Are slot and sub-wavelength grating waveguides better than strip waveguides for sensing?

Derek M. Kita, Jérôme Michon, Steven G. Johnson, and Juejun Hu

1. INTRODUCTION

Waveguide-based chemical and biological sensing is rapidly advancing as a prime application area for integrated photonics. However, there is currently no universally agreed-upon waveguide geometry that optimally enhances the detected signals in the presence of fabrication-induced scattering losses. In classical strip geometry that optimally enhances the detected signals in the core. To boost light relatively weak evanescent electric field outside the waveguide core. To boost light–molecule interactions, slot waveguides and sub-wavelength grating (SWG) waveguides have been proposed and demonstrated as alternative sensing platforms. In these waveguides, a larger fraction of the mode resides in the low-index cladding (usually the sensing medium, such as air or water) where the molecules are located. Indeed, improved refractive-index sensitivity (defined as the induced wavelength detuning per unit refractive-index change in the surrounding media) has been validated experimentally in resonator refractometry sensors based on both slot and SWG waveguides, and enhanced Raman conversion efficiency per unit device length has also been measured in slot waveguides. Nevertheless, the benefits may be offset by their increased susceptibility to optical loss induced by sidewall-roughness scattering. Many authors have studied roughness loss in dielectric waveguides both theoretically (through coupled-mode theory and the volume-current/Green's function methods) and experimentally, but these analyses focused only on minimizing the power scattered into the far field. This trade-off between mode confinement and attenuation is typically reported for plasmonic waveguide structures and some resonant sensor devices, but differences in fabrication conditions often make it difficult to perform side-by-side comparisons. Efficiently quantifying the trade-off between mode delocalization and scattering loss is critically important for the rational design of chip-scale photonic sensors (as well as electro-optic modulators and light sources), and, to date, there seems to be no work that quantifies these two effects with a single figure of merit (FOM) for strip, slot, and SWG waveguides.

In this paper, we develop easily computed FOMs (Section 2) that capture the precise trade-off between field confinement and roughness-induced scattering loss for waveguide-based sensing applications. The key to efficiently evaluating our FOMs is perturbation theory, both to evaluate the impact of mode roughness-induced scattering loss for waveguide-based sensing.
confinement on sensing (see Section 2 and Supplement 1) and to evaluate roughness scattering by computing the power radiated by equivalent current sources along the surface \cite{14, 32} (Section 3). The latter approach circumvents costly and roughness-dependent direct simulation of disordered waveguides, giving us a FOM comparison independent of the precise roughness statistics as long as the roughness correlation length is subwavelength. Although the computation of the “effective” sources to represent roughness scattering is in general quite complicated \cite{14}, we both propose a powerful general approach and demonstrate that, for typical experimental roughness, a simple “locally flat perturbation” approximation is accurate (Section 3 and Supplement 1). The results of this work indicate that a larger modal overlap with the sensing medium does not always correspond to better performance (due to increased scattering losses). By evaluating our FOMs for strip, slot, and SWG designs for bulk/surface/Raman sensing (Section 4), we obtain the results that for typical silicon-on-insulator (SOI) waveguides at a wavelength of $\lambda = 1550$ nm, (1) the simple TM-polarized strip waveguide is $>3 \times$ better than the other geometries considered here for bulk absorption sensing and refractometry, (2) the TM-polarized strip waveguide and TE-polarized slot waveguide are both $>3 \times$ better than the other geometries for surface-sensitive refractometry and absorption sensing, and (3) the TE-polarized slot waveguide is $>5 \times$ better than the other geometries for Raman sensing. Furthermore, we provide an extensive comparison (Section 5) with published experimental results, in which we applied our numerical method to compute the ratio of losses for pairs of waveguide geometries that were fabricated and reported in the literature. Our predictions exhibit good agreement within experimental accuracy. As discussed in our concluding remarks (Section 6), we believe that these comparisons
2. SENSING FIGURES OF MERIT

Here we consider three sensing techniques: (1) surface-sensitive refractometry and absorption sensing, where a sensor surface is coated with a binding agent (e.g., an antibody) that specifically attaches to a target molecule and the device monitors the change in index or optical absorption caused by monolayer or few-layer molecular binding on the surface; (2) bulk index/absorption sensing, where the sensor detects the change in bulk index/absorption in the adjacent sensing medium; and (3) waveguide-enhanced Raman spectroscopy [33–35], where the excitation light and the spontaneous Raman emission signal from molecules in the surroundings co-propagate in a waveguide. In all cases, the sensor performance is related to the external confinement factor and is divided by the attenuation coefficient. This “benefit-to-cost” ratio is described in Ref. [28] for absorption spectroscopy in silicon and silicon nitride waveguides, and in Ref. [19] for plasmonic waveguides, where there is a similar drive to maximize waveguide confinement and minimize the strong attenuation coefficients of surface plasmon polaritons [20]. Equivalently, for optical resonator refractometers, the appropriate FOMs is often stated as the ratio of refractive-index sensitivity to the resonance full width at half-maximum (FWHM) (which is proportional to the attenuation coefficient) [21–27]. For most prior literature that analyzes sensing in non-resonant waveguide devices (and also some resonant devices [2,36,37]), the quantity of interest typically reported is the fractional power change for a small increase in analyte (i.e., the numerator of our FOM) [5,38–43]. This is likely due to the difficulty in precisely measuring the attenuation coefficient or scattering loss for a single evanescent waveguide sensor or a Mach–Zehnder refractometer [15,44] (in contrast, for resonant devices, this information is immediately available in the resonance FWHM) [45].

The general goal in the first two sensing modes is to maximize the device sensitivity, which is the change in fractional optical power (ΔP/Pinput) induced by a small change in the number of analyte molecules that alter the cladding absorption coefficient or the index of refraction. Thus, the relevant metric is the sensitivity in units of inverse number density. For example, as described in Supplement 1, the absorption sensitivity is proportional to Γclad/αs, where Γclad is either the external confinement factor of the entire cladding region for bulk sensing or the surface region (approximated by a thin volume next to the interface), αs is the scattering loss per unit length, and αs is the molecular absorption coefficient (related to the bulk absorption coefficient via αs = mαo, where m is the number density of analyte molecules). In general, we find that the only geometry dependence occurs in a factor Γclad/αs, leading us to the dimensionless FOM for surface-sensitive and bulk absorption/index sensing:

\[ \text{FOM}_F = \frac{\Gamma_{\text{clad}}}{\alpha_s \lambda}, \]

where λ is the vacuum wavelength of light. In the third sensing mode (Raman sensing), the goal is to maximize the power of the Raman scattered light collected at the output of a waveguide for a given input laser power. As such, the corresponding FOM is the dimensionless quantity derived in Supplement 1:

\[ \text{FOM}_\beta = \frac{\beta}{\alpha_s}, \]

where β is the Raman gain coefficient (with units of 1/length), defined as the power of the evanescently scattered Raman light collected back into the waveguide per unit length and normalized by the input power. The external confinement factor [46–53] as well as the Raman gain coefficient are quantified via perturbation theory (see Supplement 1):

\[ \Gamma_{\text{clad}} = \frac{d n_{\text{eff}}}{d n_{\text{clad}}} = \frac{n_g}{n_{\text{clad}}} \left( \int_{\text{clad}} \frac{|\tilde{E}|^2}{\int |\tilde{E}|^2} d^2 x \right), \]

\[ \beta = \frac{\alpha_{\text{ram}}^2 \omega_m^2 m}{4c^2} \left( \frac{n_g}{n_{\text{clad}}} \left( \int \frac{1}{\epsilon(x)} |\tilde{E}(x, \omega)|^2 d^2 x \right)^2 \right), \]

where \( n_{\text{clad}} \), \( n_{\text{eff}} \), and \( n_g \) are, respectively, the cladding material, effective, and waveguide group indices; \( \epsilon \) is the permittivity; \( E \) is the electric field of the waveguide mode; \( \omega \) is the input angular frequency; the integration is over the cross section of the waveguide, restricted in the numerator to the cladding/sensing region; and \( \epsilon \) is the speed of light. We also assume that the Raman shift is relatively small so that \( |\Delta \omega| \ll \omega \). For periodic structures such as SWG waveguides, integration is performed over the volume of a unit cell with period \( \Lambda \):

\[ \beta_{\text{SWG}} = \frac{\alpha_{\text{ram}}^2 \omega_m^2 m}{4c^2} \left( \frac{n_g}{n_{\text{clad}}} \left( \int \frac{1}{\epsilon(x)} |\tilde{E}(x, \omega)|^2 d^3 x \right)^2 \right). \]

We note here that the optimized length of the waveguide for FOM\(_F\) is \( z = 1/(\alpha_s + \Gamma_{\text{clad}}\alpha_{\text{clad}}) \sim 1/\alpha_s \) and for FOM\(_\beta\) is \( z = 1/\alpha_s \). In this work, we assume that sidewall roughness scattering constitutes the dominant source of optical loss in waveguides, which is typically the case in high-index-contrast waveguide systems at moderate optical powers, such as in SOI [54–56], silicon nitride on insulator [57,58], chalcogenide glass on oxide [59], and TiO\(_2\) on oxide [60]. We also note that each FOM is generic to many different sensing device configurations, such as serpentine/spiral waveguides, ring resonators, and Mach–Zehnder interferometers.

In all three cases, the sensor limit of detection is determined by the modal overlap with the sensing region as well as the optical path length, the latter of which is limited by the waveguide propagation loss. The surface or bulk modal confinement factors and the Raman gain coefficient are readily computed from standard frequency-domain eigenmode solvers; the results are plotted in Fig. 1, which clearly show that slot and SWG structures indeed significantly enhance the modal overlap with the sensing medium.

3. SCATTERING-LOSS CALCULATIONS

In the volume-current method, waveguide perturbations can be described to first order (neglecting multiple-scattering effects) by dipole moments (polarizations) induced in the perturbation by the original (unperturbed) waveguide field \( \tilde{E}_0 \). These perturbations act as current sources that create the scattered field. For example, a small perturbation \( \Delta \epsilon \) in the permittivity acts like a current source \( \tilde{J} = \Delta \epsilon \tilde{E}_0 \) [11,61]. However, for perturbations in high-index-contrast waveguide interfaces, calculating the induced polarization (and, hence, the effective current source) is in general much more complicated [14] and requires solving a quasi-statics
problem (Poisson’s equation) [62]. For a given statistical distribution of the surface roughness, we show in Supplement 1 how we can solve a set of quasi-statics problems to compute the corresponding statistical distribution of the polarization currents. Fortunately, however, there is a simplification that applies to typical experimental regimes ($L_c > 10 \sigma$, where $L_c$ is the correlation length of the surface roughness and $\sigma$ is the root mean square (RMS) roughness amplitude). If $L_c$ (typically, 50–100 nm for SOI waveguides) is much larger than the amplitude of the roughness (0.5–2 nm in state-of-the-art devices [18,63–65]), then one can approximate the surface perturbation as locally flat. In this case, there is an analytical formula for the induced current from a locally flat interface shifted by a distance $\Delta h$ [62]:

$$\vec{J} = -\text{i}\omega \Delta h (\delta E_h - c^{\dagger} \Delta (\epsilon^{-1}) D_\perp) \delta(\vec{z}).$$  
(6)

(Note that the currents depend on the orientation of $\vec{E}_0$ relative to the interface: $E_\parallel$ is the component of $E_0$ parallel to the surface, whereas $D_\perp$ is the component of $\epsilon E_0$ perpendicular to the surface.) In this case, the correlation function of the current $J$ is simply proportional to the correlation function of the surface profile. We verify in Supplement 1 that the full quasi-static calculation reproduces this locally flat interface approximation in typical experimental regimes.

Given these currents and their statistical distribution (from the statistics of $\vec{h}$), one can then perform a set of Maxwell simulations on the unperturbed waveguide geometry (hence, at moderate spatial resolution) to find the corresponding radiated power. Again, there is a somewhat complicated procedure in general to compute the effect of currents with an arbitrary correlation function. But, in the case where the correlation length ($\sim 100$ nm) is much less than the wavelength ($\sim 1–2 \mu$m), we can approximate this “colored” noise distribution by uncorrelated “white” noise [32], effectively treating the currents as uncorrelated point sources with the same mean-squared amplitude ($\langle J^2 \rangle$). Since the power radiated by uncorrelated point sources is additive, we can then simply average the power radiated from different points on the surface to find the scattering loss per unit length $\alpha_s$ [14]. Finally, the computation simplifies even further because we only compute the ratio of the FOMs for different waveguides, in which case all overall scaling factors and units (including all dependence on the roughness statistics $L_c$ and $\sigma$) cancel out.

4. COMPARISON OF WAVEGUIDE GEOMETRIES

Using the aforementioned volume-current method, we quantified the relative scattering losses ($\alpha_s$) for a large set of strip, slot, and SWG waveguide geometries for various design parameters such as width, slot size, and duty cycle. All waveguide geometries (Fig. 1) are assumed to be fabricated from SOI wafers with Si layer thickness of 220 nm, and negligible scattering losses at the top and bottom flat surfaces, and for a target wavelength of 1550 nm.

The calculations, described in detail in Supplement 1, consist of only two computationally inexpensive steps. First, we compute the electric ($\vec{E}$) and displacement ($\vec{D}$) field profiles using a frequency-domain eigenmode solver [31], which is used to determine the confinement factors, Raman gain coefficient, and the strength of the induced dipole moments via Eq. (6). For each unique sidewall position and for each waveguide geometry, we use a single three-dimensional finite-difference time-domain (FDTD) simulation with a vertical line of dipole moments (corresponding to vertically symmetric sidewall roughness, i.e., line-edge roughness [66–68]) to compute the scattered power from both far-field radiation and reflection. Finally, the scattered power is averaged over all inequivalent sidewall positions and computed per unit length along the direction of propagation. The relative scattering losses per unit length for the considered geometries are shown in Fig. 2.

For strip waveguides, we analyzed the TE-like and TM-like polarized modes for various waveguide widths. The results for slot waveguides of TE and TM polarizations are computed for various total waveguide widths and air-slot gaps. For SWG waveguides, the width and period are fixed at 550 and 250 nm, respectively, and only TE ($x$-antisymmetric) modes are considered as functions of duty cycle, which is defined by $\text{DC} = (\Lambda - s)/\Lambda$, where $\Lambda$ is the grating period and $s$ is the size of the air gap along $z$ direction. The particular width of 550 nm and polarization was chosen for this analysis since the corresponding waveguide geometries exhibit single-mode behavior for a relatively large range of duty cycles, as confirmed by numerically computed dispersion diagrams. All waveguide dimensions are chosen such that no more than one TE and one TM mode exists.

Our results indicate that slot and SWG waveguides do in fact greatly increase scattering losses relative to standard 450 nm wide TE and TM strip waveguides [69]. The scattering loss computed

![Fig. 2.](image)

Fig. 2. Relative scattering loss, $\alpha_s/\max(\alpha_s)_{\text{strip}}$, computed by FDTD simulations of vertical ($y$ direction) lines of dipole moments and averaged over all sidewall positions. The dipole moment amplitudes were computed via Eq. (6) with incident field strength and phase determined by numerically computed mode profiles. The results for the TE (red) and TM (blue) strip (left), slot (middle), and SWG (right) waveguides are shown. For the slot waveguides, the total width is denoted as follows: $\Delta = 450$ nm, $\square = 550$ nm, and $x = 650$ nm.
by the volume-current method considers both the incident field strength at the location of the perturbation and the local density of states [70], which is determined by the surrounding geometry. (As such, current sources at the outer edges of a strip waveguide will radiate different amounts of power than current sources in the air-gap region of a slot waveguide.) Our comparison of scattering losses for different geometries, as shown in Fig. 2, is largely consistent with prior experimentally measured values of surface roughness in SOI waveguides [71–73]. However, a direct comparison between the experimental and theoretically/numerically computed values is difficult since experimental uncertainties are typically of the order of a few dB/cm. In addition, an accurate comparison of scattering losses for different geometries requires side-by-side fabrication of waveguides on the same process and material platform, since different processes introduce dramatically different roughness statistics. As such, literature for this is relatively scarce, and numerical methods provide a means for quickly evaluating and comparing new waveguide structures.

With the computed scattering losses, confinement factors, and Raman gain coefficients, we then computed the relevant FOM for each mode of sensing [via Eqs. (1) and (2)], which is presented in Fig. 3. Our results indicate that the TE slot and SWG structures do in fact provide modest improvements in sensing performance over traditional TE-polarized strip waveguides. However, the TM-polarized strip waveguides (which are here assumed to have negligible roughness on the top and bottom interfaces) exhibit significantly higher performance owing to their reduced propagation loss and longer accessible optical path length. For bulk and surface absorption sensing, the performance of the SWG waveguides gradually decreases with increase in duty cycle (as the air-slot region becomes smaller), indicating that the increase in scattering loss in small SWG gap regions outweighs the benefits provided by field localization in the air gap. On the other hand, slot-waveguide structures demonstrate improved performance with decrease in slot size (narrow-slot waveguides show 2x improvement over large-slot waveguides for surface sensing and 5x improvement for Raman sensing), with the exception of bulk absorption sensing in the air region. For bulk absorption sensing, there appears to be a critical slot size (70 nm slot for the 550 nm wide waveguides, and 130 nm slot for 650 nm wide waveguides) below which scattering losses dominate the FOM and above which the confinement factor is suboptimal.

For Raman spectroscopy, the gain coefficient is related to the fourth power of the electric field rather than the square of the electric field, and so regions of high electric field exhibit significant performance enhancements, as shown by the numerically computed values of the relative Raman gain coefficient, $\beta'$, in Fig. 1. In our computations, we find that waveguides with narrow slots do in fact tend to produce sufficiently more Raman signal than what is lost due to sidewall scattering at the silicon–air interface. Wide (550 nm) silicon waveguides with narrow (40 nm) air slots yield the highest bulk Raman FOM, by a factor of 8x over SWG waveguides and 5x over TM strip waveguide modes. Despite having an improved Raman gain coefficient, the SWG structures suffer from significantly higher scattering losses due to the increased sidewall surface area.

5. COMPARISON WITH EXPERIMENT

In order to demonstrate the utility and accuracy of this approach, we searched the literature for reports of fabricated strip, slot, and SWG silicon waveguides and their associated propagation losses.

![Fig. 3. Side-by-side performance comparison for strip (left column), slot (middle column), and SWG (right column) waveguides in the presence of slow-varying ($L_c > 10 \cdot \sigma$) Gaussian random roughness. The normalized absorption sensing figure of merit, $FOM_{\Gamma}$, and the normalized Raman gain coefficient, $FOM_{\beta}$, are calculated via Eqs. (1) and (2), respectively, as functions of relevant design parameters (total width, slot size, and duty cycle). For the slot waveguides, the total width (as depicted in Fig. 1) is denoted as follows: $\triangle = 450$ nm, $\square = 550$ nm, and $\times = 650$ nm.](image-url)
We then applied the volume-current method to compute the relative scattering losses of each of 14 reported waveguide geometries. Since the waveguides of different geometries reported within each manuscript are fabricated using identical processing protocols, which presumably result in identical or similar roughness characteristics, the reported losses are expected to accurately correspond to the different waveguide geometries. The authors in Bock et al. [76] only report the loss for a single SWG structure and reference Gnan et al. [77] to compare with strip waveguide losses. Both reported devices were fabricated using similar fabrication protocols (electron-beam lithography with hydrogen silsesquioxane resist and reactive-ion etching), and so we show in Fig. 4 the ratio of the reported propagation losses and our own numerically computed values using the reported waveguide geometries.

It is worth drawing attention to several features of Fig. 4. First and foremost, the experimental error bars associated with the measured values of propagation loss are quite large, of the order of several dB/cm. Accurately comparing the propagation losses of different waveguides requires the fabrication of all considered geometries on one single substrate and in close proximity to each other (to avoid cross-wafer variations). The waveguide losses can be characterized using either the cut-back method (many waveguides of different lengths) or resonator devices [78]. The cut-back approach is susceptible to large variability in device-to-device coupling efficiency, whereas loss measurement using resonators only computes the loss for a small waveguide length [79]. For these reasons, it is critical to measure a statistically significant number of devices. Often, information on sample size and measurement errors is omitted in the literature.

Ding et al. [75] reported the striking result that two waveguide geometries (denoted with “slot1” and “slot2” in their work) have lower losses despite a larger modal overlap with the vertical sidewall interface. As a result of the higher modal overlap, our model predicts that “slot1” and “slot2” should actually perform worse. This inconsistency may originate from other experimental sources of loss (mask defects, scattering at the top surfaces of the partially etched silicon, absorption, etc.) that were modified by the change in waveguide geometry. Otherwise, the values reported in the literature are consistent with the values predicted by our volume-current method, which validates our approach as a fast, powerful technique for precisely comparing the relative performance of arbitrary waveguide systems before investing time and resources toward fabricating real devices.

6. CONCLUDING REMARKS

In this work, we numerically computed the relative performance enhancement of slot and SWG waveguides over traditional strip waveguides for on-chip sensing applications. By explicitly computing the polarization statistics of randomly generated surface-roughness profiles (Supplement 1), we confirmed that a simple flat shifted-boundary approximation is accurate for most commonly encountered surface roughness (correlation length greater than roughly 10 times the RMS roughness amplitude). As a result, analytical formulas are available for the induced dipole moments from sidewall roughness. For situations where the roughness correlation length is significantly shorter than the wavelength of light, the volume-current method allows the far-field and reflected radiation losses from this roughness to be determined with good accuracy using inexpensive numerical methods. In particular, our approach benefits from significantly reduced simulation resolution requirements compared with brute-force simulation of waveguide structures with small perturbations, as the critical resolvable dimension is the waveguide geometry rather than the perturbation amplitude at the waveguide interface [80]. In addition, these brute-force techniques require many (or long) simulations to obtain statistically averaged effects of the randomly rough surfaces.

Our approach can be readily extended to efficiently determine optimal waveguide geometries for other sensing modes, such as stimulated Raman spectroscopy [81], with a suitably defined sensing metric. There are also a number of other material platforms to...
which this work can be extended, such as silicon nitride on silicon dioxide [33,34,82,83], titanium dioxide on silicon dioxide [35], chalcogenide glass on insulator [84,85], germanium or germanium–silicon on silicon [86,87], silicon on sapphire [88], and more, that are of great interest to the waveguide-integrated chemical sensing community. Because the waveguide geometries, materials, and FOMs can change according to sensing processes and wavelengths, conclusions about which geometries are better may also change. Lastly, there is interest in extending this work to quantify the sensing performance of additional waveguide geometries, such as photonic crystal waveguides [89] and horizontal slot waveguides [90], to name just a few.

We believe this work and the methods presented will aid in the rational design of new waveguide geometries for photonic sensing applications. Without FOMs and efficient methods for computing the loss and sensing trade-offs, it was difficult to predict whether waveguide geometries like slot and SWG waveguides will enhance the sensing performance (due to increased field overlap with the sensing medium) or decrease the sensing performance (from increased scattering losses). With the techniques presented, it is possible to quantify the precise trade-off between these two competing factors for arbitrary waveguide geometries and material platforms. This will drive future research in the area of on-chip sensing and in other areas where propagation loss in waveguides plays an important role.

**Funding.** National Science Foundation (NSF) (1709212); Army Research Office (ARO) (W911NF-13-D-0001).

**Acknowledgment.** We would like to thank Alexander McCauley for originating the polarization-statistics approach for roughness with short correlation lengths.

See Supplement 1 for supporting content.

**REFERENCES**


