Ultra-compact plasmonic waveguide modulators

Babicheva, Viktoria

Publication date:
2013

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):
Abstract

Metal-dielectric interfaces can support the waves known as surface plasmon polaritons, which are tightly coupled to the interface and allow manipulation of light at the nanoscale. Plasmonics as a subject which studies such waves enables the merge between two major technologies: nanometer-scale electronics and ultra-fast photonics. Plasmonic technologies can lead to a new generation of fast, on-chip, nanoscale devices with unique capabilities. In particular, it could offer a higher bandwidth and reduced power consumption.

Recently, there have been efforts towards addressing the challenge of developing new material platforms for integrated plasmonic devices. Furthermore, novel plasmonic materials such as transparent conductive oxides and transition metal nitrides can offer a variety of new opportunities. In particular, they offer adjustable/tailorable and nonlinear optical properties, dynamic switching and modulation capabilities, low cost, and stability being mean time fully CMOS-compatible. The plasmonic devices utilizing new materials could easily be integrated with existing nanophotonic and nanoelectronic devices into complex device geometries, bringing new levels of integration and functionalities. Similar to the advances in silicon technologies that led to the information revolution worldwide, the development of new plasmonic devices could revolutionize the field of hybrid photonic/electronic devices.

To manipulate light in hybrid photonic/electronic circuits based on CMOS-compatible materials, both passive and active plasmonic waveguide components are important. Among other proposed plasmonic waveguides and modulators, the structures where the dielectric core is sandwiched between metal plates have been shown as one of the most compact and efficient layout. Because of the tight mode confinement that can be achieved in metal-insulator-metal structures, they provide a base for extremely fast and efficient ultra-compact plasmonic devices, including modulators, photodetectors, lasers and amplifiers.

The main result of this thesis is a systematic study of various designs of plasmonic modulators based on ultra-compact waveguides with different active cores. Plasmonic modulators with the active core such as indium phosphides or ferroelectrics sandwiched between metal plates have promising characteristics. Apart from the speed and
dimensions advantages, the metal plates can serve as electrodes for electrical pumping of the active material making it easier to integrate. Including an additional layer in the plasmonic waveguide, in particular an ultrathin transparent conductive oxide film, allows the control of the dispersive properties of the waveguide and thus the higher efficiency of the plasmonic modulator. The main focus of the thesis is how to increase the extinction ratio of plasmonic devices decreasing mean time their propagation losses. Detailed simulations of different configurations are performed and working characteristics are compared to identify the most effective regimes and layouts of the ultra-compact plasmonic modulators.
Resumé


At manipulere lys i hybride fotoniske/elektroniske kredsløb baserede på CMOS-kompatible materialer kræver både passive og aktive plasmoniske bølgeledere. Blandt øvrige foreslåede plasmoniske bølgeledere og modulatorer, har layoutet på strukturer, hvor den dielektriske kerne er anbragt mellem metalplader, vist sig at være mest effektivt og kompakt. Pga. det tætte mode confinement, der kan opnås i metal-isolator-metal-strukturer, udgør disse strukturer en platform for ekstremt hurtige og effektive ultrakompakte plasmoniske komponenter, inklusive modulatorer, fotodetektorer, lasere og forstærker.

Hovedresultatet i denne afhandling er et systematisk studie af forskellige plasmoniske modulatordesign baserede på ultra-kompakte bølgeledere med forskellige aktive kerner. Plasmoniske modulatorer med aktive kerner såsom indiumfosfid eller ferroelectrics anbragt
mellem metalplader har lovende egenskaber. Udover fordelene
relateret til hastighed og størrelse, kan metalpladerne anvendes som
elektroder til elektrisk pumpning af det aktive materiale, hvilket
simplificerer integration. Ved at inkludere et ekstra lag i den
plasmoniske bølgeleder, særligt en ultratynd transparant ledende oxid-
film, muliggøres en kontrol af de dispersive egenskaber af
bølgelederen og således en højere effektivitet af den plasmoniske
modulator. Hovedfokus af afhandlingen omhandler hvordan man øger
extinction forholdet og samtidigt sænker propagationstabene.
Detaljerede simulationer af forskellige konfigurationer er udført, og
deres karakteristiske egenskaber bliver sammenlignede med henblik
på at identificere de mest effektive regimer og layouts for de
ultrakompakte plasmoniske modulatorer.
Preface and acknowledgements

This thesis is based upon studies conducted at the Department of Photonics Engineering, Technical University of Denmark, Lyngby, Denmark. From the very beginning my work was dedicated to ultra-compact plasmonic modulators based on metal-insulator-metal waveguides. It was an individual PhD project closely related to my undergraduate study on multilayered metal-dielectric structures. While metal-insulator-metal structures provide efficient functionality, later I decided to analyze other designs which utilize various plasmonic layers. Proof-of-concept device fabrication was considered, but later I decided to dedicate all time to the analysis and comparison of different designs.

Many people have contributed to the work presented in this thesis. I would like to start with special thanks to my supervisors Prof. Andrei Lavrinenko and Prof. Aleksandra Boltasseva for their support and guidance of the project.

I am also grateful to Dr. Radu Malureanu, Prof. Kresten Yvind and Irina Kulkova for our collaboration and work on our joint publications. Our numerous discussions have been very educational and invaluable for my research.

I highly appreciate the fruitful collaboration with Dr. Sergei Zhukovsky, his encouragement, nearly infinite support and positive attitude.

I would like to thank Søren Schou Gregersen, who introduced me basic concepts of clean room fabrication processes and plasmonic waveguide characterizations at the beginning of my project.

I am thankful to other present and former members of Plasmonics & Metamaterials group: Dr. Andrei Andryieuski, Dr. Aliaksandra Ivinskaya, Claudia Gritt, Oleg Lysenko, Dr. Andrei Novitsky, Morten Siwertsen, and Maksim Zalkovskij for sharing enthusiasm on various topics of plasmonics and critical suggestions during group meetings.

I have been fortunate to collaborate with a team of Purdue University where I spent seven months working on the part of my project. I want to express my appreciation of fruitful discussions with Dr. Gururaj Naik, Nathaniel Kinsey, Dr. Marcello Ferrera, and Mikhail Shalaginov cornering alternative plasmonic materials and practical issues of plasmonic waveguides as well as thank to Prof. Vladimir Shalaev and...
Prof. Alexander Kildishev for creating a stimulating work environment.

A great number of people in DTU are thanked for their personal support during my work on the project. In particular, Nadezda Kuznetsova, Prof. Niels Gregersen, Dr. Alexey Savenko, Anton Dogadaev, Alexander Lebedev, and Oleksii Kopylov for sharing common interests outside the field of work.

All above mentioned people influenced on my work in academia. I also would like to thank Prof. Dmitry Chigrin, Prof. Alexander Uskov, Prof. Martijn Wubs, and Prof. Evgenii Narimanov for useful and important discussions of subjects related to career path and personal development.

The work was performed as individual PhD project on DTU Fotonik scholarship. I also had financial support from several foundations such as Otto Mønsteds, Oticon and Thomas B. Thriges foundations as well as SPIE society by means of Optics and Photonics Education Scholarship 2012 that considerably helped me in conference participation and external research stay.
List of publications

This thesis is partly based on the work presented in the following publications and talks.

Journal Publications


Conference Proceedings and Other Contributions


During my PhD study, I also worked on some topics on localized and delocalized surface plasmon resonances that resulted in the following publications.


# List of acronyms

1D, 2D, 3D  one-, two-, three-dimensional  
AZO  Aluminum-doped Zinc-Oxide  
BaTiO$_3$  Barium Titanate  
CMOS  complementary metal–oxide–semiconductor  
ER  extinction ratio  
F-MSM  finite-thickness metal-semiconductor-metal  
FoM  figure of merit  
GZO  Gallium-doped Zinc Oxide  
IMI  insulator-metal-insulator  
InP  Indium Phosphide  
ITO  Tin-doped Indium Oxide  
LR-SPP  long range surface plasmon polariton  
MIM  metal-insulator-metal  
MSM  metal-semiconductor-metal  
PF  Purcell factor  
QD  quantum dot  
QW  quantum well  
RC  resistor–capacitor  
Si  Silicon  
SiN  Silicon Nitride  
SPP  surface plasmon polariton  
TiN  Titanium Nitride  
TCO  transparent conducting oxide  
VCSEL  vertical cavity surface emitting laser  
wg  waveguide
Contents

Abstract iii
Resumé v
Preface and acknowledgements vii
List of publications ix
List of acronyms xiii
Contents xv
Chapter 1. Introduction 1
  1.1. Plasmonic modulators: State of the art ......................... 1
  1.2. Absorption modulators: Metrics of performance .......... 7
  1.3. Plasmonic modulators based on TCO ......................... 7
  1.4. Towards CMOS-compatible nanophotonics: alternative
       plasmonic materials .................................................. 11
  1.5. Thesis outline ...................................................... 13
Chapter 2. Ultra-compact modulators using alternative plasmonic materials 15
  2.1. Multilayer structures .............................................. 15
  2.2. Performance of the modulators ................................. 21
  2.3. Waveguide and modulator integration ....................... 25
  2.4. Conclusion .......................................................... 29
Chapter 3. Plasmonic modulator optimized by patterning of active
            layer and tuning permittivity 31
  3.1. Multilayer structures .............................................. 31
  3.2. Effect of ITO permittivity changes ......................... 34
  3.3. Eigenmodes of four-layer system ............................... 36
  3.4. Periodic patterning of ITO layer ............................... 39
  3.5. Bragg grating ...................................................... 45
3.6. Discussions and conclusions ................................................... 47

Chapter 4. Plasmonic modulator based on gain-assisted metal-semiconductor-metal waveguide 51

4.1. Gain materials in plasmonic structures ........................................ 51
4.2. Simulation model ......................................................................... 53
4.3. MSM waveguide arrangements ................................................... 55
4.4. MSM waveguide with bulk gain medium ..................................... 57
4.5. MSM waveguide with quantum dots .......................................... 61
4.6. MSM waveguide with quantum wells ........................................ 65
4.7. Discussion and conclusion .......................................................... 68

Chapter 5. Plasmonic finite-thickness metal-semiconductor-metal waveguide as ultra-compact modulator 71

5.1. F-MSM waveguide modes .......................................................... 71
5.2. Field confinement in the F-MSM waveguide ............................... 74
5.3. Absorption coefficient and modulator’s performance ................. 77
5.4. Influence of n- and p-doped layers ............................................. 82
5.5. Finite-length F-MSM waveguide ............................................... 85
5.6. Finite-width F-MSM waveguide: 3D case ................................. 86
5.7. Other practical issues ................................................................. 88
5.8. Discussion and conclusion .......................................................... 89

Chapter 6. Plasmonic modulator using metal-insulator-metal waveguide with barium titanate core 93

Chapter 7. Summary ................................................................. 99

Bibliography .............................................................................. 101
Chapter 1. Introduction

1.1. Plasmonic modulators: State of the art

Plasmonics enables the merging between two major technologies: nanometer-scale electronics and ultra-fast photonics [1-21]. Metal-dielectric interfaces can support the waves known as surface plasmon polaritons (SPPs) that are tightly coupled to the interface, and allow manipulation of light at the nanoscale, overcoming the diffraction limit. Plasmonic technologies can lead to a new generation of fast, on-chip, nanoscale devices with unique capabilities (Fig. 1.1). To provide the basic nanophotonic circuitry functionalities, elementary plasmonic devices such as waveguides, modulators, sources, amplifiers, and photodetectors are required. Various designs of plasmonic waveguides have been proposed to achieve the highest mode localization and the lowest propagation losses [11,13,22-45].

Fig. 1.1 (a) Operating speed and device size: key advantages of nanoplasmonics for telecommunication [2]. (b) Logic and data switching for integrated circuits [13].
In addition to waveguides, modulators are the most fundamental component for digital signal encoding and are paramount to the development of nanophotonic circuits. Active plasmonics and plasmonics switching devices are considered as one of the most challenging directions in nanophotonics (Fig. 1.2) and aim to combine semiconductor electronics and nonlinear optics to control the optical properties of different nanodevices [46-48].

Fig. 1.2. (a) Cross functional active plasmonic concept: A transient change in the refractive index of either the metal or dielectric component of an SPP waveguide, induced by an external control excitation (optical, thermal, electronic, etc.), can switch or modulate the intensity and/or phase of a propagating SPP [48]. (b) General concept of modulator integrated with the waveguide and utilized in hybrid photonic/electronic circuits: electrical signal changes properties of sandwiched medium and thus control propagation signal [12].
The most promising mechanisms of modulation are summarized in the Table 1.1. Thermooptic response and solid state phase transitions can provide significant response, however they time frame is high and allow operation speed on the orders of MHz. Carrier concentration change in accumulation layer is faster and thus opto-electronic modulators can be designed to achieve ultra-fast operational speeds in the 10’s of GHz (Fig. 1.3). It has been shown that some designs outperform conventional silicon-based modulators [49,50].

Various types of compact modulators that utilize SPPs have been proposed [51-69]. Both phase and absorption modulations have been exploited to achieve high speed and low footprint of a modulator. The structural phase transition in vanadium dioxide VO$_2$, which exhibits strong contrast between the optical properties of its insulating and metallic phases, is analyzed for different plasmonic modes [58]. Depending on a mode, either an index modulator (with $\Delta n > 20\%$) or absorption modulator can be utilized.

Table 1.1. Switching mechanisms and their applicability.

<table>
<thead>
<tr>
<th>Switching mechanism</th>
<th>Active material</th>
<th>Issue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carrier concentration change</td>
<td>Si</td>
<td>Weak response</td>
</tr>
<tr>
<td></td>
<td>III-V</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Graphene</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TCO (ITO, GZO, AZO)</td>
<td>Lossy</td>
</tr>
<tr>
<td>Phase transition</td>
<td>VO$_2$, Ga, BiFeO$_3$, BaTiO$_3$, etc</td>
<td>Low-speed (~1 µs)</td>
</tr>
<tr>
<td>Thermo</td>
<td>Polymers</td>
<td></td>
</tr>
</tbody>
</table>

However, most of plasmonic modulators are based on control of gap plasmonic modes by the charge accumulation layer. Two different layouts are possible. In the first design, a metal-oxide-semiconductor (MOS) stack is deposited on top of a silicon waveguide [50,56,61,63-65]. Such waveguide-integrated plasmonic modulators are reasonably simply to fabricate, their size is comparable with the silicon waveguide size, and propagation losses are originated only from one metal interface. An electro-optic modulator based on Silicon-on-Insulator waveguide with an ITO-SiO$_2$-Au stack is reported in [50] having low insertion loss -1 dB and broadband operation due to the non-resonant MOS mode are shown.
Fig. 1.3 The schematic shows how a voltage (V gate) induced free carrier change (n carrier) leads to a shift of the active material’s relative permittivity ($\Delta \varepsilon_r(\omega)$, with $\omega$ being the angular frequency) and consequently to a shift in the real and imaginary part ($\Delta \kappa$) of the propagating modal index. With these two methods, either a phase or an absorption modulator (shown here) can be designed [13].

The second proposed concept includes layers of oxide and semiconductor which are embedded in metal from both sides. Metal-insulator-metal structures keep additional advantages and possess strong ability to confine light [18-24,41,42]. Two metal surfaces serve also as electrodes, thus simplifying design. In such type of waveguides light is localized in a gap with typical sizes ~100 nm and less, that facilitates manipulation of light on the subwavelength scale. The SPP propagation length in the metal-insulator-metal waveguide can be up to ten micrometers.

Horizontally arranged metal-insulator-semiconductor-insulator-metal slot waveguides exhibit high performance [67-69]. The main principle is based on inducing a highly accumulated electron layer at the SiO$_2$/Si interface. An electro-absorption CMOS-compatible modulator was characterized: 3-dB operation on 3 $\mu$m length at ~6.5 V bias and broadband modulation is achieved [67]. Such configurations can be readily integrated in standard Si circuits (Fig. 1.4). However, high mode localization requires a very high aspect ratio of the waveguide core.
Ultra-compact efficient plasmonic modulators based on strong light localization in vertically arranged metal-semiconductor-insulator-metal waveguides have been studied recently [53,56,57]. The phase modulator PlasMOStor consists of semiconductor core with Si and SiO$_2$ layers, sandwiched between two silver plates [53]. It supports both photonic and plasmonic modes, which interfere while propagating (Fig. 1.5). The modulation is based on cutting off the photonic mode and thereby changing the integral transmittance.

Fig. 1.4. Schematic of a proof-of-concept Si nanoplasmonic EA modulator: (a) top view and (b) cross-sectional view along A-A’ [67].
Fig. 1.5. (a) Schematic of an SOI-based all-optical plasmostor [20]. (b) Geometry and tabulated modes of the plasmostor. (Top) The plasmostor is composed of an Ag gate contacting a 10-nm-thick SiO2 layer on 173 nm of n-type Si; the ground is also Ag. Here, the optical source and drain consist of an Si waveguide with transverse dimensions equal to the plasmostor dimensions. (Bottom) Tabulated mode profiles, refractive indices and losses for the plasmostor in both depletion (voltage-OFF) and accumulation (voltage-ON) states at $\lambda = 1550$ nm [53].
1.2. Absorption modulators: Metrics of performance

Signal attenuation (extinction) per device’s unit length can be defined as

\[
A_{f/L} = 10 \log (P_0 / P) / L = 8.68 \text{Im}(k_{\text{eff}}),
\]

where \( k_{\text{eff}} \) is an effective propagation constant in the device structures.

Consequently, the logarithmic extinction ratio (ER) per unit length is

\[
\text{ER} = 10 \log (P_{\text{on}} / P_{\text{off}}) / L = 8.68 (\text{Im}(k_{\text{eff}})_{\text{on}} - \text{Im}(k_{\text{eff}})_{\text{off}}),
\]

where \( \text{Im}(k_{\text{eff}})_{\text{on}} \) and \( \text{Im}(k_{\text{eff}})_{\text{off}} \) are the imaginary parts of the effective propagation constant of a waveguide mode in the on-state (voltage or electrical current switched on) and off-state (voltage or current switched off). ER shows how strong one can vary the mode propagation through the waveguide.

However, for plasmonic switching devices that possess significant propagation losses, another figure of merit (FoM) can be defined (similar to [56]):

\[
\text{FoM} = \frac{\text{ER}}{A_{f/L}} = \frac{\left| \text{Im}(k_{\text{eff}})_{\text{on}} - \text{Im}(k_{\text{eff}})_{\text{off}} \right|}{\text{Im}(k_{\text{eff}})_{\text{state}}},
\]

where denominator is either \( \text{Im}(k_{\text{eff}})_{\text{on}} \) or \( \text{Im}(k_{\text{eff}})_{\text{off}} \) depending on which state is transmittive. FoM (1.3) describes how strong one can vary the mode propagation through the waveguide in comparison with the attenuation in transmittive state. The Eq. (1.3) differs only by a constant coefficient from the one introduced in [56], which is formulated regarding the propagation lengths. Such a definition of FoM gives length-independent characteristic of the structure. A proper length of the device can be chosen according to signal level requirements and fabrication restrictions for particular geometry of a modulation problem.

1.3. Plasmonic modulators based on TCO

New intermediate carrier density materials offer the prospect of additional exotic properties beyond tailorable optical properties, lower losses and integration advantages [70-79]. TCOs can provide
extraordinary tuning and modulation of their complex refractive indices, because their carrier concentrations can be changed over several orders of magnitude by applying an electric field [77]. Therefore, they are promising candidates for adding electro-optical capabilities to plasmonic devices [50,56]. In particular, a unity-order index change in a 5 nm thin Indium-Tin-Oxide (ITO) layer was demonstrated for a metal-insulator-metal (MIM) structure [77]. Tunability is accomplished by applying a bias, resulting in an electric field across the TCO layer. The resulting electric field causes a charge accumulation, or depletion, in the TCO layer (depending on the direction of electric field) which in turn changes the plasma frequency of the TCO, and consequently, its permittivity. The modulating speed is only RC limited and is expected to exceed 10’s of GHz.

Several layouts of Si photonic modulators using a TCO as a dynamic layer were proposed [50,61,66]. An extinction ratio of 1 dB/µm was demonstrated for a plasmonic modulator utilizing a metal-oxide-ITO stack on top of a silicon photonic waveguide (Fig. 1.6) [50]. Under an applied bias, the carrier concentration is changed from $1 \times 10^{19}$ cm$^{-3}$ to $6.8 \times 10^{20}$, and the propagation length is varied from 34 to 1.3 µm. However, because this structure uses a photonic mode, the miniaturization level of such a device is limited.

![Fig. 1.6. (a) Schematic of the waveguide-integrated, silicon-based nanophotonic modulator. The MOS design features a plasmonic optical mode which concentrates the mode’s electric field and allows for a good overlap with the active ITO layer. (b) Electric field density across the active fundamental MOS region of the modulator [50].](image)

Very high ER (up to 20 dB/µm) was achieved utilizing the epsilon-near-zero properties of AZO [66]. Because of the small absolute
value, a large portion of the field is localized within the layer and provides more efficient modulation.

ITO has also been implemented in an MIM waveguide structure to demonstrate a subwavelength plasmonic modulator (Fig. 1.7), for which a five percent change in the average carrier density (from $9.25 \times 10^{20}$ to $9.7 \times 10^{20} \text{ cm}^{-3}$) was studied [56]. The authors employ analytical solutions of the SPP dispersion equation in a four-layer structure; the Thomas-Fermi screening theory to derive the carrier density distribution as well as numerical simulations with a finite element method. The structure supports SPP resonance at telecommunication wavelength 1.55 $\mu$m owing to the ITO layer which has small absolute values of permittivity in near-infrared region. The resonance is broad because of high losses in ITO. It decreases device’s performance and increases bandwidth of operation at the same time. A similar structure based on a silicon-waveguide-integrated multilayer stack was fabricated and characterized. The logarithmic extinction ratio regarding to power up to 0.02 dB was achieved. Meanwhile a theoretical analysis of the Ag-ITO-Si$_3$N$_4$-Ag structure predicts the 1dB extinction ratio on 0.5 $\mu$m length. However, power transmitted through the 0.5-$\mu$m-long device is rather low; in particular losses are 24 dB/$\mu$m for the waveguide with Si$_3$N$_4$-core and 9 dB/$\mu$m with SiO$_2$-core. Thus, the proposed design can be made extremely compact. However, there is always a trade-off between modulation depth and transmittance through the modulating system and due to the high confinement achievable in the MIM structure and the high losses associated with both metal and ITO layers, the propagation length in this system is extremely limited.

Ultra-compact designs can be achieved in a modulator layout based on an MIM waveguide with a 5-nm gap [57]. Very small gap size leads to a very efficient change of the carrier concentration in a 2.5-nm-thick ITO layer (Fig. 1.8). However, the practical implementation of the proposed device is challenging.

The summary of performance characteristics of all proposed and published TCO-based plasmonic modulators are shown in Table 1.2. The numbers are extracted or recalculated from those given in the papers. One can see that while MIM-modulators posses high compactness performance characteristic are lower in comparison to silicon-waveguide based devices.
Fig. 1.7. The structure of a surface plasmon polariton absorption modulator (SPPAM). Light is coupled from a silicon nanowire into an active plasmonic section by means of a directional coupler. The active section consists of a stack of silver (Ag), indium tin oxide (ITO), and SiO2 layers. The absorption coefficient of the SPP is modulated by applying a voltage between the two silver electrodes. The insets show how a photonic mode (a) in a silicon strip waveguide excites a SPP (c) via a hybrid mode (b) in directional coupler. The insets in (d) show the electric field Ey and the magnetic field Hx as well as the time-averaged Poynting vector distributions in the active plasmonic part [56].

Fig. 1.8. Design of field-effect nanowire-based modulator. (a) SPP mode in a nanowire-MIM waveguide along with the vertical and horizontal cross sections. The nanowire size is 25 x 25 nm². (b) Electron density at Vg = 2 V along with a zoom of the active region structure [57].
Table 1.2. Plasmonic modulators based on TCO.

<table>
<thead>
<tr>
<th>Device</th>
<th>ER, dB/µm</th>
<th>$\alpha_{min}$, dB/µm</th>
<th>FoM</th>
<th>$\lambda$, nm</th>
<th>Based on</th>
<th>Date &amp; Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPPAM</td>
<td>2</td>
<td>24</td>
<td>0.08</td>
<td>1550</td>
<td>MIM</td>
<td>04/2011, [56]</td>
</tr>
<tr>
<td>$3\lambda$-size</td>
<td>1.0</td>
<td>0.04</td>
<td>25</td>
<td>1310</td>
<td>Si-wg</td>
<td>04/2012, [50]</td>
</tr>
<tr>
<td>Tunable ENZ</td>
<td>18</td>
<td>1</td>
<td>18</td>
<td>1310</td>
<td>Si-wg</td>
<td>06/2012, [66]</td>
</tr>
<tr>
<td>Tuned $\varepsilon$</td>
<td>3</td>
<td>9</td>
<td>0.3</td>
<td>1550</td>
<td>MIM</td>
<td>07/2012, Ch.3</td>
</tr>
<tr>
<td>wire-MIM</td>
<td>13</td>
<td>11</td>
<td>1.2</td>
<td>1550</td>
<td>MIM</td>
<td>09/2012, [57]</td>
</tr>
<tr>
<td>Sub-$\lambda$-size</td>
<td>6</td>
<td>0.7</td>
<td>8.5</td>
<td>1310</td>
<td>Si-wg</td>
<td>07/2013, [61]</td>
</tr>
</tbody>
</table>

### 1.4. Towards CMOS-compatible nanophotonics: alternative plasmonic materials

The promising development of chip-scale plasmonic devices with traditional noble metals is hindered by challenges such as high losses, continuous thin film growth, and non-tunable optical properties. Moreover, noble metals as plasmonic building blocks are not compatible with the established semiconductor manufacturing processes. This limits the ultimate applicability of such structures for future consumer devices. Recently, there have been efforts towards addressing the challenge of developing CMOS-compatible material platforms for integrated plasmonic devices [75,76,80]. Similar to the advances in silicon technologies that led to the information revolution worldwide, the development of new CMOS-compatible plasmonic materials with adjustable/tunable optical properties, could revolutionize the field of hybrid photonic/electronic devices. This technology would help to address the needs for faster, smaller and more efficient photonic systems, renewable energy, nanoscale fabrication, and biotechnologies. These new materials can bring exciting new functionalities that cannot be achieved with traditional metals.

While many materials have been suggested as replacements for the traditional plasmonic metals, titanium nitride (TiN) is one of the best candidates [75,80]. Moreover, TiN is very thermally stable, biocompatible, extremely hard (one of the hardest ceramics) and chemically stable; in particular, it does not oxidize like silver or copper. It was also shown that TiN provides higher mode confinement in comparison to gold [74]. This makes TiN a very promising material for telecommunication-range plasmonic waveguides.

One important advantage of TiN is that it can be grown epitaxially on many substrates including [100]-silicon, forming ultra-smooth and ultra-thin layers [80,81]. A final benefit of transition metal nitrides is...
that they are nonstoichiometric materials. Hence their optical properties depend greatly on the preparation conditions and can be varied based on the desired performance.

For a device to be fully CMOS-compatible, both the material and the processing technique used to synthesize this material should be compatible with the standard. Currently, TiN is routinely used in CMOS processing lines, but the optical properties of this material are quite poor [82-85]. This is because the primary consideration has been the electrical properties of the material, not the optical properties. However, the high temperature sputtering process (800 °C) is not utilized in the current semiconductor manufacturing processes for TiN deposition. Thus, it is acknowledged that the entire process is not currently CMOS-compatible, while devices are based on CMOS-compatible materials. However, through an optimization process of the low temperature TiN (less than 400°) currently available in the CMOS industry, TiN which possesses the required optical properties can be made available in future CMOS production lines. This is in stark contrast to the noble metals which are not allowed in the CMOS process. A similar situation was encountered for low-loss doped silica glass which is normally obtained through high-temperature annealing. Nevertheless, in 2003 a new material platform, namely Hydex®, was synthesized to bring this glass into full CMOS-compatibility where it was subsequently used for integrated nonlinear optics experiments [86,87].

Copper has also been investigated as a potential CMOS-compatible plasmonic material [67,85,88-91]. However, the use of copper first requires a TiN buffer layer to prevent its diffusion into silicon [92]. Thus, if the low-temperature TiN is optimized with competitive optical properties, the second deposition of copper is not necessary. TiN in its own right also has many advantages over copper such as chemical stability, high temperature stability, bio-compatibility and more, which are useful for many applications beyond only CMOS chips.

The TCOs may be deposited at relatively low temperatures (less than 300°C), which makes it possible to integrate them as a final stage in the standard silicon process [80]. Due to their low temperature deposition they will not impact the CMOS produced structures below. Similar nondestructive methods of integration with CMOS circuitry have been utilized to include lithium niobate crystals and electro-optic polymers on CMOS produced photonic chips [93,94].
1.5. Thesis outline

On my PhD study, I was focused on developing active plasmonic devices and suggested a variety of plasmonic modulator structures using transparent conducting oxides and indium phosphide based semiconductor materials. In the design of the structures, both the device performance and its fabrication complexity are taken into account. I numerically studied modulator geometries and compare their performance from different points of view. Modulation depth, propagation losses, mode size and their trade-off is analyzed. Compatibility with CMOS processing techniques to allow for easy integration with current nanoelectronic devices and the integration of the modulator geometries with plasmonic waveguides are investigated.

In Chapter 2, several planar layouts of ultra-compact plasmonic waveguide modulators are proposed and utilization of alternative plasmonic materials is studied. The modulation is efficiently achieved by tuning the carrier concentration in a transparent conducting oxide layer, which may serve as both the plasmonic material and as a dynamic element. Tuning the waveguide either in plasmonic resonance or off-resonance is analyzed. Resonance significantly increases the absorption coefficient of the plasmonic waveguide, which enables larger modulation depth. It is shown that an extinction ratio of 86 dB/µm can be achieved, allowing for a 3-dB modulation depth in less than one micron at the telecommunication wavelength.

In Chapter 3, the surface plasmon polariton absorption modulators with one and two ITO layers in MIM layout are considered. Further, the influence of the ITO permittivity under varying it on several units is discussed. Eigenmodes of the modulator structure are analyzed following by study of the periodic patterning in the system. Another concept based on a Bragg reflector grating is proposed for modulation as well.

In Chapter 4, plasmonic modulator based on metal-semiconductor-metal (MSM) structure is designed and its performance for plasmonic switching applications is studied. The semiconductor core is considered consisting of a bulk gain medium, quantum wells or layers with quantum dots. Theoretical model and numerical approach to MSM waveguide simulations as well as possible MSM waveguide arrangements are described and analyzed.
In Chapter 5, plasmonic waveguide with thin metal layers and semiconductor gain core is studied to be utilized as plasmonic modulator. Characteristics of all eigenmodes in a F-MSM waveguide are compared and the best one to be exploited in the plasmonic modulator is specified. Analytical calculations of a F-MSM relative effective index and field confinement factors are presented as well as a subsequent increasing of the absorption coefficient and possibility to control wave propagation. Influence of n- and p-doped layers, finite-length and finite-width waveguides are analyzed.

In Chapter 6, we study a design of plasmonic modulator with ferroelectric core. The active material, barium titanate (BaTiO₃), is sandwiched between metal plates and changes its refractive index under applied voltage. Some degree of switching of ferroelectric domains from the in-plane to out-of-plane orientation provides the change of the refractive index, which can be exploited for effective light modulation. By numerical analysis we prove that the π phase change can be achieved with a 12…15 μm length device having propagation losses 0.05…0.2 dB/μm.
Chapter 2. Ultra-compact modulators using alternative plasmonic materials

In this chapter, several layouts of compact plasmonic waveguide modulators with alternative plasmonic materials are proposed and the possibility of their integration in photonic circuits is discussed. Several planar multilayer designs are studied for structures where noble metal is replaced by titanium nitride. A thin layer of transparent conducting oxide (TCO) can serve as both a section of the plasmonic waveguide and as an active material used for modulation. Both low- and high-index cladding materials are considered. By tuning the carrier concentration in the TCO layer, and thereby adjusting the TCO permittivity to appropriate value, a surface plasmon resonance can be achieved in a multilayer stack. Applying a dynamic bias to the TCO layer allows for the permittivity to be tuned out of the resonance, resulting in a change of the absorption coefficient of the waveguide. Therefore, active control of the surface plasmon polaritons propagation is achieved. The usage of titanium nitride as a plasmonic material in modulator configurations allows achieving significantly increased modulation depth. It is shown that less than one micron of the active section is required to achieve 3 dB modulation. Such multilayer stacks can be easily integrated with existing plasmonic and conventional photonic waveguides as well as novel, semiconductor-based hybrid photonic/electronic circuits.

2.1. Multilayer structures

Example geometry of the proposed electro-optic plasmonic modulator integrated with stripe waveguides is shown in Fig. 2.1, along with the symmetric, long ranging SPP (LR-SPP) excitation mode, its direction of propagation through the structure, and the applied voltage for modulation. The waveguide sections are composed of a lower dielectric cladding (Fig. 2.1 shown in red) with a metal strip (grey) and a top dielectric cladding (red). This configuration allows for the LR-SPP mode to be utilized for low loss connections to and from the modulator. The modulator section is composed of a lower dielectric layer (Fig. 2.1 shown in red), a TCO layer used as both a plasmonic layer and an electrode (light blue), a dielectric layer to provide electrical insulation (light orange), and a top electrode (pink).
This proposed geometry provides several benefits when compared to other potential structures. Due to the plasmonic nature of the modulator, the device can achieve a very small footprint that cannot be realized in traditional photonic elements. The utilization of metal layers for electrical control provides additional benefits in term of compactness. CMOS-compatible materials allows for the device to be easily integrated into traditional manufacturing. The multilayer structure which forms the modulation, can achieve an extremely high absorption coefficient when this layer is modulated into resonance by an applied field. Therefore, a very small length is required to achieve 3-dB modulation of the signal. Due to these advantages, the proposed modulator configuration shows great promise for a CMOS compatible, on-chip electro-optic modulator.

Fig. 2.1. General scheme of an ultra-compact modulator integrated with plasmonic waveguides. In this geometry, a stripe waveguide (grey) is used to bring a long ranging SPP mode to the modulator structure. Inside the modulator structure, the waveguide is replaced by a TCO (light blue) which is plasmonic at the wavelength of interest. A modulating voltage is applied between the TCO layer and a top electrode, which are separated by a thin dielectric spacer. This voltage alters the carrier concentration in the TCO resulting in change in the absorption which modulates the SPP. A second strip waveguide is used to propagate the modulated signal to the next component.

Due to fabrication and integration advantages we consider modulators based on the strip waveguide geometry. Stripe waveguides have low loss but have poor mode localization [25,28]. However, the relatively simple planar fabrication process provides an advantage to utilize the structure in realistic devices. Here we consider only the one-
dimensional planar layout as an estimate of a stripe waveguide. In realistic structures, a finite-wide stripe waveguide will be used, and the propagation losses will depend on the geometrical parameters of the stripe. However, the performance is only marginally different from the one-dimensional structure, and the main dispersive features will be captured [95]. Therefore, for the purposes of this study, the estimates provided by assuming a one-dimensional structure are suitable for the comparison of the various modulator geometries suggested.

The main element of the structure is the TCO layer. Since the TCO can possess plasmonic properties at telecom range, a thin TCO layer can guide the plasmonic mode as well as control the signal propagation. TCOs such as ITO, Gallium Zinc Oxide (GZO), and AZO have very similar properties and allow for an efficient change of carrier concentration. We chose GZO as it has shown the ability to achieve the highest plasma frequency of the three [96]. The permittivity of the GZO layer was taken from [96] and a carrier concentration in the GZO was determined using a Drude-Lorentz model fitting: \( N_0 = 9.426 \times 10^{20} \text{cm}^{-3} \). TCO properties depend strongly on fabrication conditions such as the annealing environment and temperature [71,75]. Since a significant progress was shown on increase of carrier concentration [76,96-99], the optimization can be used to achieve better device performance. As the value \( 1.46 \times 10^{21} \text{cm}^{-3} \) for GZO was reported recently [99], here we assume that the film with carrier concentration up to \( 2N_0 \) \( (1.88 \times 10^{21} \text{cm}^{-3}) \) can be fabricated and depleted to lower value under applied voltage. Thus, we will study properties of the structures within the range \( N = 0.5N_0 \ldots 2N_0 \). The calculated permittivity of GZO for these carrier concentrations is shown in Fig. 2.2a.

Including a TiN layer increases the mode localization, which influences the modulator’s performance. Moreover, the TiN layer can also serve as a second electrode to apply bias to the GZO active layer. In this study, we use the optical constants of TiN films optimized for plasmonic applications. The permittivity of TiN is taken from [76]: \( \varepsilon_{\text{TiN}} = -83.3 + 21.3i \) at \( \lambda = 1.55 \mu\text{m} \) (Fig. 2.2b) for films which were deposited at high temperature (800 °C) using reactive DC magnetron sputtering. The TiN film was deposited at 800°C and the optical properties of 20 nm thin film was measured using spectroscopic ellipsometer (J.A. Woollam Co). High deposition temperature poses some fabrication and integration restrictions which must be taken into account. The materials beneath the TiN layer must withstand the TiN-
deposition and etch conditions without degradation. Since the properties of the TCO degrade at high temperatures, the TCO layer must be deposited only after the deposition and patterning of the TiN layer.

Fig. 2.2. (a) GZO permittivity versus its carrier concentration, $\lambda = 1.55 \mu m$. The permittivity of the GZO layer was taken from [96] and a carrier concentration in the GZO was determined using a Drude-Lorentz model fitting: $N_0 = 9.426 \times 10^{20} \text{ cm}^{-3}$ (black dotted line). (b) TiN permittivity extracted from spectroscopic ellipsometry measurements.
The simplest structure is shown in Fig. 2.3i-a where modulation is achieved by applying a bias across the GZO layer. For this structure, the zinc oxide (ZnO) layer serves as one electrode while the GZO film itself is the second. A dielectric layer between the TCO film and second conductive layer is required to provide electrical insulation. This layer should be made as thin as possible to reduce the voltage required to modulate the GZO carrier concentration.

For the design in Fig. 2.3i-a, a thick Si$_3$N$_4$ layer can be utilized deposited, for example, on a silicon substrate. It is also preferable to have materials with similar indices on the top and bottom of the plasmonic layer. In this case, the conditions are similar to those required for long-range SPP mode propagation, and the mode losses are lower [28].

Furthermore, we studied designs that include TiN layers (Fig. 2.3b,c). The addition of the layer allows for the modulator to be easily integrated with external strip waveguides. It also provides tighter field confinement, which will result in a larger attenuation of the signal during modulation. Both layouts with a thick and thin TiN layer are studied. On the structures shown in Fig. 2.3, the central layers (TiN, silicon nitride insulation, and GZO) remain in the same configuration.
Fig. 2.3. Multilayer structures. Waveguide layers (GZO and TiN) are sandwiched or covered with (i): low-index materials, (ii): silicon as a high-index material. Further in the text, we will refer to designs (a) as “without TiN”, (b) as “thick TiN”, (c) as “thin TiN”. (d) Absorption coefficients $\alpha$ for various carrier concentrations of GZO. Structures with high-index cladding (ii) show much higher absorption than structures with silicon nitride cladding (i). Notation “sym” and “asym” correspond to quasi-symmetric and quasi-asymmetric SPP modes respectively. (f) Mode size of the modulator structures versus carrier concentration in the GZO film. The absorption maximum is accompanied by highest mode localization. At lower carrier concentrations of GZO, modes are more spread-out because of the smaller magnitude of real permittivity of GZO.
Here we consider two main groups of the devices, one with low-index cladding (Fig. 2.3i) and another with high-index cladding (Fig. 2.3ii). Further in the text, ZnO, LP-CVD Si$_3$N$_4$ and PE-CVD silicon nitride (denoted in subsequent text by SiN), will be referred to as low-index materials. As we are interested in operation at the telecom wavelength of \(\lambda = 1.55\ \mu m\), the refractive indices used in the calculations are the following: \(n_{\text{ZnO}} = 1.93\) [100], \(n_{\text{SiN}} = 1.76\) (experimental characterization of samples after PE-CVD process) and \(n_{\text{Si}_3\text{N}_4} = 1.97\) (after LP-CVD process). It should be mentioned that LP-CVD Si$_3$N$_4$ requires high temperature deposition which may degrade the properties of TCO layer. Hence, only PE-CVD SiN can be deposited after the TCO layer.

Furthermore, a high-index material can be utilized as a cladding. In this work, we consider silicon as a high-index cladding \(n_{\text{Si}} = 3.48\) [101]. While amorphous silicon would be deposited as the upper cladding layer, for this investigation, we consider crystalline silicon and amorphous silicon to be identical in their optical properties at \(\lambda = 1.55\ \mu m\). In particular, the next three structures (Fig. 2.3ii) are similar to the first three but with silicon layers as a top and bottom cladding. In these cases, either the silicon or TiN layer can be used as a second electrode. In all cases we neglect optical losses in the silicon as they are much lower than plasmonic materials bring.

### 2.2. Performance of the modulators

With all these considerations, six basic geometries were chosen as templates for modulator designs, operating at the telecom wavelength of \(\lambda = 1.55\ \mu m\). The dispersion equation was solved for the one-dimensional multilayer structures with varying carrier concentrations in the GZO. To define attenuation of signal in decibels, we calculated absorption coefficient as \(\alpha = 8.68\text{Im}(\beta_{\text{eff}})\), where \(\beta_{\text{eff}}\) is complex propagation constant of plasmonic wave in the multilayer structures. The thickness of all thin layers (GZO, TiN, SiN, Si$_3$N$_4$) is 10 nm. The top and bottom layers are assumed to be infinitely thick. Thus, we will refer to the structures in Fig. 2.3a as “without TiN”. The structures in Fig. 2.3b are referred to as “thick TiN” and Fig. 2.3c as “thin TiN”.

At a particular carrier concentration, the GZO permittivity value satisfies the condition to achieve a plasmonic resonance in the multilayer structures, resulting in a significant increase in the absorption coefficient \(\alpha_{\text{max}}\) (Fig. 2.3d). Therefore, the absorption
coefficient $\alpha$, in the waveguide structure, strongly depends on carrier concentration $N$: $\alpha = \alpha(N)$. $\alpha_{\text{max}}$ is lower for structures without TiN, higher for a thick TiN layer and the highest with a thin TiN film (Table 2.1).

Modes of the structures with a GZO layer only (in Fig. 2.3a) have a quasi-symmetric electric field distribution. Adding a TiN layer increases the absorption coefficient of the multilayer structure because of the ohmic losses. Moreover, the structure with a thin TiN layer and low-index cladding (Fig. 2.3i-c) supports only the mode with a quasi-asymmetric electric field profile, while the structure with the high-index cladding (Fig. 2.3ii-c) supports both quasi-symmetric and quasi-asymmetric. In the quasi-asymmetric mode, the field is mostly localized near the waveguide (in contrast to quasi-symmetric mode, where the field is mostly spread outside the waveguide) such that the losses are even higher. In this case, $\alpha_{\text{max}}$ reaches 28 and 132 dB/µm for the low- and high-index cladding, respectively.

Furthermore, we analyzed the size of the mode in all the proposed structures. In the case of a single interface, the 1/e point of the electric field corresponds to an 86% localization of electrical energy ($1-e^{-2}$ portion). To estimate the mode size of our multilayer structures, which have a complicated field profile, we define the mode size such that 86% of electrical energy is localized within the region (Fig. 2.4). Similar to other long-range SPP based waveguides, the structure suffers from low mode localization. It can be seen from Fig. 2.3f that structures without TiN have lower mode localization in comparison to those with TiN. Moreover, utilizing a high-index cladding significantly decreases mode size. In all cases, a decrease of $\alpha$ at lower $N$ is accompanied by significant increase of mode size.
Fig. 2.4. Depiction of the mode profile for the geometry Fig. 2.3i-c showing the definition of mode size. Due to the complexity of the structure and high concentration of electrical energy in the GZO layer, the traditional definition of mode is cannot be utilized as this would simply define the mode size as the thickness of the GZO layer. Here we define mode size as the distance which encompasses 86% of the electric field energy, a condition similar to that of the 1/e definition for single interface waveguide.

At some particular carrier concentration in GZO, the absorption coefficient reaches a maximum value $\alpha_{\text{max}}$. The extinction ratio ER, or modulation depth, can be defined as

$$ER = \alpha_{\text{max}} - \alpha_{\text{min}},$$

where $\alpha_{\text{min}}$ is propagation loss in the transmissive state. Here we do not specify which states correspond to voltage on and voltage off states. It depends on how GZO layers are deposited and which carrier concentration is chosen as an initial value.

Either increasing or decreasing $N$ from $N_{\text{on}}$ results in modulation of $\alpha$ and accompanied by mode extension. However, as one can see from Fig. 2.3f, for $N < N(\alpha_{\text{max}})$, mode size increases more dramatically than for $N > N(\alpha_{\text{max}})$. For the same $\alpha < \alpha_{\text{max}}$, mode size is larger for lower $N$ than for higher. In case of coupling to plasmonic waveguide with large mode size, operation at lower $N$ can be more beneficial. However, we are interested in decreasing modulator mode size and thus it is more preferable to operate at higher $N$. Thus, $\alpha_{\text{min}}$ is defined by

$$\alpha_{\text{min}} = \alpha(2N_0).$$

(2.2)
Eq. (2.2) is valid for all the proposed structures apart from Si₃N₄/GZO/SiN/ZnO (Fig. 2.3i-a). This is because with the low-index cladding and absence of TiN, a localized mode exists only for a narrow range of \( N = 5 \times 10^{20} \) to \( 8 \times 10^{20} \) cm\(^{-3}\). Modes larger than 10 \( \mu \)m are considered delocalized and the corresponding values for \( N \) will not be used in subsequent calculations. Thus, for this layout, \( \alpha_{\text{min}} = \alpha (N = 8 \times 10^{20} \text{ cm}^{-3}) \) is defined. While operation in this narrow range of \( N \) can be more preferable as it does not require a large change of \( N \), it provides less tolerance to fabrication or design imperfections.

Similar to the value of \( \alpha_{\text{max}} \), the ER is lower for the structures without TiN, higher for a thick TiN layer and the highest with a thin TiN film (see Table 2.1 for a comparison of values). The ER is 1.8-16 dB/\( \mu \)m for a silicon nitride cladding, and 24-86 dB/\( \mu \)m for a silicon cladding. In the latter case, less than a 35-nm-length active section is required to achieve 3 dB modulation.

A figure of merit FoM for such multilayer modulator structures can be defined as

\[
\text{FoM} = \frac{\text{ER}}{\alpha_{\text{min}}}. \tag{2.3}
\]

It reflects a trade-off between the modulation depth and the loss of the signal in the transmittive state (\( \alpha_{\text{min}} \)). While the structures with a thin TiN layer provide the strongest resonance, \( \alpha \) is also relatively high at large \( N \). Such structures give the lowest performance.

The highest FoM is provided by structures without TiN. However, the lowest absorption in the transmittive state is accompanied by lowest mode localization (up to 10 \( \mu \)m). In Table 2.1 we summarize the ranges of mode extensions. In most cases, the minimum value corresponds to plasmonic resonance and the maximum to a carrier concentration \( N = 2N_0 \) (structure on Fig. 2.3i-a is an exception).

Calculations show that the high-index cladding designs possess the highest performance. Working with a 2x change in the carrier concentration of GZO, the studied plasmonic modulator can outperform previously proposed designs. For example, deeply subwavelength MIM structures were analyzed and a corresponding ER up to 12 dB/\( \mu \)m was theoretically predicted [56,57]. However, such high values are accompanied by high losses in transmittive state. The ratio of the absorption coefficients in the two states, FoM, is on order of 1. In our case, because of the possibility to detune from the plasmonic resonance, the absorption coefficient in the transmittive
state can be relatively low. Utilizing a high index cladding makes the resonance more pronounced and the required change of carrier concentration is smaller. This structure also achieves transmittive state losses down to 0.06 dB/µm, which produced the highest FoM = 400. Thus, the Fig. 2.3ii-a geometry with high-index silicon claddings and without TiN, provides the highest performance for an ultra-compact plasmonic modulator.

Table 2.1. Summary of the characteristics of different structures. Performance comparison for planar modulator designs. Designs utilizing high-index materials show the highest performance.

<table>
<thead>
<tr>
<th>Structure (layers bottom to top)</th>
<th>N(α_{max}), 10^{20} cm^{-3}</th>
<th>α_{max}, dB/µm</th>
<th>α_{min}, dB/µm</th>
<th>ER, dB/µm</th>
<th>FoM</th>
<th>Mode size, µm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si_{3}N_{4}/GZO/SiN/ZnO (Fig. 2.3i-a)</td>
<td>6.13</td>
<td>1.95</td>
<td>0.114</td>
<td>1.8</td>
<td>16</td>
<td>1.6 – 10</td>
</tr>
<tr>
<td>TiN/Si_{3}N_{4}/GZO/SiN (Fig. 2.3i-b)</td>
<td>6.13</td>
<td>8.4</td>
<td>0.55</td>
<td>8</td>
<td>15</td>
<td>0.5 – 1</td>
</tr>
<tr>
<td>Si_{3}N_{4}/TiN/Si_{3}N_{4}/GZO/SiN (Fig. 2.3i-c)</td>
<td>6.41</td>
<td>28</td>
<td>12.2</td>
<td>16</td>
<td>1.3</td>
<td>0.2 – 0.3</td>
</tr>
<tr>
<td>Si/Si_{3}N_{4}/GZO/Si (Fig. 2.3ii-a)</td>
<td>6.79</td>
<td>24</td>
<td>0.060</td>
<td>24</td>
<td>400</td>
<td>0.09 – 6</td>
</tr>
<tr>
<td>TiN/Si_{3}N_{4}/GZO/Si (Fig. 2.3ii-b)</td>
<td>7.82</td>
<td>60</td>
<td>4.2</td>
<td>56</td>
<td>13</td>
<td>0.03 – 0.2</td>
</tr>
<tr>
<td>Si/TiN/Si_{3}N_{4}/GZO/Si (Fig. 2.3ii-c, asym)</td>
<td>9.00</td>
<td>132</td>
<td>46</td>
<td>86</td>
<td>1.9</td>
<td>0.04 – 0.09</td>
</tr>
<tr>
<td>Si/TiN/Si_{3}N_{4}/GZO/Si (Fig. 2.3ii-c, sym)</td>
<td>6.60</td>
<td>46</td>
<td>0.29</td>
<td>46</td>
<td>160</td>
<td>0.06 – 1.3</td>
</tr>
</tbody>
</table>

The extinction ratio for the investigated structures was found to be 1.8-16 dB/µm for a silicon nitride cladding, and 24-86 dB/µm for a silicon cladding. It is one of the highest values reported so far for both theoretical predictions and experimental demonstration for plasmonic modulators.

2.3. Waveguide and modulator integration

Efficient modulators allow a 3 dB modulation depth within a one-micron length plasmonic modulator. Thus, these devices can be very short and considered as a small section of a larger plasmonic waveguide. To couple into these devices, several possible integration schemes can be studied. Modulator structures can be fabricated on top of a plasmonic waveguide. Here we consider 10-nm thin TiN layer, which supports the LR-SPP. In a silicon nitride cladding, the
propagation length in such waveguides is 5.5 mm. By limiting the GZO to only a small section required for modulation, the added propagation losses in the GZO dynamic layer are avoided in the remainder of the waveguide. To achieve this, the GZO can either be added directly on top of the TiN layer or used as the plasmonic material in replacement of the TiN (Fig. 2.5). In the latter case, GZO serves as both a waveguide and dynamic element.

Fig. 2.5. Schematic of plasmonic modulators integrated with TiN strip waveguides providing long range SPP propagation to and from the modulator (side view). Within the modulator section of the designs, the TiN strip waveguide is replaced by a GZO layer which is plasmonic at the wavelength of interest. The carrier concentration of this GZO film is altered by the application of an electric field across the layer. In geometry (i) the voltage is applied between the GZO and the ZnO layers, (ii) the voltage is applied between the upper and lower silicon layers.

The designs with thick TiN layers (Fig. 2.3b) could be integrated with a single-interface waveguide. However, such waveguides have significantly higher losses in comparison with thin TiN layers. Therefore, their applications are limited. For this reason, these geometries will not be considered in the following analysis.

Since the mode in the modulator section in Fig. 2.3c is quasi-asymmetric, it must be excited by the asymmetric mode of a strip-waveguide. Besides the challenge of excitation of asymmetric mode, it has much higher propagation losses. Furthermore, for the design in Fig. 2.3b,c, a thin layer of Si$_3$N$_4$ on top or beneath TiN is needed to insulate the GZO layer. However, realization of the designs with Si/TiN/Si$_3$N$_4$/Si waveguides encounters an issue. Because of the drastic difference between refractive indices of Si and Si$_3$N$_4$, it does not support the symmetric mode. It cannot be easily replaced by a Si/TiN/Si waveguide, as the thin Si$_3$N$_4$ layer (or a more advanced method of electrical isolation such as p-n junction doping) is required to maintain electrical isolation in the active section.
Thus, from the integration point of view, the best modulator structures are those without TiN (Figs. 3a and 5). As we showed in the previous subsection, these structures also give the highest performance in terms of modulation depth and propagation losses.

Similar to the previous subsections, we perform calculations for one-dimensional structures as their properties are close to those of finite-width. The coupling losses $\gamma$ for a single interface was calculated by following equation

$$\gamma = \frac{4\beta_1\beta_2}{(\beta_1+\beta_2)^2} \left[ \int_{-\infty}^{\infty} E_{1z}E_{2z}^* dz \right]^2, \quad (2.4)$$

where $\beta_1$ ($E_1$) and $\beta_2$ ($E_2$) are the mode indices (electric field) of the waveguide and waveguide modulator, respectively. Eq. (2.4) it takes into account both the mode overlap integral and the Fresnel coefficients at the boundary region.

We calculated the coupling losses for the two designs shown in Fig. 2.5, and the results are shown on Fig. 2.6. With regards to the scheme in Fig. 2.5i (the silicon nitride cladding), there is an interplay of two major effects. First, coupling losses are increased at the plasmonic resonance because of the high field localization in the GZO layer. This greatly reduces the mode overlap (Fig. 2.7i). Second, coupling losses are increased at lower and higher carrier concentration in GZO due to the mode extension outside the waveguide where GZO no longer supports a plasmonic mode (similar to Fig. 2.3i-f for the geometry “without TiN” where the usable carrier concentration is between approximately $5 \times 10^{20}$ cm$^{-3}$). As a result, coupling losses vary by only 5 dB across the modulation range.

For the scheme employing a high-index silicon cladding (Fig. 2.5ii), the coupling loss in the transmittive state monotonically increases as the carrier concentration decreases towards the maximum in the modulator absorption (similar to the first effect for the low-index cladding). However, because of the mode mismatch at the low-to-high index interface between the waveguide and modulator (Fig. 2.7i), coupling losses are higher at the GZO plasmonic resonance. This effect can be beneficial for modulator performance in specific applications as it provides additional losses in the resonant state and fewer losses in the transmittive state.
Fig. 2.6. Single interface coupling loss (between waveguide and modulator sections) versus carrier concentration in the GZO layer. This is shown for both the low-index silicon nitride and high-index silicon cladding. The significant increase of coupling losses corresponds to mode extension outside the waveguide. For the high-index cladding, the coupling losses are large because of the mode size mismatch between the Si$_3$N$_4$/TiN/Si$_3$N$_4$ waveguide and modulator with silicon cladding. For the low-index cladding, we consider only the values of N which correspond to mode size less than 10 $\mu$m.
Fig. 2.7. Example mode profiles in the two integrated modulator geometries: (i) low-index and (ii) high-index claddings. Note that the field decay outside the strip waveguide is slow and therefore appears constant in this graph. The carrier concentration in the GZO layer used for the calculations corresponds to the maximum absorption in the modulator, i.e. plasmonic resonance in the layer. Under these conditions the majority of the field is localized within the GZO layer.

2.4. Conclusion

In this chapter, we have analyzed several multilayer structures with alternative plasmonic materials to be utilized in ultra-compact plasmonic modulators. Applying an electric field across the TCO layer allows for the permittivity to be tuned, resulting in a change of the absorption coefficient of the waveguide. Therefore, active modulation is achieved. Numerous modulator layouts are investigated and the typical trade-off between compactness and propagation loss is analyzed. Amongst all the reported structures, one stands out with a remarkable FoM = 400. This figure of merit takes into account both
the modulation depth ($ER = 24 \text{ dB/µm}$) and the propagation losses in the transmittive state ($\alpha = 0.06 \text{ dB/µm}$). The corresponding geometry may allow for ultra-compact modulation with effective length much less than 1 µm. The proposed approach based on the cost-effective planar fabrication processes and the ability to easily integrate with existing semiconductor systems could enable new devices for applications in on-chip optics, sensing, optoelectronics, data storage, and information processing.
Chapter 3. Plasmonic modulator optimized by patterning of active layer and tuning permittivity

In this chapter, an ultra-compact plasmonic modulator based on metal-insulator-metal waveguide with an additional ultra-thin layer of indium tin oxide (ITO) is studied. Bias is applied to the multilayer core by means of metal plates that serve as electrodes. External field changes carrier density in the ultra-thin ITO layer, which influences the permittivity. The metal-insulator-metal system possesses a plasmon resonance, and it is strongly affected by changes in the permittivity of the active layer. To improve performance of the structure, several optimizations are proposed. We analyze effects of varying the ITO permittivity as a result of different annealing conditions, examine its influence on the modulator performance and point out appropriate values. We analyze eigenmodes of the waveguide structure and specify the range for its efficient operation. We show that substituting the continuous active layer by a one-dimension periodic stripes increases transmittance through the device and keeps the modulator’s performance at the same level. The dependence on the pattern size and filling factor of the active material is analyzed and optimum parameters are found. Patterned ITO layers allow us to design a Bragg grating inside the waveguide. The grating can be turned on and off, thus modulating reflection from the structure.

3.1. Multilayer structures

The modulator consists of 8-nm-thick ITO layer and a 70-nm thick silicon nitride (Si$_3$N$_4$) core with $\varepsilon_{\text{core}} = 4$ (Fig. 3.1a). These layers are sandwiched between two silver plates acting as electrodes. The thicknesses are considered as optimal based on the trade-off between the propagation length and performance [56]. Data for the ITO parameters are taken from [102,103]. We assume that the ITO permittivity is approximated by the Drude formula (see [102]):

$$\varepsilon = \varepsilon_\infty - \frac{\omega_{\text{pl}}^2}{\omega^2 + i\gamma\omega},$$

where $\varepsilon_\infty = 3.9$, plasma frequency $\omega_{\text{pl}} = 2.9 \cdot 10^{15}$ s$^{-1}$ and collision frequency $\gamma = 1.8 \cdot 10^{14}$ s$^{-1}$. The carrier density in the ITO layer changes under applying modulating voltage between the electrodes. Here, we follow the approach of [56] to calculate these
changes. It employs the Thomas-Fermi screening theory and averaging of the carrier density over the whole ITO layer. Thus, 5\% increasing of the average carrier density alters the plasma frequency $\omega_{pl}$ from $2.9 \cdot 10^{15}$ to $2.9716 \cdot 10^{15}$ s$^{-1}$, that shifts the ITO permittivity from $-1.67 + i0.825$ to $-1.95 + i0.867$ at telecommunication wavelength 1.55 $\mu$m. Data [103] for a silver permittivity are approximated by the Drude formula as well, where parameters are following: $\varepsilon_{\infty}(Ag) = 1$, plasma frequency $\omega_{pl}(Ag) = 1.38 \cdot 10^{16}$ s$^{-1}$ and collision frequency $\gamma(Ag) = 3.22 \cdot 10^{13}$ s$^{-1}$. It gives $\varepsilon_{Ag} = -128.7 + 3.44i$ at $\lambda = 1.55$ $\mu$m.

We reduce dimension of the problem and consider the modulator as a two-dimensional four-layer metal-insulator sandwich, as shown in Fig. 3.1b. We solve numerically the SPP dispersion equation for the two-dimensional structure the same way as in [56,58]. The reference system supports two SPP modes in the transverse magnetic (TM) polarization.

The first mode with the lowest absorption has $\alpha_{\text{off}} = 1.08$ $\mu$m$^{-1}$, $\alpha_{\text{on}} = 0.83$ $\mu$m$^{-1}$ and corresponding FoM$_1 = 0.30$ at $\lambda = 1.55$ $\mu$m. The second mode has a very high absorption coefficient, i.e. $\alpha_{\text{off}} = 5.47$ $\mu$m$^{-1}$ and $\alpha_{\text{on}} = 5.87$ $\mu$m$^{-1}$, that gives a fairly low FoM$_2 = 0.07$. However, because of such high absorption of the second mode we study performance of the device only for the first one.

We calculate also a transmission coefficient of the reference modulator. It exhibits significant attenuation of propagating waves due to the ITO layer. The transmitted signal amplitude in the voltage-off state is only 12\% of the incoming wave for modulator’s length $L = 2$ $\mu$m and approximately 1\% for $L = 4$ $\mu$m. These numbers specify a problem which the reference design experiences, namely a rather low level of transmitted signals. Our optimization aims to increase this level.

For the purposes of further designing steps we also analyze the symmetric system: it consists of the Si$_3$N$_4$ core embedded between two 8-nm-thick ITO layers and sandwiched between silver plates (Fig. 3.1c). In this case the bottom layer experiences a carrier depletion and decreasing of the plasma frequency ([on−]-state), that is opposite to the top layer ([on+]-state). While reference voltage results in the 5\% increasing of the average carrier density in the top ITO layer, the plasma frequency $\omega_{pl}$ for the bottom ITO layer changes from $2.9 \cdot 10^{15}$ to $2.8267 \cdot 10^{15}$ s$^{-1}$ due to the 5\% decreasing of the average carrier
density under the same voltage. Correspondingly, the ITO permittivity is shifted from \(-1.67 + i0.825\) to \(-1.95 + i0.867\) (top layer) and to \(-1.39 + i0.784\) (bottom layer) at \(\lambda = 1.55\ \mu m\). In total, the whole five-layer system possesses a very low \(\text{FoM}_{5\text{-layers}} = 0.05\), i.e. the controllability of the system is lost. The effective permittivity of the three layers sandwiched between the silver plates under applied voltage is turned out to be approximately the same as without the voltage. Meanwhile, the additional ITO layer doubles maximal absorption in plasmonic resonance of the system, and the effect is even more pronounced on the resonance slope. Thus, for \(\lambda = 1.55\ \mu m\) the five-layer system has approximately four times higher absorption coefficient than the reference asymmetric case, in particular \(\alpha_{\text{off}} = 4.05\) and \(\alpha_{\text{on}} = 4.25\). However, further we show some advantages of using two ITO layers with patterning.

Yet another merit in the performance of the four-layer modulator can be achieved by applying bias in both directions. In this case the change of the plasma frequency for the voltage [on+] state and [on–] state is between \(2.9716 \cdot 10^{15}\) and \(2.8267 \cdot 10^{15}\) s\(^{-1}\), that gives varying of absorption coefficient between \(\alpha_{\text{on–}} = 1.47\ \mu m^{-1}\) and \(\alpha_{\text{on+}} = 0.83\ \mu m^{-1}\), and corresponding \(\text{FoM}_{\text{flip}} = 0.78\). The efficiency is better than in case of applying the double voltage in one direction: the average carrier density changes on 10\%, the ITO plasma frequency becomes \(3.0415 \cdot 10^{15}\) s\(^{-1}\), giving \(\varepsilon_{\text{ITO}} = - 2.23 + i0.908\), \(\alpha_{\text{on}} = 0.657\) and finally \(\text{FoM}_{10\%} = 0.64\). The flipping of voltage direction can be also preferable for systems with a low break-through voltage threshold.

Fig. 3.1. a) The reference modulator structure: 8-nm ITO layer and 70-nm Si\(_3\)N\(_4\) layer are embedded between the silver plates [56]. b) Schematic two-dimensional view of the equivalent four-layer system. c) The symmetric five-layer system with two 8-nm-thick ITO layers and 62-nm-thick Si\(_3\)N\(_4\) core.
3.2. Effect of ITO permittivity changes

To improve modulator's performance we study variations of the ITO permittivity. In practice it is possible to vary the optical properties of transparent conducting oxides by different anneal conditions, i.e. temperature and environment, during the fabrication process. The range of the permittivity variations can be several units or even more [71]. We assign the variations in the ITO permittivity to parameter $\varepsilon_\infty$ in Drude formula. Thus, in our study, we sweep the real part of the permittivity keeping the change of the plasma frequency $\omega_{pl}$ (from $2.9 \cdot 10^{15}$ s$^{-1}$ to $2.9716 \cdot 10^{15}$ s$^{-1}$). We assume the collision frequency $\gamma$ being also fixed and therefore the imaginary part of the permittivity is kept as before, namely 0.825 and 0.867 for the off-state and on-state, respectively. So the difference between the off-state and on-state permittivities ($\varepsilon_{\text{off}} - \varepsilon_{\text{on}} = 0.28 - \text{i}0.042$) is fixed too.

The results of calculations for the absorption coefficients and FoM are shown in Fig. 3.2. The FoM reaches the first maximum on the steepest slope of the absorption coefficient curve. It then nullifies in the point of equal absorption coefficients in both states and then increases again up to the biggest FoM$_{\text{max}} = 0.37$ for Re($\varepsilon_{\text{off}}$) $\approx$ 0.83 that corresponds to $\varepsilon_\infty$ = 6.4. In principle such values of Re($\varepsilon_{\text{off}}$) are available with the state-of-the-art fabrication technology [71]. Consequently, our further study is carried out with this optimized value of $\varepsilon_\infty$. We note that Re($\varepsilon_{\text{off}}$) $\approx$ -1.27 ($\varepsilon_\infty$ = 4.3) gives high FoM = 0.36 as well. Fig. 3.2 indicates that the absorption coefficient for positive Re($\varepsilon_{\text{off}}$) is higher in the on-state than in the off-state. It means that by applying voltage we decrease transmittance of the system.

Further, we analyze a five-layer system (Fig. 3.1c) in the situation with the reversed sign of the changes in losses. Re($\varepsilon_{\text{off}}$) of different signs can be obtained under different anneal conditions for the bottom and top ITO layers, as they are deposited at different stages of fabrication. The same way as for the four-layer system we vary Re($\varepsilon_{\text{off}}$) of the bottom layer in broad range and keep Re($\varepsilon_{\text{off}}$) $\approx$ 0.83 for the top layer. Calculations show that the additional layer does not provide any advantages and none of Re($\varepsilon_{\text{off}}$) can result in FoM exceeding FoM$_{\text{max}} = 0.37$ for the optimal four-layer structure.
Fig. 3.2. The absorption coefficients $\alpha$ and FoM of the four-layer system and various permittivities of the ITO layer. Maximal values $\text{FoM}_{\text{max}} = 0.37$ and $\text{FoM} = 0.36$ are achieved for $\text{Re}(\varepsilon_{\text{off}}) \approx 0.83$ and $\text{Re}(\varepsilon_{\text{off}}) \approx -1.27$ respectively.

Fig. 3.3. The absorption coefficients $\alpha$ and FoM of the five-layer system with various permittivities of the bottom ITO layer. The permittivity of the top layer is fixed $\text{Re}(\varepsilon_{\text{off}}) \approx 0.83$. FoM values are not exceeding $\text{FoM}_{\text{max}} = 0.37$ for the four-layer system.
3.3. Eigenmodes of four-layer system

In this subsection we analyze eigenmodes of the two-dimensional four-layer structure with the optimized ITO permittivity $\varepsilon_\infty = 6.4$ (Fig. 3.1b). We solve numerically the SPP dispersion equation in the frequency range between 130 THz and 220 THz. For frequencies below 220 THz the system supports only two modes: the propagation constants and absorption coefficients are shown in Fig. 3.4 and Fig. 3.5 respectively.

We consider three states: the off-state with the ITO plasma frequency $2.9 \cdot 10^{15} \text{ s}^{-1}$ (without any voltage); [on+] state under applied voltage with the ITO plasma frequency $2.9716 \cdot 10^{15} \text{ s}^{-1}$; and [on–] state under the reverse bias with the plasma frequency $2.8267 \cdot 10^{15} \text{ s}^{-1}$. We also plot the propagation constant for the three-layer system with the 78-nm-thick Si$_3$N$_4$ core (without ITO layer).

The propagation constants of both modes exhibit monotonous increase with frequency apart from the region 160-200 THz, where $\beta$ of the first mode decreases being strongly dependent on changes in ITO permittivity. Consequently, the absorption coefficient of the first mode has maximum in that frequency region (SPP resonance). The absorption coefficient of the second mode monotonically decreases.

Plots in Fig. 3.4 and Fig. 3.5 indicate that for the first mode the modulator working region should be between 160 THz and 220 THz, where a change of the ITO plasma frequency strongly affects the propagation constant and absorption coefficient of the system. Therefore, both mechanisms of modulation: the absorption change and the propagation constant variation can be applied. The latter is illustrated for a grating system further. Despite that the propagation constant has a large response on applied voltage around 180 THz, the effect causes an increase of absorption. Designing device, this circumstance should be taken into account.

Fig. 3.6 shows field distributions of two modes in the waveguide cross-section at frequency 160 THz. Both modes are plasmonic and quasi-symmetric (the ITO layer removes symmetry). They can be undesirably excited at the same time. However, in contrast to [50] or [53], the second mode damps much faster and consequently neither pronounced interference nor beatings can appear. We characterize the quality of a mode by factor $Q = \beta_{\text{off}}/\alpha_{\text{off}}$. 
Fig. 3.4. Propagation constants $\beta$ for two modes “1” and “2” without voltage (off-state), under applied direct voltage ([on+] state) and under applied reverse voltage ([on–] state) for different frequencies. Green dot-dash line indicates the light line in $\text{Si}_3\text{N}_4$ with refractive index $n = 2$.

Fig. 3.5. Absorption coefficients $\alpha$ and quality factor $Q = \beta_{\text{off}}/\alpha_{\text{off}}$ for two modes “1” and “2” without voltage (off-state), under applied direct ([on+] state) and reverse ([on–] state) voltage for different frequencies.
Because of the high absorption and small propagation constant of the second mode, the Q-factor for the second mode is much lower than for the first one, e.g. $Q_1 \approx 10$ and $Q_2 \approx 2.5$ for 193.4 THz (see Fig. 3.5). Furthermore, the second mode has much lower FoM than the first one. Thus, the second mode does not support efficient performance.

Fig. 3.6. Field distributions (absolute values of electric field $E$ (a) and transverse magnetic component $H_y$ (b)) of two plasmonic modes in the vertical cross-section. Calculations are performed for the off-state system. Fields are normalized such that $|H_y| = 1$ at the interface between ITO and Si$_3$N$_4$.

In the considered frequency range the second mode has smaller propagation constant than that of light in the dielectric core (the mode is above the light line). It allows the exponential increase of the electric field from the interface between the ITO film and dielectric core (see Fig. 3.6). Thus, it is a quasi-leaky mode [105], and it possesses high modal losses. Similar to [43] the system has an antenna mode region ($Q_2 > 1$) and a reactive mode region ($Q_2 < 1$). As one can see in Fig. 3.5 the transition occurs at 180 THz. At the transition frequency the system changes its response on applied voltage: in the reactive region the absorption coefficient is mostly affected and the propagation constant is not, while in the antenna region it is vice
versa. Calculations show that the second mode crosses the light line at 300 THz (assuming the same optical parameters of silver and ITO) and becomes a bound mode. The small propagation constant of the second mode can provide efficient coupling from free space or another waveguide.

3.4. Periodic patterning of ITO layer

The effective modulation performance means the deep modulation range and improved transmittance, which we aim implementing optimizations in the reference system. To increase transmittance we replace the continuous ITO film by periodic stripes (Fig. 3.7a). The patterning is characterized by a filling factor defined as the ratio of the stripe width \( w \) to the period \( P \) of the structure, \( f = w/P \). \( f = 1 \) corresponds to the continuous ITO film.

The system with patterned layer requires numerical solution of Maxwell’s equations. We perform simulations with the commercial software package CST Microwave Studio in the frequency domain [106]. The silver plates are 120-nm-thick, which is enough to keep the domain-termination error at a negligible level.

We analyze several systems with different period \( P \) and total length \( L \). Results of simulations for transmission coefficient in the voltage-off state and FoM versus the filling factor are shown in Fig. 3.7b (\( \lambda = 1.55 \mu m \)). As we foresaw the transmission increases dramatically with the fall of the filling factor. The finer is sampling (less period \( P \)) the higher is the transmission coefficient for the same total length and filling factor. For all three periods (\( P = 250 \) nm; 500 nm and 4 \( \mu m \)) the FoM is almost constant in the broad range of filling factors. However, for \( f \leq 0.2 \) FoM falls down abruptly and the fall is steeper for larger periods.

In the first approximation transmission through one period can be considered as transmission through two successive layers: a part of the waveguide with the ITO stripe (length \( w \) and absorption coefficient \( \alpha_{\text{state}} \)), and a part of the waveguide without the ITO stripe (length (\( P-w \)) and absorption coefficient \( \alpha_0 \)). Thus, total transmission coefficient of the patterned waveguide with length \( L \) is:

\[
Tr_{\text{state}} = \exp(-\alpha_{\text{state}} fL - \alpha_0 (1-f) L),
\]  

(3.1)
where “state” is either “on” or “off”, and $\alpha_0 = 0.027 \, \mu\text{m}^{-1}$ for 78-nm-thick Si$_3$N$_4$ core.

In this case FoM can be found as:

$$\text{FoM}(f) = \frac{\log(Tr_{\text{on}}) - \log(Tr_{\text{off}})}{\log(Tr_{\text{off}})} = \frac{\text{FoM}_{\text{max}}}{1 + (1 - f)\alpha_0/f\alpha_{\text{off}}}. \quad (3.2)$$

Fig. 3.7b shows good agreement of approximation (3.1) with numerical simulations. It completely coincides with results for $P = 4 \, \mu\text{m}$ and $L = 4 \, \mu\text{m}$. Eq. (3.2) explains the constant value of the FoM in the broad range of $f$. For $f = 0.2$ the ratio $(1 - f)\alpha_0/f\alpha_{\text{off}} \approx 0.1$, and for $f \leq 0.2$ the additional term in denominator (3.2) has an essential contribution. It means that losses in the part of waveguide without ITO influence the total transmission coefficient, and meanwhile the fraction of ITO is insufficient for the effective change in the propagation.

The transmission coefficient has a slight dependence on the period (for samples of equal total length $L$), and the smallest periods give highest transmission coefficient. Nevertheless, for the certain FoM the systems with highest period have highest transmission (compare two systems marked with crosses in Fig. 3.7b). So, partial removing of the lossy ITO layer with as big as possible period can be recommended aiming the higher transmittance with the same performance. For the following analysis we take filling factor $f = 0.2$ as optimal patterning that gives high transmittance.

To study performance of the patterned system we consider one period of the waveguide, which contains one ITO stripe and has length $P$. Throughout numerical simulations two outbound metal corners of the stripes are rounded with the curvature of 4 nm (apart from the 5-nm stripes with the radius of curvature 2.5 nm). Such curvature value is chosen to suppress any spurious field enhancement at the sharp corners and in the same time it is not too large to cause serious deviations in the structure response. To compare transmission efficiency in the case of different periods (the filling factor is fixed) the transmission coefficients are renormalized to the chosen structure length 4 $\mu\text{m}$ using the scattering matrix formalism (see e.g. [107]).

We plot amplitude transmission spectra (Fig. 3.8) for various sizes of the stripes from 5 nm to 800 nm as well as for the continuous homogeneous film with the dielectric function.
\[ \varepsilon_{\text{eff}} = f \varepsilon_{\text{ITO}} + (1-f) \varepsilon_{\text{core}}, \]  

(3.3)

where \( f = 0.2 \). This dielectric function effectively approximates the permittivity of the patterned ITO layer [108].

Fig. 3.7. a) The improved modulator design with periodic ITO stripes. 
b) Transmission coefficient (Tr) and FoM of the periodically patterned systems versus filling factor. Samples have different lengths (\( L = 2 \) and 4 \( \mu \)m) and periods (\( P = 250 \) nm, 500 nm and 4 \( \mu \)m). The structure with larger period outperforms among two patterned structures with equal FoM (marked by “x”). Transmission coefficients (A) correspond to approximation (3.1). FoM(A) is calculated with formula (3.2).
All spectra exhibit a pronounced minimum in the transmission with clear dependence of its position on the stripe width. Spectrally the minimum dramatically shifts for stripes with sizes below 100 nm and reaches the outward-red minimum for the continuous film with permittivity $\varepsilon_{\text{eff}}$ (3.3).

We explain such spectral shift by excitation of a hybridized plasmonic mode on single stripe [109]. The stripe of width $w \leq 100$ nm and thickness $h = 8$ nm with rounded corners can be considered as a nanowire with the quasi-elliptical cross section and aspect ratio $w/(2h)$. The silver plate connected to the ITO stripe serves as a mirror and effectively doubles the stripe thickness. Such quasi-elliptical nanowire in the electric field polarized along the minor axis supports dipole and quadrupole modes [109]. Their excitation causes intensive decrease in the transmission coefficient. So the overlap of their responses produces the transmission spectra we observe in Fig. 3.8.

To reveal the exact positions of both resonances we perform simulations for a hypothetic material identical to ITO but with diminished losses. The spectral profiles became sharper and deeper for 4-times-reduced $\gamma$ and splitting of the peak is clear in Fig. 3.9a. Further decreasing of losses makes resonances even narrower. Simulations also show that for systems with fixed width $w \geq 30$ nm the positions of both dips are independent on filling factor in the range $f \leq 0.5$.

It should be mentioned that applying the corners smoothing is especially important in the simulations with reduced losses. While transmission spectra for system with realistic ITO parameters are only slightly affected by the rounded concerns of stripes, the results are strongly affected in the case of diminished losses. A system with sharp corners and small losses has larger numbers of dips in the transmission spectra. Moreover, the approximation by an elliptical nanowire is not applicable in this case.

Simulations show that one dip is almost unaffected by the stripe width variation. It is fixed near 130 THz (see Fig. 3.9a). It corresponds to the quadrupole mode. Another dip, which belongs to the dipole mode, red-shifts from approximately 175 THz to 150 THz with changing stripe width from 800 nm to 15 nm. By further decreasing of the stripe width the dip shift becomes more pronounced and goes even below 100 THz (see Fig. 3.8). Frequencies of the transmission minima are
well correlated with the nanorods resonances obtained with the help of analytics from [109].

![Graph of Transmission Coefficient vs Frequency](image)

**Fig. 3.8.** The transmission spectrum has absorption maximum which shifts towards lower frequencies. Stripes have width $w$ in the range from 5 nm to 800 nm. The continuous film is calculated in the effective medium approximation (3.3). Simulations are performed for the off-state.

**Fig. 3.9a** presents changes in the transmission spectrum for turned on bias as well. Changes are the most pronounced in the region of the highest slope of the spectra. The shift of the absorption maximum affects position of the maximal FoM. Decrease of stripe sizes shifts the maximal FoM top toward lower frequencies (Fig. 3.9b). Having in mind telecommunications applications (193.4 THz), the absorption maximum of the modulator has to be around 180 THz. Therefore, it is important to work with stripes above 100 nm.
Fig. 3.9. (a) The amplitude transmission spectra for stripes of various widths $w$ in the off- and on-states. Decreasing of the collision frequency $\gamma$ in 4 times clearly shows the absorption maxima. (b) FoM for different stripe sizes.
3.5. Bragg grating

Based on mode analysis one can conclude that there is a frequency range where the pronounced change in the mode propagation constants is accompanied by the absorption coefficient variations. So the absorption modulation performance can be further improved by the phase effect.

It was shown that a plasmonic Bragg reflector can be formed by the metal-insulator-metal waveguide with periodic changes of the insulator material (index-modulated reflectors) or/and insulator thickness (thickness-modulated reflectors) [110-113]. Waves propagation in the structures with various shapes (step-profile, s-shaped, sawtooth-profile, triangular-shaped and so on) is described in terms of effective indices of the guided modes. We study the system shown in the inset of Fig. 3.10. The bottom ITO stripes are placed periodically in anti-phase to the top stripes. We assume that there are no any gaps or overlapping between the stripes after their projection on a horizontal plane. The off-state system is supposed to be a uniform waveguide with the complex propagation constant $\beta_{\text{off}} + i\alpha_{\text{off}}$, which we can find in the first approximation from the band diagrams (Fig. 3.4 and Fig. 3.5). Therefore, we neglect all multiple reflections inside off-state structure. Under applied voltage the part of the waveguide with the top (bottom) ITO stripes supports the mode with propagation constant $\beta_{\text{on}} + (\beta_{\text{on}} - )$. The concept of modulation is similar to that proposed in [114]: propagating waves experience Bragg reflection because of the periodic variation of effective mode indices along the waveguide under applied voltage.

The grating with the top stripe width $w_{\text{top}}$ and the bottom stripe width $w_{\text{bot}}$ obeying condition

$$w_{\text{top}} \beta_{\text{on}} = w_{\text{bot}} \beta_{\text{on}} = \pi / 2$$

(3.4)

satisfies the Bragg condition and causes enhanced reflection from the system.

For a quantitative analysis of the Bragg grating performance we chose two particular frequencies: $\nu_1 = 177.5$ THz and $\nu_2 = 210$ THz. At frequency 177.5 THz $\beta_{\text{off}} = 9.4$ $\mu$m$^{-1}$, $\beta_{\text{on}} = 10.0$ $\mu$m$^{-1}$, $\beta_{\text{on}} = 8.9$ $\mu$m$^{-1}$, and according to (3.4) $w_{\text{top}} = 157$ nm and $w_{\text{bot}} = 177$ nm (the first set). At frequency 210 THz $\beta_{\text{off}} = 9.4$ $\mu$m$^{-1}$, $\beta_{\text{on}} = 10.3$ $\mu$m$^{-1}$, $\beta_{\text{on}} =
10.6 $\mu$m$^{-1}$, and correspondently $w_{\text{top}} = 152$ nm and $w_{\text{bot}} = 148$ nm (the second set). We simulate the 6 periods-long systems that correspond approximately to a 2-$\mu$m-long waveguide. The reflection spectra are shown in Fig. 3.10. By applying voltage both reflection maxima grow up in 2-3 times and shift towards higher frequencies. For the first set of parameters there is the enhancement of reflection near 177.5 THz, however, the maximum reflection is shifted towards 200 THz. For the second set one additional peak at 210 THz is clearly seen. The reason for such complex response of reflection on voltage is that waves with propagation constants 10.0 $\mu$m$^{-1}$ and 8.9 $\mu$m$^{-1}$ are excited also on higher frequencies for both branches “on+” and “on−”, as shown in Fig. 3.4. Therefore, the zone of enhanced reflection is extended from 160 THz to 220 THz.

To cross-check the phase modulation performance we calculate reflection of the system by the scattering matrix formalism (see e.g. [107]). For the on-state case we assume a sequence of parts with thicknesses $w_{\text{top}}$ and $w_{\text{bot}}$ and complex propagation constants $\beta_{\text{on+}} + i\alpha_{\text{on+}}$ and $\beta_{\text{on−}} + i\alpha_{\text{on−}}$. The effective mode indices are defined as $n_+ = (\beta_{\text{on+}} + i\alpha_{\text{on+}})/k_0$ and $n_- = (\beta_{\text{on−}} + i\alpha_{\text{on−}})/k_0$, where $k_0 = 2\pi v/c$ is the wave number in vacuum. The whole transmission matrix is a product of transmission matrices for single interfaces and propagation matrices of single layers. The single interface transmission and reflection coefficients in the transmission matrix are defined as $t = 2\sqrt{n_+ n_-}/(n_+ + n_-)$ and $r = (n_+ - n_-)/(n_+ + n_-)$.

Values of $\beta_{\text{on+}}$, $\beta_{\text{on−}}$, $\alpha_{\text{on+}}$ and $\alpha_{\text{on−}}$ are taken as the first mode solutions of the SPP dispersion equation for the four-layer system (represented in Fig. 3.4 and Fig. 3.5). Reflection spectra calculated with the transfer matrix approach are shown in Fig. 3.10 with the legend sign “matrices”. The agreement with the finite element method simulations is both qualitative and quantitative. It validates the approximation of the waveguide as a sequence of parts with different propagation constants. As a result, the 2-$\mu$m system designed according to (3.4) gives higher and broader reflection peaks under applied voltage. Further increase in the number of periods, unfortunately, does not improve the reflection. The reason is that losses in the longer system compensate the enhancing of the Bragg grating reflection.
In principle, such changes in reflection should influence the transmission coefficient. However, actual transmission changes are quite small because of low values of the reflection coefficients. Nevertheless, the Bragg grating configuration with modulation of reflection bears some potential for active elements.

![Reflection coefficients for the Bragg grating systems for enhanced reflection at 177.5 THz (\(w_{\text{top}} = 157 \text{ nm and } w_{\text{bot}} = 177 \text{ nm}\) and at 210 THz (\(w_{\text{top}} = 152 \text{ nm and } w_{\text{bot}} = 148 \text{ nm}\)). Thickness of the stripes is 8 nm and total structure thickness between silver plates is 78 nm. Inset shows schematic view of the structure.](image)

**3.6. Discussions and conclusions**

Current state-of-the-art plasmonic modulators outperform among Si-based electro-optic modulators [50,51]. However, plasmonic modulators suffer very high propagation losses. Losses are lower in MOS waveguide-integrated designs at the expenses of larger dimensions. The structure which contains an active layer sandwiched between two metal plates allows more compact layout.

We have studied optimization options for the plasmonic absorption modulator, which consists of a metal-insulator-metal waveguide with
an ITO controlling layer. The reference system with the continuous ITO film has 2.2 dB/µm extinction ratio and 7.2 dB/µm propagation losses. The main target is to increase transmittance. The proposed improvements are based on changes in the ITO permittivity, which can be realized under different fabrication conditions, and periodic patterning of the ITO layer. First, varying of the ITO layer permittivity (with the fixed doping level and change of the carrier density) increases the modulation depth of the device; specifically the FoM rises from 0.30 to 0.37. In particular, optimized system has 3.2 dB/µm extinction ratio and 8.7 dB/µm propagation losses. Thus, the 3-dB modulation can be achieved with the 1-µm-long and ~0.5-µm-wide device. Second, the patterned device exhibits an improved functionality. We have elaborated a theoretical model to explain the transmission and FoM features. Losses of the modulator with periodic ITO stripes are significantly decreased in comparison with the reference modulator design with the continuous ITO layer. To illustrate this improvement we have calculated the transmission coefficients and shown that they increase from 12% (for the reference case) up to 70% in the voltage-off state of the structure with $L = 2$ µm without any significant drop in the FoM.

Meanwhile we have found some peculiarities in the transmission spectra in dependence on the width of the stripes. For the narrow stripes we explain the pronounced transmission spectra minima by the excitation of nanorod plasmonic resonances. We observe good correlation between simulations and analytic results. The alternative modulation configuration has been considered through utilization of the phase control of the modulator’s mode. By elaborating periodic patterning similar to the Bragg grating we have achieved the 2-3 times grow in the maximum of reflection coefficient under applied voltage.

In perspective we would like to point out that not all the resources for the modulator optimizations have been committed. In this work we utilize results from the Thomas-Fermi screening theory. Accordingly to it the carrier density was changed from $9.25 \cdot 10^{26}$ m$^{-3}$ to $9.7 \cdot 10^{26}$ m$^{-3}$ under the 12 V applied bias [56]. The modulator’s performance can be drastically enhanced if we take into account possible concentration changes reported in other references. For example, results of [66] are based on the concentration increase from $4.19 \cdot 10^{26}$ m$^{-3}$ to $7.08 \cdot 10^{26}$ m$^{-3}$ (1 V bias is mentioned). Moreover, in the recent work [50] the authors analyze the dramatic almost 70-fold change in the
concentration from $0.1 \cdot 10^{26} \text{ m}^{-3}$ to $6.8 \cdot 10^{26} \text{ m}^{-3}$ (up to 4 V in experiments). The final mode index change also strongly depends on the concentration averaging over the ITO layer. We average over the whole 8-nm-thick layer. Half-layer averaging (similar to [66]) or even more fine resolution give higher changes in the permittivity. Another potential direction of optimization is connected with the resonances, e.g. an SPP resonance [56] or Fabri-Perot resonance in finite-length metal-dielectric-metal cavities [12,115], and provides additional advantages in transmission. As a drawback, resonances restrict the operational bandwidth.
Chapter 4. Plasmonic modulator based on gain-assisted metal-semiconductor-metal waveguide

In this chapter, performance of metal-semiconductor-metal (MSM) waveguides with InGaAsP-based active material layers as ultra-compact plasmonic modulators is analyzed. The modulation is achieved by changing the gain of the core that results in different transmittance through the waveguides. A MSM waveguide enables high field localization and therefore high modulation speed. Bulk semiconductor, quantum wells and quantum dots, arranged in either horizontal or vertical layout, are considered as the core of the MSM waveguide. Dependences on the waveguide core size and gain values of various active materials are studied. The designs consider also practical aspects like n- and p-doped layers and barriers in order to obtain results as close to reality. The effective propagation constants in the MSM waveguides are calculated numerically. Their changes in the switching process are considered as a figure of merit. It is shown that a MSM waveguide with electrical current control of the gain incorporates compactness and deep modulation along with a reasonable level of transmittance.

4.1. Gain materials in plasmonic structures

The main obstacle in the implementation of plasmonic nanocomponents is the presence of high losses in connection with highly-confined plasmonic waves propagation. A way to compensate losses is the introduction of gain materials [116-118]. Recently, loss compensation by gain in hybrid dielectric loaded plasmonic waveguides has been studied [31,32]. Garcia-Blanco et. al. [31] analyzed long-range dielectric-loaded surface plasmon-polariton waveguides and loss compensation by the rare-earth-doped double tungstate crystalline material as a gain medium. A hybrid plasmonic waveguide with a nano-slot was examined in [32], where also potential gain media were discussed.

Loss compensation by a gain semiconductor core in MIM waveguides led to the concept of metal-semiconductor-metal (MSM) waveguide. Gain-assisted propagation in a subwavelength MSM waveguide was theoretically analyzed [118]. The thickness of the active core of high-
index semiconductor was varied from 50 nm to 500 nm. For a gold MSM waveguide the required critical gain coefficient for a lossless propagation is achievable at telecommunication wavelengths. More recent systematic theoretical investigation of giant modal gain and amplified propagation in MSM structures with silver plates and the 100-200 nm semiconductor cores is presented in [119]. The optimum gain level, which gives the strongest plasmonic resonance and slowing down of the group velocity is demonstrated.

Encapsulation of InP-based semiconductors in metal and configuring compact MSM waveguides was discussed regarding lasing application. Lasing in metallic-coated nanocavities, namely in the MSM waveguide with rectangular cross-section InP/InGaAs/InP pillars is realized [120-122]. Theoretical analysis of a MSM waveguide with the InGaAs core [123] shows the possibility of realization of a plasmonic semiconductor nano-laser. Semiconductor electrons-holes dynamics in conduction and valence bands is described numerically. The influence of intermediate low index buffer layers in a MSM sandwich is studied as well.

Recently an all-semiconductor active plasmonic system based on InAs heterostructures has been proposed for mid-infrared operation wavelengths [124]. An InGaAsP quantum wells stack was tried to improve properties of negative index materials [125,126]. Whereas the goal of the work was to design the fast optical modulation (tens picoseconds time scale) in the Si spacer layer of such metamaterials, gain produced rather low changes in transmission because of metal screening.

From another point of view, changing the permittivity and thus controlling the propagation in a MIM waveguide is an attractive subject of active plasmonics. A straightforward way is to vary the transmission through the waveguide by changing the permittivity of a sandwiched medium. This idea led to theoretical description of switches using gain-assisted MIM structures [127,128], where the proposed solutions utilize a quantum-dots-doped semiconductor. It is also shown that Fabry-Perot resonances in the active core layer enhance the switching effect.

Another possibility to control propagation of light in a MIM waveguide by changing permittivity of the medium in side-coupled cavities was theoretically analyzed in [127, 129-130]. In particular, it
was shown that implementation of gain media in parts of the waveguide can compensate losses along the device [129].

Despite the fact that photonic switches based on either bulk semiconductors or quantum dots have been studied, there is lacking of comprehensive research on MSM compact devices with the active core from well-defined materials. All proposed MSM structures with loss compensation and switching are analyzed in the assumption of some model gain material uniformly distributed along the dielectric core. Here we bring the subject closer to reality considering InGaAsP-based semiconductors in the MSM core.

4.2. Simulation model

We performed frequency domain simulations using the commercial software package CST Microwave Studio [106]. The material gain $g$ is connected with the imaginary part of the semiconductor permittivity: $\varepsilon'' = -\frac{gn'}{k_0}$, where $k_0$ is the free-space wave-number and $n'$ is the real part of the refractive index. So, for positive gain values the imaginary part of permittivity is negative. Due to the CST constrains, materials with the negative imaginary part of permittivity cannot be implemented directly. To circumvent this constrain and perform CST numerical simulations with gain materials we follow the suggestion from [132]. It consists in editing the permittivity values, in particular changing the sign of $\varepsilon''$, in the Visual Basic script of the history list file. CST simulations were compared with analytical calculations to validate this approach.

Fabry-Perot resonances on a finite-length semiconductor core can significantly increase the effect of the total transmission change in a MSM waveguide (see e.g. [127,128]). However, here we imply that plasmonic waveguides are uniform in the propagation direction and study device characteristics independently from its length.

We are interested in transmission properties of MSM structures at the telecom wavelength 1.55 $\mu$m. A MSM waveguide supports a surface plasmon polariton (SPP) wave in the transverse magnetic (TM) polarization. The silver plates (Fig. 4.1) have the thickness of 120 nm, which is enough to keep the domain-termination error at negligible level. The silver permittivity is approximated by the Drude formula:

$$\varepsilon_{Ag} = \varepsilon_{\infty} - \frac{\omega_{pl}^2}{\omega^2 + i\gamma\omega},$$

where $\varepsilon_{\infty} = 1$, plasma frequency $\omega_{pl} =$
1.38\cdot10^{16} \text{s}^{-1} \text{ and collision frequency } \gamma = 3.22\cdot10^{13} \text{s}^{-1}. \text{ It gives } \varepsilon_{\text{Ag}} = -128.7 + 3.44i \text{ at } \lambda = 1.55 \text{ \mu m}. \\

The high gain core in a MSM device can completely compensate plasmonic losses and cause wave amplification. In case of \text{Im}(k_{\text{eff}})_{\text{on}} \cong 0 \text{ we can receive infinitely high FoM (1.3) that does not have physical meaning. Most of structures studied in this chapter allow complete loss compensation. Therefore, characterizing MSM devices we present results for ER (1.2). Spontaneous emission noise is not included in our model.} \\

The on-state refers to a MSM waveguide, whose core (an InGaAsP-based semiconductor layer) is active and exhibits gain. Such active material can be designed to have a band gap at the wavelength of interest 1.55 \text{ \mu m}. In the off-state the core material exhibits high losses interpreted as negative material gain. Zero-current negative gain is equal to the maximum material gain that can be theoretically achieved in the semiconductor material [107]. \\

We also compared switching to an “insulator-state” of a MSM waveguide, which refers to the waveguide with a passive sandwiched medium. The passive medium means that the medium has a band gap at wavelengths shorter than the wavelength of interest (1.55 \text{ \mu m}). So precisely at 1.55 \text{ \mu m} it is transparent, i.e. it has neither gain nor loss caused by electron-hole recombination. Plasmonic related losses exceed those of the semiconductor core by far. So, in the “insulator-state”, the MSM device has losses caused by the silver electrodes only. \\

Depending on the active region structure, i.e. bulk semiconductor, quantum wells (QWs) or quantum dots (QDs), and carrier density the spectral range of gain of the InGaAsP-based material can vary from tens to hundreds of nanometers around the telecom wavelengths. This characteristic influences the device operation bandwidth. The energy band gap of the gain medium is controlled by the fraction of each component in the In_{x}Ga_{1-x}As_{y}P_{1-y} material, and therefore defines the wavelength of the maximum gain. Material should also be engineered based on operating carrier density. Components with higher bandgap, which is closer to wavelength of interest 1.55 \text{ \mu m}, should be used for lower carrier density and consequently achievable gain (e.g. In_{0.558}Ga_{0.442}As_{0.95}P_{0.05} has bandgap \sim 1.58 \text{ \mu m}). However, we took data from [107] and studied bulk gain material In_{0.53}Ga_{0.47}As that has
bandgap ~1.69 µm and requires a high carrier density to achieve high gain at 1.55 µm.

The infrared refractive index for different In$_x$Ga$_{1-x}$As$_y$P$_{1-y}$ structures is approximately $n' = 3.1+0.46y$ [133] and varies only slightly with respect to the wavelength. Therefore, we neglected the material dispersion.

In order to simulate an as-realistic-as-possible system we included the n- and p-doped layers at the top and bottom of the active layer (see Fig. 4.1b) to prevent the non-equilibrium distributions and recombination of carriers at the contacts. The n- and p-doped layers should have somewhat different composition to obtain the band gap at higher frequency. However, in calculations we fix the real part of permittivity $\varepsilon' = 12.46$ (which corresponds to the real part of refractive index $n' = 3.4...3.5$ for a small gain) for all studied semiconductors apart from permittivity of the InP layer fixed as $\varepsilon_{\text{InP}} = 10$, which corresponds to $n_{\text{InP}} = 3.16$. The imaginary part of effective propagation constant $\text{Im}[k_{\text{eff}}]$ and modal gain $g_m$ in the waveguide are proportional to each other: $g_m = -2\text{Im}[k_{\text{eff}}]$.

### 4.3. MSM waveguide arrangements

Realization of MSM waveguides based devices encounters several problems and can be accomplished in different ways. One way is to deposit metal plates as horizontal layers (Fig. 4.1a) and most of the fabricated devices are based on this procedure (e.g. [22,53]). However, InGaAsP-based heterostructures can be deposited only on a semiconductor native substrate, what means that some additional techniques, e.g. bonding or membrane etching and coating, should be applied to encapsulate semiconductor layers in metal. In such horizontal design n- and p-doped layers should be deposited between metal plates and gain medium while metal plates serves as electrodes. The horizontal layout is preferable for systems when a very narrow gap between metal plates is required. As a MSM waveguide can have width about $w \sim 1$ µm (see Fig. 4.1a), borders influence can be neglected, and the system can be simulated as two dimensional with $w$ taken as infinitely long. However, as the noble metals exhibit poor adhesion to III-V semiconductor materials this possibility may prove to be a challenge to fabricate without using e.g. Ti or Cr adhesion layers. Such lossy layers can significantly affect device operation.
Ultra-compact plasmonic waveguide modulators

![Diagram of MSM waveguides with gain core](image)

**Fig. 4.1.** Cross sections of MSM waveguides with the gain core. a) Basic MSM waveguide. Horizontal arrangement. b) Refined structure of the MSM waveguide which includes the gain core and n- and p-doped layers. Horizontal arrangement. c) Vertical arrangement with top pumping. d) Vertical arrangement with metal plates as electrodes.

The other option is to etch narrow ridges and to coat them with metal. This technique is applied in the case of fabrication of nanolaser structures [121,123]. In such vertical design, the high aspect ratio (that is $H/d$, see Fig. 4.1c) is challenging to achieve in practice, despite the InP-based materials are relatively strong because of their crystalline structure. The high aspect ratio is required due to the mode confinement necessities. Because of finite aspect ratio $H/d$ the mode profile in such waveguide deviates from the two-dimensional case. Two cases can be considered here (Fig. 4.1c and d). The structure shown in Fig. 4.1c with very high aspect ratio was realized for lasing by Hill et. al. [121]: the core thickness was $d = 100...350$ nm, while the whole ridge height $H > 1$ µm. In Fig. 4.1c n- and p-doped layers are horizontal and can be deposited during the growth of the gain medium. In this case, metal plates do not serve as electrodes anymore and passivation layers (e.g. Si$_3$N$_4$ or InP) on both sides of the ridge are required. Optical or electrical pump should be applied in the vertical direction. That assumes large distance of current propagation and can slow down the modulation. Electrically driven fast modulation can be
realized in the design of Fig. 4.1d, which requires n- and p-doped layers depositing after ridge etching. However, it is very challenging for practical realization.

We performed numerical simulations of a MSM waveguide of finite height $H$ and various core thicknesses $d$ (correspond to Fig. 4.1c in the passive state). Fig. 4.2 shows propagation length $L (L = 1/\text{Im}[k_{\text{eff}}])$ for different waveguides in comparison with the ideal case (infinitely high waveguide). The waveguide is sandwiched between high-index material, the same as the core ($\varepsilon = 12.46$); metal corners are rounded for simulation reason with radius of curvature 4 nm. Even for $H = 1 \mu m$ the propagation length halves because of mode spreading outside waveguide (see inset of Fig. 4.2).

**Fig. 4.2.** Propagation length for MSM waveguide of finite height $H$ in passive state. Inset: Electric field distribution for waveguide’s cross section with $H=200$ nm and $d=40$ nm.

### 4.4. MSM waveguide with bulk gain medium

In the first approximation the semiconductor core can be considered as a layer of a uniformly distributed gain material (Fig. 4.1a). For the bulk gain core we set material gain $g_b$ in the range from $-0.6 \cdot 10^4$ cm$^{-1}$ to $0.6 \cdot 10^4$ cm$^{-1}$ for In$_{0.53}$Ga$_{0.47}$As composition at the telecom wavelength (equivalent to $\varepsilon'' = 0.5225...-0.5225$) with carrier density
up to $6.7 \cdot 10^{19}$ cm$^{-3}$ [107]. The core thickness $d$ is varied from 20 nm to 400 nm. The main advantage of the double-sided metal-clad waveguide is the strong field localization between the plates. The inset in Fig. 4.3 shows the effective refractive index of the waveguide, growing from the refractive index of sandwiched material $n' = 3.53$ for a wide core up to 6.7 for a very thin core.

Fig. 4.3 shows Im[$k_{\text{eff}}$] versus the MSM core thickness for various gain values. We consider: the “insulator-state” for the passive semiconductor core, “off-state” for the maximum loss in semiconductor as well as several gain values. To validate the CST simulations with the negative imaginary part of permittivity we compare results with the ones obtained by analytical solution of the SPP dispersion equation for a three-layer system (Fig. 4.3). The results match perfectly.

As was discussed above the system has very high losses in the off-state. With the increase in current, the material gain becomes higher and, at some point, compensates losses. For example, for the gain value $g_{b1} = 0.1 \cdot 10^4$ cm$^{-1}$ ($\epsilon'' = -0.0871$) complete loss compensation in the MSM waveguide (Im[$k_{\text{eff}}$] = 0) can be achieved in the case of a core thickness $d = 60$ nm. Further increase of gain gives light amplification in such system. For gain values more than $g_{b2} = 0.34 \cdot 10^4$ cm$^{-1}$ ($\epsilon'' = -0.2961$) the strong field localization within the gap between metal plates starts to play a crucial role and the MSM design gives an essential benefit. In the case of high gain the Im[$k_{\text{eff}}$] dependency on the core thickness is monotonically increasing, opposite to the case of lower gain (see Fig. 4.3). In other words, for high gain values the thinner the MSM core the longer propagation can be achieved.

In Fig. 4.4 we show Im[$k_{\text{eff}}$] in the MSM waveguide with a bulk gain medium for various carrier densities. For better comparison, we converted the material gain data [107] into the semiconductor permittivity and added in Fig. 4.4 (dotted line). The imaginary part of semiconductor permittivity has nearly linear behaviour on the logarithmic scale as follows from carriers and photons rate equations, thus the imaginary part of the effective propagation constant exhibits the same tendency.

Further we consider an advanced structure, which includes n- and p-doped layers on both sides of the active layer (see Fig. 4.1b). We assume the n- and p-doped layers being 10 nm thick. These
intermediate layers are approximated as insulator layers ($\varepsilon'' = 0$) having the same permittivity as the semiconductor core, $\varepsilon' = 12.46$. The partial removal of the gain material causes an increase in $\text{Im}[k_{\text{eff}}]$. In such MSM waveguides complete loss compensation with gain value $g_{b1} = 0.1 \cdot 10^4 \text{ cm}^{-1}$ can be achieved for core thicknesses more than 100 nm (see Fig. 4.3).

Fig. 4.3. Imaginary part of the effective propagation constant for various thicknesses of the bulk gain semiconductor core. All results are from CST numerical simulations apart from $"g_{b1A}"$ which is the analytical solution of the SPP dispersion relation for a three-layer system. Notation: “No” for passive waveguide (insulator-state); “Off” for switched off current $g_b = -0.6 \cdot 10^4 \text{ cm}^{-1}$; “Off(p)” the same for refined structure with n- and p- doped layers; “$g_{b1}$” on-state with $g_{b1} = 0.1 \cdot 10^4 \text{ cm}^{-1}$; “$g_{b1A}$” for analytical calculation same as “$g_{b1}$”; “$g_{b1}(p)$” on-state with $g_{b1} = 0.1 \cdot 10^4 \text{ cm}^{-1}$ in refined structure with n- and p- doped layers; “$g_{b2}$” on-state with $g_{b2} = 0.34 \cdot 10^4 \text{ cm}^{-1}$; “$g_{b3}$” on-state with $g_{b3} = 0.6 \cdot 10^4 \text{ cm}^{-1}$; “$g_{b3}(p)$” on-state with $g_{b3} = 0.1 \cdot 10^4 \text{ cm}^{-1}$ and refined structure with n- and p- doped layers. Inset: Effective refractive index in the MSM waveguide filled with In$_{0.53}$Ga$_{0.47}$As bulk gain medium.

Calculated ER of both systems versus the core thickness is shown in Fig. 4.5 for various transition states. Introduction of n- and p-doped layers for thin core waveguides decreases ER as expected.
As an additional option we also consider the difference between the on-state and passive state (“insulator-state”) of the MSM waveguide. The “insulator-state”, that is $g = 0$, can be achieved in the studied material for relatively low carrier density, namely $1.25 \cdot 10^{18}$ cm$^{-3}$ [107].

![Graph showing imaginary part of the effective propagation constant in the MSM waveguide with the bulk gain medium for various carrier densities. Imaginary part of semiconductor permittivity is shown for comparison.](image)

Fig. 4.4. Imaginary part of the effective propagation constant in the MSM waveguide with the bulk gain medium for various carrier densities. Imaginary part of semiconductor permittivity is shown for comparison.
Fig. 4.5. ER of MSM device with the bulk gain core. Notation: “\( g_{bi} \)-Ins” shows switch between insulator-state with \( g_b = 0 \) and on-state with \( g_b (g_{b1} = 0.1 \cdot 10^4 \text{ cm}^{-1} \) and \( g_{b3} = 0.6 \cdot 10^4 \text{ cm}^{-1} \)); “\( g_{bi}(p) \)-Ins” the same, but for the refined structure; “\( g_{bi} \)-Off” shows switch between off-state with \( g_b = -0.6 \cdot 10^4 \text{ cm}^{-1} \) and on-state with \( g_b (g_{b1} = 0.1 \cdot 10^4 \text{ cm}^{-1} \) and \( g_{b3} = 0.6 \cdot 10^4 \text{ cm}^{-1} \)); “\( g_{bi}(p) \)-Off(p)” the same, but for the refined structure.

4.5. MSM waveguide with quantum dots

The second realistic option for the core design in MSM structures can be an InGaAsP layer with InAs QDs. Significant progress in QDs devices has been done at 1.3 \( \mu \text{m} \) wavelength [134,135] while efficient emission on 1.55 \( \mu \text{m} \) wavelength is still under investigation [136, 137].

Despite on the small volume, QDs exhibit material gain more than one order of magnitude larger than QWs under the same injection current density because of the strong confinement and quantization of energy levels [138,139]. However, such gain can be obtained only when the electric field is in the QDs plane. Thus, in the horizontal design (Fig. 4.6a) QDs layer gives quite low gain, e.g. material gain \( g_{TM} = 0.4 \cdot 10^3...1.2 \cdot 10^3 \text{ cm}^{-1} \) in conversion of the QDs to an equivalent bulk layer [140]. Such low bulk gain we discussed in the previous
subsection. Therefore depositing of QDs in horizontal design does not provide any additional benefits.

Further we consider the system in Fig. 4.6b. The regular vertical arrangement of QDs appears in Stranski–Krastanov growth [141] as well as grown by metalorganic vapour-phase epitaxy of columnar QDs (vertically aligned, closely stacked QDs) [140].

Parameters for numerical simulations of QDs are the following. A single QD has an ellipsoid shape; the in-plane cross-section is circular with 30 nm diameter, while the shortest ellipsoid axis is 5 nm. QDs are arranged in a regular array with period 50 nm in both in-plane directions. The described geometrical arrangement corresponds to \(4 \cdot 10^{10}\) cm\(^{-2}\) QDs density [136]. An InGaAsP matrix containing one QDs layer has thickness \(s = 10\) nm and \(\varepsilon' = 12.46\), the same as in the previous analysis. We performed simulation for a single QDs layer with periodic boundary conditions in vertical directions, that equivalent to the two-dimensional waveguide with an infinite number of QDs layers. We used \(g_{d1} = 1 \cdot 10^4\) cm\(^{-1}\) and \(g_{d2} = 5 \cdot 10^4\) cm\(^{-1}\).

The QDs volume ratio in the 10-nm thick stack layer is approximately 9\%. For simulations we consider the equivalent system with bulk semiconductor (see Fig. 4.1a) having the gain value defined by the volumetric ratio and gain of the QDs. Thus, the average gain is \(b_{d1} = 0.9 \cdot 10^3\) cm\(^{-1}\) and \(b_{d2} = 4.5 \cdot 10^3\) cm\(^{-1}\).

The total thickness of semiconductor between the metal plates \(d = 50 \cdot N\) nm, where \(N=1\ldots5\) is a number of columns. Up to 5 columns containing QDs were considered for simulations. The imaginary part of the effective propagation constant and ER of MSM devices with QDs are shown in Fig. 4.7 and Fig. 4.8, respectively. Results for QDs and an effective bulk medium are perfectly matched, so the averaging procedure is justified. We assume that the perfect periodic arrangement of QDs in simulations does not influence system propagation properties, which are the same as for a core with randomly distributed QDs. While to achieve \(b_{d2} = 4.5 \cdot 10^3\) cm\(^{-1}\) for bulk material requires a very high current, it is feasible for the QDs effective medium.

For the highest gain \(g_{d2} = 5 \cdot 10^4\) cm\(^{-1}\) in the on-state Im[\(k_{eff}\)] monotonically increasing with the number of columns, which causes decreasing of ER. Average gain \(b_{d2} = 4.5 \cdot 10^3\) cm\(^{-1}\) > \(g_{b2}\), so we observe
the same trend in $\text{Im}[k_{\text{eff}}]$ vs. size dependence as for the bulk material (see Fig. 4.3).

However, as was discussed above, the vertical design has weak mode confinement. In practice, up to 10-20 QDs layers with the same properties can be fabricated. It means that with vertical spacing $s = 10$ nm the total height of the structure cannot exceed $H_A \sim 200$ nm that is not enough for mode localization. So apart from this $H_A \sim 200$ nm of gain region the sandwiched structure should be complemented to $H \sim 1$ µm by another high-index material. Similar case was realized in [121], where MSM waveguide has $H_A = 300$ nm of active gain region, while the whole ridge height $H > 1$ µm. It increases $\text{Im}[k_{\text{eff}}]$ of the whole waveguide, therefore, any partial filling by active material increases propagation losses and suppresses efficiency of modulator.

Fig. 4.6. Cross sections of the MSM waveguide with QDs layers. a) Horizontal design. b) Vertical design.

Another solution is to increase vertical spacing $s$ and obtain relatively uniform distribution of gain medium along waveguide’s cross section. However, in practice, it is hard to obtain vertical alignment at large spacing. Thus if we assume, in order to achieve $H_A \sim 1$ µm, spacing $s = 50$ nm, the average gain will fall down to $b_{d2} = 0.9 \cdot 10^3$ cm$^{-1}$. 
Ultra-compact plasmonic waveguide modulators

Fig. 4.7. Imaginary part of the effective propagation constant in the MSM waveguide with QDs. Notation: “Insulat” is passive waveguide, $g_{d1} = 1 \cdot 10^4 \text{ cm}^{-1}$, $g_{d2} = 5 \cdot 10^4 \text{ cm}^{-1}$, “$b_{d1}$” and “$b_{d2}$” corresponds to the MSM waveguide with uniform distribution of bulk gain material with $b_{d1} = 0.9 \cdot 10^5 \text{ cm}^{-1}$ and $b_{d2} = 4.5 \cdot 10^3 \text{ cm}^{-1}$.

Fig. 4.8. ER of the MSM device with QDs. “On-Ins $g_{d1}$” shows transition between insulator-state ($g=0$) and on-state with $g_{d1}$. “On-Off $g_{d1}$” shows transition between off-state with -$g_{d1}$ and on-state with $g_{d1}$ ($g_{d1} = 1 \cdot 10^4 \text{ cm}^{-1}$ and $g_{d2} = 5 \cdot 10^4 \text{ cm}^{-1}$).
4.6. MSM waveguide with quantum wells

As an alternative design to the MSM waveguide with a bulk semiconductor core we consider the MSM core, which consists of several QWs separated by barriers (Fig. 4.9a). Tensile strained quantum wells are used to ensure a strong interaction with the electrical field of the mode. We took $\text{In}_{0.466}\text{Ga}_{0.534}\text{As}$ QWs and $\text{In}_{0.51}\text{Ga}_{0.49}\text{As}_{0.86}\text{P}_{0.14}$ barriers with thicknesses 5 nm. The composition is chosen such that to obtain the bandgap at 1.55 $\mu$m [142]. The wells are strain compensated which will allow stacking of many wells. However, due to the position of the barrier material in the miscibility gap we decided to limit the number of wells to 25. We performed simulations taking this number into account and defining the total core thickness based on it. In a QWs MSM system it is necessary to have at least one different medium with a higher energy band gap for the n- or p-doped layer. Therefore, we choose 10 nm-thick $\text{In}_{0.67}\text{Ga}_{0.23}\text{As}_{0.71}\text{P}_{0.29}$ layer and 10 nm-thick InP layer. As it was discussed in [123], adding a thin layer with low refractive index in between the waveguide and the metallic plates decreases the net gain in the system. This is due to confining the field outside the gain material, despite of lowering the fields in the metallic layer.

The results of numerical simulations for $\text{Im}[k_{\text{eff}}]$ and ER are shown in Fig. 4.10 and Fig. 4.11, respectively. We studied two gain levels, $g_{w1} = 0.4 \cdot 10^4$ cm$^{-1}$ and $g_{w2} = 1 \cdot 10^4$ cm$^{-1}$, which correspond to $\epsilon'' = -0.348$ and $\epsilon'' = -0.87$ respectively. The former $g_{w1}$ is currently achievable, e.g. in tensile QWs [107,143,144], while the latter $g_{w2}$ is feasible [145]. In such MSM waveguides with QWs complete loss compensation appears either for more than three QWs with gain $g_{w1}$ or already with one QW with gain $g_{w2}$. In contrast to QDs system, for considered $g_{w1}$ and $g_{w2}$ values, ER of the QWs system is growing monotonously with the number of QWs and starting at some thicknesses (approximately $d=90$ nm) can be higher than in the case of the bulk semiconductor with $g_{b1} = 0.1 \cdot 10^4$ cm$^{-1}$.

Vertical design can be considered as well (Fig. 4.9b). The estimation of structure efficiency can be conducted as following. One period of the structure has thickness $s = 10$ nm and averaged gain $b_{w1} = 0.2 \cdot 10^4$ cm$^{-1}$ or $b_{w2} = 0.5 \cdot 10^4$ cm$^{-1}$, such that results for high gain of bulk medium can be applied. However, the total thickness of the gain region will be $H_{A} \sim 250$ nm, as the number of QWs is limited. To ensure mode localization, more than 400 nm thick layers of n- and p-
doped semiconductor should be added as well, which increase $\text{Im}[k_{\text{eff}}]$ and decrease device efficiency. Another option is the same as in the QDs system: the uniform distribution of QWs along the waveguide cross section can be applied by increasing the spacing between QWs.

Fig. 4.9. Cross sections of the MSM waveguide with QWs. The MSM waveguide is composed of $\text{In}_{0.466}\text{Ga}_{0.534}\text{As}$ quantum wells stack, separated by $\text{In}_{0.51}\text{Ga}_{0.49}\text{As}_{0.86}\text{P}_{0.14}$ barriers with n-doped InP layer and p-doped $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ layer. a) Horizontal design. b) Vertical design.
Fig. 4.10. Imaginary part of the effective propagation constant in the MSM waveguide with the QWs core. Notation: “Insulat” is passive waveguide, $g_{w1} = 0.4 \cdot 10^4 \text{ cm}^{-1}$, $g_{w2} = 1 \cdot 10^4 \text{ cm}^{-1}$.

Fig. 4.11. ER of the MSM device with QWs. “On-Ins $g_{di}$” shows transition between insulator-state ($g = 0$) and on-state with $g_{di}$. “On-Off $g_{di}$” shows transition between off-state with $-g_{di}$ and on-state with $g_{di}$ ($g_{w1} = 0.4 \cdot 10^4 \text{ cm}^{-1}$ and $g_{w2} = 1 \cdot 10^4 \text{ cm}^{-1}$).
4.7. Discussion and conclusion

Three types of plasmonic modulators based on metal-semiconductor-metal waveguides with incorporating gain material are considered. The switching principle is based on suppression of losses in plasmonic waveguides by gain material. The semiconductor core is studied in one of three options: bulk semiconductor, QWs and layers with QDs. In the numerical simulations, realistic materials and applied gain parameters achieved in InGaAsP-based structures are studied. Different gain parameters including such that required very high carrier density were taken into account. One of the reasons to do this is the further application of bulk gain medium results for effective approximation of QDs and QWs structures.

We observe promising behaviour in all cases: in principle it is possible not only to compensate losses, but also get waveguide mode amplification. The performance of the proposed modulators is greatly improved in comparison with plasmonic waveguides with controllable losses through electrically driven indium tin oxide layers. In practice, a passive MIM waveguide with \( d = 120 \text{ nm} \) core thickness has \( \text{Im}[k_{\text{eff}}] = 0.043 \mu \text{m}^{-1} \). This means a 15-dB-attenuation on 40 \( \mu \text{m} \) length. Following the fabrication requirements we have considered different layouts of the modulators with horizontal or vertical arrangements of gain layers. We can compare horizontal design with bulk semiconductor core and with QWs layers. A MSM waveguide with gain value \( g_{\text{b1}} = 0.1 \cdot 10^4 \text{ cm}^{-1} \) of bulk material and 10 nm n- and p-doped passive layers gives slight amplification, i.e. \( \text{Im}[k_{\text{eff}}] = 0.007 \mu \text{m}^{-1} \). Therefore, the dynamic range of the signal between passive and gain states in the 40 \( \mu \text{m} \)-long structure is estimated as 17.5 dB, while between absorbing (\( \text{Im}[k_{\text{eff}}] = 0.093 \mu \text{m}^{-1} \)) and gain states is doubled to 35 dB. Assuming the same length for a QWs-based MSM waveguide in the horizontal arrangement (\( g_{\text{w1}} = 0.4 \cdot 10^4 \text{ cm}^{-1} \) and \( \text{Im}[k_{\text{eff}}] = -0.055 \mu \text{m}^{-1} \)) we obtained 34 dB depth of signal modulation between passive and gain states and 52 dB between absorbing and gain states. These results show that the QWs-based waveguides, in the horizontal arrangement are preferable to the bulk based ones. Aside from that, QWs and QDs cores can be compared in vertical arrangement. QWs with \( g_{\text{w1}} = 0.4 \cdot 10^4 \text{ cm}^{-1} \) give effective bulk gain \( b_{\text{w1}} = 0.2 \cdot 10^4 \text{ cm}^{-1} \) of 10 nm thick active layer and QDs with \( g_{\text{d1}} = 1 \cdot 10^4 \text{ cm}^{-1} \) give effectively \( b_{\text{d1}} = 0.09 \cdot 10^4 \text{ cm}^{-1} \) because of small
volume. Thus, both structures have approximately the same efficiency.

The FoM of the InGaAsP-based plasmonic modulator is several times higher than the one for MIM plasmonic modulators with the silicon nitride and indium tin oxide multilayered core [56]. However, utilizing the studied MSM devices for modulation purposes encounters certain obstacles. The deep modulation requires a full switch of the current. Without enhancement of carrier recombination the response time is in the range of nanoseconds [107]. Nevertheless, modulation with speed up to 10 GHz might be feasible, because of the inherent advantages of the MSM design, i.e. the small dimensions, strong coupling to the plasmonic structure, and enhancement of spontaneous emission due to the tight confinement of modes between two metal plates [19,20] that serves as electrodes.

Summarizing, all three designs of the MSM core exhibit effective switching. The preference in choosing the optimal design diverts from the pure numerical analysis of the device performance to the question of feasibility of fabrication such metal-semiconductor sandwiches. In addition, decision on vertical or horizontal arrangement of gain layers depends also on the waveguiding parameters, e.g. polarization and coupling efficiency. The considered structure with electrical control possesses a high performance and can efficiently work as a plasmonic component in nanophotonic architectures.
Chapter 5. Plasmonic finite-thickness metal-semiconductor-metal waveguide as ultra-compact modulator

In this chapter, properties of a finite-thickness metal–semiconductor–metal (F-MSM) waveguide are explored with the point of view to be utilized as an ultra-compact and fast plasmonic modulator. The study of the device was made using both analytical approaches for planar two-dimensional case as well as numerical simulations for finite-width waveguides. The eigenmodes of the F-MSM waveguide, propagation constant, confinement factor, Purcell factor, absorption coefficient, and extinction ratio of the structure are calculated and analyzed. It is shown that using thin metal layers instead of thick ones one can obtain higher extinction ratio of the device.

5.1. F-MSM waveguide modes

Various types of plasmonic waveguides have been proposed and analyzed. Stripe waveguides based on a single thin metal film provide a relatively high propagation length, i.e. low modal losses, but suffer from low field confinement [28]. As a consequence, there was a turn toward hybrid plasmonic waveguides that combine thin metal stripes with semiconductor and dielectric layers of different refractive index [35-40]. Furthermore, these waveguides were improved by adding gain material and thus making possible to mitigate the modal losses [31-34].

From another side, a structure where a dielectric core is sandwiched between two metal layers is very promising as it can provide even smaller dimensions and higher field localization [23,24]. A metal–insulator–metal (MIM) waveguide does not exhibit cutoff even at very small core thickness and as the result allows unprecedented thin layouts of tens of nanometers [41,42]. Planar MIM plasmonic waveguides with thin metal claddings were studied and optimized for various purposes [43-45,146]. They possess several symmetric and asymmetric modes as well as metal-clad and quasi-bound ones, depending on the frequency range [43]. Here we combine both characteristic features and propose a finite-thickness MSM (F-MSM) structure with the active semiconductor core sandwiched between thin metal layers.
As the first step we determine the eigenmodes in a F-MSM waveguide. We solve the SPP dispersion equation for the five-layer system that corresponds to the two-dimensional (2D) F-MSM waveguide shown schematically in Fig. 5.1a. Calculations are performed for the telecommunication wavelength 1.55 µm. The InP-based gain core has different thickness \( d \) and permittivity \( \varepsilon = 12.46 + i\varepsilon'' = (n' + in'')^2 \), where \( n' \) and \( n'' \) are the real and imaginary parts of the refractive index, respectively. The imaginary part of the permittivity can be defined as \( \varepsilon'' = -gn'/k_0 \), where \( g \) is the material gain and \( k_0 \) is the free-space wave-number. We analyzed the system with various metal-layer thicknesses \( t \) assuming a metal permittivity \( \varepsilon_{Ag} = -128.7 + 3.44i \), corresponding to the one of silver at this frequency. The silica layers with permittivity \( \varepsilon_d = 2.34 \) are assumed to extend infinitely in the z-direction (see Fig. 5.1a). Further, we introduce the relative effective index \( n_{eff} = \beta/(n'k_0) \), where \( \beta \) is the SPP propagation constant. The relative effective index is defined as the ratio between the effective mode index and the material refractive index. In our case, material InP has refractive index \( n_{InP}' = 3.53 \).

A F-MSM waveguide with \( d = 30 \) nm and \( t = 5 \) nm supports three modes (Fig. 5.1). Two modes have symmetric electric field inside the core and symmetric/asymmetric fields inside the metal layers (notation Ss and Sa, respectively). The third mode has asymmetric fields inside all layers (Aa mode). The mode with the asymmetric field inside the core and symmetric one in the metal layers (As mode) is not supported by the waveguide with such thin core.

The characteristics of the modes are summarized in Table 5.1. It can be seen that the absorption coefficient \( \alpha \) increases together with increasing of the propagation constant \( \beta \). The field of the Ss mode is mostly located outside the waveguide core, residing mostly in the low loss dielectric surrounding the waveguide. Due to this, the absorption coefficient of the mode is much lower than the ones for Aa and Sa modes. Furthermore, the modes that possess the highest absorption also have the most pronounced response \( \Delta\alpha \) when changing the gain of the semiconductor core. Characteristic value \( q \) shows at which distance the field amplitude drops \( e \) times and can be found as \( q = 1/\text{Im}(k_z) \), where \( k_z = (\varepsilon_dk_0^2-(\beta+i\alpha)^2)^{1/2} \) is the z-component of the propagation vector.

The field confinement factor \( \Gamma \) can be defined with respect to either field intensity \( E^2 \) or Poynting vector component \( P_x \) along the
propagation direction. Also, based on how the waveguide is defined, it can be calculated as the ratio of the electric field energy (flux) inside the core (notation “core”) or electric field energy (flux) confined inside the core and both metal plates (notation “core + 2Me”) to the total amount of electric field energy (flux) The confinement of Ss and Aa modes is very low due to the field extension outside the waveguide. It is also fully consistent with rather weak response $\Delta\alpha$ and large characteristic distance $q$.

Fig. 5.1. (a) Finite-thickness MSM waveguide. (b), (c), (d) $E_z$ field profile in the waveguide with $d = 30$ nm and $t = 5$ nm for Ss, Aa and Sa modes, respectively. The field is normalized such that $E_z = 1$ a.u. at the interface between gain core and metal layer.

The highest confinement of the Sa mode leads to the largest response $\Delta\alpha$. So, our further analysis of the F-MSM waveguide as plasmonic modulator is performed only for the Sa mode, the one showing the best ability to be controlled.
Table 5.1. Characteristics of different modes in the F-MSM waveguide with \( d = 30 \text{ nm} \) and \( t = 5 \text{ nm} \)

<table>
<thead>
<tr>
<th>Type</th>
<th>Ss</th>
<th>Aa</th>
<th>Sa</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta ), ( \mu \text{m}^{-1} )</td>
<td>6.22</td>
<td>8.11</td>
<td>42.88</td>
</tr>
<tr>
<td>( \alpha ), ( \mu \text{m}^{-1} )</td>
<td>1.8x10^{-4}</td>
<td>0.127</td>
<td>0.59</td>
</tr>
<tr>
<td>( \Delta \alpha = \alpha_{\text{off}} - \alpha_{\text{on}}, \mu \text{m}^{-1} ) ( (g_{\text{on}} = 2000 \text{ cm}^{-1}) )</td>
<td>1.16 x10^{-4}</td>
<td>0.0198</td>
<td>0.32</td>
</tr>
<tr>
<td>( q ), ( \mu \text{m} )</td>
<td>1.86</td>
<td>0.19</td>
<td>0.024</td>
</tr>
<tr>
<td>( \Gamma (E^2, \text{core}) )</td>
<td>5.0 x10^{-4}</td>
<td>0.044</td>
<td>0.65</td>
</tr>
<tr>
<td>( \Gamma (E^2, \text{core} + 2\text{Me}) )</td>
<td>5.4 x10^{-4}</td>
<td>0.058</td>
<td>0.72</td>
</tr>
<tr>
<td>( \Gamma (P_x, \text{core}) )</td>
<td>2.5 x10^{-3}</td>
<td>0.0017</td>
<td>0.95</td>
</tr>
<tr>
<td>( \Gamma (P_x, \text{core} + 2\text{Me}) )</td>
<td>2.6 x10^{-3}</td>
<td>0.0020</td>
<td>0.96</td>
</tr>
</tbody>
</table>

5.2. Field confinement in the F-MSM waveguide

One possibility to increase the effective index is to reduce the metal thickness \( t \). We studied properties of waveguides taking several thicknesses of the core \( (d = 30 \ldots 200 \text{ nm}) \) and metal films. In the latter case we investigated two routes. One is changing the thickness of both metal films simultaneously. Another one is to fix the thickness of one film \( (t_1 = 70 \text{ nm}) \) and vary only the thickness of the second film.

The structures with relatively thick metal layers possess modal properties similar to a 3-layered infinite MSM waveguide: the relative effective index increases with decreasing of \( d \) and almost does not depend on the metal thickness (Fig. 5.2a). However, for thin metal layers the relative effective index shows a strong dependence on the metal thickness. Relative effective index \( n_{\text{eff}} \) increases several times for small metal thicknesses \( t < 15 \text{ nm} \).

In some calculations we considered very thin metal layers, in particular down to 5 nm. At such small thickness the metal properties can be different from the bulk ones. Moreover, non-local and quantum effects start to play a role. Nevertheless, we believe that our approximation shows the general qualitative tendency even if the actual numbers may be quantitatively different.
Fig. 5.2. Relative effective indices (a) and Purcell factor (b) of F-MSM waveguides significantly increase for thin metal layers. “S” corresponds to the case when thickness of both metal films is varied simultaneously (symmetric structure); “A” corresponds to the case when thickness of one metal film is fixed ($t_1 = 70$ nm) and another varied (asymmetric structure, shown on the inset of (b)).

The trend can be explained in the following way. The F-MSM waveguide with thin metal layers can be considered as two coupled metal stripe waveguides. Each metal strip supports an asymmetric mode often referred to as the short-range SPP (see Fig. 5.1d). This
mode has higher effective index and absorption coefficient than the long-range SPP and its effective index dramatically increases with the decrease of the metal thickness [28,147]. Therefore, the effective index of two coupled stripe waveguides is substantially increased (see Fig. 5.2a).

In the case of an active MSM waveguide, we can define the PF to show the increase of the radiative rate. It was shown that MIM and MSM waveguides with active materials possess quite high PF [19-21,148]. Moreover, the PF scales roughly as the fifth power of the relative effective index \( n_{\text{eff}}^5 \) [148]. We calculated the PF of the F-MSM waveguide following the approach and approximations of [148]. In this case the PF is defined by waveguide properties, in particular, field profile and dispersive curve. Results show that PFs above one hundred can be achieved with sufficiently thin metal layers (Fig. 5.2b).

While the relative effective index of the Sa mode increases with decreasing the metal thickness, the confinement factor in a F-MSM waveguide has the opposite tendency. Both confinement factors defined via \( E^2 \) and \( P_x \) either in the waveguide core only or in the waveguide core and two metal plates are decreasing (Fig. 5.3a). Such behavior is abnormal for a conventional optical waveguide, where the field confined in a higher index structure causes higher relative effective index. To solve the contradiction we calculated the confinement of \( P_x \) and \( E^2 \) only in two metal plates (Fig. 5.3b). \( \Gamma(P_x) \) in two metal plates is also decreasing. However, \( \Gamma(E^2) \) is increased and it agrees with the increase of the relative effective index. Nevertheless, the magnitude of these changes is much lower than magnitude of changes in waveguide’s core and surrounding SiO₂.

Fig. 5.4 shows a comparison of \( P_x \) and \( E^2 \) profiles for two thicknesses of metal layers. As the thickness decreases, for each metal stripe, the field is localized closer to the metal surfaces. This, in turn, generates less coupling of the two metal stripes and, therefore, less field is confined inside the waveguide core. These changes in the field distribution inside the waveguide core are more pronounced than at the interface and as a result they have more influence on the confinement factor.
5.3. Absorption coefficient and modulator’s performance

Although the increase of the effective index is a desired effect, doing it by thinning the core and metal claddings has a drawback. The fields penetrate more into metal and cause an increase of losses in the system. As a consequence, the absorption coefficient of the F-MSM waveguide with thinner metal layers is higher (Fig. 5.5a). Therefore, loss compensation requires higher gain. We calculated the critical gain needed for the full loss compensation in F-MSM waveguides with thin metal layers. In Fig. 5.5b we plot the relative effective index values that can be achieved with certain gain levels and core thicknesses. As it can be seen, thinner cores allow higher effective indices. The required material gain is feasible for InP-based semiconductors utilizing either bulk material properties or a stack of quantum wells [149].

We considered that the active core possesses gain in the on-state and absorption in off-state with the same absolute value of $\varepsilon$". We performed calculations of the $\text{ER} = 8.68 \cdot (\alpha - \alpha_{\text{on}})$ of the F-MSM waveguide for $d_1 = 30$ nm and $d_2 = 50$ nm, and various thicknesses of the metal layers. Calculations for the uniform gain core (Fig. 5.1a) show that for the same amount of gain, the thinner active core exhibits better performance (Fig. 5.6). Thus, $\text{ER} = 5 \text{ dB/\mu m}$ and higher can be achieved. The results for symmetric (Fig. 5.6a) and asymmetric (Fig. 5.6b) structures are similar. It shows that at least one thin metal film is needed to achieve high performance.
Fig. 5.3. (a) Confinement of electric field ($E^2$) and Poynting vector ($P_z$) for symmetric waveguides with core $d = 30$ and 50 nm inside the core (C) or inside core and two metal plates (C+2Me). (b) Confinement for the same waveguides but inside only two metal plates (2Me).
Fig. 5.4. (a) Poynting vector $P_x$, along the propagation direction, and (b) $E^2$ field profile in the symmetric waveguides with $d = 30$ nm and different metal thicknesses: $t = 5$ nm and 10 nm. All profiles are normalized such that the total integrals over whole cross-section are the equal.
Fig. 5.5. (a) Absorption coefficients of F-MSM waveguides (symmetric (S) and asymmetric (A) structures). It significantly increases for thin metal layers along with the increase of the relative refractive index (see Fig. 5.2a). (b) Relative effective index that can be achieved in a symmetric waveguide with the provided level of the critical gain (metal losses are fully compensated). The metal layer thickness is varied along each line.
Fig. 5.6. Extinction ratio of F-MSM waveguide ($d_1 = 30$ nm and $d_2 = 50$ nm) for different gain values of the InP-based active core: $g_1 = 1000$ cm$^{-1}$, $g_2 = 2000$ cm$^{-1}$, and $g_3 = 3000$ cm$^{-1}$. (a) Thickness of both metal films is varied simultaneously. (b) Thickness of one metal film is fixed ($t_1 = 70$ nm) and another is varied. Insets: FoMs of two cases respectively.
In contrast to ER, figure of merit ($\text{FoM} = (\alpha - \alpha_{\text{on}})/\alpha_{\text{on}}$) increases with increasing of thickness of metal films (Fig. 5.6, insets). It is due to the significant decrease of $\alpha_{\text{on}}$ for thicker metal layers. The FoM of this MSM device is several times higher than for MIM plasmonic modulators with the silicon nitride and indium tin oxide multilayered core [56].

5.4. Influence of n- and p-doped layers

To analyze a realistic structure in case of electrical pump, we consider a F-MSM waveguide with n- and p-doped layers (Fig. 5.7a). These layers are required to inject carriers into the gain core. In calculations we consider 10-nm-thick layers with permittivity $\varepsilon = 12.46$. In our approximation, n- and p-doped layers do not possess any gain or loss. We study the same parameters of the waveguide as for the uniform waveguide core.

However, in case of the F-MSM waveguide with the n- and p-doped layers, the sample with total core thickness $d_1 = 30$ nm gives lower performance than waveguide with $d_2 = 50$ nm (Fig. 5.7b). In this case, the extinction ratio is lower, $\text{ER} = 1-2$ dB/$\mu$m. According to the calculated field distribution (see Fig. 5.4b), in the core of the waveguide the field is strongest next to the metal interface, and so active properties of the materials are the most pronounced there. In the same time, the doped layers are placed in that part and thus limit the effect of the gain material.
Fig. 5.7. (a) F-MSM waveguide with n- and p-doped layers and metal films of equal thicknesses \( t \). Core thickness \( d \) includes the gain part and both doped layers. (b) Extinction ratio of the waveguide \((d_1 = 30 \text{ nm} \text{ and } d_2 = 50 \text{ nm})\) for different gain of the InP-based active core: \( g_1 = 1000 \text{ cm}^{-1} \), \( g_2 = 2000 \text{ cm}^{-1} \), and \( g_3 = 3000 \text{ cm}^{-1} \). Calculations are performed for 10-nm-thick doped layers.

Further, we analyze a waveguide with a 5-nm-thick metal layer and varied core thickness \( d \). Absorption coefficients show a non-monotonic dependence with a minimum around \( d \approx 100 \text{ nm} \) (Fig. 5.8a). However, the ER of the system is monotonically rising with increasing of the waveguides core, i.e. the thicker the core the larger is ER (Fig. 5.8b). In other words, taking into account the n- and p-doped layers thicker waveguides are preferable. These results are consistent with ones obtained for the MSM waveguide with infinitely thick metal layers.
Fig. 5.8. (a) Absorption coefficient and (b) extinction ratio of F-MSM symmetric waveguide with n- and p-doped layers at various gain values: $g_0 = 0$, $g_1 = 1000$ cm$^{-1}$, $g_2 = 2000$ cm$^{-1}$, and $g_3 = 3000$ cm$^{-1}$. Thickness of metal layers is fixed, $t = 5$ nm. Calculations are performed for 10-nm-thick doped layers.
5.5. Finite-length F-MSM waveguide

To study wave propagation in such F-MSM waveguide, we performed frequency-domain simulations using the commercial software CST Microwave Studio [106]. Effectively 2D system layout is presented in Fig. 5.9a. It consists of an active and two identical non-active parts. A mode is excited at the left port and sequentially propagates through the first non-active part, active part and the second non-active part. As far as the wave propagates it is attenuated and reflected at the interfaces between the active and non-active parts. We calculated the dependence of the reflection and attenuation values with the length of the active part $L_a$ (Fig. 5.9b,c). We performed calculations for two systems. The first one is a waveguide where both active and the non-active parts of the core have $\text{Re}(\varepsilon) = 12.46$ (referred as “U”-waveguide in Fig. 5.9b,c), thus providing low reflectivity at the interface (see Fig. 5.9b,c). In the second, the waveguide core material of the non-active part differs from the one of the active part (referred as “C”-waveguide). In this case we considered the permittivity of the non-active part as $\varepsilon_{na} = 4$.

We also employ the transfer-matrix approach [107] to find an analytic solution for the mode propagation reflection and attenuation values. Complex propagation constants in the active ($\beta_a + i\alpha_a$) and non-active ($\beta_{na} + i\alpha_{na}$) parts of the waveguide are derived from dispersion relations and further substituted in expressions for propagation in layers with different thicknesses. The effective mode indices are defined as $n_a=(\beta_a+i\alpha_a)/k_0$ and $n_{na}=(\beta_{na}+i\alpha_{na})/k_0$, where $k_0 = \frac{2\pi\nu}{c}$ is the wave number in vacuum. The whole transmission matrix is a product of transmission matrices for single interfaces and propagation matrices of single layers. We consider the single interface transmission and reflection coefficients in the transmission matrix defined using the classical formulae where the refractive indices are replaced with the effective mode ones: $t = 2\sqrt{n_a n_{na}}/(n_a + n_{na})$ and $r = (n_a - n_{na})/(n_a + n_{na})$.

Numerical and analytical results coincide well (see Fig. 5.9b,c). This agreement leads to the conclusion that the analytic transfer-matrix approach using effective mode indices can be applied for plasmonic waveguide systems. In the same time, it shows that no other waves are
excited at the waveguide interface so the symmetry of the Sa mode is preserved.

Fig. 5.9. (a) Instantaneous electric field distribution in a complex F-MSM waveguide with active and two non-active parts. Thicknesses of metal layers is $t = 5$ nm and core $d = 30$ nm. Logarithmic attenuations (b) and reflection coefficients (c) in different structures. Notations: “A” stands for results obtained as analytic solutions based on the transfer-matrix approach; “N” stands for numerical simulations; “C” for complex waveguides in contrast to uniform ones “U”; “off” for the core with absorption $\varepsilon'' = 0.174$; “on” for the active core with $\varepsilon'' = -0.174$ (corresponds to gain $g_2 = 2000$ cm$^{-1}$); “no” for the passive core $\varepsilon'' = 0$ (neither gain, no loss).

5.6. Finite-width F-MSM waveguide: 3D case

To compare 2D approximations with a more realistic structure we analyzed a three-dimensional (3D) F-MSM waveguide of finite width $w$ in the y-direction (see Fig. 5.1a). We perform numerical simulations for various widths from 20 nm to 3 $\mu$m studying two cases: the first layout with the sharp corners and the second one with the corners rounded with 2.5-nm-radius curvature. We considered on-, off- and insulator-state (Fig. 5.10). Insulator-state corresponds to core without gain ($g = 0$). For relatively small widths ($w \leq 500$ nm) both propagation constant and attenuation coefficient differ significantly
from results obtained in the 2D approximation (Fig. 5.10). The trend shows that the propagation constant can be increased but with the price of an increase in losses. Confinement of the mode in two dimensions causes even higher field localization and thus increasing losses (Fig. 5.11a).

![Graph](image)

**Fig. 5.10.** Characteristics of the F-MSM waveguide of finite width $w$. The thickness of metal layers is $t = 5$ nm and core $d = 30$ nm. (a) Propagation constant and (b) signal attenuation in the F-MSM waveguide for various waveguide widths.

With the increase of the width, the parameters tend to those of the 2D approximation with the infinite width. However, it can be noticed from Fig. 5.10 that the calculated values do not converge exactly to 2D ones. Similar results can be found in [56] for other configurations of MSM waveguides. Moreover, the effect was studied experimentally for a metal stripe waveguide and was ascribed to contribution of
radiation damping due to field scattering on edges [95]. Indeed, at the waveguide edges the fields are significantly different, even for wide waveguides, from the ones in the middle of the waveguide that can be considered similar to the ones of a 2D waveguide (see Fig. 5.11b for a waveguide width $w = 3$ µm). Nevertheless, the quantitative difference between the 2D and 3D results with $w \geq 500$ nm is not large: 0.7% for propagation constant and 1.5%, 3% and 7% for attenuation coefficients in off-, insulator-, and on-states, respectively. We conclude that 2D analysis can be applied for F-MSM waveguides with widths bigger than 500 nm.

Difference in sharp and rounded corners is noticeable for small $w$, but hardly visible with large $w$. As the result, for large $w$, the rounding of the corners does not influence the waveguide’s parameters so does not increase the convergence to 2D values.

Fig. 5.11. Mapping of instantaneous electric field distribution in the cross-section of the F-MSM waveguide with $w = 20$ nm (a) and 3 µm (b).

5.7. Other practical issues

The fabrication of InP-based semiconductor devices brings several challenges. In particular, crystalline structure can be grown only on another InP-based material. Thus additional techniques, e.g. flip-chip and bonding, to place the active layer between the metal layers are needed. However, namely the InP-based devices provide the highest gain and consequently have the highest potential in compensating metal losses. It should be also mentioned that the considered structure is completely planar and, as the result, it does not contain any features inside waveguide, which can significantly complicate the fabrication process.

The structure can be controlled both electrically and optically. The electrical pumping can be realized using the metallic plates as electrodes, thus one can say that it is intrinsically built-in. However, in
this case we need the n- and p- layers and as the result the device’s performance is decreased. Furthermore, an additional speed limitation comes from the RC-circuit set up by the structure. In contrast, there is no need of doped layers and no frame of RC-circuits in the layout with the optical pumping, as studied materials form Schottky junction on the interface metal-semiconductor. Thus, more efficient and faster modulation can be realized. However, the optical pumping scheme requires a more sophisticated layout to get access to a source and deliver the modulating signal.

Including additional layers or designing a periodic structure could enhance the performance of the system even more (see for example [150]). In this case, instead of InP based semiconductors alternative gain materials, like colloidal quantum dots or dye solutions, may be used. In contrast to crystalline semiconductors, colloidal quantum dots and dye solutions can be deposited layer by layer and, thus, arranged in multilayered structures.

Ultra-compact and ultra-fast modulators are among the main requirements for modern photonic integrated circuits. A surface plasmon polariton modulator supplemented with loss-compensation material provides such possibilities. Their potential applications range from direct laser modulation to on-chip optical routing and computation. For this, various in- and out-coupling schemes are needed. While for out-coupling to free space the solution can be straightforward the in-coupling from a VCSEL device is not necessary an easy task. A general in- and out-coupling scheme is difficult to suggest but we believe that, together with the development of plasmonic circuitry in photonics, the coupling problem will be solved.

Due to the strong field localization, MSM structures are featured as ultra-compact devices and show advances to utilize ultra-thin gain nanostructures embedded in metal layers. For instance, an ultimate size semiconductor nanolaser has been recently designed based on a single quantum well sandwiched between thin silver layers [151].

### 5.8. Discussion and conclusion

We have showed that utilizing thin finite metal films in a MSM waveguide can increase the relative effective index in several times compared to a MSM waveguide with thick layers. The relative effective index increase results also in increase of the Purcell factor and thus enhancing the spontaneous emission rate. The main control
mechanism is to modulate stimulated emission of semiconductor. Therefore, from one side, the higher speed of spontaneous emission the faster stimulated emission can be switched off. Thus, achieved Purcell factor (above 100) may lead to orders of magnitude increase of the modulation speed.

On the other side, a higher Purcell factor leads to increased spontaneous emission thus extra noise in the device. Nevertheless, for device’s analysis we can neglect the increase in the noise level and consider that the signal is significantly higher. This is due to two factors. First, in usual photonic devices, noise is accumulated along the waveguide (tens or even hundreds of microns) and results in undesirable effect. However, the active part of the considered plasmonic modulator is only several microns long and several tens of nanometer thick, and does not allow rising of the noise level. Second, noise is subject to plasmonic losses in the waveguide and thus it is greatly diminished.

The high Purcell factor significantly increases the current density that is required for achieving loss compensation in the waveguide. It was pointed out as the main obstacle for utilizing the MSM structures [148]. Indeed, very high current intensively heats the structure. However, again, because of small dimensions of the structure and its crystalline nature, efficient outward heat transfer can be implemented, for example, with the help of thermoconducting materials as the waveguide claddings.

The main advantage with respect to the silicon based modulators [152,153] is the footprint of our device. In particular, it can show 10-dB modulation depth with a footprint of about $2 \times 0.5 \ \mu\text{m}^2$ (length × width). We have not yet investigated the bandwidth characteristics but the modulation concept is non-resonant, so the operating bandwidth can be high and such investigation is the part of my future work. Modulation with speeds up to 500 GHz is feasible, as tight mode confinement between metal plates increase the rate of active material response [149,154].

In summary, we proposed to exploit an MSM waveguide with finite-thickness metal layers for modulation of plasmonic waves. In this work we were interested only in the symmetric low energy dispersive branch as it gives the highest field confinement, propagation constant and absorption coefficient. The absorption coefficient is even higher than the ones of the MSM waveguide with thick metal layers.
However, particularly this property provides higher extinction ratio of the modulator based on the F-MSM waveguide. The thinner the metal layers the higher the extinction ratio, namely 3-8 dB/µm with a feasible gain level. We believe that, using this property, the plasmonic F-MSM structures can outperform earlier proposed layouts.
Chapter 6. Plasmonic modulator using metal-insulator-metal waveguide with barium titanate core

Barium titanate (BaTiO₃) was shown to provide high performance for photonic thin film modulators [155-158]. Here we propose implementation of BaTiO₃ as the active core of a MIM waveguide. Under applied voltage, some part of ferroelectric domains can be switched from the in-plane (with an ordinary refractive index $n_o$) to out-of-plane orientation (with extraordinary index $n_e$) [159-160]. Thus, the refractive index for a field polarized along one axis can be changed, and control of propagating signal is achieved. Different voltage provides different degree of domain switching, and thus the required level of modulation can be realized. The similar option for electrooptic modulation was studied in plasmonic interferometers [15].

Schematic view of the MIM waveguide with BaTiO₃ is shown in Fig. 6.1. We compare three metals: silver, gold, and aluminum. We are interested in modulation on the telecom wavelength only, $\lambda_0 = 1.55 \mu m$, so metal parameters are fixed: $\varepsilon_{Ag} = -129 + 3.3i$ [104], $\varepsilon_{Au} = -115 + 11.3i$ [104], and $\varepsilon_{Al} = -240 + 49i$ [161]. Silver and gold conventionally show the best plasmonic performance. However, they are not compatible with CMOS fabrication facilities. Aluminum possesses higher losses, but it is CMOS-compatible and can be adopted in manufacturing.

BaTiO₃ is highly birefringent with refractive index difference $\Delta n = 0.05$. Thus, we consider $n_o = 2.35$ and $n_e = 2.3$ [158]. We solve the dispersion equation of a three-layer structure with $\varepsilon_{xx} = n_o^2$ and $\varepsilon_{zz} = n_e^2$ tensor components of the permittivity of the core (see for example [162]). It corresponds to the device off-state. Under applied voltage, domains are switched and both tensor components become equal $\varepsilon_{xx} = \varepsilon_{zz} = n_e^2$. 
We solved the dispersion equation for different core thicknesses $d = 50\ldots500$ nm. Figs. 6.2 and 6.3 show changes of propagation constant $\beta$ and absorption coefficient $\alpha$ within the switch. The extinction coefficient of BaTiO$_3$ is negligible, and consequently, absorption coefficients for on- and off-states are almost the same (see Fig. 6.3). In contrast, the variations of $\beta$ are highly pronounced: $\beta$ varies in the range $9.5\ldots11.5$ $\mu$m$^{-1}$ with corresponding effective index $n_{\text{eff}} = 2.34\ldots2.84$. Thus such mode is suitable for the signal phase control.

Fig. 6.1. Metal-insulator-metal waveguide as plasmonic modulator.

Fig. 6.2. Propagation constant for MIM waveguides with different metals.
We calculated the length of the waveguide required to achieve the $\pi$ phase change between the on- and off-states:

$$L_\pi = \frac{\pi}{(\beta_{\text{off}} - \beta_{\text{on}})}.$$  \hspace{1cm} (6.1)

It has values around 13…15 $\mu$m in the broad range of $d$ (Fig. 6.4). The propagation losses are 0.05…0.2 dB/$\mu$m. Thus, with such short device length $L_\pi$ the modulator has relatively high transmittance:

$$T = \exp(-\alpha L_\pi).$$  \hspace{1cm} (6.2)

Aluminum has the highest absolute value of real and imaginary parts, therefore it possesses the lowest $\beta$ and highest $\alpha$ among three metals under study. Thus, the Al modulator has a bit higher $L_\pi$ and much lower transmittance.

$L_\pi$ decreases with the decrease of core thickness $d$ (see Fig. 6.4). However, transmittance $T$ drops as well. In terms of transmittance, the modulator with silver claddings significantly outperforms others, while gold and aluminum claddings show approximately the same values for $d \geq 300$ nm (see Fig. 6.5).
Fig. 6.4. Device length $L_\pi$ which is required to achieve $\pi$ phase difference with two states.

Mach-Zehnder-interferometer-based modulators with the silicon core were theoretically studied and experimentally realized in the horizontal slot waveguide configuration [67-69]. These devices demonstrate 3-dB modulation depth having the typical length of several microns only. However, transmittance through the structures is relatively low.

Fig. 6.5. Transmittance of the $L_\pi$-long device.
Thus, adopting BaTiO$_3$ an ultra-compact plasmonic switch can be realized. The signal phase control can be implemented via the Mach-Zehnder interferometer layout with the device length down to 12 µm. Mode characteristics do not have pronounced variations with different core thicknesses. The operation bandwidth is large as the BaTiO$_3$ refractive index only slightly varies with the wavelength, and the effect of mode index change is essentially non-resonant.

In summary, we studied characteristics of a MIM plasmonic waveguide with the BaTiO$_3$ core aiming to utilize it as an effective plasmonic modulator. To achieve the 3-dB modulation depth a 12…15-µm-long device is required. Such arm’s length results in -2…-4 dB transmittance. These parameters indicate the high potential of such devices.
Chapter 7. Summary

Recently plasmonics has grown in a highly developed and advanced field with a lot of interesting effects and useful applications. It allows combining nanoscales of devices and sensors with very fast response, so extending their functionality over the whole optical range. One of the central problems in plasmonics is localization and transfer of optical energy on a nanoscale. Advanced techniques to manipulate SPPs as information carriers for integrated circuits are been developing.

Pioneering works in the search for new plasmonic materials have suggested new intermediate carrier density materials such as transparent conducting oxides and transition metal nitrides as promising plasmonic building blocks with low loss, extraordinary tuning and modulation capabilities, and compatibility with standard semiconductor technology.

The concept of utilizing transparent conducting oxides as an active material with varied carrier density has been studied. TCO’s properties depend strongly on fabrication conditions such as the doping level, annealing environment and temperature. It allows to fabricate material with the adjusted permittivity and hence to design devices with the most efficient characteristics.

Different types of plasmonic modulators have been studied during last few years. A plasmonic waveguide based on an active core sandwiched between thick metal plates has promising characteristics. Apart from the speed and dimensions advantages, the metal plates can serve as electrodes for electrical pumping of the active material making it easier to integrate. Including an additional thin layer in a MIM waveguide, e.g. silicon oxide, vanadium oxide or transparent conductive oxide, allows the control of the dispersive properties of the waveguide and thus increasing the efficiency of a plasmonic modulator.

The carrier density in the transparent conducting oxide layer changes under modulating voltage, in turn it changes the effective plasma frequency of the layer and thus a significant change of effective mode index is achieved in a multilayered stack.

Plasmonic modulators with a gain material to be implemented as ultra-compact and ultra-fast active nanodevices in photonic integrated
circuits are investigated. The InP-based semiconductor core allows electrical control of signal propagation. By pumping the core we can vary the gain level and thus the transmittance of the whole system. Bulk semiconductor medium, quantum wells and quantum dots as active material were studied for different layouts of the modulator. The analysis shows that an extinction ratio of several decibels per micron can be achieved, depending on the gain level.

The suggested plasmonic modulator can be part of a metal-insulator-metal waveguide or coupled to another type of plasmonic waveguide with high mode localization. There are various possibilities to couple metal-insulator-metal waveguides to conventional photonic waveguides [85,163-164], and similar concepts can be utilized for plasmonic waveguide modulator structures. A detailed study of various couples is one of the targets of my future work.
Bibliography


integrated plasmonic devices based on Semiconductor-Insulator-Metal strip waveguides,” Optics Express 18, 18945-18959 (2010).


106. CST Microwave Studio, http://www.cst.com/


121. Martin T. Hill, Milan Marell, Eunice S. P. Leong, Barry Smalbrugge, Youcai Zhu, Minghua Sun, Peter J. van Veldhoven, Erik Jan Geluk, Fouad


architecture for WDM applications,” Optics Express 20, 27420-27428 (2012).


