of 1.45. As can be seen, the mode is tightly confined to the bottom of the PMMA stripe and part of the electric field is out of it in the lateral direction. Note that the lateral confinement is due to the contrast between the SPP effective mode index of the three-layer structure metal/dielectric/air in region I, \( n_{SPP1} \), and that of the metal/air in region II and III, \( n_{SPP2} \), so it is still
diffraction limited. Actually not only the mode effective index of the DLSPPW, but also its propagation length is affected by $n_{SPP1}$, which is determined by the dielectric stripe height and width. So one can change the geometrical parameters of the PMMA stripe to tune the DLSPPW properties. As is pointed out in section 2, the SPP propagation loss at the infinite metal/dielectric interface is larger when the dielectric contacting the metal has a higher index. Since the dielectric region above the metal layer can be regarded as an effective medium with the effective index affected by the dielectric thickness $H$, we can speculate that for DLSPPWs with larger $H$, the mode confinement will be stronger but the propagation loss will be higher because both the real and imaginary parts of $n_{SPP1}$ will be larger; while for DLSPPWs with smaller $H$, the propagation length will be longer. These speculations have been experimentally demonstrated and verified [75].

When we divide the cross section of DLSPPW in figure 9(a) into three regions, we are actually using the idea of an effective index method (EIM). The application of EIM in the modeling of DLSPPWs has been numerically investigated and is found to be quite effective when the DLSPPW mode is away from being cut-off [76]. Actually the validity of EIM has been verified in many plasmonic waveguides and the use of this approach helps one interpret the mode properties of different plasmonic waveguides both qualitatively and quantitatively. We will also discuss the use of this later when we come to different plasmonic waveguides.

Despite the presence of a diffraction limit in DLSPPWs, very good confinement can still be achieved thanks to the fact that the real part of $n_{SPP1}$ is larger than the dielectric index, resulting in a typical subwavelength dimension $300 \, \text{nm} \times 300 \, \text{nm}$ for the near-infrared wavelengths and $600 \, \text{nm} \times 600 \, \text{nm}$ for the telecom wavelengths, if we assume DLSPPW consists of a PMMA stripe on a thin gold layer. The typical propagation lengths for the two cases are around $11 \, \mu\text{m}$ (at $\lambda \approx 780 \, \text{nm}$) and $50 \, \mu\text{m}$ (at $\lambda \approx 1500 \, \text{nm}$) respectively. In addition to the mode properties, another advantage of DLSPPW over many other plasmonic waveguides results from its geometry. Besides working as an essential part in a DLSPPW to support plasmonic modes, the metallic part can also be used to carry electrical signals, thus the dream of merging photonics and electronics on the same platform can be realized. Researchers have made some advancement in this direction; compact fiber-coupled DLSPPW-based plasmonic components, including MZIs, waveguide ring resonators (WRRs) and directional couplers (DCs) whose outputs were controlled via the thermo-optic effect by electrically heating the gold stripes supporting the polymer ridges of DLSPPWs have been designed, fabricated and characterized recently at telecom wavelengths [77]. Furthermore, due to the convenience of doping some optically active medium into the polymer, the DLSPPWs have also shown great potential of all-optical functionalities. Partial loss compensation in the DLSPPWs can also be realized relying on the stimulated emission of the active dopant. Experimental results on the partial loss compensation of DLSPPWs with quantum dots as the gaining medium working both at telecom [78] and near-infrared wavelengths [79] have been reported; roughly a 30% increase in the propagation length in both cases has been achieved.

In addition to the tight mode confinement at the bottom of the PMMA stripe, also visible in figure 9(b) is that there is some electric field in the substrate. From the calculation, we know that the real part of the effective index for the DLSPPW is smaller than the refractive index of the substrate, so the DLSPPW is actually a leaky mode and the leakage degree is determined by the metal thickness. This kind of leakage into the substrate, although introducing an additional loss to the DLSPPWs, provides an effective means of imaging the SPP propagation in the DLSPPWs. This technique, known as LRM, can provide the information both in the image space and in the Fourier space and has been broadly used in the investigation of DLSPPW components [78].

Whilst discussing DLSPPWs, it is worth mentioning the recent progress in the experimental realization of one special kind of DLSPPWs featuring both good mode confinement and a much lower propagation loss [80]. The so-called long-range DLSPPWs (LR-DLSPPWs) inherit the advantages of having a very long propagation length from long-range SPP waveguides (LR-SPPWs) and of being technologically simple from the DLSPPWs. The LR-SPPWs, which will be discussed in detail in section 4.2, have propagation losses (due to the absorption by the metal involved) which are significantly reduced, due to the fact that the longitudinal component of the electric field in the metal is minimized [81]. However, the LR-SPPW mode confinement is weak, with the mode size being much larger than the light wavelength, a circumstance that prevents such waveguides from being used for highly compact devices. Recently, the idea of LR-DLSPPWs was proposed and
Figure 10. (a) Schematic layout for the cross section of the fabricated LR-DLSPPW structure, with a PMMA ridge on top of a gold stripe deposited on an underlying Ormoclear buffer layer. The whole structure is supported by a low-index layer of Cytop (which ensures mode confinement to the buffer and ridge region). The inset (time-averaged electric-field distribution) represents the mode structure of the fundamental LR-DLSPPW mode at the wavelength of 1500 nm; topographical (time-averaged electric-field distribution) represents the mode ensures mode confinement to the buffer and ridge region. The inset (b) of different sizes: (b), (c) 21.3 × 4.3 μm² and (d), (e) 38.4 × 13.6 μm². SNOM images were recorded at different wavelengths with shear-force feedback: (c) 1500 nm and (e) 1450 nm. The LR-DLSPPW mode propagates from left to right. Adapted with permission from [80]. Copyright 2011 The Optical Society of America.

numerically investigated [82, 83]. Numerical calculations suggest that the LR-DLSPPW configuration guiding of an SPP mode over mm-long distances, while retaining a relatively strong mode confinement.

Figure 10(a) gives the schematic for the cross section of the LR-DLSPPW, with all the dimensions indicated. As can be seen from the mode profile in the inset of figure 10(a), the LR-DLSPPW mode is tightly bound to the metal stripe, which is embedded between the PMMA ridge and the buffer layer. The same as in DLSPW, the mode confinement in the lateral direction is due to index guiding. For the vertical direction, the guiding is similar to LR-SPPWs, except that it is a five-layer structure of air/PMMA/metal/Ormoclear/Cytop. As will be shown in section 4.2, it is very important for LR-SPPWs to have symmetric claddings with a low index to achieve the long propagation length. So one has to adjust the geometrical parameters of the PMMA stripe and buffer layer thickness to achieve an equivalent symmetric environment in the vertical direction, in order to get the optimal mode properties including both confinement and propagation loss [83]. The two outer layers of materials in LR-DLSPPWs, air and Cytop, have a lower index compared to the two inner claddings, PMMA and Ormoclear, which can provide two advantages: to enhance the confinement in the vertical direction; to provide an environment with a lower effective index above and below the gold stripe to reduce the propagation loss.

The fabrication was started by spinning a 255 nm thick layer of a high refractive index UV-curable organic-inorganic hybrid material (Ormoclear, \( n_b = 1.53 \)) onto a 4 μm thick amorphous fluoro-polymer (Cytop, \( n_s = 1.34 \)) coated silicon wafer. Straight and bent 500 nm wide and 15 nm thick gold stripes were patterned with EBL, metal evaporation and lift-off. The waveguide fabrication was completed using the second step of EBL for patterning 1 μm thick ridges of PMMA \( (n_t = 1.49) \) on top of the gold stripes. Scanning near field optical microscopy (SNOM) results of the light propagation along the straight LR-DLSPPW find an almost linear increase in the LR-DLSPPW mode propagation length with an increasing wavelength (reaching \( \sim 520 \mu m \) for \( \lambda \approx 1545 \text{ nm} \)). The confinement of the LR-DLSPPW modes has been further investigated by characterizing abruptly bent LR-DLSPPWs, whose topographical and corresponding near-field optical images are shown in figures 10(b) and (c). Using a simple Gaussian approximation for the fundamental mode field distribution in width and taking into account the experimentally estimated values for the transmission through an abrupt bend, the averaged mode width of the investigated LR-DLSPPW can be estimated to be around 1.02 μm at 1550 nm, which is consistent with the numerical results shown in the inset of figure 10(a). Compact S-bend structures based on sine curves [80], allowing for continuous bend curvature and thereby adiabatic modification of the LR-DLSPPW mode throughout the bend, were also characterized demonstrating excellent performance (see figures 10(d) and (e)) with overall losses of \( \sim 1.8 \text{ dB} \) (which should also be partially attributed to the defect observed in the middle bend section (see figure 10(d))), confirming the strong subwavelength confinement of the LR-DLSPPW mode.

The fact that the LR-DLSPPW mode can be both confined to the subwavelength cross sections and guided with low dissipation (compared to most SPP waveguide configurations) holds promise for the realization of complex highly integrated plasmonic circuits, whose performance could even be further improved by applying the possibility of losses compensation by the stimulated emission of LR-DLSPPW modes.

4.1.2. Hybrid SPP waveguides. Another plasmonic waveguide whose geometry shows some resemblance to that of DLSPW, with some dielectric architecture on top of the planar metal layer, is the hybrid SPP waveguides (HSPPWs). HSPPW has recently attracted much research interest because it possesses a confinement capability beyond the diffraction limit while retaining a relatively long propagation length compared to other plasmonic waveguides. The HSPPW was first proposed in the form of a dielectric nanowire with a higher refractive index separated from a metal surface by a nanoscale dielectric gap with a lower index [84], schematically shown in figure 11(a) where \( \varepsilon_2 > \varepsilon_1 > \varepsilon_3 \). It was originally believed that the mode was due to the coupling between the photonic mode supported by the nanowire and the plasmonic mode at the metal surface. Later, theoretical analysis showed that the mode is actually based on a 2D conductor-gap-dielectric (CGD) mode [15], whose analysis can be simplified using the 2D geometry shown in figure 11(b). The CGD mode is due to the high permittivity discontinuity between the central gap layer and the outer metal/high-index dielectric layer, which renders the electric-field component perpendicular to the metal surface highly enhanced in the central layer. The real part of the mode effective index of this CGD mode, \( Re(n_{CGD}) \), will decrease as the gap layer thickness is increasing; so from the effective index point of view, the HSPPW shown...
in figure 11(a) can be regarded as a waveguide with a graded index change in the lateral direction, which means that it is still an index guiding waveguide. It is worth noting that for the 2D geometry shown in figure 11(b), theoretically there is a critical gap layer thickness \( t_c \) at which the CGD mode has a minimal loss approaching zero [15]. This can be easily interpreted. As is known, the real part of the mode effective index of SPP on an \( \varepsilon_m/\varepsilon_2 \) interface is larger than the index of \( \varepsilon_2 \), i.e. \( \text{Re}(n_{\text{spp}}) > \sqrt{\varepsilon_2} \). When the thin gap dielectric layer with a lower index is inserted between them, \( \text{Re}(n_{\text{CGD}}) \) will decrease and keep this trend as the gap layer thickness \( t \) increases until it reaches a critical point \( t_c \) where \( \text{Re}(n_{\text{CGD}}) = \sqrt{\varepsilon_2} \). When \( 0 < t < t_c \), \( \text{Re}(n_{\text{CGD}}) > \sqrt{\varepsilon_2} \), so the CGD mode is steadily sustained by the geometry in figure 11(b). But when \( t \) reaches \( t_c \) and \( \text{Re}(n_{\text{CGD}}) = \sqrt{\varepsilon_2} \), the CGD mode will be leaking dramatically to the dielectric, so the overall loss will be minimized. However, numerical simulation results conducted for a 3D modified HSPPW on the silicon-on-insulator (SOI) platform [85], with the cross section shown in figure 11(c), or for the original HSPPW shown in figure 11(a), do not show any evidence of a critical thickness \( t_c \) where the mode loss is significantly lower, although there is indeed \( t \) value where the loss is the minimum for the HSPPW shown in figure 11(a). This is probably because the high-index dielectric is also patterned in these waveguides into the form of a practical structure (it cannot be infinite) so that the ideal case of the CGD mode can never be constructed. The high-index dielectric being patterned usually also results in a photonic mode supported in the dielectric, which as well as the existence of the pure CGD mode renders the waveguide with the name HSPPWs, although in principal the hybridization nature is not necessary, as the CGD mode is a well sustained bound mode. We also need to note that \( t_c \) is usually a few nanometers or at most a few tens of nanometers depending on the index contrast between \( \varepsilon_1 \) and \( \varepsilon_2 \), which is really challenging for practical applications with structures realized with nanofabrication facilities.

The HSPPW is attractive in that it can provide both subwavelength confinement and relatively low propagation loss. A plasmon laser at a deep subwavelength scale was experimentally demonstrated using chemically synthesized high-gain cadmium sulfide semiconductor nanowire, separated from a silver surface by a 5 nm thick MgF2 gap [86]. Another advantage of this waveguide is that it can contain high refractive index material, which is not in direct contact with the metal so that the problem of the high loss associated with SPPs on the interface between metal and high-index material can be circumvented. As a consequence, one can combine the plasmonic component with the important photonic material of Si and then realize the seamless integration of plasmonic devices and functionalities with silicon photonics on the same platform. For example, it is possible to achieve ultrafast all-optical switching using the nonlinear properties of silicon [87]. Then the problem of being highly lossy with silicon-based plasmonics [88] can be alleviated. Recently, a modified form of subwavelength HSPPW composed of a thin SiO2 gap between the top gold layer and the under silicon core on a SOI platform, schematically shown in figure 11(c), has been experimentally realized. The typical mode profile for such a HSPPW at the wavelength of 1550 nm is shown in figure 11(d), in which 340 nm thick Si residing on top of a silica layer originally from the SOI platform is covered with a 50 nm thick SiO2 layer and another 50 nm thick capping layer of gold. The width for all the three top layers is 200 nm and one can see that the mode is tightly confined in the low-index SiO2 layer between Si and Au. Some mode power is also visible in the Si region since the lateral refractive index contrast is quite high. Experimental results demonstrated that this waveguide with subwavelength confinement of light and a relatively long propagation length of 40 \( \mu \)m was achieved at telecom, as well as the compact S-bends and Y splitters based on the same waveguide [85].

### 4.2. SPP guided by metal stripes

Now we will discuss plasmonic waveguides with more-than-one metal/dielectric interface. The SPs on those interfaces usually couple to each other to form some more sophisticated supermodes, one of which is usually preferable in terms of mode confinement, propagation length and ease of excitation. The first waveguide geometry we would like to discuss is the metal stripe, which supports an important mode termed as LR-SPPWs when it is embedded in a homogenous dielectric environment. The name originates from the fact that this fundamental mode has a significantly lower propagation loss than the regular SPP on the infinite interface between the same metal and dielectric. LR-SPPW gained broad research attention prior to the time when plasmonics started to become a hot topic and the main branch of nanophotonics [23, 89]; researchers at that time worked on LR-SPPWs aiming at a much lower propagation loss. Considerable effort can be seen in the literature describing the research progress on this topic both in the experimental and theoretical respects; various photonic components based on LR-SPPWs have been
been discussed in the \[81, 93\]. For the purpose of radiation symmetry of the modes can be divided into four main families according to the electric field and magnetic field exist. We assume the field symmetry. For a 3D waveguide, all six components of the metal stripe extends along the central plane according to the magnetic index and is increasing as the metal thickness increases. This means that the \(ss^0\) mode is always a pure bound mode in the symmetric environment. What is noteworthy is that the \(ss^0\) mode has no cut-off metal thickness for this symmetric environment. Figures 12(b) and (c) illustrate the mode profiles of \(E_x\) for the two cases when the metal thickness is 10 nm and 60 nm respectively. It is straightforward that for a very thin metal stripe, the mode extends to the outer dielectric, with the overall size much larger than the wavelength. When the metal thickness approaches zero, \(Re(n_{eff})\) of the \(ss^0\) mode gradually approaches the index of the dielectric and the mode evolves to the TEM mode in the outer dielectric. This is why the mode has a pretty low loss when the metal thickness is small, since a large proportion of the power is in the dielectric region. When the metal thickness increases, the mode begins to shrink and becomes more localized at the metal stripe, as can be seen in figure 12(c) compared to figure 12(b). More mode localization around the metal implies that the loss is larger. This is confirmed with the propagation length as a function of the metal thickness shown in figure 13(a), from which one can see that the propagation length drops dramatically as a function of the metal thickness. Note that when the metal is very thin \((t < 20 \text{ nm})\), a very long propagation length of a few hundreds of micrometers can be achieved. For a longer wavelength the metal dissipation is less lower and the propagation length will be even longer. In practice, due to the limit of the minimum metal layer thickness with a smooth surface that can be achieved with state of the art nanofabrication facilities, a propagation length up to a few millimeters can still be realized for the telecommunication wavelengths [90]. The extremely low propagation loss and the symmetry of the mode, which can easily be excited with a focused laser beam or the mode from an optical fiber, render the \(ss^0\) mode very promising to transmit light signals over a relatively long distance. The drawback is that the mode of the \(ss^0\) mode extends to the outer dielectric, with the size much larger than the wavelength, making this waveguide not a good candidate for subwavelength photonic integration. Nevertheless, the LR-SPPW can still find broad applications in those areas where the confinement is not crucial, e.g. in sensing. One may also notice that the

Figure 12. (a) Sketch of the metal stripe embedded in a symmetric environment (when \(\varepsilon_1 = \varepsilon_2\)) or asymmetric environment (when \(\varepsilon_1 \neq \varepsilon_2\)). (b) and (c) \(E_x\) profile at 633 nm when \(\varepsilon_1 = \varepsilon_2 = 4\) for the silver stripe with the width kept as 1 \(\mu\text{m}\) and the thickness being 10 nm in (b) and 60 nm in (c).

4.2.1. Symmetric environment (LR-SPPs). When the dielectric above the metal stripe has the same permittivity as the one below the stripe, i.e. \(\varepsilon_1 = \varepsilon_2\), the environment around the metal stripe is homogenous. We assume the working wavelength is 633 nm and the metal is silver whose permittivity of \(\varepsilon_m = -19 + 0.53\) at this wavelength. The dielectric is assumed to have a permittivity of \(\varepsilon_1 = 4\). First we keep the metal width to be constantly 1 \(\mu\text{m}\) and change the metal thickness. Figure 13(a) shows the real part of the mode effective index and the propagation length when the metal thickness is increasing from 10 to 100 nm. As can be seen, the trend is very similar to those results obtained from a 2D IMI structure shown in figure 4. For all these thickness values the effective index of the mode is larger than the environment index and is increasing as the metal thickness increases. This means that the \(ss^0\) mode is always a pure bound mode in the symmetric environment. What is noteworthy is that the \(ss^0\) mode has no cut-off metal thickness for this symmetric environment. Figures 12(b) and (c) illustrate the mode profiles of \(E_x\) for the two cases when the metal thickness is 10 nm and 60 nm respectively. It is straightforward that for a very thin metal stripe, the mode extends to the outer dielectric, with the overall size much larger than the wavelength. When the metal thickness approaches zero, \(Re(n_{eff})\) of the \(ss^0\) mode gradually approaches the index of the dielectric and the mode evolves to the TEM mode in the outer dielectric. This is why the mode has a pretty low loss when the metal thickness is small, since a large proportion of the power is in the dielectric region. When the metal thickness increases, the mode begins to shrink and becomes more localized at the metal stripe, as can be seen in figure 12(c) compared to figure 12(b). More mode localization around the metal implies that the loss is larger. This is confirmed with the propagation length as a function of the metal thickness shown in figure 13(a), from which one can see that the propagation length drops dramatically as a function of the metal thickness. Note that when the metal is very thin \((t < 20 \text{ nm})\), a very long propagation length of a few hundreds of micrometers can be achieved. For a longer wavelength the metal dissipation is less lower and the propagation length will be even longer. In practice, due to the limit of the minimum metal layer thickness with a smooth surface that can be achieved with state of the art nanofabrication facilities, a propagation length up to a few millimeters can still be realized for the telecommunication wavelengths [90]. The extremely low propagation loss and the symmetry of the mode, which can easily be excited with a focused laser beam or the mode from an optical fiber, render the \(ss^0\) mode very promising to transmit light signals over a relatively long distance. The drawback is that the mode of the \(ss^0\) mode extends to the outer dielectric, with the size much larger than the wavelength, making this waveguide not a good candidate for subwavelength photonic integration. Nevertheless, the LR-SPPW can still find broad applications in those areas where the confinement is not crucial, e.g. in sensing. One may also notice that the
Figure 13. (a) The real part (solid line) of the mode effective index and propagation length (dashed line) of the $s_{0}^{b}$ mode as a function of the metal stripe thickness when $w$ is kept constantly being 1 µm and the dielectric index is 2; (b) increase of the real part of the mode effective index compared to the surrounding medium index (solid line) and the propagation length (dashed line) of the $s_{0}^{b}$ mode when the environmental medium index changes.

4.2.2. Asymmetric environment (leaky and SR-SPPs). In real situations, the materials below and above the metal stripe are usually different, so the environment is asymmetric. For instance, a gold stripe is fabricated on a glass substrate but is covered by another polymer layer from the top. As an example, we assume the index of the substrate is still 2, but is 1.95 for the superstrate. Figure 14 gives the evolution of $Re(n_{eff})$ and the propagation length at 633 nm when the metal stripe width changes, while its thickness is kept as 20 nm. One can see that as the metal width decreases, both $Re(n_{eff})$ and the propagation length show the same trend as the symmetric case. It is worth noting that when the width is below a certain value, $Re(n_{eff})$ is smaller than the substrate index, which means that the mode has become leaky in this case. So there is a cut-off width for the $s_{0}^{b}$ mode with this asymmetric geometry, as opposed to the symmetric case. Further numerical simulations show that the cut-off width increases when the index contrast between the substrate and the superstrate is larger. The power leaking into the substrate will introduce more losses to the mode, but from figure 14 we know that the mode loss is higher when the width is larger, so the long-range mode can only exist in asymmetric structures only near the cut-off. Away from the cut-off width, the $s_{0}^{b}$ mode is lossier due to either the leakage into the high-index cladding or the dissipation in the metal, so it is actually a SR-SPP mode. When the mode is near the cut-off, its field extends deeply into the high-index cladding since the transverse wave-vector in it is close to 0; on the low-index side, the mode is more localized to the metal stripe. Since the mode profile is not symmetric, there will be a decrease in the

width changes, while its thickness is kept as 20 nm. One can see that as the metal width decreases, both $Re(n_{eff})$ and the propagation length show the same trend as the symmetric case. It is worth noting that when the width is below a certain value, $Re(n_{eff})$ is smaller than the substrate index, which means that the mode has become leaky in this case. So there is a cut-off width for the $s_{0}^{b}$ mode with this asymmetric geometry, as opposed to the symmetric case. Further numerical simulations show that the cut-off width increases when the index contrast between the substrate and the superstrate is larger. The power leaking into the substrate will introduce more losses to the mode, but from figure 14 we know that the mode loss is higher when the width is larger, so the long-range mode can only exist in asymmetric structures only near the cut-off. Away from the cut-off width, the $s_{0}^{b}$ mode is lossier due to either the leakage into the high-index cladding or the dissipation in the metal, so it is actually a SR-SPP mode. When the mode is near the cut-off, its field extends deeply into the high-index cladding since the transverse wave-vector in it is close to 0; on the low-index side, the mode is more localized to the metal stripe. Since the mode profile is not symmetric, there will be a decrease in the

$Re(n_{eff})$ mode discussed in [81] has an even smaller propagation loss than the $s_{0}^{b}$ mode; however, this mode is not preferable for signal transmission in that the mode is anti-symmetric in the $x$ direction and then has a low overlap with a focused laser beam or another waveguide mode.

We also keep the thickness of the metal stripe constant and vary its width. Similar results to those shown in figure 13(a) have been obtained. Further numerical results show that as the metal thickness or width increases while the other dimension is fixed, $Re(n_{eff})$ of the $s_{0}^{b}$ mode has an asymptote of that for a plasmon-polariton mode supported by an isolated corner [81]. This is different to the case of a 2D metal film, whose $a_{0}$ and $s_{b}$ modes have an asymptote relevant to the SPP on an infinite metal/dielectric interface.

We also investigate the influence of the environmental index to the $s_{0}^{b}$ properties. Using a metal stripe with width 1 µm and thickness 20 nm as an example, we calculate the difference between $Re(n_{eff})$ and the dielectric index as well as the propagation length versus the environmental medium index, shown in figure 13(b). Consistent to the properties of the regular SPP on infinite metal/dielectric interface [15], $Re(n_{eff})$ deviates from $n$ to a larger extent and the propagation length decreases, when $n$ becomes larger. These trends also result from the fact that as the surrounding dielectric index increases, a larger proportion of the mode energy resides in the metal region. The results also indicate that the index of the surrounding medium around the metal stripe should be as small as possible to achieve a minimum propagation loss; that is why the use of some polymers with an extremely low refractive index, like Cytops in the fabrication of LR-SPPWs, has been explored in the literature [95].
coupling efficiency when the end-fire approach is used. All these factors should be considered in real applications since a certain degree of asymmetry always exists.

4.3. Slot and gap waveguides

The LR-SPPW mode has a low propagation loss; however, one can see from the above subsection that the confinement is rather poor. One of the main reasons that people are exploring the realization of photonic components based on SPS is that they have the potential of achieving true subwavelength confinement while retaining the super broad bandwidth information transmission capacity from photonics. For this objective, the G-SPP mode supported by the MIM structure discussed in section 2.3 is of great interest since the mode is non-diffraction-limited. The MIM geometry can only provide 1D confinement; for real applications, 2D confinement is required. Based on the effect that the propagation constant of the G-SPP mode is affected by the metal cladding permittivity, the insulator index or width, one can tune one or more of these parameters to realize some real 3D waveguides with subwavelength 2D confinement. Figure 15 illustrates the cross section of some examples of the plasmonic waveguides based on the G-SPP mode that have been investigated in the literature in recent years. However, some of these waveguides, including the metal heterowaveguide [96] (figure 15(a)), the SPP gap waveguide (figure 15(b)) [36] and the index-guided SP waveguide (figure 15(c)) [97], suffer from the problem of either fabrication infeasibility or characterization difficulty. Among all the plasmonic waveguides based on the G-SPP mode proposed to date, the plasmonic slot waveguide schematically shown in figure 16(a) has attracted considerable attention. In addition to its fabrication feasibility, another reason is that under symmetric situations it supports a fundamental mode with no cut off in the width and low bending loss, even when the bend radius is really small [98].

Liu et al [98] and Veronis et al [99] proposed the plasmonic slot waveguides separately and investigated their properties in the visible and at telecom wavelengths respectively. Since the mode properties are more or less the same in the two regimes, we will have a brief overview of them here at the wavelength of 1550 nm. The structure is built on a substrate with permittivity \( \varepsilon_{d1} \), on top of which a gold layer \( \varepsilon_m = 0.55 + 11.5j \) [18] with thickness \( h \) is deposited. A slot, with the width denoted as \( w \), is formed within the gold layer and filled with the same material as the superstrate, whose permittivity is \( \varepsilon_g \). For simplicity, at the beginning we assume the substrate and the superstrate are composed of the same material with a refractive index of 1.45. We first keep the height of the metal at 100 nm and investigate the dependence of the mode effective index and the propagation length on the slot width. The results are shown as red solid lines in figure 17. Also shown in figure 17 are the same two attributes for the G-SPP mode in the 2D MIM waveguide with the same width as the plasmonic slot waveguide, and for the edge mode with the same metal height supported by the structure when half of the metal is replaced with the superstrate material, schematically shown in figure 16(b). One can see from the results, especially from the propagation length, that when the slot width is small compared to the height, the plasmonic slot mode behaves more like a G-SPP mode, with the mode effective index lower and the propagation length smaller than those of the G-SPP mode respectively. As the width is increasing and larger than the height, the behavior of the plasmonic slot mode deviates from the G-SPP mode gradually, with the mode effective index first approaching to that of the MIM waveguide and then exceeding it after the width is larger than a certain value. Further numerical results show that as the width increases, both the mode effective index and the propagation length approach asymptotically to those of the edge mode respectively, shown as the blue dotted lines in figure 17. These results suggest that when the slot width is large, the plasmonic slot mode is actually due to the coupling of two edge modes, which is so weak that the behavior of the single edge mode dominates in the plasmonic slot mode. In this case, the plasmonic slot mode demonstrates an edge-mode-like characteristic. When the width is small, the coupling is strong and the plasmonic slot mode behaves similarly to a G-SPP mode, demonstrating a G-SPP-like mode characteristic.

The insets in figure 17(a) illustrate the \( E_x \) (the major component of the electric field) profile for the width of 50 nm and 100 nm respectively, from which one can see straightforwardly the characteristics of the modes for these two different widths. For the 50 nm width, the mode is tightly confined within the slot and the lateral distribution is like a G-SPP mode while for the 100 nm width, most of the power is around the corners and one can see clearly the feature of the edge modes. Another property we need to mention here is that in the G-SPP-like mode, using the mode profile for the 50 nm width shown in the inset of figure 17(a) as an example, \( E_x/H_y \) is the major component of the electric/magnetic field respectively and they distribute more uniformly within the slot compared to the case when the width is large. In this case, the plasmonic slot mode is a quasi-TEM mode. In the edge-mode-like region, the optical power is mainly concentrated at the four metal corners; \( E_x/H_y \) cannot be considered as the major component any more, so the quasi-TEM characteristic no longer holds.

The characteristics of the plasmonic slot mode being G-SPP-like when the width is small compared to the height and edge-mode-like when the width is larger are more obvious when the thickness of the metal layer is changed while the slot width is fixed as 100 nm. Shown in figure 18, when the metal thickness decreases from the value of the slot width,
Figure 16. (a) Cross section of the plasmonic slot waveguide with the metal ($\varepsilon_m$) sitting atop the substrate ($\varepsilon_d$). The metal slot is filled with the same material as the superstrate ($\varepsilon_d$); (b) when half the metal is replaced with the superstrate materials, the structure supports an edge mode; (c) effective index view of the plasmonic slot waveguide shown in (a), note that the mode propagates to the $z$ direction for both the two cases.

Figure 17. Mode effective index (a) and propagation length (b) as a function of the slot width when the height is kept as 100 nm. Also shown are the same two quantities for the G-SPP mode in the MIM waveguide with the same width and for the edge mode with the height of 100 nm. In both figures, red solid lines are for the plasmonic slot waveguide, green dashed lines for the MIM waveguide and blue dotted lines for the edge mode. Insets in (a) show the $E_x$ profile of the plasmonic slot mode at 1550 nm wavelength with the slot width of 50 nm and 100 nm respectively.

Figure 18. Mode effective index (a) and propagation length (b) as a function of the slot height when the width is kept as 100 nm. Also shown are the two attributes for the G-SPP mode in the MIM waveguide with the same width. In both figures, red solid lines are for the plasmonic slot waveguide and green dashed lines are for the MIM waveguide.

the mode effective index increases and the propagation length drops significantly, following the characteristics of the edge mode. When $h$ is small enough, the mode effective index is even higher than the G-SPP mode effective index. In the other direction, when the height is increasing, the plasmonic slot mode behaves asymptotically to the G-SPP mode, because the higher the slot is, the more its profile is approaching the 2D MIM structure.
We have discussed the plasmonic slot mode properties when the substrate and the superstrate have the same permittivity, i.e. the symmetric case. For the case when the substrate is different from the superstrate, since the plasmonic slot mode effective index is roughly decreasing when the slot width or height increases, as shown in figure 17 and figure 18, there exists a width or height value above which the plasmonic slot waveguide mode effective index will be below that of the edge mode or the SR-SPP mode supported by the substrate/metal/superstrate structure. In that case, the plasmonic slot mode is leaky. Different to a regular optical waveguide, the plasmonic slot waveguide works as a bounded mode when the width (height) is smaller than the cut-off width (height). Note that one exploits the use of the plasmonic slot waveguide for subwavelength confinement, so the width (height) is usually far smaller than the cut-off value; one does not need to be concerned with the cut-off for plasmonic slot waveguides. Note that the mode profile is no longer symmetric along the y direction for the asymmetric environment; this fact has to be considered when one uses the mode from another waveguide to excite the plasmonic slot mode.

When the plasmonic slot mode works in the G-SPP-mode-like case, one can interpret the mode from the EIM point of view. For the cross section in figure 16(a), one can consider the central layer in the y direction as an effective medium with the index $n_{\text{eff}}$ from the G-SPP mode supported by the MIM geometry. This value can be obtained analytically by solving the dispersion equation of the G-SPP mode. In this way, the plasmonic slot waveguide can be regarded as a 2D slab waveguide, shown in figure 16(c), with the cladding indices being those of the substrate and superstrate in figure 16(a). By solving the dispersion equation of the TE mode of the slab waveguide (the $E_x$ component is perpendicular to the slab plane and note that the coordinates change between figure 16(a) and (c)) analytically, one can get the mode effective index for the plasmonic slot waveguide. Results from this method demonstrates a considerably good agreement with those from the pure numerical results from 2D mode solvers, even for higher order modes in the y direction [35]. The EIM modeling provides a simple but nevertheless effective and relatively accurate means of solving the plasmonic slot modes. Note that the EIM no longer holds its accuracy when the slot height is larger than its width. In that case, the G-SPP mode effective index is even smaller than that of the plasmonic slot mode. The reason why EIM loses its validity can be interpreted from the vectorial property of the mode. As is known, only one component of the electric field (TE case) or the magnetic field (TM case) is considered in the 2D slab structure. So the EIM is valid only when one component of the electric/magnetic field is the major component, which is also the transverse field considered in the slab structure. When the plasmonic slot mode works as G-SPP-like, $E_y$ is the major component and is tightly confined within the slot, then EIM works fine. However, when the width is large and the slot works in the edge-mode-like way, both $E_x$ and $E_y$ are around the metal corner and neither of them can be considered as the major component any more; thus the EIM loses its validity. This point will be discussed again when we talk about the modes supported by metal grooves and edges later.

Besides the tight confinement, it is worthy of note that, due to the small penetration depth of light into the lateral metals, light can propagate in the plasmonic slot waveguide with low cross talk and through the sharp corners with low bending loss (the loss is mainly due to the propagation loss and out-of-plane scattering loss) [98]. These characteristics make the plasmonic slot waveguide a good candidate for planar photonic integration over a short range.

Regarding the fabrication, due to the small dimension of the metal slot, quite a lot of fabrication effort has been explored to realize the geometry shown in figure 16(a). For the slot in the lateral direction, the milling of metals including the use of a focused ion beam (FIB) to directly mill the metal layer [100] and argon ion milling to pattern the metal layer with the help of a resist mask from the EBL [101], are the main approaches. Vertically aligned plasmonic slot waveguides should be easier to fabricate with an EBL and lift-off process, although in the literature [32, 102] FIB was still used to pattern the metal for input/output coupling. The recently investigated continuous layer gap plasmon structures [103] require the patterning of only one metal layer, thus further relieving the fabrication difficulty and providing a new platform for the G-SPPs. Note that these vertical plasmonic slot waveguides are not preferable compared to the lateral counterparts in terms of cross talk and bending loss, because for planar photonic integration the packing of multiple waveguides and the bending occur in a lateral plane, in which the mode has a larger extension for the vertical plasmonic slot waveguides.

4.4. Channel plasmon polaritons

From the EIM point of view, the plasmonic slot waveguide discussed in section 4.3 can be regarded as a step index waveguide when the slot height is larger than its width. The confinement in the y direction is due to the effective index contrast between the central layer and the claddings. Based on the same idea, one can also realize a gradual index waveguide by etching a groove with varying width from the top [33]. One popular structure is to have the width decreasing linearly with the increase of the depth. The Eigen mode supported by this V groove, schematically shown in figure 19(a), is termed as the channel plasmon-polariton (CPP) mode. The CPP mode has received quite a lot of attention in the past few years and quite a few of plasmonic devices based on this mode have been experimentally realized [104, 105].

With the EIM approach, the V groove can be considered in the vertical direction as a stack of MIM waveguides with infinitesimal height and position-dependent insulator width. The effective index of each stack is determined by the insulator width and is thus decreasing from the V groove bottom to the top, if we consider the G-SPP mode here for the MIM structures. Compared to its application in the plasmonic slot waveguide, the EIM approach is a little trickier here because the structure of the multilayer medium needs to be solved to get the final mode effective index. Nevertheless, the EIM approach can still give a quite qualitative and also quantitative understanding of the CPP mode properties, including the dependence of the mode effective index, profile
and propagation loss on the groove angle and height. The use of EIM on CPP is also discussed in the literature [35]; here we choose to rely on the pure numerical technique provided by COMSOL software to calculate the CPP mode properties.

The dispersion of CPP as a function of the working wavelength is discussed in [106]; it is shown that for a shorter wavelength the mode is confined at the groove bottom while with an increase of the wavelength the fundamental CPP mode shifts progressively toward the groove opening, ceases to be guided at the groove bottom and becomes hybridized with wedge plasmon-polariton (WPP) modes, which are supported along the edges at the groove opening. This can also be understood from the EIM point of view because for the same waveguide geometry the requirement for an index contrast is higher for longer wavelengths.

For comparison, we present here the dependence of the mode property on the geometrical parameters. The results are calculated at the wavelength of 1550 nm for a groove milled into gold with a fixed groove angle $\theta$ and variable depth. The dielectric within the groove and above the metal surface is assumed to be air. In the numerical calculations, a filleting process is used at all the sharp corners of the groove. The bottom of the V groove is assumed to have a curvature radius of 50 nm. These curvatures remove numerical singularity and also provide a better fitting between the numerical model and the real fabricated structures. As one example, when $\theta$ is 20° and the groove depth is 3.0 $\mu$m, the mode profile, including both the amplitude and vectorial distribution of the electric field, is shown in figure 19(b). One can see that the mode is highly confined within the groove and the field becomes stronger toward the bottom. Also visible in the same figure is that some optical field exists at the two corners of the opening, illustrating a certain degree of hybridization. To be strict, one can refer the mode supported by a V groove with infinite depth to be CPP ($\infty$) and that by an individual top corner of the groove to be WPP ($\infty$).

Figure 19(c) gives the electric-field amplitude and the electric-field vector distribution of the WPP ($\infty$) mode supported by an individual corner of the same V groove in figure 19(b), obtained by leaving only the left boundary of the groove and replacing the metal to the right of it with air. One can see that both $E_x$ and $E_y$ exist around the corner since the electric field is perpendicular to the metal surface. For the CPP mode supported by a V groove with finite depth, a certain degree of hybridization between the CPP ($\infty$) mode at the bottom and the two coupled WPP ($\infty$) modes at the top always exists. More results for different groove angels show that when the groove angle is small, the mode tends to be tightly confined to the bottom and as $w$ increases, the confinement becomes looser and the mode shifts gradually to the groove opening with the hybridization to the WPP modes in a similar way to the case when the working wavelength is increasing. This is due to the fact that the index contrast is lower in the vertical direction as the angle increases. Numerical results also demonstrate an important feature of the CPP mode that when it is tightly confined, the electric field within the groove is mainly along the $x$ direction. This can be seen from the vector distribution of the electric field in figure 19(b). Recalling the requirement for the validity of EIM discussed in the last subsection, one can expect that EIM works fine in this case. As the groove widens, the hybridization degree between the CPP mode at the bottom and the two WPP modes becomes higher. More of the $E_y$ component will appear in the mode as well as $E_x$, then the semi-vectorial EIM approach gradually loses its accuracy.

Figure 20 shows the dependence of the mode effective index and propagation length for the fundamental CPP mode as a function of the groove depth when the angle is set to be 20° and 15° respectively. The mode effective index for the WPP ($\infty$) mode at one single groove corner when the groove angle is 20° is also shown as well as that for the SPP mode at infinite gold/air interface at the same wavelength. One can see from the results that as $d$ increases, for the groove angle of 20° the CPP mode effective index also increases, while for the angle of 15° the CPP mode effective index first increases and then slightly decreases. Note that for the angle of 20° when the depth is smaller than a certain value, the effective mode index for the CPP mode is smaller than that of the WPP ($\infty$) mode or even the SPP mode on the infinite gold/air interface, which indicates that the CPP mode has a cut-off depth $d_c$. This can

---

**Figure 19.** (a) Schematic of a triangular V groove with angle $\theta$ and depth $d$ milled into the metal substrate. The material in the groove and above the metal is assumed to be air in this paper. (b) CPP mode profile including the electric-field amplitude and the electric-field vector distribution inside a V groove with $\theta$ of 20° and depth of 3000 nm. (c) Electric-field amplitude and the electric-field vector distribution of the WPP mode supported by barely the left corner of the V groove in (b). The wavelength in both (b) and (c) is 1550 nm.
be easily understood considering the mode guiding mechanism from the effective index contrast in the vertical direction. For the 15° angle, the cut-off depth is not seen in figure 20(a) since the effective contrast is higher. From figure 20(b) one can see that the propagation length is increasing as the depth increases when the depth is larger than \(d_c\). When the mode is cut off, the loss will be lower as the depth further decreases, since more mode power is coupled to the SPP mode at the infinite gold/air interface.

The CPP mode characteristics, including the confinement and propagation loss, are determined by both the groove angle and depth. From the EIM point of view, the V-groove angle determines how fast the width of the MIM stacks and then the effective index change in the vertical direction. A sharper angle and larger depth will result in a tighter mode confinement at the V groove bottom, but will also introduce some more fabrication challenges. In practice, one usually changes the angle to be within the range between 15° and 25° for gold with air in the groove to get a highly confined mode with a reasonable propagation length for telecom wavelengths.

In the experimental studies of CPP based plasmonic waveguides and components, FIB milling is the main fabrication technique reported to date to carve the triangular groove into the metal, and SNOM imaging is usually used for characterization. By scanning the SNOM probe along the sample surface at a constant distance of a few nanometers maintained by shear-force feedback, the near-field information of the CPP mode can be collected, from which the main CPP characteristics, including the mode effective index, width and propagation length, can be determined directly. Good agreement between the experimental results obtained in this way and the theoretical results from COMSOL for V grooves supporting strongly confined (\(\lambda/5\)) CPP modes has been achieved [107].

4.5. SPP guiding by nanowires

In order to not lose generality, we also have a short review of the cylindrical metal nanowires as plasmonic waveguides. This type of plasmonic waveguide is different from those discussed previously in section 4 in the geometrical form and also in that the metallic nanowire is usually chemically synthesized but not fabricated with clean room facilities. The special approach of chemical synthesis renders these nanowires single-crystalline and not so vulnerable to surface roughness [108]. The cross section of the nanowire is circular thus the mode property can be solved analytically, just like the regular optical fibers. We are not going to discuss nanowires with rectangular or square cross sections here, which can be classified as the metal stripes reviewed in section 4.1.

Using the boundary conditions in the cylindrical coordinates, the dispersion equation for different order modes supported by the metallic nanowires can be solved analytically. For example, the fundamental mode is a pure TM mode and its characteristic equation is [109]:

\[
\frac{\varepsilon_1 I_1(\xi_1) + \varepsilon_2 K_1(\xi_2)}{\xi_1 I_0(\xi_1) + \xi_2 K_0(\xi_2)} = 0. \tag{15}
\]

In the equation \(\varepsilon_1\) and \(\varepsilon_2\) are the permittivity for the nanowire and the surrounding cladding respectively, \(I_j\) and \(K_j\) are the \(j\)th order modified Bessel functions, \(\xi_1 = \gamma a (i = 1, 2)\) where \(a\) is the radius of the nanowire and \(\xi_2\) is defined by \(\xi_2 = \sqrt{\beta^2 - \varepsilon_2 k_0^2}\) in which \(\beta\) is the mode propagation constant and \(k_0\) is the free-space wave-vector.

A similar dispersion equation for higher order modes can also be derived. Note that the higher order modes have both the electric-field component and magnetic field component in the longitudinal direction; they are no longer TM modes. Using these dispersion equations, one can solve the mode property analytically, including the mode effective index, profile and the propagation loss. Alternatively, one can also use the numerical technique to calculate the mode properties directly. As an example, we calculate the mode properties at 633 nm for a silver (\(\varepsilon = -19 + 0.53j\) [18]) nanowire cylinder with air as the cladding. When the nanowire has a subwavelength radius, it supports two modes which do not have a cut-off radius. The fundamental mode is a TM mode, whose electric-field magnitude and vectorial distribution are shown in figure 21(a). One can see that the azimuthal order is 0 and the electric field
is normal to the nanowire surface and decays exponentially away from it. Although simple, this mode provides a tight confinement of the mode to the nanowire surface; it has been explored extensively in recent years for subwavelength radiation guiding. The next mode is a hybrid mode and has azimuthal distributions, whose mode profile is shown in figure 21(b).

In figure 22 we plot both the mode effective index and the propagation length for the TM mode and the 1st order mode supported by a silver nanowire at 633 nm as a function of its radius. One can see that for the TM mode, there is no cut-off radius and as the nanowire radius decreases, the mode effective index increases. This implies a larger field decaying constantly in the surrounding medium and a stronger confinement of the mode to the nanowire surface. As the radius decreases, the propagation length becomes shorter because more mode power is localized in the metal nanowire. The property of this mode is quite similar to the SR-SPP mode supported by a thin metal film in a symmetric environment. As a contrast, the 1st order mode is more like the LR-SPP mode supported by the metal film, with the mode effective index approaching the cladding index and the propagation length diverging as the nanowire radius decreases. One can imagine that when the nanowire is thin enough, it can almost be ignored and the mode field evolves into a free-space light.

Because of the asymmetric characteristic profile which is hard to excite, the SR-SPP mode by a thin metal film is not preferred for information transmission; however, the TM mode here has a radial polarization and therefore has a large overlap with the mode from a cylindrical dielectric nanowire or nanofiber. Theoretical results from the body-of-revolution finite-difference time-domain (FDTD) calculations show that coupling efficiencies above 95% in the visible and close to 100% in the near-infrared can be achieved with realistic parameters for a metal nanowire butt-coupled to a dielectric nanowire [110]. In practice, this perfect end-to-end scheme is hard to achieve; the metal nanowire is usually placed in close vicinity with another dielectric nanofiber with a certain length of overlap in the propagation direction. Experimental results demonstrate that a photon–plasmon coupling efficiency up to 80% with the coupling length down to the 200 nm level is achieved between individual Ag and ZnO nanowires at the wavelength of 650 nm [111]. The high coupling efficiency makes it possible to realize the hybrid integration between nanowire plasmonics and dielectric nanowire photonics, and then circumvent the problem of high propagation loss at the metallic nanowire part. Using a similar approach with a waveguiding nanoscale fiber taper for highly efficient launching, the propagating SPPs are excited in the silver nanowire with high efficiency and the propagation loss can be measured [112]. A typical value of 0.41 dB \( \mu \text{m}^{-1} \) in a 260 nm diameter silver nanowire at 633 nm wavelength is obtained therein, which is even lower than that for the TM mode obtained by theoretical calculations. This discrepancy may be attributed to the fact that part of the optical power excites higher order modes as well as the fundamental TM mode. Characterization of the pure bending loss in curved crystalline silver nanowire plasmonic waveguides by decoupling the power loss caused by bending and propagation is reported; the experimental bending loss agrees with the theoretical simulation results quite well [113]. Hybrid nanophotonic components fabricated out of coupled Ag and ZnO nanowires, including polarization splitters, MZIs and microring cavities [111], as well as controllable routers and multiplexers based on branched silver nanowires [114], have also been experimentally reported, demonstrating the potential application of nanowire plasmonics to realize various plasmonic components with different functionality.

The chemical approach of metal nanowire synthesis offers the possibility to realize nanowires with high quality. The drawback of this approach is that it cannot be controlled to achieve complicated plasmonic structures at will, as opposed to nanofabrication techniques. Thus one has to use the manual micromanipulation technique to assemble the nanowires into the required geometry, which is rather tricky. The lack of a stable and mass production technique somehow limits the
application of metallic nanowires in further applications of subwavelength photonic integration.

4.6. SPP guiding along metal wedges

The last plasmonic waveguide geometry we would like to include is the metal wedge atop the metal substrate, whose cross section is schematically shown in figure 23(a). Although the geometry is completely complementary to the V-groove structure, the WPP supported by the wedge structure has a distinctive property as the CPP mode.

Similar to what we have done for V-groove modeling, considering that for real structures, the top tip can not be perfectly sharp and also in order to not generate singularity in the numerical simulations, in all our calculations we truncate the sharp end to be round with a radius of 10 nm and assume the wedge bottom edges have the curvature radius of 100 nm. Figure 23(b) illustrates the mode profile including the electric-field amplitude and the electric-field vector distribution supported by a gold \( n = 0.55 + 11.5j \) wedge in air at 1550 nm with a wedge angle of 20° and height of 1300 nm. One can see that the polarization is very different from that of the CPP mode. In the tightly confined CPP mode shown in figure 19(b), the major component of the electric field is normal to the V-groove sidewalls, or approximately parallel to the metal surface. But for the WPP mode, the polarization is radially distributed at the end, more like that of the TM mode supported by a metallic nanowire; there is no major component in the electric field or in the magnetic field. According to our discussion in section 4.3 on the applicability of the EIM technique, we can not expect that EIM works with high accuracy for the WPP modes.

We also note that the CPP mode profile is symmetric along the central plane, but the WPP mode is actually anti-symmetric. The underlying physics can be easily understood by considering the fact that the V groove can be considered as a stack of MIM structures in the vertical direction and the
CPP mode is based on the G-SPP mode; while the wedge structure should be regarded as a stack of metal stripes whose SR-SPP modes constitute the WPP modes. The LR-SPP mode supported by the metal stripes cannot be used here because its mode effective index is lower than that on an infinite metal/air interface. Thus the CPP and WPP modes inherit the symmetry characteristics of the G-SPP mode of the MIM structure and the SR-SPP mode of the metal film respectively. Keeping in mind that the SR-SPP mode effective index is more sensitive to the metal stripe thickness than the G-SPP to the insulator depth, one can expect that the WPP mode is more localized at the wedge tip than the CPP mode at the groove bottom.

Figure 24 shows the evolution of the mode effective index and the propagation length of the WPP mode as a function of the gold wedge height when the angle is fixed as 20°. The mode effective index of the SPP at the infinite gold/air interface at the same wavelength of 1550 nm is also given as the dotted line.

4.7. Comparison of SPP waveguides

Since a variety of plasmonic waveguide geometries have been proposed and investigated to date, it is natural to compare the main waveguide configurations in terms of their key characteristics, such as waveguide mode confinement and propagation length, which are expected to exhibit significant differences. Understanding these differences is very important for one to choose appropriate plasmonic waveguides for specific applications in different situations. In this part we will compare arguably the most developed and investigated plasmonic waveguide configurations: LR-SPPW, DLSPPW, LR-DLSPPW, CPP and HSPPW. Due to the use of Si in the considered HSPPW and in order to provide a fair comparison of these waveguides, the working wavelength is set at 1.55 μm. Since the mode characteristics can be very different depending on the waveguide cross section dimensions, unless otherwise stated we choose one typical configuration for each of these waveguides. The standard dimensions for the considered waveguide configurations include Au film (2 μm by 30 nm) in LR-SPPW (figure 12(a)), PMMA (600 nm by 600 nm) and Au film (70 nm thick) in DLSPPW (figure 9(a)), V groove (1500 nm deep and 300 nm opening width) in CPP (figure 19(a)), high-index Si stripe (width 200 nm and height 340 nm), low-index SiO2 (width 200 nm and height 50 nm) and gold layer (width 200 nm and height 60 nm) in HSPPW (figure 11(c)). The dimension of the LR-DLSPPW is the same as in figure 10(a) except that the buffer layer thickness has been changed from 255 to 360 nm. In the calculations of CPP modes, a filleting process is used (as in section 4.4) to set the curvature radius at the V groove bottom and opening corners, which should be 10 nm and 50 nm respectively. Although this process does not affect the groove angle, it reduces the actual depth to 1430 nm and increases the groove opening width to roughly 350 nm.

4.7.1. Mode confinement. We use the power distribution, i.e. the Poynting vector component \( P_z(x, y) \) which is normal to the waveguide cross section, to characterize the mode confinement. There are two quantities which are often used to characterize the mode confinement, the mode area \( A_c \) and the lateral (in the surface plane) mode width \( w_0 \). \( A_c \) is defined as the area size encompassing half of the total power. In the definition of \( w_0 \) one has to average the mode...
light intensity distribution along the vertical coordinate, thereby also eliminating the influence of strong fields around the metal corners. We realize this concept by first integrating $P_z(x, y)$ along the $y$ direction and thus obtaining the lateral power distribution $P_z(x) = \int P_z(x, y) \, dy$, from which $w_0$ is defined as the separation between two lateral points where $P_z(x)$ decays to $1/e^2$ of its maximum. Table 1 presents the mode area, lateral width and the propagation length for the five plasmonic waveguides.

If we compare the mode confinement of five plasmonic waveguides in terms of the mode area, from the results one can see that the HSPPW employing the CGD mode in the vertical direction has the best mode confinement. In contrast the LR-SPPW has the worst mode confinement with the mode area much larger than $\lambda^2$ although it has the longest propagation length. Among the considered waveguides, the LR-DLSPPW exhibits remarkably both the subwavelength mode confinement and the propagation length of up to several millimeters, which is close to that of the LR-SPPW. Note that for planar photonic integration, in most cases, it is the lateral mode width and not the mode area that determines the available component integration density. One can see that the CPP based on the G-SPP mode has the smallest $w_0$ among the five waveguides.

Although the mode characteristics are dependent on the waveguide dimension, the data shown in table 1 provides one with a rough idea of those waveguide properties. We can also see straightforwardly the general trade off found in plasmonic waveguides between the mode confinement and propagation loss.

### 4.7.2. Cross talk and bend loss.

To reach a high photonic integration, the key point is the component density achievable in the sample plane, i.e. the number of components with proper functionalities that can be placed together in a unit area. From this perspective, the cross talk between adjacent waveguides, determining how closely two waveguides can be placed, and the waveguide bend loss experienced by the mode propagating around a bend with a given curvature radius are really crucial characteristics.

The cross talk between two waveguides laterally placed in parallel can be characterized using the coupling length $L_c$, the distance after which the optical power is completely transferred from one waveguide to the other. $L_c$ can be calculated by considering the mode effective index difference between the two supermodes when two plasmonic waveguides are involved in the mode calculations, using the formula $L_c = 0.5w_0/|n_{\text{eff}s} - n_{\text{eff}t}|$ where $n_{\text{eff}s}$ and $n_{\text{eff}t}$ are the effective mode index for the symmetric and anti-symmetric supermodes respectively. Figure 25 presents the coupling length versus the center-to-center distance between adjacent waveguides for these five plasmonic waveguides. Since these waveguides have different propagation losses, it makes more sense to compare the absolute values of $L_c$ with the propagation length of SPPs in a single plasmonic waveguide. We increase the center-to-center distance until a coupling length larger than 10 times the propagation length for a single plasmonic waveguide is achieved. Two plasmonic waveguides can be considered isolated when $L_c$ is much larger than the propagation length because within the distance of the propagation length, the coupling from one waveguide to the other is very weak.

The results are more or less consistent with the mode area values of different waveguides. A large $A_e$ usually results in a large overlap between the modes of adjacent waveguides when one waveguide is placed into the mode field of the other. Then a higher coupling or larger cross talk can be expected. From the results one can see that for two LR-SPPWs to be isolated, the waveguide center-to-center distance should be more than 35 $\mu$m. This value decreases to be around 6 $\mu$m for LR-DLSPPWs and 2 $\mu$m for DLSPPWs since their mode sizes are much smaller. For two CPP waveguides or two HSPPWs to be isolated, the waveguide center-to-center distance only needs to be less than 1 $\mu$m due to their tight confinement. Although from table 1 one can see that the CPP waveguide mode area is larger than that of the HSPPW, the isolation distance is quite close. This is due to the mode property of the CPP, which extends in the vertical direction and thus has a small overlap in the lateral plane.

The bend loss is another important characteristic to evaluate a waveguide’s performance. The smallest radius of a bend that the light can propagate through with acceptable efficiency gives straightforwardly the footprint of the device composed of bent waveguides. To calculate the transmission through a plasmonic waveguide bend, one has to consider both the radiation loss due to deformed waveguide structures and the propagation loss through the bend. The former increases while the latter decreases with the decrease of the bend radius for most plasmonic waveguide configurations, resulting in an optimal bend radius at which the trade-off between radiation loss and propagation loss leads to the highest transmission.

The rigorous calculation of light transmission through a curved plasmonic waveguide bends requires heavy 3D computational electrodynamics effort. However, when the bending radius is larger than the mode width, the conformal transformation technique [118] can be adopted to convert the curved plasmonic waveguides into equivalent straight waveguides (ESWs). Then the overall loss, including both the radiation loss and propagation loss, can be roughly estimated by solving the mode effective index for the eigen modes of the ESWs. Since a spatially varying refractive index in the horizontal direction is used in the ESWs to incorporate the bend radius, perfectly matched layers (PMLs) are necessary in the COMSOL calculations to absorb the leaking wave from the waveguide core. We have compared the result for the light transmission through a 90° LR-SPPW bend at the 633 nm wavelength from this approach to that from a mode solver using the method of lines (MoLs) formulated in cylindrical coordinates [119], as well as the results for light transmission.

<table>
<thead>
<tr>
<th>Waveguide</th>
<th>LR-SPPW</th>
<th>DLSPPW</th>
<th>LR-DLSPPW</th>
<th>CPP</th>
<th>HSPPW</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_e$ ($\mu$m$^2$)</td>
<td>62.3</td>
<td>0.241</td>
<td>1.16</td>
<td>0.182</td>
<td>0.0588</td>
</tr>
<tr>
<td>$w_0$ ($\mu$m)</td>
<td>12.3</td>
<td>0.96</td>
<td>1.51</td>
<td>0.27</td>
<td>0.367</td>
</tr>
<tr>
<td>$L_c$ ($\mu$m)</td>
<td>7451</td>
<td>49.4</td>
<td>3125</td>
<td>19.7</td>
<td>26.7</td>
</tr>
</tbody>
</table>

Table 1. Mode area and propagation length for the five plasmonic waveguides.
through a 90° DLSPPW bend, both from this method and from 3D FEM calculations [120]. In both cases, a quite good agreement has been achieved. Light transmissions through a 90° bend based on LR-SPPW, DLSPPW and LR-DLSPPW are calculated using this approach. However, there are some restrictions for the ESW model to be acceptably accurate [121], e.g. there is a limitation of the smallest bending radius for a fixed dimension in the horizontal direction of the computational domain. For CPP and HSPPW bends, the bending radius can be so small that the horizontal dimension of the computational domain must be quite small as well. Then the PMLs must be placed quite close to the waveguide mode and then the overall loss may be slightly overestimated. For light transmission through 90° bends based on these two waveguides, we then performed full wave simulations using the FEM method with uniaxial PMLs used at the waveguide output port to truncate the simulation domain.

Figure 26 presents the calculated light transmission through a 90° bend as a function of the bending radius for the five plasmonic waveguides. From the results one can see that for a LR-SPPW bend, the bending loss will be quite high when the radius is lower than 10 mm, which is consistent with the results from the experimental investigation of LR-SPPW bends [25, 91]. Even at the optimal bending radius of 18 mm for a LR-SPPW bend, the transmission is still extremely low. This explains why in the experimental work on LR-SPPW devices no 90° bend has been used. Instead, S-bends with a fixed optical path and variable off-set distance are chosen. Beyond the optimal bending radius, light transmission begins to reduce because of a higher propagation loss through the longer optical path. The transmission through a 90° bend based on DLSPPW in figure 26(b), LR-DLSPPW in figure 26(c) and HSPPW in figure 26(e) also firstly increases and then decreases as a function of the bend radius, showing that the bending loss at a smaller bend radius or the propagation loss at a larger bend radius dominates the overall loss. Note that light can propagate through all those bends with a transmission higher than 75%; the optimal optical bend radius for the DLSPPW, LR-DLSPPW and HSPPW is around 6 µm, 150 µm and 1.0 µm respectively, showing the different ability of these waveguides to route light through sharp bends. For the CPP waveguide, we can see in figure 26(d) that the transmission through a 90° bend of it monotonically increases as the bend radius is decreasing even to a value close to the width of the V groove, demonstrating that the radiation loss is much lower than the propagation loss.

We also need to note that for the ESW model the real part of the effective mode index for a curved waveguide deviates from that of a straight waveguide. So when a straight plasmonic waveguide is connected to a curved waveguide there is additional coupling loss, or in other words transition loss, due to the reflection at the interface, which is not taken into account in the above calculations.

Having discussed cross talk and bend loss, we note that the two characteristics can be linked to the waveguide mode confinement ability. While the cross talk between adjacent plasmonic waveguides is dependant more on the mode area $A_e$, which determines the overlap between adjacent waveguide modes, the bending loss is more related to the lateral mode width $w_0$, which shows the extent of the waveguide mode in the lateral dimension.

4.7.3. Figures of merit. Different plasmonic waveguides have different mode confinements and propagation losses as well and it is not intuitive to compare them directly due to the general rule between the two quantities in plasmonics. So it is more convenient to propose a figure of merit (FOM) which takes into account both quantities and then one can use the FOM to assess a plasmonic waveguide.

Berini et al proposed a simple FOM [122], which takes into account both the waveguide propagation length, $L_p$, and the mode area, $A_e$. As we mentioned above, the key point of realizing high photonic integration is to accommodate as many plasmonic devices as possible within a certain work
plane. This requires the waveguide to have minimum lateral crosstalk and bending loss. Here we adopt the new FOM proposed in [83], which considers the lateral mode width, $w_0$, as well as the guided wavelength and the propagation length. This new FOM, defined in equation (16), has a direct relation to the maximum number of components that can be integrated on a chip within an area limited by the corresponding SPP waveguide mode propagation length, thereby taking into account the integration potential along with the insertion (propagation and bend) loss:

$$\text{FOM} = \frac{L_p^2 \lambda_0}{n_{\text{eff}} w_0^3}. \quad (16)$$

Table 2 gives the new FOM values for all these five plasmonic waveguides. From the results one can see that the LR-DLSPPW among the considered five waveguides has the largest value of the new FOM. This can be attributed to its low propagation loss and good mode confinement.

The new FOM provides a direct comparison among different plasmonic waveguides in terms of the device integration density. Note that the choice of plasmonic waveguide under specific situations is application oriented. For example, CPP waveguides or other plasmonic waveguides based on G-SPP modes should be considered first when sharp bends are required.

5. Excitation of SPP waveguide modes

Having discussed the mode properties of some representative plasmonic waveguides, in this section we will briefly review some general techniques used to excite those plasmonic modes. The guideline to excite the SPP waveguide mode is either to provide the wave-vector matching required for SPP propagation in the relevant direction, usually along the interface between the metal and a low-index ($n_2$) dielectric. Since the effective index of the SPP mode $n_{\text{eff}}$ is larger than $n_2$, the incident angle $\theta$ in the prism should be larger than $\theta_c = \arcsin(n_2/n_1)$, where $\theta_c$ is the critical angle for the TIR between media $n_2$ and $n_1$. This means the prism coupling scheme works in the TIR condition for SPP excitation. When the SPP is successfully excited, the reflection from the prism sees a significant attenuation due to energy conservation. So this scheme also derives the name of attenuated total reflection (ATR). Depending on the position of the metal film with respect to the prism and the low-index medium, there are historically two different configurations, namely Kretschmann and Otto configurations respectively. Note that there is actually a destructive interference occurring in the reflection when the SPP is excited. One knows that for a perfect destructive interference, both the amplitude and phase conditions should be fulfilled. While the phase condition is actually the wave-vector matching, which can be fulfilled by tuning the incident angle, the amplitude condition should be achieved by changing the metal film thickness in the Kretschmann configuration or the gap between the prism and the metal layer in the Otto configuration.

The prism coupling scheme can also be used to excite the plasmonic waveguide modes, provided the same wave-vector matching condition as for the excitation of the SPP on a
planar interface can be met. The plasmonic waveguides with a relatively short length can be placed on one surface of the prism where the characterization of SPP propagation is also convenient and accessible. This has been widely used in the early study of LR-SPPW modes [23, 123, 124]. Nowadays this technique is also heavily utilized for investigating DLSPPWs, some tapering structures are preferred to enhance the coupling efficiency [125]. Short metallic nanowires working as Fabry–Perot resonators have also been investigated by placing the nanowires on a glass prism and illuminating the sample under TIR through the prism [126].

Due to the difference between the relatively large incident laser spot and the small transverse dimension of DLSPPWs, some tapering structures are preferred to enhance the coupling efficiency [125]. Short metallic nanowires working as Fabry–Perot resonators have also been investigated by placing the nanowires on a glass prism and illuminating the sample under TIR through the prism [126].

As we have mentioned in section 3, there are some drawbacks with the prism coupling technique. It is limited by the material refractive index which is naturally available, so the plasmonic waveguide mode with a really high mode effective index cannot be excited with this scheme. The prism is usually bulky, which is OK for investigations of plasmonic component properties, but cannot be incorporated into subwavelength photonic integration. What is also worth noting is that the plasmonic waveguide mode excited by the prism coupling can also leak to the substrate because of the presence of the high-index prism.

Another way of converting free-space optical radiation into the SPP waveguide modes is to use a grating to circumvent the wave-vector mismatch between the impinging optical wave and the SPP waveguide mode. The superimposition of the grating reciprocal vector with the projection of the free-space light wave-vector along the waveguides provides the required propagation constant for the plasmonic waveguides. This can be easily demonstrated using the mathematical equation:

\[ k_0 \sin \theta + mg = \beta \]

in which \( \theta \) is the incident angle, \( \beta \) is the plasmonic waveguide propagation constant, \( m \) is an integer and \( g = 2\pi / \Lambda \) where \( \Lambda \) is the grating period.

As a general phase matching technique, grating coupling can find broad applications. For instance, it is used for the excitation of conventional optical waveguides and the recent SOI waveguides [127]. The use of it for the excitation of SPP on an infinite interface has also been discussed in section 3.1. For SPP waveguide modes, the grating area is usually large to collect more power from the free-space excitation while the SPP waveguide transverse dimension is in the subwavelength dimension. Despite the fact that grating coupling still works with the large dimension mismatch, as can be seen in the excitation of organic nanofiber-loaded SPP waveguides with metallic gratings [128], it will increase the coupling efficiency significantly if some adiabatic mode transformation configuration can be used [129].

### 5.2. Excitation with strongly focused radiation

As another option besides prism and grating coupling, a strongly focused optical beam, usually from a microscope objective of high numerical aperture, when incident onto a SPP waveguide, can excite the SPP waveguide modes, because the subwavelength dimensional SPP waveguide can also be regarded as a defect and thus can provide the Fourier component with the high spatial frequencies necessary for wave-vector matching [130]. The guided mode excitement in a DLSPPW was successfully demonstrated using the illumination from a 20× objective with 0.35 numerical aperture [74]. Note that the polarization of the incident focused beam should be aligned properly with regard to the SPP waveguide to excite the mode.

For specific SPP waveguides, some special structures in waveguides can still be explored to further harvest the focused optical beam. For example, CPP modes can be excited efficiently using tapered terminations of V-shaped groove waveguides working as nano-mirrors [131]; numerical simulations show that the optical nano-antenna can also be used to drastically enhance the reception efficiency for plasmonic slot waveguides [132].

### 5.3. Fiber end-fire coupling

Besides the excitation from free-space optical radiation, it is also possible and sometimes preferable to use the guided photonic mode from another waveguide to excite the SPP modes. The spatial mode profile similarity, or in other words the mode overlap degree between the photonic and plasmonic modes, is now crucial to get a high coupling efficiency. We focus on fiber end-fire coupling exclusively in this subsection because an optical fiber is the most accessible optical waveguide in the lab. The fiber or fiber taper is usually suspended in the air with such a short distance from the plasmonic waveguide that the guided mode profile in the fiber is still maintained when the light is incident onto the plasmonic waveguide end.

End-fire coupling from a polarization maintaining fiber has been broadly used in the experimental investigation of LR-SPPW modes supported by metal stripes with micrometer-size widths [25, 90, 91, 133]. Note that the transverse distribution of the LR-SPPW mode shown in figure 12(b) is by nature not Gaussian, but consists of two exponentially decaying tails extending from the metal film surface into the surrounding dielectric cladding layers. However, the overlap integral of the symmetric exponents with a Gaussian distribution can be even better than 97% [90]. This ensures that a high coupling efficiency for LR-SPPW modes can be achieved using the fiber end-fire approach.

The DLSPPWs working at telecom wavelengths usually have sub-micrometer dimensions while the typical diameter of

<table>
<thead>
<tr>
<th>Waveguide</th>
<th>LR-SPPW</th>
<th>DLSPPW</th>
<th>LR-DLSPPW</th>
<th>CPP</th>
<th>HSPPW</th>
</tr>
</thead>
<tbody>
<tr>
<td>New FOM</td>
<td>(3.2 \times 10^4)</td>
<td>(3.4 \times 10^3)</td>
<td>(3.2 \times 10^6)</td>
<td>(2.9 \times 10^4)</td>
<td>(1.1 \times 10^4)</td>
</tr>
</tbody>
</table>
the light spot from a tapered fiber is still a few micrometers. This dimension difference makes it hard to use fiber end-fire coupling to directly excite the DLSPPW modes. What is more, the propagation length of DLSPPWs at telecom is typically a few tens of micrometers and the mode signal may be buried in the background noise from the fiber taper if the end-fire approach is used directly. In practice, the DLSPPWs can be connected via funnel structures with access to polymer waveguides extending outside the metal covered regions of DLSPPWs all the way up to the substrate edges. The polymer waveguide has a photonic mode with the dimension close to that of the single mode optical fiber, thus the DLSPPW mode can be excited through the fiber-to-photon-to-DLSPPW transition [77].

5.4. Hybrid integration (coupling via dielectric waveguides)

Plasmonic waveguides can provide advantages over photonic waveguides in many aspects, e.g. subwavelength mode confinement, high transmission through sharp bends and the capacity of carrying both electric and photonic signals simultaneously. However, the main problem is the propagation loss, especially for highly confined modes, which restricts the application of plasmonic waveguides for information transmission over a relatively long distance. One may imagine that if the seamless integration between a photonic waveguide and a plasmonic waveguide can be realized on the same platform, then the advantages of both waveguides can be combined and many existing problems can be circumvented. For example, the photonic waveguide can be used for the rest of the chip while for places where sharp bends are present or a large electric field is required to enhance the nonlinear response, plasmonic waveguides should be used. In essence, the realization of practical on-chip plasmonic devices cannot get rid of the use of photonic waveguides on the same chip. In addition, the hybrid integration between plasmonic waveguides and dielectric waveguides also provides a new means of SPP waveguide mode excitation, which facilitates the experimental investigation of plasmonic components.

In recent years quite a lot of research effort has gone into exploring the possibility of this hybrid integration for various plasmonic waveguides, both theoretically and experimentally. The Si photonic waveguide constructed on the SOI platform is a very important dielectric waveguide representative and achieving a seamless transition between the Si and plasmonic waveguide systems is attractive, because it allows both plasmonic and Si photonic functionalities to be incorporated on the same substrate, paving the way to the eventual monolithic integration of photonics, plasmonics and electronics all on the same chip. A lot of work has been done to explore the possibility of integrating various plasmonic waveguides on the SOI platform.

Efficient, broadband light transmission between a SOI waveguide and a plasmon waveguide, as shown schematically in figure 16(b), supporting an edge plasmon mode was experimentally demonstrated on the SOI platform and 3.4 dB of insertion loss was achieved [134]. This is one of the earliest attempts demonstrating hybrid integration between SOI waveguides and plasmonic waveguides. For the G-SPP mode, the numerical analysis predicts at a 1.55 μm wavelength a coupling loss of less than 2 dB from a micrometer-size dielectric waveguide to a 50 nm wide plasmon gap waveguide by a 6 μm short optimized tapered plasmon waveguide coupler [135]. The same level of coupling efficiency was also obtained by simply placing a high-index contrast dielectric slab waveguide terminated flat at the exit end of a plasmonic waveguide [136]. However, these simulations are just 2D and the confinement in the third dimension is not considered, so the results may be too optimistic. For 3D plasmonic slot waveguides of the asymmetric case (silica as the substrate and air in the slot), a surface plasmonic coupler composed of a tapered silicon strip waveguide and a subwavelength scale metal gap waveguide was experimentally demonstrated using mainly FIB milling; a coupling efficiency of 35% was achieved [100]. For a similar horizontal plasmonic slot waveguide but with PMMA as the top cladding and in-slot material, to provide index matching in the vertical direction and obtain mode matching between the SOI and plasmonic slot waveguides, the SOI-taper-plasmonic-slot-funnel coupler was fabricated using a combination of EBL and Ar ion milling, demonstrating a direct coupling efficiency around 30% between the two waveguides systems [101]. The plasmonic slot waveguide mode presents a higher resemblance to the Si slot waveguide than the SOI waveguide. A coupling efficiency between the Si slot waveguide and quasi-MIM plasmonic waveguide up to 43% was obtained [137].

The hybrid plasmonic waveguide with Si working as the high-index material shown in figure 11(e) can be naturally built on the SOI platform. Theoretical results from COMSOL show that a coupling efficiency above 90% at 1.55 μm can be achieved for a SOI waveguide evanescently coupled to an Ag–SiO₂–Si hybrid plasmonic waveguide with quite a short overlap length [138], demonstrating the potential of this hybrid plasmonic waveguide to be integrated into a silicon-based platform. Experimental investigation of the SOI waveguide butt-coupled to a Au–SiO₂–Si waveguide with a tapering structure based on the same hybrid plasmonic waveguide also presented a high level of coupling efficiency, around 80% [85].

Among the various plasmonic waveguides proposed to date, the DLSPPW is the most extensively investigated. A comprehensive theoretical analysis of the coupling between DLSPPWs and rib/wire silicon photonics waveguides has been presented and the effect of practical issues in the fabrication on coupling efficiency is also discussed [139]. An optimal coupling efficiency around 80% from the SOI rib waveguide or 85% from the SOI wire waveguide to the DLSPPW can be achieved theoretically. The experimental coupling efficiency results agree pretty well with the theoretical predictions for the corresponding structures. With the high coupling efficiency achieved in the experiment, a new way of measuring the propagation loss in DLSPPWs other than the use of LRM or SNOM is provided and the properties of some plasmonic components based on DLSPPWs can be evaluated directly [139]. Those results demonstrate that DLSPPWs can be efficiently interfaced with optical systems fabricated on a SOI chip to combine the advantages of plasmonics with the
low propagation loss of Si photonic circuits. Relying on efficient coupling between low-loss SOI waveguides and the DLSPP waveguides, the transmission spectra of waveguides-resonator resonators at telecommunication wavelengths can then be measured directly [140].

Besides the SOI platform, the semiconductor nanowire from chemical synthesis [141] or the manually drawn silica nanofiber [142] utilizing the high-index contrast between the core material and the air cladding, can also work as efficient photonic waveguides with a subwavelength diameter. The fundamental mode supported by these cylindrical photonic waveguides and the TM mode in a metal nanowire demonstrate an even higher resemblance compared to those waveguide pairs discussed above, and thus one can expect a high coupling between the two modes. The coupling between dielectric nanowire and metal nanowires has been discussed in section 4.

There are also some reports in the literature on the coupling between the dielectric waveguide mode and plasmonic mode with less confinement. For example, the coupling between the LR-SPP mode and the TM mode of a conventional dielectric waveguide which is placed in parallel with the metal nanowire has been discussed in section 4.7, the pure bending loss of a waveguide bend based on this mode usually needs a large bend radius to have a low bending loss. For a 2 \( \mu \text{m} \) wide and 25 nm thick gold film in the symmetric environment with a SiO\(_2\) substrate and a gel upper cladding having the refractive index matched to SiO\(_2\), the experimental results show that at the working wavelength of 1550 nm, the S-bends with a fixed overall length of 4 mm; radii of curvature smaller than 12.5 mm will suffer from significant radiation losses [91]. This is consistent with the numerical result of SPP transmission through a LR-SPPW bend in figure 26(a). This large bend radius makes the symmetric Y splitter have the same level of radius of curvature as well. For a thinner (15 nm) and wider (8 \( \mu \text{m} \)) gold stripe waveguide embedded in the BCB polymer, since the confinement is even poorer, quantitative experimental investigation of the bending loss as a function of bend radius shows that when the bend radius is about 20–25 mm, the additional loss due to the S bend is around 2 dB [25]. This large bend radius also results in a large footprint (length: 10 mm, width: 250 \( \mu \text{m} \)) for a Y splitter based on the same S bend.

DLSPPWs with subwavelength transverse dimensions have a tighter mode confinement and shorter propagation length. Numerical simulations demonstrate that at the working wavelength of 1550 nm for a DLSPPW with PMMA cross section 600 nm \( \times \) 600 nm, the optimum bend radius of a 90° circular bend with both propagation loss and bending loss considered, is around 5 \( \mu \text{m} \) [120]. The Y splitter based on the same DLSPPW with the highest total output around 75% can have a footprint of 8 \( \mu \text{m} \) \( \times \) 2.6 \( \mu \text{m} \). For a DLSPPW consisting of a thinner SiO\(_2\) core (cross section 700 nm \( \times \) 60 nm), because the lateral confinement is worse, a larger bend radius is expected. Experimental results show that at a shorter wavelength of 800 nm, the optimum bend with the highest transmission has a radius as large as 10 \( \mu \text{m} \) [72].

Horizontal plasmonic slot waveguides and CPP waveguides employing G-SPP modes have a really tight mode confinement; light cannot penetrate metal which is aligned in the lateral plane. So for these waveguides one can expect a high light transmission through a bend with a small bend radius. Theoretically light can propagate with almost no additional loss through a sharp bend based on a 2D MIM waveguide [144] or 3D CPP waveguide [145] when the metal gap width is much smaller than the wavelength. The overall loss through a bend will be more affected by the propagation loss and out-of-plane scattering loss. Experimental results also confirmed the high transmission of light using 0.6 \( \mu \text{m} \) wide and 1.1 \( \mu \text{m} \) deep grooves bent with the smallest curvature radius of \( \approx \)0.83 \( \mu \text{m} \) [146]. Compact wideband Y splitters based on horizontal plasmonic slot waveguides [147] and CPP waveguides [104] have also been experimentally demonstrated and in all these devices the propagation loss rather than the bending loss dominates the overall loss.

For other plasmonic waveguides with a tight mode confinement, a small bend radius can also be realized. A sine-curved S bend (which connects two 10 \( \mu \text{m} \) off-set waveguides over a distance of 20 \( \mu \text{m} \)) based on LR-DLSPPW, is observed to show excellent performance with overall losses around 1.8 dB at 1450 nm [80]. Conductor-gap-silicon S-bend structures consisting of two 90° arcs with a radius
of 1 μm and 2 μm, as well as Y-splitters by interposing two plasmonic S-bends, were also fabricated and experimentally measured [85].

6.2. Wavelength selective SPP components

Passive SPP components which work with wavelength dependence provide further functionality to manipulate SPPs. These phase sensitive components, including gratings, MZIs, WRRs, DCs and multimode interference (MMI) devices, are very important and essential in application areas like plasmonic switching, wavelength multiplexing and demultiplexing.

The mode effective index of a waveguide is dependent on the width of the core. This effect is more evident for plasmonic waveguides compared to the conventional dielectric waveguides because the plasmonic mode is so sensitive to the core width, as stated in section 2. Even a minor change in the width may introduce a significant variation of the mode effective index. Based on this principle, the periodic variation of the plasmonic waveguide width along the propagation direction will lead to a modulation of the mode effective index, and the device of plasmonic gratings, in a similar way to the fiber Bragg gratings formed in regular optical fibers, can be realized. Plasmonic Bragg gratings like these have been proposed and numerically investigated based on the G-SPP mode in 2D MIM structures [148] and experimentally realized using CPP structures [105]. Introducing a defect into the MIM plasmonic gratings will form a Fabry–Perot cavity, which has shown potential applications in the semiconductor core of MIM nano-lasers [149]. The mode effective index is also related to the metal stripe width in LR-SPPWs and the PMMA stripe width in DLSPPWs; in-line Bragg gratings for wavelength selection in the telecommunication range have also been experimentally demonstrated [150, 151].

A WRR composed of a straight waveguide laterally coupled to a ring resonator usually provides more pronounced wavelength selection ability and is therefore another important component in integrated optics. Note that the extinction ratio, or in other words the contrast in the transmission through a WRR between resonant and non-resonant wavelengths, is affected by both the coupling efficient and the internal loss within the resonator. For a plasmonic WRR, the coupling efficient is usually quite low due to the tight confinement in plasmonic waveguides, while the internal loss is typically quite high as a result of the propagation loss. So the condition of critical coupling can hardly be realized to get a high extinction ratio [152]. Different schemes have been proposed to increase the extinction ratio for WRRs based on various plasmonic waveguides, including aperture-coupling [153, 154] for the G-SPP mode, in which the lateral coupling efficiency is extremely low due to the small skin depth of light in the metals, and the racetrack resonator scheme [152, 155], with which the length of the coupling region is increased intentionally.

Other characteristics including the quality factor, the bandwidth and the free spectral range are affected by the roundtrip length and the internal loss, both of which are related to the ring radius. The determination of the ring radius in WRRs depends on the roundtrip loss inside the resonator, including both the bending loss and the propagation loss, which are further determined by the trade-off between mode confinement and propagation loss of the plasmonic waveguide. Due to the low coupling, the high roundtrip loss as well as the fabrication difficulty, experimental realization of plasmonics WRRs with acceptable performance is quite challenging. To date plasmonic WRRs have only been fabricated and demonstrated using CPP structures [104, 105] and DLSPPWs [151]. A typical radius for these WRRs working at telecom is around 5 μm, which lead to a footprint of the WRRs to be less than 100 μm².

2 × 2 MZIs and DCs also show some wavelength selection behavior due to the fact that the coupling length between adjacent waveguides is wavelength dependent as a result of waveguide dispersion. The DCs should be designed with a proper length and waveguide gap so that the coupling is strong enough for the DCs to accommodate a complete transfer of power from one waveguide to the other within the length of the coupling region. Similar designing criteria should also be fulfilled for MZIs. Under these circumstances, only plasmonic waveguides with a relative low propagation loss can be used for constructing these two components. Realizations of MZIs and DCs pose less challenges compared to WRRs and have been experimentally demonstrated based on CPP structures [104, 156], LR-SPPWs [25, 91] and DLSPPWs [72, 73].

Some plasmonic waveguides can also support a multimode when the width is large enough and MMI devices based on the self-imaging property in the multimode region can also be achieved. Some numerical results [157–159] in the literature have shown the feasibility of MZIs devices based on certain plasmonic waveguides. Experimental realization of MMI couplers has been demonstrated using LR-SPPWs [25]. An efficient thermo-optic DLSPPW 2 × 2 switch has been recently experimentally demonstrated using a plasmonic–photonic dual mode interference configuration and a high thermo-optic coefficient polymer [160], showing the potential of these multimode devices in the application of plasmonic switches.

6.3. Radiation modulation with SPP components

Besides the manipulation of SPPs using those passive devices described above, plasmonic circuits should also incorporate active components capable of modulating, switching or re-routing the transmitted plasmonic signal. The active control of SPPs requires special material properties like electro-optic, thermo-optic, magnetic-optic and optical nonlinear effects, so the modulation depth and bandwidth are restricted by the naturally available strength and response time of these effects. Although these effects may be enhanced by the high electric field at the metal surface, they are inherently weak and also suffer from the short interaction length due to the high propagation loss of plasmonic waveguides. To date the successful realization of SPP modulation and switching in integrated plasmonics using SPP waveguides described in section 5 is achieved mainly based on the thermo-optic effect of polymers in LR-SPPWs [133, 161] and DLSPPWs [77, 160]. Besides the relatively low propagation loss in
the cavity has been modified to that of free space with a cladding for LR-SPPWs has been proposed [162] and ultrafast all-optical active nanoplasmonic devices based on optical free carrier generation and absorption in five-layer silicon nanoplasmonic waveguide structures has been investigated [87]. The experimental implementation of these effects in nanoscale plasmonic waveguides, however, poses great challenges in both fabrication and characterization which have yet to be solved.

Researchers have also made some progress in the active control of the planar SPP mode on an infinite metal film surface, which has been reviewed in section 3. A thin layer of active medium e.g. CdSe quantum dots [163] or photochromic molecules [164] is placed on top of the metal film; two gratings or slit structures described in section 3 can be used to convert between the free-space light and the propagating SPP mode on the metal surface. Another pump laser is incident into the region between the two gratings or slit structures to modify the optical property of the active medium and then the property of the planar SPP mode, so that the output from the second grating or slit structure back to free-space light will change correspondingly. This process can be quite fast depending on the electron-hole recombination time of the active medium. A novel ultrafast SPP modulator has also been realized by direct ultrafast optical excitation of the metal using a femto-second layer pulse incident onto the metal surface to disturb the equilibrium in the energy–momentum distribution of electrons [165], demonstrating the response time on the femto-second timescale and unprecedented terahertz modulation bandwidth.

### 6.4. SPP generation and detection

The SPP waveguides can be excited using those methods described in section 5. However, it will benefit the development of nanoplasmonics if coherent plasmonic fields can be generated directly within the waveguide.

When a nano-emitter is placed within the mode field of a SPP waveguide, the spontaneous emission from it may couple into the waveguide and lead to the generation of a SPP mode. As has been experimentally demonstrated [166], the emission from an optically excited single CdSe quantum dot in close proximity to a silver nanowire couples directly to guided SPs in the nanowire, offering many exciting prospects like the realization of single-photon transistors. Very recently propagating SPPs in thin metal films was reported to be launched using individual semiconducting single-walled carbon nanotubes upon excitation in the visible regime [167]. Stimulated emission of radiation from emitters may also couple directly into the form of a SPP mode, leading to the amplification of SPP. The idea of SP amplification by the stimulated emission of radiation (SPASER) [168], was firstly experimentally realized using a gold nanoparticle surrounded by a silica shell containing dye molecules that exhibit optical gain [169]. When pumped by short optical pulses, the dye molecules can create sufficient gain to overcome the losses in the gold, generating a coherent source of highly localized SPPs. Plasmonic nano-lasers from structures in the form of SPP waveguides have also been reported in which the end of the SPP waveguides can be used as reflectors to form Fabry–Perot cavities for wavelength selection. Examples of these include plasmonic nano-lasers in MIM type waveguides filled with electrically pumped semiconductor cores [170], or in HSPPWs with optically pumped CdS nanowires [86].

SPP amplification can also be achieved using the stimulated emission of a gain medium, which may lead to partial or full SPP loss compensation. This is especially favorable when the emission is coupled to strongly confined SPP modes, which usually exhibit high propagation loss. Examples of emitters that have been used for SPP amplification in the literature include quantum dots [78], dye molecules [171] and erbium ions which are doped in phosphate glasses [172].

Direct detection of the SPP signal in the plasmonic circuit without converting it back to light will further stimulate the realization of hybrid integration of plasmonics and electronics on the same platform. All-electrical detection of the SPP signal usually measures the tiny electrical voltage brought by the SPP field and requires a highly sensitive setup incorporated directly into the plasmonic circuit. There has been some progress demonstrated in this field of exploration, and highly confined SPPs in a MIM waveguide were detected and characterized by incorporating semiconductors as the insulator, forming an integrated metal–semiconductor–metal photodetector [173]. A similar plasmon diode, converting the electromagnetic field of the plasmons into a direct current, was also experimentally realized using an organic material between two silver layers [174]. The near-field coupling between guided SPPs and a Ge nanowire field-effect transistor was also used to electrically detect the plasmon emission from an individual colloidal quantum dot coupled to the metal nanowire SPP waveguide, demonstrating a nanoscale SPP detector with high efficiency [175].

### 6.5. Quantum plasmonics

Research on the interaction between a single emitter and the surrounding electromagnetic field is of great interest in terms of both fundamental research and device applications, constituting a major content in the discipline of quantum optics. Conventionally an optical micro-cavity [176] with a high quality factor and low mode volume is usually used to enhance the spontaneous emission rate of an emitter inside the cavity. The local density of optical states (LDOS) in the cavity has been modified to that of free space with a
quantity known as the Purcell factor, which is proportional to the ratio between the quality factor and the mode volume of the cavity. The unprecedented diffraction-unlimited mode confinement provided by plasmonics, albeit with certain losses, has opened new perspectives for the interactions between single emitters and surrounding electromagnetic fields, leading to a new branch of quantum plasmonics [177].

The enhancement of the spontaneous recombination rate in a quantum well placed in a SPP field supported by a silver film has been reported [178] and studied theoretically, demonstrating the strongest enhancement occurs near the SP frequency, at which the SP LDOS is at its maximum. Fluorescence enhancement of quantum dots or dye molecular was also reported when the dyes are placed in the vicinity of metal nanoparticles [179, 180] or other metal nanostructures like nano-prisms [181] which exhibit LSP resonances. Plasmonic emission enhancement from the emitters can also be spectrally controlled by changing the dimension or the environmental index of the metal nanostructures to tune the LSP resonances. Besides the influence of emission enhancement to neighboring emitters due to the strongly confined LSP field from metallic nanoparticles, those metallic nanostructures themselves also exhibit some quantum effects when their dimensions are decreased to the quantum size regime. Recent research showed that when a nanoparticle diameter is less than two nanometers, the LSP resonance shifts to higher energy by 0.5 electronvolts, a substantial deviation from classical predictions [182]. Understanding and prediction of the underlying mechanisms will make these metallic nanoparticles with quantum sizes find unprecedented application opportunities in various areas.

SPP waveguides with tight mode confinement can naturally provide ultra-small mode volumes for non-resonant wavelengths, suggesting the spontaneous emission enhancement in a broadband spectrum. In this respect, SPP waveguides based on the G-SPP mode, which has no cut-off width, are favorable in terms of both confinement ability and propagation loss. The spontaneous emission process of an optical, dipolar emitter in 2D MIM slabs and 3D plasmonic slot waveguide structures has been theoretically studied, exhibiting strong emission enhancements at non-resonant conditions [183]. CPP plasmonic waveguides and ring cavities also working based on the G-SPP mode can also provide an extremely small optical cavity mode volume. Theoretical calculations [184] show that Purcell factors well above 2000 can be observed over a broad spectral range with CPP ring cavities based on small angle V grooves, demonstrating the potential application of these cavities in cavity quantum electrodynamics.

In addition to its success in electronics, the fantastic material graphene with atom-size thickness has proved itself to be a suitable alternative to those noble metals widely used as SPP materials in the optical regime, because graphene also has a negative permittivity in certain frequencies and thus it can work as the SPP material. Theoretical calculations show that graphene plasmons exhibit a much larger confinement and a relatively longer propagation length compared to conventional SPP [185]. Graphene plasmons also inherit the properties of regular SPPs, e.g. field and emission enhancement, thus assisting the exploration of fundamentally new regimes of quantum plasmonic interactions at the nanoscale. Another advantage of graphene is that the property of the graphene plasmons can be easily tuned by varying the graphene carrier density with the electrical gating method, which provides an unprecedented new way to achieve electric control of light [186, 187]. Although still in its infancy, the application of graphene in plasmonics and photonics has demonstrated great potential.

7. Outlook

The ever increasing pace in the development of SPP-based waveguides and waveguide components has already resulted in many important achievements, demonstrating not only the unique capability of SPP waveguides for efficient concentration and manipulation of light in nanoscale regions but also their suitability for realizing dynamic and active light controlling functionalities. The SPP-based structures have also demonstrated their ability to deliver light energy to nanoscale optical and electronic devices, quantum dots and even separate molecules. The possibility of strong subwavelength localization of SPP waveguide modes makes these structures particularly useful for the future design and development of highly integrated and efficient nano-optical signal-processing devices and circuits. They are also expected to provide an essential and efficient link between conventional optical communication components and nanoelectronic systems for data and information processing.

Current research on SPP waveguide configurations seems to be developing along two interrelated and yet distinctly different directions dealing with (i) SPP waveguides that support extremely confined SPP modes and (ii) relatively low-loss SPP waveguides that can be efficiently controlled with electrical signals. In the first case, one takes advantage of the unique possibility offered by SPP waveguides to confine the transmitted radiation (in the lateral cross section) beyond the diffraction limit, thus allowing the realization of ultra-compact plasmonic circuitry [10]. Moreover, these SPP waveguides can be very efficiently coupled with individual emitters, opening thereby exciting perspectives for quantum plasmonics [178]. At the same time, their high propagation loss makes it difficult to exploit these configurations for active control, since material effects are typically very weak requiring long interaction lengths. On the other hand, SPP waveguides also offer another unique possibility—seamless interfacing of photonics and electronics with the same metal circuitry being used to both support the SPP propagation and control SPP guiding with electrical signals. Here one takes advantage of control electrodes being placed at the SPP field maximum, a feature that allows one to minimize the required electrical power and thereby decrease heat dissipation. The second direction is concerned with the exploitation of this unique feature using diffraction-limited SPP waveguides, whose modes are relatively weakly confined, thus being amenable to fiber coupling, e.g., via access silicon waveguides [140] and exhibiting relatively long propagation lengths.
Recent investigations have demonstrated that active plasmonic components based on dielectric-loaded SPP waveguides can be used in real data traffic environments for wavelength division multiplexed switching applications with an ultra-low power-response time product [188].

Realization of intrinsic difficulties in identifying the overall best SPP waveguide configuration alleviates the burden of improving the performance of SPP waveguides and waveguide components while making further research better defined and more focused. At the same time, this understanding suggests that one might try to combine both types of SPP waveguides in a plasmonic circuit containing both single-photon emitters/detectors and dynamic (electrically controlled) switches and routers. Important future practical applications of SPP-based integrated circuits in optical signal processing, sensing and imaging, including the realization of their unique features and benefits, are strongly dependent on further theoretical advances in this area. However, these future applications are even more dependent on the successful development of experimental methods and techniques for the reliable fabrication of optimally designed structures, the efficient generation and detection of SPP waveguide modes. Plasmon-based photonics is coming of age, but the practical integration of all plasmonic components designed for efficient plasmon generation, manipulation and detection remains being a significant and formidable challenge to be addressed in future research.

Acknowledgment

The authors acknowledge financial support from the Danish Council for Independent Research (FTP project No. 09-072949 ANAP).

References

[16] Raether H 1988 Surface Plasmons on Smooth and Rough Surfaces and on Gratings (Berlin: Springer)
[29] Lezec H J, Dionne J A and Atwater H A 2007 Negative refraction at visible frequencies Science 316 430–2


[38] Volkov V S, Bozhevolnyi S I, Rodrigo S G, Martín-Moreno L, García-Vidal F J, Devaux E and Ebbesen T W 2009 Nanofocusing with channel plasmon polaritons Nano Lett. 9 1278–82


[51] Chen J, Li Z, Yue S and Gong Q 2010 Efficient unidirectional generation of surface plasmon polaritons with asymmetric single-nanoslit Appl. Phys. Lett. 97 041113


[54] Li X, Tan Q, Bai B and Jin G 2011 Experimental demonstration of tunable directional excitation of surface plasmon polaritons with a subwavelength metallic double slit Appl. Phys. Lett. 98 251109


[56] Shao D B and Chen S C 2005 Surface-plasmon-assisted nanoscale photolithography by polarized light Appl. Phys. Lett. 86 253107


[59] Liu Z, Steele J M, Lee H and Zhang X 2006 Tuning the focus of a plasmonic lens by the incident angle Appl. Phys. Lett. 88 171108


[90] Nikolajsen T, Leossson K, Salakhutdinov I and Bozhevolnyi S I 2003 Polymer-based surface-plasmon-polariton stripe waveguides at telecommunication wavelengths Appl. Phys. Lett. 82 668–70


[121] Lui W W, Hirono T, Yokoyama K and Huang W 1998 Full-vectorial wave propagation in semiconductor optical bending waveguides and equivalent straight waveguide approximations J. Light. Technol. 16 910–4


[127] Han Z and Bozhevolnyi S I 2009 Ultracompact plasmonic racetrack resonators in integrated plasmonic devices Appl. Phys. Lett. 94 051111


[142] Han Z 2010 Efficient coupling between dielectric and plasmonic slot waveguide and aperture-assisted coupler based on sub-diffraction modulated ultracompact Fabry–Perot resonators in integrated plasmonic devices Opt. Express 17 12678–84


[147] Han Z and He S 2007 Two-dimensional model for three-dimensional index-guided multimode plasmonic waveguides and the design of ultrasmall multimode interference splitters Appl. Opt. 46 6225–33


[171] Leon I D and Berini P 2010 Amplification of long-range surface plasmons by a dipolar gain medium Nature Photon. 4 382–7

[172] Ambati M, Nam S H, Ulm-Avila E, Genov D A, Bartal G and Zhang X 2008 Observation of stimulated emission of surface plasmon polaritons Nano Lett. 8 3998–4001


[177] Jacob Z and Shalaev V M 2011 Plasmonics goes quantum Science 334 463–4


