Optical Material Characterization Using Microdisk Cavities

Thesis by
Christopher P. Michael

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

2009
(Defended May 18, 2009)
Basic research is like shooting an arrow in the air and, where it lands, painting a target.

— Dr. H. B. Adkins, American chemist
Acknowledgments

For the reasons that you may know and those you may not . . .

Thank you, Oskar.


My gratitude goes to the Gordon and Betty Moore Foundation, the National Science Foundation, and the Bill and Melinda Gates Foundation.

Thank you, Mom, Dad, and Jessie.

Thank you for everything, Sarah.
Abstract

Since Jack Kilby recorded his “Monolithic Idea” for integrated circuits in 1958, microelectronics companies have invested billions of dollars in developing the silicon material system to increase performance and reduce cost. For decades, the industry has made Moore’s Law, concerning cost and transistor density, a self-fulfilling prophecy by integrating technical and material requirements vertically down their supply chains and horizontally across competitors in the market. At recent technology nodes, the unacceptable scaling behavior of copper interconnects has become a major design constraint by increasing latency and power consumption—more than 50% of the power consumed by high speed processors is dissipated by intrachip communications. Optical networks at the chip scale are a potential low-power high-bandwidth replacement for conventional global interconnects, but the lack of efficient on-chip optical sources has remained an outstanding problem despite significant advances in silicon optoelectronics. Many material systems are being researched, but there is no ideal candidate even though the established infrastructure strongly favors a CMOS-compatible solution.

This thesis focuses on assessing the optical properties of materials using microdisk cavities with the intention to advance processing techniques and materials relevant to silicon photonics. Low-loss microdisk resonators are chosen because of their simplicity and long optical path lengths. A localized photonic probe is developed and characterized that employs a tapered optical-fiber waveguide, and it is utilized in practical demonstrations to test tightly arranged devices and to help prototype new fabrication methods. A case study in Al$_x$Ga$_{1-x}$As illustrates how the optical scattering and absorption losses can be obtained from the cavity-waveguide transmission. Finally, single-crystal Er$_2$O$_3$ epitaxially grown on silicon is analyzed in detail as a potential CMOS-compatible gain medium due to its high Er$^{3+}$ density and the control offered by the precise epitaxy. The growth and fabrication methods are discussed. Spectral measurements at cryogenic and room temperatures
show negligible background losses and resonant Er\(^{3+}\) absorption strong enough to produce cavity-polaritons that persist to above 361 K. Cooperative relaxation and upconversion limit the optical performance in the telecommunications bands by transferring the excitations to quenching sites or by further exciting the ions up to visible transitions. Future prospects and alternative applications for Er\(_2\)O\(_3\) and other epitaxial rare-earth oxides are also considered.
# Contents

Acknowledgments iv

Abstract v

Glossary of Acronyms xii

Preface xiv

1 Microdisk Optical Cavities 1

1.1 Quasi-normal Modes .................................................. 3

1.2 Coupled-Mode Theory .................................................. 8

1.2.1 Coherent Backscattering ......................................... 10

1.2.2 Microdisk-Waveguide Coupling ................................. 12

1.3 Cavity Parameters ..................................................... 17

2 "Dimpled" Fiber-Taper Probe 21

2.1 Taper Pulling and Molding ........................................... 22

2.2 Microphotonic Testing with a Dimple-Taper Waveguide .......... 25

2.2.1 Noise Measurements ................................................ 27

2.2.2 Dense Si Microdisk Array .......................................... 29

2.2.3 Planar Si-SiO$_2$ Microring via RIE Processing ................. 31

2.2.4 Planar Si-SiO$_2$ Microring via LOCOS Processing .......... 32

2.3 Conclusions ............................................................ 35

3 Linear Absorption in AlGaAs and GaAs Microcavities 36

3.1 $\text{Al}_x\text{Ga}_{1-x}\text{As}$ Samples and Processing ....................... 37

3.2 Estimation of Scattering and Absorption Rates .................... 39
4 Optical Processes in Epitaxial c-Er$_2$O$_3$ on Silicon

4.1 Rare-Earth Ions in the R$_2$O$_3$ Bixbyite Lattice .......................... 44
  4.1.1 Growth .................................. 47
  4.1.2 Processing ................................. 50

4.2 Optical Properties ............................................. 52
  4.2.1 Absorption and Emission Spectra for the $^4$I$_{13/2}$$^4$I$_{15/2}$ Transition 53
  4.2.2 Upconversion ...................................... 57
  4.2.3 Effective $^4$I$_{13/2}$ Lifetime ......................... 60
  4.2.4 Power-Dependent Radiative Efficiency ......................... 63
  4.2.5 Cavity Polaritons ................................... 67
     4.2.5.1 Low-Temperature Polaritons ..................... 69
     4.2.5.2 High-Temperature Polaritons .................... 76
  4.2.6 Rate Equation Estimates ............................. 82

4.3 Conclusions .................................................. 84

A Experimental Supplement ........................................... 86
  A.1 Mach-Zehnder Interferometer ............................... 86
  A.2 Pulse Optimization for EOSPACE Modulators ................. 90
  A.3 Effective Index for the HE$_{11}$ Mode ....................... 92
  A.4 Coherent Coupling to a Uniform Density of Dipoles ........... 92

B Calculation of the $\beta$-Factor for Er$_2$O$_3$ Microdisks ................. 95

Bibliography ..................................................... 98
# List of Figures

1.1 Microdisk geometry .................................................. 4
1.2 FEM simulation of field components for TE and TM microdisk modes .... 7
1.3 Comparison of effective-index and FEM models for microdisk QNMs .... 9
1.4 Illustration of microdisk-waveguide coupling .......................... 13
1.5 Power conservation and phase for coupling to a doublet resonance ....... 14
1.6 Comparison of cavity loading for singlet and doublet modes ............. 17

2.1 Process for producing dimpled fiber-taper couplers .................... 23
2.2 The “dimpled” fiber-taper waveguide probe ............................ 25
2.3 Dimple taper profile as a function of fiber tension ...................... 26
2.4 Basic arrangement for microcavity transmission measurements .......... 26
2.5 Reducing mechanical noise through higher taper tension ............... 28
2.6 Testing demonstration: dense microdisk array .......................... 30
2.7 Testing demonstration: planar microring .............................. 31
2.8 Testing demonstration: LOCOS process development .................... 33

3.1 Al$_x$Ga$_{1-x}$As microdisk resonators ................................ 38
3.2 Summary of measured resonance parameters for Al$_x$Ga$_{1-x}$As microdisks ... 38
3.3 Sample surface roughness measurement on a AlGaAs microdisk ........... 40

4.1 Rare-earth ions in the R$_2$O$_3$ lattice ................................ 45
4.2 Er$_2$O$_3$ grown via ALE on Si(111) .................................. 48
4.3 Processing Er$_2$O$_3$ microdisk cavities ................................ 51
4.4 Er$_2$O$_3$ microdisk transmission spectrum ($R \approx 20 \mu$m) ............... 53
4.5 Emission and absorption spectra for the $^4I_{13/2} \rightarrow ^4I_{15/2}$ transition .... 54
4.6 Sample fits to cavity transmission resonances .......................... 55
List of Tables

2.1 Average cavity mode parameters for microdisk array . . . . . . . . . . . . . . . 30
3.1 Summary of GaAs and Al$_x$Ga$_{1-x}$As material absorption rates. . . . . . 41
4.1 Lattice constant and mismatch for R$_2$O$_3$ bixbyite crystals on (111)Si . . . . 49
4.2 Summary of observed cavity-polaritons at $T = 8.6$ K . . . . . . . . . . . . . 71
4.3 Summary of cavity-polariton parameters at $T > 300$ K . . . . . . . . . . . . . 81
# Glossary of Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFM</td>
<td>Atomic force microscopy</td>
</tr>
<tr>
<td>ALE</td>
<td>Atomic layer epitaxy</td>
</tr>
<tr>
<td>APD</td>
<td>Avalanche photodiode</td>
</tr>
<tr>
<td>BOX</td>
<td>Buried oxide</td>
</tr>
<tr>
<td>CCD</td>
<td>Charge-coupled device</td>
</tr>
<tr>
<td>CMOS</td>
<td>Complementary metal-oxide-semiconductor</td>
</tr>
<tr>
<td>cQED</td>
<td>Cavity quantum electrodynamics</td>
</tr>
<tr>
<td>CW</td>
<td>Continuous wave</td>
</tr>
<tr>
<td>DBR</td>
<td>Distributed Bragg reflector</td>
</tr>
<tr>
<td>DCA</td>
<td>Digital communications analyzer</td>
</tr>
<tr>
<td>EDFA</td>
<td>Erbium-doped fiber amplifier</td>
</tr>
<tr>
<td>EOM</td>
<td>Electro-optic modulator</td>
</tr>
<tr>
<td>FEM</td>
<td>Finite-element method</td>
</tr>
<tr>
<td>FPC</td>
<td>Fiber polarization controller</td>
</tr>
<tr>
<td>FSR</td>
<td>Free spectral range</td>
</tr>
<tr>
<td>GVD</td>
<td>Group velocity dispersion</td>
</tr>
<tr>
<td>LOCOS</td>
<td>Local oxidation of silicon</td>
</tr>
<tr>
<td>LPF</td>
<td>Long-pass filter</td>
</tr>
<tr>
<td>MBE</td>
<td>Molecular beam epitaxy</td>
</tr>
<tr>
<td>MZI</td>
<td>Mach-Zehnder interferometer</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td><strong>OSA</strong></td>
<td>Optical spectrum analyzer</td>
</tr>
<tr>
<td><strong>PD</strong></td>
<td>Photodiode</td>
</tr>
<tr>
<td><strong>PECVD</strong></td>
<td>Plasma-enhanced chemical vapor deposition</td>
</tr>
<tr>
<td><strong>PL</strong></td>
<td>Photoluminescence</td>
</tr>
<tr>
<td><strong>PLC</strong></td>
<td>Photonic lightwave circuit</td>
</tr>
<tr>
<td><strong>PM</strong></td>
<td>Polarization maintaining</td>
</tr>
<tr>
<td><strong>PML</strong></td>
<td>Perfectly matched layer</td>
</tr>
<tr>
<td><strong>QD</strong></td>
<td>Quantum dot</td>
</tr>
<tr>
<td><strong>QNM</strong></td>
<td>Quasi-normal mode</td>
</tr>
<tr>
<td><strong>RHEED</strong></td>
<td>Reflection high-energy electron diffraction</td>
</tr>
<tr>
<td><strong>RIE</strong></td>
<td>Reactive ion etch</td>
</tr>
<tr>
<td><strong>RS</strong></td>
<td>Russell-Saunders</td>
</tr>
<tr>
<td><strong>SEM</strong></td>
<td>Scanning electron microscopy</td>
</tr>
<tr>
<td><strong>SOI</strong></td>
<td>Silicon-on-insulator</td>
</tr>
<tr>
<td><strong>SPF</strong></td>
<td>Short-pass filter</td>
</tr>
<tr>
<td><strong>TEM</strong></td>
<td>Transmission electron microscopy</td>
</tr>
<tr>
<td><strong>VOA</strong></td>
<td>Variable optical attenuator</td>
</tr>
<tr>
<td><strong>WGM</strong></td>
<td>Whispering gallery mode</td>
</tr>
<tr>
<td><strong>XRD</strong></td>
<td>X-ray diffraction</td>
</tr>
<tr>
<td><strong>ZDWL</strong></td>
<td>Zero-dispersion wavelength</td>
</tr>
</tbody>
</table>
Preface

When I first joined Oskar Painter’s group in early 2005, his first students were beginning to accelerate into their thesis projects and reap the benefits from their hard work setting up the labs. Matt Borselli and Tom Johnson’s attempts at Raman lasing in silicon were being thwarted by dynamic free-carrier effects, and they were addressing surface-state absorption and passivation issues during their drive toward microlasers with a silicon core and erbium-doped cladding. Except for a few late processing runs, Paul Barclay had largely disappeared from Steele (along with his 852-nm laser) to work in the Mabuchi lab, and Kartik Srinivasan was studying microdisk lasers containing quantum dots while biding his time to build up the cQED experiment.

During that summer, I joined Kartik on the quantum dot project, and I began to learn how to pull fiber-tapers and how to test microresonators with them. Luckily, Colin Chrystal was Matt’s SURF student that summer, and the taper-pulling program he developed has been invaluable to me ever since. With my vision, I rely entirely upon the interference in the fiber transmission because I have trouble seeing the taper. If it were not for one afternoon when I ripped the acrylic box off the puller and installed a camera, I still might not believe Oskar and Paul about the fiber “bowing.” Throughout that first year, I pulled a lot of tapers, and I tried to adapt Paul’s epoxy microjoint technique to make pigtailed taper-coupled resonators for the cryostat. However despite exotic cryo- and space-ready epoxies, we were never able keep the taper attached below ∼200 K. The various manufacturers were fairly defensive of their products and utterly flabbergasted when we told them we were using less than one picoliter per joint. After we realized we would have to buy low-temperature piezoelectric stages, I moved to developing the dimple tapers, which have become quite successful and reduce the processing time for our typical devices by ∼40%.

In the spring of 2006, Kartik and I began a project in collaboration with Kevin Hennessy at UCSB to understand optical losses in III-V microcavities. Both of our groups were
working towards the strong coupling regime in cQED using quantum dots. The original objective was to compare Kartik’s RIE (Ar:Cl₂) with Kevin’s wet chemical etches (HBr) of AlₓGa₁₋ₓAs and to understand why both methods seemed to underperform at shorter wavelengths. Unfortunately, the only method to assess the procedures (and not just lone exceptional devices) was to study many modes on many resonators at many wavelengths and identify statistical trends in the resulting heap of data. Since Kartik and Oskar were pretty sure this was a material issue, we reduced the scope to study only the HBr process (Kevin also gained a processing assistant so he would not actually have to make these devices). I got a lot of practice taper testing and quickly learned about microdisks, especially once we realized the edge roughness was different on our two material samples. Through the early summer, I worked to pile up data between classes and being hammered by the APh 24 lab, located in the lifeless cave that would shortly become the Kavli Nanoscience Institute. During that time I also decided to switch off the quantum dot project. Kartik was running past microlasers and master equations while I was still dealing with resonator coupling and holography labs. Ironically, Matt was simultaneously developing some more advanced single-shot techniques to separate scattering and absorption that would have saved me a lot of time and effort—I found out about these while reading his thesis a few months after I had finished taking my data.

So I packed my bags and landed in silicon photonics. After doing some basic silicon testing for the dimple paper, our intention was that I would continue work with the Er³⁺-doped silica from InPlane Photonics that Matt had left after his graduation. However, that fall we were presented with the opportunity to switch our DARPA program to work with Translucent in developing Er₂O₃. Oskar and I made the switch realizing the big-risk/big-reward potential. At the very least, it would temporarily free me from low-loss surface-sensitive processing and give us exclusive access to new material (unlike the crowded field of erbium-doped glasses). I had already broken my rule #1 for choosing thesis topics (no cryogenics) so breaking rule #2 (no exotic materials) did not seem that bad—especially since the growers (Vijit Sabnis and Homan Yuen) were ready and willing to ship me more material than I could ever deal with. I am still waiting to break rule #3 (no ultrahigh vacuum) and give MBE a shot.

In retrospect, we should have given more credence to some of our initial minor concerns, but we jumped in. Tom gave me a great head start on the processing by throwing the Er₂O₃
films at a battery of chemicals, of which most did nothing. I proceeded optimistically as Homan and Vijit addressed material problems at about the same pace I was figuring out the processing and testing aspects. The first disks I made with $H_2SO_4$ produced weak, but measurable, PL and exhibited the transmission features that would later be identified as cavity-polaritons at room temperature. By March of 2007, we were beginning to become concerned with the upconversion and our inability to see any absorption saturation. In June, I figured out the Ar$^+$ milling, made a bunch of samples, and then did not go back into the clean room for over a year (except to train people and pester Raviv Perahia). Once free from the clean room, I began gaining dominion over equipment in the test lab, which has provided me far too much contact with the New Focus laser repair department.

That fall, after a year on the project, we were beginning to get frustrated, but I keep myself busy with absorption spectra, radiative efficiency, upconversion, and polaritons at high and low temperatures. Vijit and Homan had moved on to work on starting a new company, but Translucent was bringing on new people. At this point, I also began some adventures into cQED by helping Kartik with some of his quantum dot measurement and testing my own disks in the cryostat. I worked throughout 2008 investigating the lifetime and polariton behaviors. Some of the data could have been obtained more quickly with unprocessed films and free-space optics. However, my unorthodox measurements fit within our microcavity infrastructure, and we were always optimistic that we would uncover some physics that would allow us to quickly proceed. Although with the rare-earth studies from the 1970s and a few rate equation parameters in hand, a coherent picture was emerging that $Er_2O_3$ is not a suitable laser material, and the weak coherent coupling to each ion made the ubiquitous Rabi splitting depressingly classical. I began a few last-ditch efforts to achieve lasing that pushed our equipment to the limits and the material beyond its damage threshold because we had invested a lot of time and energy. A Λ-system laser between carefully chosen Stark levels at 8 K is a viable option, but the benefits probably do not outweigh the effort. I decided to leave many questions unanswered concerning upconversion lasing at green wavelengths in the current material after burning standing-wave gratings into three disks and outright vaporizing a fourth. There is a subtle lesson here about knowing when to terminate an experiment.

While $Er_2O_3$ did not live up to its promise, these rare-earth oxide films still have considerable potential especially for controlling the emission spectrum and converting absorbed
photons to other wavelengths using $\text{Y}_2\text{O}_3$ or $\text{Gd}_2\text{O}_3$ alloys. I was exposed to a wide range of photonics and atomic physics trawling for immediate applications and results to utilize the current material, which really is an initial step in a much more ambitious material science and engineering program. Even though much of this thesis is focused on absorption rather than emission, I take solace in the difficult road to Maiman’s first ruby laser and that a commercially acceptable solution still does not exist to generate gain for optical interconnects on CMOS microelectronics.

After presenting the formalism for coupling to WGM cavities used throughout these measurements, this thesis is divided according to material systems. Silicon microresonators are presented in chapter 2 with details of the dimpled taper probes, which initially saw their greatest use in prototyping Si devices and fabrication methods. A sample analysis of loss in $\text{Al}_x\text{Ga}_{1-x}$ is presented in chapter 3, and chapter 4 concentrates on $\text{Er}_2\text{O}_3$. Ancillary experimental and theoretical details are reserved for the appendices.

C. P. Michael
Pasadena, CA
May 2009
Chapter 1

Microdisk Optical Cavities

The resonant circulation of electromagnetic waves enables a variety of macroscopic systems from lasers to interferometric gravitational wave detectors, but recent pressure from the telecommunications and microelectronics industries, along with more fundamental research programs, has driven the development of wavelength-scale optical cavities, waveguides, and other components. Microphotonic cavities can be designed to give a large electric field strength per photon, a sparse spectrum of optical modes, and great sensitivity to the structure’s surfaces [1, 2]. By combining these attributes with a high circulating intensity (i.e., low field decay rate for the optical mode), microcavity physics is central to many subjects, including low-threshold and “thresholdless” lasers [3], control of spontaneous emission through the Purcell Effect [4, 5], nonperturbative light-matter interactions [2, 6], low-power nonlinear optics [7, 8], chemical and biological “lab-on-a-chip” devices [9], and cavity optomechanics [10].

Several microresonator geometries exist beyond the canonical Fabry-Pérot cavity that can be optimized for different applications. Planar cavities with distributed Bragg reflectors (DBRs) require little etching and are commonly used with epitaxially grown material, as in exciton-polariton measurements with III-V and II-VI quantum wells [11, 12]. Similar studies with organic emitters, such as cyanine dyes in a J-aggregate structure [13], are possible by spin coating the organic material onto a wafer with a single DBR and then depositing a metal film to form the top mirror. Vertically etching a planar cavity to form a micropillar offers greater lateral mode confinement and more directional coupling through the top mirror into a single output mode, which has been used to efficiently collect single photon pulses [14]. Photonic crystals [15, 16] utilize Bragg reflection in multiple dimensions to open electromagnetic band gaps and to achieve a high electric field strength (per photon) [17]
without a dielectric discontinuity at the center of the cavity [18], and whispering-gallery microresonators (e.g., spheres [19], disks [20], and toroids [21]) provide a rich spectrum of modes with low intrinsic loss.

While significant work is focused on cavity design, material properties are of equal concern in many applications. These properties include both bulk quantities (absorption coefficient, radiative and nonradiative relaxation rates, etc.) and surface quantities (absorption by surface states, surface recombination velocity, etc.). Microcavities are a convenient tool for assessing material quality because they offer a long effective interaction length and measurable losses associated with input/output coupling. Device attributes associated with material processing (surface roughness, passivation layers, etc.) can also be investigated once bulk effects are adequately understood.

To maximize their utility for material studies, microcavities should be trivial to design with low intrinsic loss, easy to fabricate in a variety of materials, and straightforward to identify fabricated imperfections without optical measurements. Compared to other geometries, microdisk resonators offer an attractive combination of simplicity and material flexibility. Disks with diameters (thicknesses) $\gg \lambda_o/n_d \ (>\lambda_o/2n_d)$ support a large number of modes with very low radiation loss and require only a single material layer,\(^1\) where $\lambda_o$ is the free-space wavelength and $n_d$ is the refractive index of the disk material. The quality of the fabrication process for these devices can be easily evaluated by the surface roughness, unlike photonic crystals that require fault-tolerant designs [22] or precise simulation of the as-fabricated structures [23]. Microdisks are also robust against subtle leaky geometric variations as in micropillars [24, 25] and microrings [26].\(^2\) Microspheres and toroids offer many of the same advantages along with the lowest optical loss for any microcavity structure, which makes these devices extremely sensitive to refractive index changes in their modes’ evanescent tail [28–30], but the glass reflow procedure limits their material range to doped silica (e.g., through sol-gel deposition [31, 32] or ion implantation [33, 34]) or postreflow surface coatings [35, 36]. While replica molding techniques extend the toroid’s versatility to cured polymers [37], processing methods for crystalline toroids (e.g., silicon [38]) are not as mature as for microdisks. In this chapter, the quasi-normal whispering gallery modes

---

\(^1\)Excluding a sacrificial substrate that must be undercut to form the disk’s pedestal.

\(^2\)The whispering gallery modes of a microdisk with low refractive index can couple to modes with the same momentum in a high-index pedestal, which then radiate down into the substrate. However, this loss channel can be easily ameliorated by further undercutting as in Ref. [27].
WGMs of a microdisk resonator are discussed, and the coupled-mode theory and cavity parameters used throughout this work are presented.

1.1 Quasi-normal Modes

The normal modes of a system are solutions to the appropriate time-independent wave equation with the condition that the wave amplitude vanishes at the system boundary (either at infinity or the boundary of a more suitable domain), as for the optical field in an ideal Fabry-Pérot cavity or the elastic vibrations of a perfectly isolated rigid body. Under this constraint, the field magnitude is conserved, and the eigenvalues are real. The quasi-normal modes (QNM) of an open system are solutions to the same wave equation with an outgoing traveling-wave boundary condition that causes the field to decay [39]. In this classical analysis, the nature of the coupling to the bath of radiation modes is concealed unlike the quantum treatment of dissipation where the bath’s degrees of freedom are included and later removed using specific assumptions [40].

The microdisk structure consists of a planar slab with radius $R$ and height $h$ as in Fig. 1.1(a)—we assume the pedestal and substrate are undercut far enough to have little overlap with the cavity field. While analytic solutions for this geometry do not exist, the cavity modes can be found with reasonable accuracy using a 2D effective-index model. In Refs. [41, 42], the QNMs are further simplified by using exponential decay to describe fields outside the cavity. These real normal-mode eigenvalues do not describe radiation loss, but this analysis can be extended to include radiating fields and complex eigenvalues.$^3$

Further insight into the modes’ behavior can also be gained from a conformal coordinate transformation and the WKB approximation [44, 45], but semianalytic solutions without this transform have been most useful for quickly evaluating dispersion characteristics and providing accurate starting values for targeted numerical eigenvalue solvers.

Starting with the electromagnetic wave equation for a linear piecewise-homogeneous medium in the absence of free charges and currents,

$$\nabla^2 \vec{F} - \frac{n^2(\vec{r})}{c^2} \frac{\partial^2 \vec{F}}{\partial t^2} = 0,$$

(1.1)

where $\vec{F} \in \{\vec{E}, \vec{H}\}$, the standard linear constitutive relations are $\vec{D} = \epsilon(\vec{r})\vec{E}$ and $\vec{B} =$

$^3$This QNM analysis was explored prior to its publication by Heebner et al. [43].
Figure 1.1: (a) Perspective sketch of a microdisk and the $e^{\pm im\phi}$ field dependence for a second-order mode ($m = 35$, $p = 2$). (b) SEM image of a microdisk array fabricated from silicon-on-insulator (SOI)—i.e., Si disks with SiO$_x$ pedestals on a Si substrate.

$$\mu(\vec{r})\vec{H} \approx \mu_o\vec{H}, \quad c = 1/\sqrt{\varepsilon_o\mu_o}$$
is the speed of light in vacuum, and $n(\vec{r}) = \sqrt{\varepsilon(\vec{r})/\varepsilon_o}$ is the refractive index. By considering only oscillatory solutions $\vec{F}(\vec{r}, t) = \vec{F}(\vec{r})e^{-i\omega t}$, the time-independent wave equation in cylindrical coordinates becomes

$$\left( \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \phi^2} + \frac{\partial^2}{\partial z^2} + k_o^2 n^2(\vec{r}) \right) \vec{F}(\vec{r}) = 0,$$

where physical fields are given by $\Re(\vec{F})$, and $k_o = \omega/c$ is the desired eigenvalue representing the mode’s wavevector in free space. To account for material dispersion, $n(\omega)$ is calculated by fitting the Sellmeier formula to published data for Si and SiO$_x$ [46] and by using a simplified interband-transition model for Al$_x$Ga$_{1-x}$As [47]. Since data is unavailable, we model Er$_2$O$_3$ as nondispersive ($n = 2.0$); Sellmeier coefficients are available for Y$_2$O$_3$ but would not account for the optical transitions in the Er$^{3+}$ ions [48].

The high refractive index contrast along $\hat{z}$ for thin microdisks restricts the out-of-plane momentum for WGMs and allows the system to be simplified using a self-consistent effective-index model. This method also builds an intuitive picture of WGM behavior based on how the mode’s total momentum is roughly partitioned along $\hat{r}$, $\hat{\phi}$, and $\hat{z}$. By assuming the mode’s momentum along $\hat{z}$ is the same as in the analogous slab waveguide, the 3D disk with refractive index $n_d$ (cladding index $n_o$) can be treated as a 2D circle with a refractive index $n_{\text{eff}}$, where $n_{\text{eff}}$ is the effective index for the fundamental mode of a symmetric slab waveguide with the same $h$ and $n_d$ following the calculation in Ref. [49]. Under this assumption which is usually valid except for $h \gg \lambda_o/2n_d$, the Maxwell’s Equations relating the field components partially decouple giving two orthogonally polarized cavity modes. The TE
modes have non-zero $\{E_r, E_\varphi, H_z\}$ fields, and the TM modes consist of $\{H_r, H_\varphi, E_z\}$, where the $\hat{r}$-component identifies the TE-TM character similar to slab waveguides. These QNMs can be obtained by solving Eq. (1.2) for $F_z$, where $F_z = H_z$ and $E_z$ for the TE and TM modes, respectively. Using separation of variables, Eq. (1.2) can be rewritten for solutions of the form $F_z = R(r)A(\varphi)Z(z)$:

\[
\frac{\partial^2 Z}{\partial z^2} + k_0^2 [n^2(z) - n_{\text{eff}}^2]Z = 0, \quad (1.3a)
\]
\[
\frac{\partial^2 A}{\partial \varphi^2} + m^2 A = 0, \quad (1.3b)
\]
\[
\left( r^2 \frac{\partial^2}{\partial r^2} + r \frac{\partial}{\partial r} + r^2 k_0^2 n_{\text{eff}}^2(r) - m^2 \right)R = 0, \quad (1.3c)
\]

where $n_{\text{eff}}(r) = n_{\text{eff}}$ at $r < R$ and $n_{\text{eff}}(r) = n_0$ at $r > R$. Equation 1.3a effectively accounts for the out-of-plane propagation ($Z \approx e^{\pm i\beta_\perp z}$) with

\[
\beta_\perp^2 + k_0^2 n_{\text{eff}}^2 = k_0^2 n^2(z), \quad (1.4)
\]

where $\beta_\perp$ is the propagation constant normal to the disk plane and $k_0 n_{\text{eff}}$ is proportional to the in-plane momentum. The solutions $A = e^{\pm im \varphi}$ to Eq. (1.3b) describe clockwise and counterclockwise traveling waves [Fig. 1.1(a)] with angular momentum proportional to the mode index $m$, which is an integer since the field must be single valued for $\Delta \varphi = 2\pi$. These traveling-wave modes are degenerate due to rotational symmetry [$\varphi \to -\varphi$ in Eq. (1.2)].

The radial solution to Eq. (1.3c) is

\[
R = \begin{cases} 
J_m(k_0 n_{\text{eff}} r) & \text{for } r \leq R, \\
H_m^{(1)}(k_0 n_0 r) & \text{for } r \geq R,
\end{cases} \quad (1.5)
\]

where $J_m(x)$ and $H_m^{(1)}(x)$ are the Bessel and Hankel functions of the first kind.\(^4\) Matching the fields tangential to the disk boundary at $r = R$ gives a transcendental equation for the complex QNM eigenvalues ($k_o$)

\[
\frac{m}{k_0 n_{\text{eff}} R} \left( 1 - \xi \frac{n_0}{n_{\text{eff}}} \right) = \frac{J_{m+1}(k_0 n_{\text{eff}} R)}{J_m(k_0 n_{\text{eff}} R)} \xi \left( \frac{n_0}{n_{\text{eff}}} \right) \frac{H_{m+1}^{(1)}(k_0 n_0 R)}{H_m^{(1)}(k_0 n_0 R)}.
\]  

\(^4\)Bessel and Hankel functions of the second kind [$Y_m(x)$ and $H_m^{(2)}(x)$] are divergent as $x \to 0$ and $x \to \infty$, respectively.
where $\xi = n_{\text{eff}}^2/n_o^2$ for TE modes and $\xi = 1$ for TM modes.

Equation 1.6 can be solved numerically in the complex plane using a 2D minimization routine starting near the real axis $[\Re(k_o) \gg \Im(k_o)]$. There will be multiple solutions to Eq. (1.6) for a given angular momentum ($m$) corresponding to modes of different radial order $p$ [number of antinodes for $H_z$ or $E_z$ along $\hat{r}$ as in Fig. 1.1(a)]. The index $p$ also qualitatively describes the momentum along the radial direction. Hence, the fundamental ($p = 1$) radial mode for a given $m$ will have the lowest $|k_o|$ because higher $p$ modes will have a larger total momentum if the angular component is fixed. In general for two modes with the same $|k_o|$ if $p_1 < p_2$, then $m_1 > m_2$. Microdisks also support modes with higher vertical order (more than one $F_z$ antinode along $\hat{z}$) which can be found with Eq. (1.6) by using $n_{\text{eff}}$ for the appropriate order slab mode; however, the devices in this work only support the fundamental slab modes. For completeness, the condition to cut off the second-order mode is

$$\frac{2h}{\lambda_o} \sqrt{n_d^2 - n_o^2} < 1,$$

which corresponds to $h = 233\,\text{nm}$ in Si ($n_d = 3.476$) at $\lambda_o = 1550\,\text{nm}$; the fundamental TE and TM modes are never cut off in a symmetric slab waveguide. Throughout this thesis, microdisk modes are commonly organized in families according to polarization and radial order (e.g., TE$_{p=2}$); within these families, consecutive modes ($\Delta m = \pm 1$) have a relatively small free spectral range, comparable radiation loss, and similar phase velocity at the disk edge—which is pertinent to WGM-waveguide coupling.

While the effective-index solution provides an efficient method for finding the QNM eigenvalues with modest computational resources, numerical models using the finite element method (FEM) are more flexible for variable structures and calculations involving the full 3D geometry. Finite-element simulation of axially symmetric resonators can be accomplished using Comsol’s Multiphysics/Femlab platform by incorporating the $e^{imp\theta}$ dependence into a fully vectorial Cartesian waveguide solver. In these models, the 3D cavity and cladding are represented by a cross section in the $\hat{r}$-$\hat{z}$ plane, and a targeted eigenvalue solver finds numerically exact field profiles for the QNMs at a given $m$ value, as in Fig. 1.2. Full 3D simulations provide no additional information for these studies and are too mem-

---

5This code was originally developed by S. Spillane [50, 51]. M. Borselli and T. J. Johnson reconfirmed its findings and significantly extended its scope. A more detailed and independent description of this implementation can be found in Ref. [52]; axially symmetric electromagnetic structures are included as a standard application mode in Comsol Multiphysics v3.5.
Figure 1.2: Sample FEM simulations of field components for (a) TE and (b) TM modes in a Si microdisk with $R = 7.5 \, \mu m$ and $h = 210 \, nm$ showing the region $5 \leq r \leq 8 \, \mu m$. Simulations in the $\hat{r}$-$\hat{\phi}$ plane utilize the same $n_{eff}$ approximation as Eq. (1.3a), but they are less accurate than the solutions to the $\hat{r}$-$\hat{z}$ FEM models or Eq. (1.6). The notation “$i(\hat{\phi})$” indicates the $F_{\phi}$ components are $\pi/2$ out of phase.

ory intensive for typical desktop computers. When setting the geometry, the disk and cladding must be bound by perfectly matched layers (PMLs) that allow energy to escape the simulated domains in order to properly account for radiation losses [53]. These PMLs have complex anisotropic permittivity and permeability that quickly attenuate fields along a single direction without producing spurious reflections. By including PMLs thick enough to fully absorb the radiated waves, both highly confined and highly lossy modes can be found with $\Im(k_o)$ varying by more than $10^{14} \times$ as in Fig. 1.2. The quantity $\Re(k_o)/2\Im(k_o)$ is approximately $2\pi$ times the number of optical cycles that occurs while the field decays—this value is later identified as the radiation limited quality factor $Q_{rad}$.

These quasi-3D FEM models can be used to address two important aspects of real

---

6Dual-core 2–3 GHz processor, 64-bit operating system, 4 GB RAM.
microdisks that are not captured by the effective-index approximation: 1) the disk edge is not perfectly vertical and 2) surface effects. First, the primary objective during disk fabrication is minimizing surface roughness to reduce scattering losses, and the optimized anisotropic etches do not generally produce vertical sidewalls. Slanted edges couple the TE and TM polarizations by breaking the disk's vertical symmetry, which can occasionally produce hybrid modes if a TE and a TM mode are nearly degenerate and have the same $m$. The TE/TM labels are no longer rigorously accurate with slanted walls because significant longitudinal components may exist, but modes are conventionally identified as “TE” if $|E_t| > |E_z|$ at the disk’s midplane. For most general considerations, the bevelled edge only marginally increases the radiation losses [54, 55]. Second, Fig. 1.2 illustrates TM modes are significantly more sensitive to the top and bottom surfaces while the TE modes have greater overlap with the sidewall. This difference has practical consequences that will be discussed later, such as absorption by surface states [56].

Comparing approximate effective-index solutions with FEM results in Fig. 1.3 for two disk microdisk devices indicates the effectiveness of Eq. (1.6) to quickly explore wide areas in parameter space. The approximate solutions take $\lesssim 0.5$ s to compute without previous information about the QNMs. Alternatively, FEM simulations take $\sim 1$–2 min with a reasonable mesh resolution [$<0.05\lambda/n(\vec{r})$] and a good initial guess for the desired eigenvalue ($\Delta\lambda_o \lesssim 5$ nm); however, they can be accelerated ($\sim 10$ s) during parametric studies if better guesses can be supplied ($\Delta\lambda_o \lesssim 0.2$ nm). In Fig. 1.3, the effective-index solutions overestimate $\lambda_o$ found using the FEM by 3.6 nm and underestimate $\Im(k_o)$ by 1.9× on average.

1.2 Coupled-Mode Theory

Coupled mode analysis is a general framework to describe the evolution of optical fields due to perturbations that do not significantly alter the system’s eigenmodes [57]. As is common in perturbation theories, a total field in a system (driven with a laser at $\omega = \omega_\ell$) is expanded as

$$\vec{E}(\vec{r}, t) = e^{-i\omega_\ell t} \sum_j a_j(t) \vec{E}_j^0(\vec{r}),$$

in terms of slowly varying amplitudes $a_j(t)$ and the system’s unperturbed QNM field profiles $\vec{E}_j^0(\vec{r})$. For small dielectric perturbations [$\epsilon(\vec{r}) \rightarrow \epsilon(\vec{r}) + \delta\epsilon(\vec{r})$] as treated in Refs. [42, 58],
Figure 1.3: Comparison of the eigenvalues \( (k_0) \) found using effective-index and finite-element models for (a) TE modes in a \( \text{Er}_2\text{O}_3 \) disk and (b) TM modes in a Si disk plotted against free-space wavelength \( \lambda_0 = 2\pi/\Re(k_0) \). The solutions for the two models giving the same QNM are connected by grey lines.
the equations of motion for \( \{ a_j(t) \} \) to first order are

\[
\frac{da_j}{dt} = -i\Delta\omega_j a_j + i \sum_k \kappa_{jk} a_k,
\]  

where \(|a_j|^2\) is normalized to the energy in the \( j \)th optical mode, \( \Delta\omega_j = \omega_j - \omega_\ell \), and

\[
\kappa_{jk} = \frac{\omega_\ell}{2} \left( \frac{\int \delta \epsilon (\vec{E}_j^0)^* \cdot \vec{E}_k^0 d\vec{r}}{\epsilon |\vec{E}_j^0|^2 d\vec{r}} \right). 
\]

These equations are completely general, but they can easily be applied to microdisk cavities by using the QNMs in §1.1 as the basis for the fields in and around the resonator. Additional terms for coupling to other degrees of freedom can simply be added to Eq. (1.9) because the modal coupling is entirely described by the \( \kappa_{jk} \) coefficients:

\[
\frac{da_j}{dt} = -i\Delta\omega_j a_j - \frac{\gamma_{ij}}{2} a_j + i \sum_k \kappa_{jk} a_k,
\]

where \( \gamma_{ij} = \gamma_{\text{rad}} + \gamma_{\text{ss}} + \gamma_a + \cdots \) is the intrinsic loss rate of the \( j \)th mode due to radiation (\( \gamma_{\text{rad}} \)), surface scattering (\( \gamma_{\text{ss}} \)), absorption (\( \gamma_a \)), etc.\(^7\) The nature and magnitude of these terms can be found by considering the appropriate physics; the radiation loss is most easily obtained from the complex eigenvalue: \( \gamma_{\text{rad}} = 2c\Im(k_o) \). This approach has been the basis for analyzing many static and dynamic effects in microdisks, including those due to temperature, free carriers, four-wave mixing, laser oscillation, and coherent light-matter interactions [59–63]. For practical testing of low-loss microdisk devices at low power (i.e., without nonlinear effects), mode coupling from two sources must be addressed: 1) coupling of disk modes through coherent backscattering and 2) coupling between the resonator and an external waveguide.

### 1.2.1 Coherent Backscattering

As discussed previously, the QNMs of an ideal microdisk consist of pairs of degenerate traveling-wave WGMs that propagate clockwise and counterclockwise around the disk periphery. In real structures, etch-induced sidewall roughness lifts this degeneracy by break-

\(^7\)Conventionally, coefficients for field amplitude decay or coupling are denoted with “\( \kappa \)” while magnitude coefficients are denoted with “\( \gamma \).” This thesis will primarily use “\( \gamma \)” because of its association with energy and power, but “\( \kappa \)” will be used when it is more common in the relevant literature. Both coefficients are angular rates (rad/s) like \( \omega_j \).
ing the rotational symmetry and couples the counterpropagating WGMs through coherent Rayleigh backscattering. Cursory inspection of Eq. (1.10) gives $\kappa_{m,-m} = \kappa_{-m,m}^* \equiv |\kappa_{cb}| e^{i\phi_o}$ because the two modes only differ in their angular dependence ($e^{\pm im\phi}$); for convenience the perturbation $\delta \varepsilon(\vec{r})$ is real and chosen such that $\kappa_{m,m} = 0$. In addition, the integral in the numerator of Eq. (1.10) makes $|\kappa_{cb}|$ proportional to the $\pm 2m$ spatial Fourier component of the roughness: $\int \delta \varepsilon(\varphi) (\vec{E}_m^o)^* \cdot \vec{E}_m^o d\varphi = \int \delta \varepsilon(\varphi) e^{i(2m\varphi)} d\varphi$. Modes of different angular order ($m_1, m_2$) will also be coupled through these imperfections, but the amplitude of the $\pm (m_1 - m_2)$ component must be large to overcome poor vertical and radial field overlap. Without intentionally etching a grating into the resonator [64, 65], efficient backscattering only occurs between the paired $\pm m$ WGMs, and their amplitudes are governed by Eq. (1.11):

\[
\frac{da_{cw}}{dt} = -\left( i\Delta \omega + \frac{\gamma_i}{2} \right) a_{cw} + i|\kappa_{cb}| e^{-i\phi_o} a_{ccw},
\]

\[
\frac{da_{ccw}}{dt} = -\left( i\Delta \omega + \frac{\gamma_i}{2} \right) a_{ccw} + i|\kappa_{cb}| e^{i\phi_o} a_{cw},
\]

where the coherent backscattering ($\kappa_{cb}$) terms do not introduce loss. Surface roughness, as with any dielectric inhomogeneities, will also scatter light out of the cavity (i.e., couple to radiation modes), but this loss is accounted for by the $\gamma_{ss}$ component of $\gamma_i$. The scattering-loss rate can be estimated for real devices with a induced-current method [41, 42] rather than explicitly coupling to the radiation modes through Eq. (1.10).

Since $\kappa_{cb}$ and $\gamma_{ss}$ have different dependence on the roughness amplitude (which is discussed further in §3.2), resonators can enter a regime with $|\kappa_{cb}| \gtrsim \gamma_{ss}$ where it is convenient to express the cavity modes in a sine/cosine basis:

\[
a_c = \frac{1}{\sqrt{2}} \left( a_{cw} + e^{-i\phi_o} a_{ccw} \right) \quad \text{and} \quad a_s = \frac{1}{\sqrt{2}} \left( a_{cw} - e^{-i\phi_o} a_{ccw} \right),
\]

corresponding to standing waves with amplitudes $\sqrt{2} \cos(m\varphi - \varphi_o/2)$ and $\sqrt{2} \sin(m\varphi - \varphi_o/2)$, respectively [58, 66–69]. The phase of the backscattering parameter ($\varphi_o$) determines the standing waves’ azimuthal orientation relative to the roughness with the higher-frequency $a_s$ mode (lower-frequency $a_c$ mode) residing more in the low (high) index regions. In this basis, the evolution of the mode amplitudes decouple and produce a doublet resonance at
\[
\omega = \omega_0 \pm |\kappa_{cb}|:
\]
\[
\frac{da_c}{dt} = - \left( i \Delta \omega - i |\kappa_{cb}| + \frac{\gamma_{c,i}}{2} \right) a_c, \quad (1.14a)
\]
\[
\frac{da_s}{dt} = - \left( i \Delta \omega + i |\kappa_{cb}| + \frac{\gamma_{s,i}}{2} \right) a_s. \quad (1.14b)
\]

While mathematically the standing waves have the same loss rate ($\gamma_i$) as the traveling waves, the physical $\{a_c, a_s\}$ modes may decay differently $\{\gamma_{c,i}, \gamma_{s,i}\}$ because they are $\pi/2$ out of phase and will not have the same overlap with scattering and absorbing sites.

### 1.2.2 Microdisk-Waveguide Coupling

The optical modes of microphotonic devices are commonly excited by either generating light within the structure or introducing light from an external source through a waveguide coupler. While the first method is attractive in many cases because its collection optics usually consist of a single objective lens with high numerical aperture, its spectral resolution is typically limited to \(>1\) GHz by a spectrometer grating, and it cannot be applied to passive structures. The second method may require additional coupling elements, but it can probe active and passive devices with a resolution typically limited to \(<1\) MHz by a tunable laser source. Coupling to an external waveguide is also generally more efficient [70].

Following the formalism of Haus [57], evanescent coupling between the cavity and a single traveling-wave mode in the waveguide, as in Fig. 1.4(a), can be described by

\[
\frac{da_{cw}}{dt} = - \left( i \Delta \omega + \frac{\gamma_c + \gamma_s}{2} \right) a_{cw} + i |\kappa_{cb}| e^{-i\varphi_o} a_{ccw} + \kappa_e s, \quad (1.15a)
\]
\[
\frac{da_{ccw}}{dt} = - \left( i \Delta \omega + \frac{\gamma_c + \gamma_s}{2} \right) a_{ccw} + i |\kappa_{cb}| e^{i\varphi_o} a_{cw}, \quad (1.15b)
\]

where $s$ is the amplitude of the “source” waveguide mode (with $|s|^2$ normalized to power), $\kappa_e$ is the waveguide-cavity coupling coefficient, and $\gamma_e$ is the extrinsic loss rate associated with cavity energy leaking out through the waveguide. Using Eq. (1.13), the standing wave amplitudes obey

\[
\frac{da_c}{dt} = - \left( i \Delta \omega - i |\kappa_{cb}| + \frac{\gamma_{c,i} + \gamma_e}{2} \right) a_c + \frac{\kappa_e}{\sqrt{2}} s, \quad (1.16a)
\]
\[
\frac{da_s}{dt} = - \left( i \Delta \omega + i |\kappa_{cb}| + \frac{\gamma_{s,i} + \gamma_e}{2} \right) a_s + \frac{\kappa_e}{\sqrt{2}} s. \quad (1.16b)
\]
Given the linearity of the coupled mode theory presented in §1.2, $\kappa_e$ can be calculated using Eq. (1.10) and integrating the eigenmodes of the waveguide and cavity over a finite 3D coupling region.$^8$ Evaluating $\kappa_e$ is difficult because of the oscillatory integrand for the superimposed waves propagating along $\hat{\varphi}$ and $\hat{y}$; however, insight into efficient waveguide-disk can be gained by examining the numerator of Eq. (1.10): $\int \delta\epsilon(\vec{E}^0_\delta)^* \cdot \vec{E}^0_{wg} d\vec{r}$. First, $\kappa_e$ depends on the geometric overlap of the evanescent fields of the cavity and waveguide modes. Hence, thinner cavities couple strongly as the WGMs are not buried deeply in the slab; the waveguide must also be small enough to give large evanescent tails to the input mode. For thin cavities and waveguides with small cross-sectional area, the second and more stringent requirement for efficient coupling is phase matching between the two modes due to the angular dependence of the integrand: $(e^{-im\varphi})^* e^{-i\beta_{wg} y} \approx e^{im\varphi} e^{-i\beta_{wg} (R + \delta x) \varphi}$ where $\beta_{wg}$ is the propagation constant of the waveguide mode and $\delta x$ is the gap between the waveguide and disk edge. Qualitatively, the coupling to different WGMs can be quickly assessed by comparing the effective index for propagation tangent to the disk edge for the cavity mode ($n_{cav}^{\text{eff}} \approx m/Rk_o$) and waveguide ($n_{wg}^{\text{eff}} = \beta_{wg}/k_o$).

Assuming the coupling is lossless and obeys time-reversal symmetry, a scattering matrix analysis requires $\kappa_e = i\sqrt{\gamma_e}$, and the fields transmitted ($t$) and reflected ($r$) by the coupling

---

$^8$The different normalization for $a_{cw}$ and $s$ (to energy and power, respectively) must also be addressed.
Figure 1.5: Power conservation and phase shift for the transmitted field (t) as a function of detuning for a waveguide coupled to a doublet resonance with \( \{\gamma_e, \gamma_{c,i}, \gamma_{s,i}, |\kappa_{cb}|/2\pi\} = \{0.5, 1.25, 2.5, 2.5\} \) GHz. The total power dissipated by the \( j \)th mode is \( P_j = \gamma_j \frac{|a_j|^2}{s} \).

The arbitrary phase of \( \kappa_e \) is chosen so \( t = s \) when the modes are uncoupled. This theory provides an effective method for extracting \( \{\gamma_i, \gamma_e\} \) for individual WGMs by fitting the normalized cavity-waveguide transmission \( T = \frac{|t/s|^2}{1 + i\sqrt{\gamma_e a_{cw}}} \) with Eq. (1.17a) and the steady-state solutions for the mode amplitudes [Eq. (1.15) or (1.16)]; the reflected power is given by \( R = \frac{|r/s|^2}{1 + i\sqrt{\gamma_e a_{cw}}} \). For a waveguide coupling to a traveling wave mode, the transmission spectrum consists of a Lorentzian singlet,

\[
T(\Delta \omega) = \left| 1 + i\sqrt{\gamma_e} \left( \frac{a_{cw}}{s} \right) \right|^2 = \left| \frac{1}{2}\frac{(\gamma_i - \gamma_e) + i\Delta \omega}{\gamma_i + \gamma_e + i\Delta \omega} \right|^2,
\]

centered at \( \omega_\ell = \omega_j \) with a loaded full width \( \delta \omega \equiv \gamma_\ell = (\gamma_i + \gamma_e) \). Critical coupling \( (T = 0) \) occurs on resonance when the input rate \( (\gamma_e) \) equals the intrinsic loss rate \( (\gamma_i) \); under (over) coupling describes \( \gamma_i > \gamma_e \) \( (\gamma_i < \gamma_e) \). Furthermore, \( T(\Delta \omega = 0) \) depends only on the ratio \( \gamma_e/\gamma_i \), so significant power can be transferred to a low-loss cavity even with a small extrinsic coupling rate. Figure 1.5 shows \( T(\Delta \omega) \) for a waveguide coupled to a disk with resolved standing-wave modes and demonstrates how power is conserved.
When $|\kappa_{cb}| \ll \frac{1}{2}(\gamma_i + \gamma_e)$, the coherent backscattering term can be ignored in the evolution of $a_{cw}$ [Eq. (1.15a)] to first order. In this limit, the $\kappa_{cb}$ dynamics become inconsequential relative to the overall cavity decay. However, even weak backscattering will excite the $a_{ccw}$ mode and produce a reflected field; the steady-state reflection spectrum from a mode with weak Rayleigh scattering is

$$R(\Delta \omega) = \frac{\frac{\gamma_e^2 |\kappa_{cb}|^2}{\Delta \omega^2 + \left[\frac{1}{2}(\gamma_i + \gamma_e)\right]^2}}{2}. \quad (1.19)$$

As expected the reflected power is very small, but it will be non-zero for all real WGM resonators.

These coupled mode amplitudes accurately describe real resonators for a wide range of $\gamma_i$ and $\gamma_e$ values as long as the system consists of a single waveguide mode and the paired WGMs. The interference between two nearly degenerate cavity modes is frequently encountered when searching for high-$n_{\text{eff}}$ $\text{TE}_{p=1}$ modes with a fiber-taper. To help overcome the phase mismatch, the taper is placed on the disk edge to increase the modal overlap. In this position, $T(\Delta \omega)$ will display weak coupling to high-$m$ low-$\gamma_i$ modes, but it will also deeply couple to broad high-$p$ modes that are nearly phase matched to the waveguide resulting in a Fano-like response [71]. These resonances require the source terms in Eqs. (1.15) and (1.16) to account for the phase shift (Fig. 1.5) induced by other cavity modes [72, 73]. Since the source field ($s$) is typically only an excitation of the waveguide’s fundamental eigenmode, multiple waveguide modes effectively introduce a parasitic loss component ($\gamma_p$) in $\gamma_i$ to account for decay into the extra waveguide channels:

$$\gamma_i' = \gamma_i + \gamma_p = \gamma_i + \sum_{k \neq 1} \gamma_{k,e}, \quad (1.20)$$

where $\gamma_{k,e}$ is the extrinsic loss rate into the $k$th waveguide mode. The summation does not include the pump mode ($k = 1$) because $\gamma_{1,e}$ is already explicitly included in Eq 1.15, and the coupling coefficient for the source field remains $\kappa_e = i\sqrt{\gamma_{e,1}}$. Parasitic losses can significantly broaden the transmission response of a WGM resonance for waveguides supporting higher order transverse modes. To avoid parasitic loading, the waveguide should be small enough to cut off all but the fundamental mode. The single-mode criteria for a tapered optical fiber
waveguide is
\[ \frac{r_{wg}}{\lambda_0} < \frac{2.405}{2\pi\sqrt{n_{wg}^2 - n_o^2}}. \]  

where \( r_{wg} \) is the taper radius and \( n_{wg} \) \((n_o)\) is the refractive index of the waveguide (cladding), which gives \( r_{wg} < 0.56 \mu m \) for \( \lambda_0 = 1550 \text{ nm} \) [49]; no similar geometric constraint exists for partially etched ridge waveguides.\(^9\) The amount of parasitic coupling is described by the junction’s ideality \( I = \gamma_e / (\gamma_e + \gamma_p) \) which is the ratio of the desired coupling over the coupling to all modes of the waveguide; \( I \geq 0.9997 \) has been observed for a fiber-taper coupled to a microsphere cavity [75]. Parasitic loading is usually of little concern for high-\( n_{\text{eff}} \) WGMs since the higher-order waveguide modes have lower \( n_{\text{eff}} \) than the fundamental and worse phase mismatch with the WGMs.

Even for single-mode waveguides, a parasitic channel exists when coupling to a standing-wave mode. Comparing Eqs. (1.15) and (1.16) after rotating to the \( \{a_c, a_s\} \) basis, the coefficient on the source field \( s \) is reduced \( \left(i\sqrt{\gamma_c} \rightarrow i\sqrt{\gamma_e/2} = i\sqrt{\gamma_e}\right) \) as the waveguide feeds \( a_c \) and \( a_s \) equally, but the extrinsic loss term apparently remains the same. Because the standing waves couple evenly to the waveguide’s forward- (+\( \hat{y} \)) and backward-propagating (-\( \hat{y} \)) modes, the desired loading by the forward mode produces an extrinsic decay \( \gamma'_e = \gamma_e \), and coupling to the backward mode produces parasitic decay at rate \( \gamma_p = \gamma_e/2 \). In this case, \( I = 0.5 \), and \( \gamma'_i + \gamma'_e = \gamma_i + \gamma_e \). By feeding the reflection channel, pure standing-wave resonators (such as a Fabry-Pérot cavity) can never over couple to a waveguide and only asymptotically approach critical coupling as \( \gamma'_e \rightarrow \infty \). For real WGM cavities, the finite backscattering rate complicates the loaded transmission response (Fig. 1.6) since the modes can contain both traveling- and standing-wave character. With \( |\kappa_{cb}| \ll \gamma_i/2 \), \( T(\Delta\omega) \) for the traveling-wave mode follows Eq. (1.18), critically couples at \( \Delta\omega = 0 \) with \( \gamma_e = \gamma_i \), and overcouples at \( \gamma_e > \gamma_i \) when the waveguide mode becomes the cavity’s dominate loss channel. When backscattering is significant, increased waveguide loading morphs the doublet lineshape to a singlet as in Fig. 1.6(b). During this evolution, the waveguide transmission resonant with the standing-wave modes \( (\Delta\omega = \pm |\kappa_{cb}|) \) never reaches zero, and the resonance behaves like an overcoupled traveling-wave WGM when \( \gamma_e \gg \{\gamma_{c,i}, \gamma_{s,i}, |\kappa_{cb}|\} \).

The transition between these two regimes occurs when the backscattering and extrinsic loss

\(^9\)The single-mode condition generally attributed to R. A. Soref [74] for ridge waveguides in SOI only applies for large cross-sectional areas and is not a cutoff condition. Higher-order transverse modes are supported, but they leak out after propagating a short distance.
Figure 1.6: Comparison of cavity loading for (a) a singlet resonance with $\{\gamma_i, |\kappa_{cb}|\}/2\pi = \{1.25, 0\}$ GHz and (b) a doublet resonance with $\{\gamma_{c,i}, \gamma_{s,i}, |\kappa_{cb}|\}/2\pi = \{1.25, 2.5, 2.5\}$ GHz. The extrinsic coupling increases between each curve by $\Delta \gamma_e/2\pi = 0.2$ GHz.

are nearly equal: $\gamma_e/2 \approx |\kappa_{cb}|$. At $\gamma_e/2 = |\kappa_{cb}|$, the resonance reaches critical coupling at a detuning between $\pm |\kappa_{cb}|$ determined by $\gamma_{c,i}$ and $\gamma_{s,i}$.$^{10}$

1.3 Cavity Parameters

The coupled mode theory presented in the previous section can adequately describe the dynamics of the cavity fields under a wide range of conditions (especially once additional nonlinear terms and coupling to other degrees of freedom are included), but there are a number of more succinct physical parameters commonly used to characterize optical microcavities. Perhaps the most common is the cavity quality factor ($Q$), which is generally defined for any resonator (optical, electrical, mechanical, etc.) as

$$Q \equiv \frac{\omega_0 U_c}{P_d},$$

(1.22)

where $\omega_0$ is the oscillation frequency, $U_c$ is the stored energy (i.e., the optical cavity energy), and $P_d$ is the power dissipated by the resonator. The number of photons in the cavity is then simply $n_{ph} = U_c/\hbar\omega_0$. In a waveguide-cavity system, the dissipated power is the same as power “dropped” into the cavity from the waveguide: $P_d = [1 - T(\Delta\omega)]P_{in}$, where $P_{in}$

$^{10}$Recall $\gamma_e$ describes the cavity’s energy decay while $\kappa_{cb}$ describes coupling for a field. Because of this convention, $\gamma_e/2 = |\kappa_{cb}|$ corresponds to the field amplitude coupling into the waveguide and counterpropagating WGM at the same rate.
is the input power in the waveguide. By simply depending on \( T(\Delta \omega) \), the dissipated power includes parasitic loss into the reflected channel but omits power coupled from the cavity back into the transmitted channel—this portion is just “delayed.” Because \( P_d = -dU_c/dt \), Eq. (1.22) can be recast as a differential equation solved by \( U_c(t) = U_c(0)e^{-\omega_o t/Q} \), and the cavity energy response in the time- and frequency-domain identify \( Q \) in terms of the cavity photon lifetime (\( \tau_{\text{ph}} \)) and loaded cavity decay rate (\( \gamma_\ell \)),

\[
Q = \omega_o \tau_{\text{ph}} = \frac{\omega}{\gamma_\ell},
\]

where \( \tau_{\text{ph}} \) is a “linear” time if \( \omega_o \) is an “angular” frequency. As before, \( \gamma_\ell \) includes all loss mechanisms (\( \gamma_i + \gamma_e + \gamma_p \)) and is the full width at half maximum of \( U_c(\omega) \). The cavity’s intrinsic \( Q_i \) is determined by the cavity’s inherent loss mechanisms independent of the waveguide (\( Q_i = \omega/\gamma_i \)), whereas Eq. (1.23) is conventionally labelled the loaded \( Q_\ell \). Because \( Q \) is a dimensionless quantity, it is often partitioned among the various loss mechanisms rather than separating \( \gamma_\ell \). For examples, \( Q_{\text{ss}} = \omega_o / \gamma_{\text{ss}} \) would be the cavity \( Q \) if surface scattering were entirely responsible for the decay of \( U_c(t) \). Since the loss rates are additive, the components of \( Q_i \) add reciprocally:

\[
\frac{1}{Q_i} = \frac{1}{Q_{\text{rad}}} + \frac{1}{Q_{\text{ss}}} + \frac{1}{Q_a} + \ldots,
\]

where the summation does not include \( Q_e \) or \( Q_p \), which are associated with extrinsic coupling to the desired and parasitic waveguide modes, respectively.

While \( \{Q, \tau_{\text{ph}}, \gamma_\ell\} \) quantify how long the photons remain in the cavity, the phase and group velocities are necessary to establish how they circulate around the resonator. Unfortunately, the linear propagation constant (\( \beta \)) is not rigorously defined because it varies with radius—converting the angular propagation to Cartesian coordinates gives:

\[ e^{\pm im\varphi} \approx e^{\pm imx/r}. \]

Since the phase velocity (\( v_p \)) tangent to the disk edge can vary with radius (\( r \)), it can be found according to its usual definition

\[
v_p \equiv \frac{\omega}{\beta} = \frac{\omega_o r}{m}.
\]

More care must be taken with the linear group velocity (\( v_g \)), which is the velocity of the spatially extended optical wave packet. The group velocity is approximately constant across
the WGM for all but the smallest microdisks and is given by

\[ v_g = \frac{\partial \omega}{\partial \beta} \approx \frac{\omega_{m+1} - \omega_m}{m+1 - m} = \delta \omega_{\text{fsr}} R_{\text{avg}} \]

(1.26)

where \( \delta \omega_{\text{fsr}} \) is the free spectral range (FSR) between two adjacent modes in the same family and \( R_{\text{avg}} \) is the average radius of the mode’s energy density; the group index is then \( n_g = c/v_g = c/(\delta \omega_{\text{fsr}} R_{\text{avg}}) \). For the \( p = 1 \) WGMs studied in this work, \( R_{\text{avg}} \) is nearly equal to the disk’s radius. The group velocity is particularly important when dealing with loss (\( \alpha \)) or gain coefficients expressed as an inverse length. The processes producing the loss/gain occur in the time domain and are analyzed using transition rates (e.g., Fermi’s Golden Rule). These rates are then converted to an inverse length using the group velocity (e.g., by considering how long it takes for a pulse to traverse an absorbing region of fixed length) rather than the phase velocity. While free-space optical studies in bulk materials may be weakly dispersive, the guided waves in microphotonics can experience significant and dynamic modal dispersion (e.g., Ref. [76]). For this reason, the various optical processes will be mainly described using rates (\( \gamma_x \) as in §1.2), and they will only be converted to inverse lengths for comparison to published sources: \( \alpha_x = \gamma_x/v_g \).

The group velocity can be used to find several other quantities. The effective interaction length (\( L_{\text{eff}} \)) is the mean distance a photon will travel in the cavity before it escapes and is given by: \( L_{\text{eff}} = v_g \tau_{\text{ph}} \). For a low-loss micron-scale resonator, \( L_{\text{eff}} \) can easily be \( >1 \text{ cm} \)—demonstrating why microcavities are sensitive to small perturbations. The round-trip time for a photon to circle the cavity is \( \tau_{\text{rt}} = 2\pi R_{\text{avg}}/v_g \approx 2\pi/\delta \omega_{\text{fsr}}, \) and the number of round trips a photon makes is \( \tau_{\text{ph}}/\tau_{\text{rt}} \approx \delta \omega_{\text{fsr}}/(2\pi \gamma_\ell) = \mathcal{F}/2\pi \) where \( \mathcal{F} \equiv \delta \omega_{\text{fsr}}/\gamma_\ell \approx Q/m \) is the cavity’s finesse. These relations then relate the dropped power to the power circulating in the WGM (\( P_{\text{cir}} = \mathcal{F} P_d/2\pi \)), which can exceed 10 W when the cavity is pumped continuous wave (CW) or pulsed (pulse width \( \gtrsim 3\tau_{\text{ph}} \)) with 1 mW from an external waveguide.

Depending on the application, the strength (per photon) and distribution of the cavity fields are often significant. In cavity quantum electrodynamics (cQED), the electric field strength per photon is

\[ |\vec{E}| = \sqrt{\frac{\hbar \omega}{2n^2 \epsilon_0 V_{\text{eff}}}} \]

(1.27)
where $V_{\text{eff}}$ is the effective mode volume given by

$$V_{\text{eff}} = \frac{\int n^2(\vec{r})|\vec{E}(\vec{r})|^2dV}{\max \left[ n^2(\vec{r})|\vec{E}(\vec{r})|^2 \right]}, \quad (1.28)$$

This mode volume can be found using the FEM and roughly corresponds to the volume occupied optical fields. While $V_{\text{eff}}$ generally scales with the cavity size, this behavior can be misleading for slotted structures [18]. Other definitions of $V_{\text{eff}}$ exist for nonlinear processes that employ higher moments of the field (e.g., Refs. [62, 77]), but Eq. (1.28) will be used throughout this thesis unless otherwise noted. In many hybrid photonic structures, the fraction of the optical mode’s energy in a given volume ($\delta V$) or on a surface ($\delta A$) is more important than $V_{\text{eff}}$. For example, a hybrid microcavity laser consisting of an active layer stacked on a passive waveguide layer can exhibit significantly lower gain if the optical mode largely contained in the passive material. The volumetric overlap factor can be calculated with the FEM using

$$\Gamma = \frac{\int_{\delta V} n^2(\vec{r})|\vec{E}(\vec{r})|^2dV}{\int n^2(\vec{r})|\vec{E}(\vec{r})|^2dV}, \quad (1.29)$$

and a similar areal overlap is defined as

$$\Gamma' = \frac{\int_{\delta A} n^2(\vec{r})|\vec{E}(\vec{r})|^2dA}{\int n^2(\vec{r})|\vec{E}(\vec{r})|^2dV}, \quad (1.30)$$

where both denominators are integrated over all space. The $\Gamma'$ factors are commonly used when analyzing issues of surface sensitivity (sensing, absorption by surface states, etc.), and they can be converted to a volumetric energy fraction by assuming the surface has some effective interaction depth ($d'$): $\Gamma = \Gamma' d'$. 
Chapter 2

“Dimpled” Fiber-Taper Probe

In microelectronics manufacturing, nondestructive parametric testing using metal probe tips greatly increases fabrication yield through statistical process control [78]. For testing of glass and semiconductor photonic lightwave circuits (PLCs), many methods exist for the coupling of light into and out of on-chip waveguides [79]. However, no simple, local probe exists for wafer-scale, nondestructive, optical characterization of on-chip components. Traditional optical coupling methods include end-fire or butt coupling [80–82] and prism-based coupling [83, 84]. End-fire coupling from free-space or optical fibers can be made highly efficient, even to high-index contrast semiconductor waveguides, through the use of tapered waveguide sections [80–82] or other non-adiabatic mode converters [85, 86], but they are limited to coupling at the periphery of the chip where a cleaved facet can be formed. Evanescent-coupling methods involving conventional prism couplers, angled-fiber tip couplers [87], eroded-fiber couplers [88], and optical fiber-tapers [89–91] can provide effective coupling to and from on-chip waveguides, but these probes are less suited to wafer-scale coupling to micron-scale photonic elements due to their macroscopic extent in one or both in-plane dimensions. Evanescent coupling techniques also rely on phase matching to obtain highly efficient coupling [70, 75, 77, 92], which can be difficult (although not impossible [77, 93]) to satisfy for semiconductor-based microphotonic chips. Other methods of coupling light onto photonic chips for characterization purposes involve dedicated on-chip testing structures such as in-plane grating couplers [94]. These couplers typically also involve specialized processing to achieve high coupling efficiency: blazed gratings [95], a combination of lateral and vertical Bragg reflectors [96], or additional overlayers [97].

We present a variant of the silica optical fiber-taper evanescent-coupler that is designed for rapid, wafer-scale diagnostic testing of on-chip photonic components such as waveguides...
and resonant filters. Previous work involving straight fiber-tapers required devices to be elevated by several microns above the chip surface to prevent parasitic coupling to the surrounding substrate. Curved fiber-taper probes [92, 98–101] have been demonstrated to reduce parasitic loss into the substrate. However, they tend to be less mechanically stable than their tensioned straight-taper counterparts and suffer from noise induced by fluctuations in the taper’s position. In this work we have developed a microscopic “dimpled” fiber-taper probe which allows for low-noise local probing of individual devices on a wafer. By increasing the tension in the taper, fluctuations in the taper-chip gap can be greatly reduced to the levels present in straight fiber-taper measurements. To demonstrate the utility of the dimpled taper optical probe, we describe the characterization of two types of devices on a SOI wafer platform: a dense two-dimensional array of high-\( Q \) silicon microdisk resonators and, second, a planar microring resonator.

2.1 Taper Pulling and Molding

The dimpled fiber-taper probe is made from a standard straight fiber-taper that is pressed against a mold and heated (Fig. 2.1). We form “straight” fiber-tapers by simultaneously heating and pulling standard telecommunication fiber (specifically Corning SMF-28e and NuFern 1060-XP) using two linear motorized stages and a hydrogen torch. The fiber is positioned at the edge of the flame so only a small portion of the fiber is heated; larger hot zones produce long pliable tapers that are more difficult to mount. The diameter of the taper is constantly monitored during the pull by either watching the fiber’s mechanical deflection in the flame or by measuring the fiber’s transmission at the desired wavelength. In the final single-mode taper coupler, the fundamental core-guided HE\(_{11}\) fiber mode is adiabatically converted by slowly thinning the fiber to the HE\(_{11}\) taper mode with evanescent tails that extend significantly into the surrounding medium. However at larger diameters while it is being drawn, the optical field in the tapering region is a superposition of the LP\(_{mn}\) modes [102]. As the diameter is reduced, the higher-order components are reflected producing an oscillating signal with a frequency that increases as the pull progresses. When the fiber approaches the diameter that cuts off all higher order modes, the transmission oscillations abruptly cease [Fig. 2.1(b)]. A running standard deviation of this signal (sampling at 100 kHz and analyzing a circular buffer at 5 Hz) exhibits a sharp drop that can be used
Figure 2.1: Process for producing dimpled fiber-taper probes. (a) Taper pulling with the fiber placed at the edge of the flame to produce a small hot zone. (b) Characteristic in situ down-sampled transmission traces for 1060-XP and SMF-28e fiber during a taper pull. The running standard deviation for these traces shows the onset of single-mode behavior. (c) Forming the dimple by heating taper while it is pressed against a mold fiber. (d) Detaching the probe from the mold. (e) Schematic of a U-mount with variable tension.
to trigger the end of the pull.\textsuperscript{1} To produce tapers with the desired length ($\sim1$ cm for the region with $d \lesssim 10 \mu m$) and minimum diameter ($\sim1 \mu m$), typical pulls continue $\sim10$ s after the oscillations stop and take 120–150 s with the stages each moving at $\sim1.5$ cm/min, but there is significant variability depending on how the fiber is prepared. For instance, the single-mode jump in the running standard deviation will contain one or more discrete steps if the fiber is subjected to undue stress while the polymer coating is being removed.

After mounting the taper in a U-bracket \textsuperscript{104}, the narrowest part of the taper is pressed against a silica mold with the desired radius of curvature [Fig. 2.1(c)]; a bare optical fiber with a diameter of approximately 125 $\mu m$ is used as the mold in these experiments. The taper and mold are heated with a hydrogen torch and allowed to cool. To detach the taper from the mold, we simply pull the taper away slowly while moving it back and forth (i.e., we are inducing fatigue failure at the taper-mold joint). A thin layer of soot on the mold improves the detachment yield—an appropriate layer is usually deposited when the torch is sparked.\textsuperscript{2} After releasing the fiber from the mold, the taper retains an impression of the mold, Fig 2.2(b), which forms a global minimum with respect to the rest of the taper. The dimpling process introduces negligible additional loss, and the total loss of the dimpled taper can be less than 0.5 dB relative to the unpulled optical fiber. However, tapers typically exhibit a total loss 1–2 dB. Using a specially designed U-mount with a set screw to control the tensioning, varying the taper’s tension changes the radius of curvature of the dimple. Under high tension, the dimple becomes very shallow but never completely straightens. After dimpling, the probe is mounted onto a three-axis 50-nm encoded stage and is fusion spliced into a versatile fiber-optic setup. During testing, devices are placed in the near field of the probe, as in Fig. 2.2(a,c); adjustments to a pair of goniometers ensure the straight run of the taper is parallel to the sample surface.

Measurement of the non-resonant insertion loss as the waveguide is moved relative to nearby semiconductor microstructures gives the effective interaction length and profile of the local probe. First, we record the loss as a 1.6-$\mu m$ wide GaAs cantilever is scanned along the taper’s length ($\hat{x}$-direction) while holding the taper at a fixed height. At tensions used in standard testing, Fig. 2.3(a) shows only $\sim20 \mu m$ (full width at half maximum) of the taper

\textsuperscript{1}This method was implemented by C. Chrystal. A similar procedure has been published by F. Orucevic et al. using a running Fourier transform \textsuperscript{103}.

\textsuperscript{2}N.B. Do not spark H\textsubscript{2} torches near the fiber either before or after tapering as it will significantly reduce the transmission of the final coupler.
at the bottom of the dimple is close enough to interact with the sample. Second, the loss is measured as a function of the probe’s height ($\hat{z}$-direction) above a 11.6-µm wide GaAs mesa. By assuming an exponential vertical dependence for the insertion loss $L \propto e^{-z_t(x)/z_o}$ where $z_t(x)$ is the probe’s “near-field” profile and $z_o$ is the decay length from Fig. 2.3(b), we convert the axial dependence of the loss [Fig. 2.3(a)] into $z_t(x)$ [Fig. 2.3(c)]—i.e., the height of the taper relative to the lowest point of the dimple. Since only the lowest part of the dimple interacts with the sample, this method can only determine the taper’s profile within $\sim 1.25 \, \mu m$ of the surface. Fitting the profiles determines the effective probe radius to be 159, 228, and 498 µm at low, medium, and high tension, respectively. These radii differ from the mold radius ($\sim 62 \, \mu m$) due to tensioning of the taper and how the fiber detaches from the mold after heating.

### 2.2 Microphotonic Testing with a Dimple-Taper Waveguide

To study microresonators with a dimpled probe and the general arrangement in Fig. 2.4, precision stages are used to move the waveguide into the device’s near field giving controllable and reproducible coupling. Isolating the measurements from stray air currents is imperative, and all testing is conducted in a continuously $N_2$-purged enclosure. The $N_2$ environment also extends the probe’s lifetime since fiber-tapers are hygroscopic and can suffer significant
Figure 2.3: Non-resonant insertion loss (a) as a function of axial position ($\hat{x}$) as a narrow cantilever is moved along the taper length and (b) as a function of transverse position ($\hat{z}$) as the dimple is raised above a mesa. (c) Inferred dimple taper profile at “low,” “medium,” and “high” tension.

Figure 2.4: Basic arrangement for microcavity transmission measurements. An ancillary normalization detector can be used to remove the oscillations produced by the residual etalon between the laser diode’s facets. A fourth laser (spanning 1480–1580 nm) is used to study any resonances at the boundaries of the primary lasers’ sweep ranges (i.e., near 1495 and 1565 nm). Abbreviations: variable optical attenuator (VOA), fiber polarization controller (FPC), Mach-Zehnder interferometer (MZI), and photodiode (PD).
losses by absorption at the first overtone of the OH\textsuperscript{−} vibrational modes [105, 106].\textsuperscript{3} The HE\textsubscript{11} mode of the taper is excited using fiber-coupled swept tunable-laser sources (spanning 963–993 nm and 1420–1625 nm, linewidth of <300 kHz over the 25-ms timescale used to scan across a high-\(Q\) resonance) and a paddle-wheel polarization controller to selectively couple to the cavity’s TE-like and TM-like WGMs. To accurately measure the intrinsic quality factor, the cavities are weakly loaded by the dimpled probe to ensure the extrinsic loading is negligible (\(\gamma_\text{e} \approx 0\)) and avoid any parasitic coupling. Two complementary VOAs are used to maintain a constant optical power at the photodetector (to give a constant electronic noise level) while power at the device can be varied up to 60 dB [107]. To eliminate any nonlinear effects or absorption saturation while acquiring the transmission spectra, the lasers are usually attenuated to give \(\sim 200\) nW at the taper, of which \(\sim 10\%\) is coupled into the microdisk cavity. Without any optical amplification, the signal is detected using a high-speed InGaAs photodiode, electrically amplified using a low-noise analog preamplifier, and then is acquired by a analog-to-digital converter. For narrow spectral features (e.g., \(Q > 10^6\) corresponds to \(\delta \lambda \lesssim 1.5\) pm), the linewidth measurement can be calibrated with a fiber-optic Mach-Zehnder interferometer to an accuracy of \(\pm 0.02\%\) (see also §A.1).

2.2.1 Noise Measurements

Because evanescent coupling to fiber-tapers is exponentially dependent on position, fiber-taper measurements are very susceptible to any noise sources that produce physical displacements of the taper. For straight tapers, increasing tension to reduce these fluctuations is common, and the U-mount [104] naturally provides the appropriate tautness. Under standard testing conditions at low dimple-taper tension [Fig. 2.5(a)], coupling to the mode of a microdisk resonator [see Fig. 2.2(a) and §2.2.2] varies significantly between consecutive scans. Increasing the tension makes the coupling depth much more reproducible, as in Fig. 2.5(b). At tensions that give acceptable noise levels, the depth of the dimple is still adequate for testing densely spaced planar devices.

To quantitatively study the noise, we measure non-resonant insertion loss as a function of time. The dimple is placed above the etched GaAs mesa so that approximately 60\% of the incident power is coupled into the substrate \([z \approx 170\) nm as in Fig. 2.3(b)]. The

\textsuperscript{3}After months of use in a N\textsubscript{2} box, enough water and dust will accumulate to significantly reduce the taper’s transmission. Often this loss can be partially ameliorated by dipping the taper in trichloroethylene (TCE).
Figure 2.5: Reducing noise through higher taper tension. Without averaging multiple scans, ten consecutive traces of a microdisk “doublet” resonance [42] display greater displacement noise at low tension (a) than at higher tension (b). (c) Noise power spectra with the same tensions as in Fig. 2.3(a,c).

The mesa structure assures a constant 11.6-µm interaction length for different taper tensions. We minimize the electrical noise contribution by maximizing the incident optical power in order to decrease the needed electrical gain and filtering. We also eliminate extraneous noise sources (unused computers, monitors, overhead lights, etc.) and turn off the N₂ flow into the testing enclosure. To obtain a background spectrum that is independent of any taper displacement, the dimple is raised so no power is coupled into the substrate, and then the power is attenuated to give the same output voltage from the detector. The resulting noise power spectra in Fig. 2.5(c) reveal increased tension reduces broadband noise between approximately 10 and 1000 Hz, reflecting the relevant timescales for scanning across a high-Q resonance. The series of high-frequency peaks at ∼15.8 kHz occur at the pulse-position-modulation clock frequency of the stage motor controller. The dominant spike at low frequencies is bimodal with peaks at ∼120 Hz and ∼130 Hz with a total bandwidth of ∼20 Hz. The motor controller also contributes to noise in this band, but it is not the dominant noise source. We hypothesize that electrical noise actuates the motors and drives low-Q vibrational modes of the fiber-taper. By measuring insertion loss as a function of the dimple-substrate gap [Fig. 2.3(b)] and comparing it to noisy time-domain transmission traces under low tension, we estimate the upper bound on fluctuations in the taper height to be 7.9 ± 1.4 nm, which is consistent with our earlier measurements with straight tapers.

Unlike axially symmetric straight fiber-tapers, dimpled tapers are affected by external torques that drive torsional modes of the dimple about the straight portion of the taper. Measurements of dimpled tapers resonantly coupled to high-Q microcavities (Q ≈ 10⁶)
indicate optical forces can statically deflect the dimple up to 200 nm for 600 µW of power coupled into the cavity mode. Even with lower-\(Q\) optical resonances (\(Q \approx 10^2\)), optical forces from similar amounts of dropped power can drive mechanical oscillations with amplitudes of \(\sim 500\) nm. These results are discussed in detail elsewhere [108].

### 2.2.2 Dense Si Microdisk Array

To demonstrate the dimpled taper’s ability to test closely spaced devices, we study a 5×5 array of silicon microdisks [Fig. 2.2(c)] with disk diameters of 10 µm and periodicity of 20 µm—corresponding to an areal density of 2.5×10⁵ cm⁻². Undercut microdisks were chosen over planar resonators to ease phase matching between the cavity and taper modes. The microdisks were fabricated from silicon-on-insulator with a 217-nm device layer [(100) orientation, p-type, 14–20 Ω·cm] and a 2-µm SiO₂ buried oxide layer (BOX). The resonators were defined using electron-beam lithography, resist reflow, and C₄F₈:SF₆ reactive ion etching (RIE); then the disks were partially undercut by etching the buried oxide using dilute HF. The silicon surfaces are temporarily hydrogen passivated using repeated Piranha/HF treatments. Long-term passivation is achieved using a 3-nm dry thermal oxide cap grown in O₂ at 1000°C followed by a 3-hour anneal in N₂ at 1000°C and then a 1.5-hour slow cool down in N₂ from 1000°C to 400°C. For details on the lithography, chemical passivation, and oxide passivation, see Refs. [42], [56], and [109], respectively.

Near 1532 nm, we track three TE-like modes of different radial orders [\(p = 1–3\) in Fig. 2.6(a)] across all 25 disks in the array. One disk supported no high-\(Q\) whispering-gallery modes in the range spanning 1495–1565 nm, and we were unable to couple to the TE \(p = 1\) mode in two other disks—most likely because their \(Q\) was too low to overcome the phase mismatch with the taper mode. In Fig. 2.6(b), varying the disk-taper coupling through their separation practically demonstrates the level of displacement noise present in these measurements; each circle represents the transmission minimum for an individual scan at the given probe position. Table 2.1 summarizes the average measured wavelength (\(\lambda_o\)), quality factor, and doublet splitting (\(\Delta\lambda\)) for each mode; the distributions of wavelength and quality factor⁴ appear in Fig. 2.6(a). The highest \(Q\) for a single standing wave mode is 3.3×10⁶ with \(Q/V_{eff} = 2.3 \times 10^5\) for a calculated mode volume \(V_{eff} = 14.09 (\lambda/n)^3\). With minimal free-

---

⁴For doublet modes, the quality factor used in Fig. 2.6(c) is the average \(Q\) between the two standing-wave modes.
Figure 2.6: (a) Sample transmission spectrum for a single microdisk. (b) Coupling dependence on the disk-taper gap for a TE $p = 3$ mode of the device in Fig 2.2(c). (c) Distribution of wavelengths and quality factors for the TE $p = 1$–3 modes near 1532 nm. The solid diamond indicates the mode tested in (b). Spatial distribution for the (d) wavelength, (e) quality factor, and (f) doublet splitting of the TE $p = 1$ modes.

<table>
<thead>
<tr>
<th>Mode</th>
<th>Observed</th>
<th>$\lambda_0$ (nm)</th>
<th>$Q$</th>
<th>$\Delta\lambda$ (pm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TE $p = 1$</td>
<td>22/25</td>
<td>1531.008 ± 1.487</td>
<td>(1.73±0.93)$\times10^6$</td>
<td>11.31 ± 10.12</td>
</tr>
<tr>
<td>TE $p = 2$</td>
<td>24/25</td>
<td>1531.393 ± 1.508</td>
<td>(3.95±1.32)$\times10^5$</td>
<td>10.93 ± 5.60</td>
</tr>
<tr>
<td>TE $p = 3$</td>
<td>24/25</td>
<td>1532.429 ± 1.489</td>
<td>(2.19±0.70)$\times10^5$</td>
<td>10.70 ± 5.77</td>
</tr>
</tbody>
</table>

Carrier absorption in the bulk, the modal loss likely has significant contributions from both surface absorption and surface scattering since the ratio of the doublet splitting (related to the surface scattering rate) over the resonance linewidth varies from 3.1 to 28.1 for modes with $Q > 10^6$. The spatial arrangement of the mode parameters across the array [Fig. 2.6(d–f)] shows a systematic change in $\lambda_0$ and more random variations in $Q$ and $\Delta\lambda$. The $\lambda_0$ distribution implies the sample was slightly tilted with respect to the beam writer’s focal plane. Similar geographic patterns exist for the parameters of the $p = 2$ and $p = 3$ modes.

For silicon wafers with 14–20 Ω·cm resistivity, bulk free-carrier absorption [110] limits microcavities to $Q < 9\times10^7$–1.4$\times10^8$ at $\lambda_0 = 1532$ nm.
2.2.3 Planar Si-SiO\textsubscript{2} Microring via RIE Processing

Testing planar devices is accomplished in the same fashion. Non-undercut microring resonators, shown in Fig. 2.7(a), were fabricated from SOI with a 195-nm silicon device layer and a 3-\(\mu\)m BOX. The same lithography, resist reflow, and RIE procedure used for the microdisks [42] was used to define the microrings although without the final HF undercut. Repeated Piranha oxidations and HF dips are again used to chemically passivate the surfaces prior to thermal oxidation [56]; these treatments also slightly undercut the resonators [Fig. 2.7(b)]. Finally, a 30-nm dry thermal oxide was grown as a capping layer, and the microring sample was annealed according to the same \(N_2\) schedule as the microdisks [109].

Microrings are slightly more difficult to test with fiber-tapers than undercut microdisks. A large phase mismatch exists between the taper and microcavity because of the extra dielectric beneath the Si-core guided modes. With the taper in contact with the ring, the coupling depth is more than sufficient to assess the devices’ optical loss characteristics. However, the coupling is not adequate to efficiently excite and collect emission from active devices [70]. For applications requiring high pump/collection efficiency, photonic crystal waveguides can be used to overcome the poor phase matching between the modes in the taper and the modes in the on-chip device [77].

Figure 2.7(c) shows a transmission spectrum of a ring with an 80-\(\mu\)m diameter and 2-\(\mu\)m width after the final chemical treatments and thermal oxidation. The measured quality
factor of $4.8 \times 10^6$ (loss coefficient $\alpha < 0.15 \text{dB/cm}$) represents the highest quality factor for any planar microresonator to date. Reproducing $Q$s found previously only in relatively thick and undercut silicon disks [42] is promising for the future development of PLCs with high-$Q$ silicon microresonators integrated with bus waveguides.

### 2.2.4 Planar Si-SiO$_2$ Microring via LOCOS Processing

Wet chemical oxidation of ridge waveguides has been used to smooth silicon surfaces and reduce Rayleigh scattering [111]. However, similar treatments of high-$Q$ microdisks did not reduce the scattering rate but rather increased the measurable loss by introducing unpassivated mid-gap surface states [56]. As evident by the contradictory results for chemical oxides, significant care must be taken in low-loss fabrication procedures. Dry reactive-ion chemistries are generally preferred over wet chemical etchants because they are more anisotropic, but RIE recipes must be carefully optimized for roughness and can damage the surface (chemically and physically). Two alternatives have been to use shallow rib waveguides, in which the modes are weakly guided and have minimal overlap with the surfaces [112], or to use a diffusion-limited chemical reaction to produce the cladding layer [113]. In the latter case, dry thermal oxidation of Si device layer in SOI produces a low-loss SiO$_2$ cladding with a well-passivated interface [109] and high refractive index contrast to the underlying Si core. This section describes initial results for the fabrication and dispersion engineering of submicron waveguides formed by the local oxidation of silicon (LOCOS).

With the ultimate objective of achieving CW parametric oscillation in a Si microcavity, the final devices must be high $Q$, have a small group velocity dispersion (GVD) with a zero-dispersion wavelength (ZDWL) near the desired wavelength, and exhibit a low free-carrier lifetime ($\tau_{fc}$). Degenerate four-wave mixing can be used to generate broadband gain in silicon microphotonics because of its large $\chi^3$ nonlinearity [114], but the modal dispersion must be controlled to ensure phase matching. In WGM resonators, the phase matching is trivially satisfied between consecutive modes in a single family, but the GVD must be near zero to converse energy (i.e., the signal and idler modes must be equally spaced about the pump mode) [115]. An additional constraint when operating with a photon energy above half the material’s bandgap is that $\tau_{fc}$ must be short to minimize absorption by $e^-/h^+$ pairs generated by two-photon excitation.

The primary design goals for these initial devices are 1) $n_{eff}^{wg} \approx 2.0$ so losses can rapidly
be assessed with the dimple taper, 2) an exposed top Si surface to control $\tau_{fc}$ with passive chemical treatments [116], 3) to only partially consume the unmasked Si layer to allow cavity tuning via electrical injection of free carriers, and 4) to explore the dependence of the GVD on the waveguide dimensions [117]:

$$\text{GVD} \equiv \frac{\partial}{\partial \omega} \left( \frac{1}{n_g} \right) = \frac{\partial^2 k}{\partial \omega^2}.$$

Before the oxidation in the LOCOS procedure, a SiN$_x$ layer is deposited on the Si surface, lithographically patterned, and then etched [Fig. 2.8(a)]. During the subsequent oxidation, the SiN$_x$ acts as a local diffusion barrier and masks the desired area while lateral diffusion
rounds the Si profile at the edges [Fig. 2.8(b)]. Since the resulting “bird beak” contour may not be accurately captured by effective-index methods, the waveguide modes and dispersion are evaluated using the FEM. The effective-index requirement limits the geometries to submicron dimensions that cut off all but the $\text{TE}_{p=1}$ mode [Fig. 2.8(c)]—TM modes will be avoided initially because of lateral leakage into the partially thinned slab [26]. Preliminary simulations show the dispersion can be flattened by modifying the waveguide profile [Fig. 2.8(d)], but the GVD for modes with these dimensions may never reach zero in the telecommunications windows. Smaller Si structures having greater overlap with the BOX or a SiO$_2$ cap will have a lower GVD because silica is anomalous in this range. To proceed, we target a waveguide width of 230 nm, Si thickness of 210 nm, and a ridge height of 160 nm with $n^{\text{wg}}_{\text{eff}} = 2.02$ at $\lambda_0 = 1550$ nm. While ring resonators with these dimensions have significant $+\text{GVD}$ at 1550 nm, they are an acceptable starting point for process development. Future devices may utilize the $\text{TM}_{p=1}$ modes to simultaneously achieve the $n_{\text{eff}}$ and GVD objectives while demanding greater pattern control (i.e., waveguide width).

The process begins by using Piranha etch (3:1 concentrated H$_2$SO$_4$ to 30% H$_2$O$_2$) to remove any residual organics on a SOI wafer (217 nm Si, 2 $\mu$m BOX). After a quick HF rinse to strip the native oxide, a 10-nm dry thermal oxide is grown at 1000$^\circ$C to serve as an etch-stop layer and protect the Si surface. A 75-nm low-stress SiN$_x$ mask is then deposited by plasma-enhanced chemical vapor deposition (PECVD). This layer must be thick enough to provide an adequate diffusion barrier, but it must be compliant (thin) enough to not delaminate as the SiO$_2$ expands [Fig. 2.8(b)]. The nitride is then patterned with electron-beam lithography, resist reflow, and RIE (C$_4$F$_8$:SF$_6$). Since the selectivity between SiN$_x$ and SiO$_2$ is poor, the etch must be carefully timed to stop at the etch-stop layer. Following another Piranha clean to take off any remaining resist and etch by-products, the sample is oxidized in dry O$_2$ (electronic grade II) at 1100$^\circ$C to define the waveguides. The SiO$_2$ and SiN$_x$ are then removed with HF leaving the Si surface bare for future surface treatments.

Initial results show reasonable smoothness for the waveguide sidewalls in Fig. 2.8(e), but there is significant room for improvement throughout the fabrication flow. Since the exact process conditions vary between SiN$_x$ etches (rate $\approx$ 1.5 nm/s), it is difficult to consistently halt the RIE before penetrating the SiO$_2$ layer. Rather than adjusting the etch recipe, which should be optimized for roughness, the initial SiO$_2$ thickness should be increased even though this increases the duration of the LOCOS step. Second, the written waveguide
dimensions must be calibrated against the resulting widths to account for the resist reflow and SiN$_x$ sidewall angle, which is not perfectly vertical. Even after correcting for lateral diffusion that will undercut the SiN$_x$ by the same depth the sample is oxidized (e.g., oxidizing down 160 nm to define a 230-nm waveguide requires a mask width of $\sim$550 nm), the final waveguides were too large ($\sim$500 nm) for testing with a dimple taper. Third, the long LOCOS oxidation anneals the PECVD SiN$_x$ and increases its resistance to HF. Instead of being etched during the final HF strip, the SiN$_x$ lifts off in one large sheet as the thin SiO$_2$ etch-stop is slowly undercut. Much of surface roughness on top of the waveguides is likely due to the SiN$_x$ gradually pealing off the Si. In addition, the longer etch allows HF to penetrate defects in the silicon/oxide layers and partially undercut the BOX layer; it is uncertain whether these defects are process induced or are intrinsic to the Si layer in the original SOI wafer. A thicker SiO$_2$ etch barrier will also ameliorate these issues by expediting the liftoff. To test ring resonators from this initial run, we partially oxidize the samples to reduce $n_{\text{eff}}$ — the waveguide now consist of a smaller Si core with a SiO$_2$ cladding. These Si-SiO$_2$ rings have $Q \approx 1-2 \times 10^5$ (3.6–7.2 dB/cm) and no measurable doublet splitting. Further oxidation completely consumes the Si core, but physical inspection in Fig. 2.8(f) shows the surface roughness is reduced similar to the chemical oxidation in Ref. [111]. These preliminary results along with recent reports [118,119] show the potential for achieving low-loss dispersion-engineered Si-SiO$_2$ microrings using LOCOS processing.

### 2.3 Conclusions

Using a dimpled fiber-taper waveguide, we have demonstrated a localized optical probe capable of testing dense arrays of planar devices. Proper tensioning makes the dimpled taper more robust against fluctuations in position and decreases broadband noise. For research and development, the dimple geometry facilitates rapid prototyping and cuts processing time by $\sim$40% for cavities that are to be probed using fiber-taper waveguides. Even without dedicated test structures to ease phase-matching constraints, the local dimpled-taper probe enables nondestructive wafer-scale optical characterization for manufacturer-level statistical process control. Higher yields through low-cost testing will become increasingly important in a growing market where the burgeoning demand for bandwidth is making integrated micro-electronic-photonic solutions more attractive [120].
Chapter 3

Linear Absorption in AlGaAs and GaAs Microcavities

In recent semiconductor cavity QED experiments involving self-assembled III-V quantum dots (QDs), Rabi splitting of the spontaneous emission line from individual QD excitonic states has been measured for the first time [121–123]. Potential application of these devices to quantum networks [124] and cryptography [125] over long-haul silica fibers has sparked interest in developing QD-cavity systems with efficient light extraction operating in the telecommunication bands at 1300 and 1550 nm [126]. The initial demonstrations of vacuum-Rabi splitting in this system, a result of coupling a single QD to localized optical modes of a surrounding microresonator, have been greatly aided by prior improvements to the design and fabrication of semiconductor microcavities [23, 127, 128]. At the shorter wavelengths involved in these Rabi splitting experiments (740–1200 nm), the optical quality factors \( Q \) of the host AlGaAs microcavities were limited to \( Q \approx 2 \times 10^4 \) corresponding to a loss rate comparable with the coherent QD-cavity coupling rate. Further reduction of optical loss would increase the relative coherence of the QD-cavity system and would allow greater coupling efficiency to the cavity mode.

In previous measurements of wavelength-scale AlGaAs microdisk resonators, we have demonstrated \( Q \)-factors up to \( 3.6 \times 10^5 \) in the 1400 nm band [129] and attributed the improved performance to an optimized resist-reflow and dry-etching technique, which produces very smooth sidewalls [42]. Subsequently, we have also measured Al\(_{0.3}\)Ga\(_{0.7}\)As microdisks with similar quality factors between 1200 and 1500 nm; however, these disks exhibit a significant unexpected decrease in \( Q \) at shorter wavelengths (\( \lambda_0 \approx 852 \text{ nm} \)) [130]. In related work on silicon microdisks, methods have been developed to specifically measure and character-
ize losses due to material absorption and surface scattering \cite{42,56,69}. In this chapter\footnote{Reprinted with permission from Appl. Phys. Lett. 90, 051108 (2007). Copyright 2007, American Institute of Physics. The III-V material was grown by H. Kim and processed by K. H. Lee with K. Hennessy at UCSB.} we study the properties of GaAs and Al$_{0.18}$Ga$_{0.82}$As microdisks across three wavelength bands centered at 980, 1460, and 1600 nm. After estimating and removing the surface-scattering contribution to the cavities losses, we find the remaining absorption, composed of losses in the bulk and on the surfaces, depends significantly on both wavelength and material composition.

3.1 Al$_x$Ga$_{1-x}$As Samples and Processing

The samples were fabricated from high-quality heterostructures grown by molecular beam epitaxy (MBE) on a GaAs substrate. Two different samples were grown: a "GaAs" sample containing a 247-nm GaAs disk layer, and an "AlGaAs" sample with a 237-nm Al$_{0.18}$Ga$_{0.82}$As disk layer. In both samples the disk layer was grown nominally undoped (background doping levels $n_p \lesssim 10^{15}$ cm$^{-3}$) and was deposited on a 1.6-$\mu$m Al$_{0.7}$Ga$_{0.3}$As sacrificial layer. Microdisks with a radius of \~{}3.4-$\mu$m were defined by electron-beam lithography and etched in a 55 percent (by volume) HBr solution containing 3.6 g of K$_2$Cr$_2$O$_7$ per litre \cite{131}. The disks were partially undercut by etching the sacrificial layer in 8 percent HF acid for 45 s, prior to e-beam resist removal.

The measured $Q_i$s for all observed modes are summarized in Fig. 3.2(a); each family of modes is identified by comparing the coupling behavior and free spectral range to FEM models \cite{56}. Modes with $Q_i$s dominated by radiation loss, i.e., the measured $Q_i$ is near $Q_{\text{rad}}$ calculated using FEM simulations, are omitted. For all the microdisk modes in these measurements, the calculated $Q_{\text{rad}}$ is $\gtrsim 10^6$ and typically is $>10^8$. In the 1600-nm band, the high-$Q$ TE modes are $p = 1$–4 in GaAs and $p = 1$–3 in AlGaAs; all TM modes are radiation limited in this band. Near 1460 nm, the $TE_{p=1}$ and $TM_{p=1}$ modes in both materials are detectable and not radiation limited. In the 980-nm range, identifying modes becomes more difficult: at this wavelength families through $TE_{p=8}$ and $TM_{p=7}$ are not radiation limited, and significant spectral overlap between the modes causes Fano-like resonance features \cite{71}. In addition, we are unable to couple to the lowest order modes of both polarizations ($p \approx 1$–3) because they are poorly phase matched to the fiber-taper.
Figure 3.1: (a) SEM image of a typical GaAs microdisk. (b) Sample scan of a TE$_p=4$ doublet in a GaAs disk at $\lambda_0 = 1582.796$ nm. Fitting the data gives the linewidth $\delta \lambda = 4.57$ pm and the splitting $\Delta \lambda = 20.20$ pm. Although not visible in (a), the edge roughness is different for the (c) GaAs and (d) AlGaAs samples.

Figure 3.2: Measured (a) $Q_i$ and (b) $\Delta \lambda$ for the (+, ×) GaAs TE- and TM-polarized microdisk modes and (○, □) Al$_{0.18}$Ga$_{0.82}$As TE and TM modes, respectively. In (c) and (d), connected points represent calculated bounds on (c) $Q_{ss}$ and (d) $Q_a$ for each family of modes. The data were compiled from two disks of each material.
3.2 Estimation of Scattering and Absorption Rates

Despite efforts to produce perfectly smooth side walls, Figs. 1(c,d) indicate that significant surface roughness is still present and produces significant mode splitting ($\Delta \lambda$), shown in Fig. 3.1(b). Following the theory developed in Refs. [42] and [41], $\Delta \lambda$ and $Q_{ss}$ are both dependent on the characteristic volume of the scatterer ($V_s$):

$$\Delta \lambda = \frac{\pi^{3/4}}{\sqrt{2}} \lambda_0 V_s (n_d^2 - n_o^2) \sum_\eta \pi_s(\hat{\eta}), \quad (3.1)$$

$$Q_{ss} = \frac{\lambda_0^3}{\pi^{7/2} n_o (n_d^2 - n_o^2) V_s^2 \sum_\eta \pi_s(\hat{\eta}) G(\hat{\eta})}, \quad (3.2)$$

where $n_d$ and $n_o$ are the indices of refraction of the disk and surrounding medium, respectively, $\pi_s(\hat{\eta})$ is the spatially-averaged $\hat{\eta}$-polarized normalized electric field energy at the $s$th surface (i.e., the disk edge):

$$\pi_s(\hat{\eta}) = \frac{\epsilon_0 \left| \vec{E}(\vec{r}) \cdot \hat{\eta} \right|^2_{s,\text{avg}}}{\frac{1}{2} \int n^2(\vec{r}) \epsilon_0 \left| \vec{E}(\vec{r}) \right|^2 dV}, \quad (3.3)$$

and $G(\hat{\eta}) = \{2/3, 2, 4/3\}$ is a geometrical factor weighting the radiation contribution from the $\hat{\eta} = \{\hat{r}, \hat{\phi}, \hat{z}\}$ polarizations. The mode field profiles are calculated by the FEM. For FEM models in the 980-nm span, we treat all measured TE (TM) modes as TE$_{p=7}$ (TM$_{p=6}$) because the appropriate field parameters do not vary significantly between radial orders.

We employ two separate measurements to find rough bounds on $Q_{ss}$.$^2$ First, we use the average doublet splitting for each family to find the average $\langle V_s \rangle_p$ sampled by each mode and then calculate the $Q_{ss}$ associated with each family.

**Splitting Method:** $\Delta \lambda \Rightarrow \langle V_s \rangle_p \Rightarrow Q_{ss}$

Second, we statistically analyze the roughness of the disk edges in high resolution SEM images, as in Fig. 3.3 [69]. Fitting the autocorrelation of the roughness to a Gaussian, the roughness amplitude ($\sigma_r$) and correlation length ($L_c$) give the “statistical” scatterer volume ($\overline{V}_s = \sigma_r t \sqrt{RL_c}$ where $t$ and $R$ are the disk’s thickness and radius) for each disk, which is

$^2$Since this work was completed, a more direct method of measuring $\gamma_a$ has been developed utilizing a cavity’s thermo-optic bistability [107].
Figure 3.3: (a) Sample surface roughness measurement on a AlGaAs microdisk showing the disk edge extracted from a high-resolution SEM image and a fit to a circle. (b) Autocorrelation of the disk roughness (i.e., difference from perfectly circular edge); the Gaussian fit gives $\sigma_r = 2.4\,\text{nm}$ and $L_r = 27.8\,\text{nm}$. Since the resist was not reflowed prior to etching, the (anti)correlation at larger displacements may be due to raster effects during the beam write.

used to estimate $Q_{ss}$.

**Statistical Method:** $\sigma_r, L_r \Rightarrow V_s \Rightarrow Q_{ss}$

The average $\{\sigma_r, L_c\}$ for the GaAs and AlGaAs disks are $\{0.6, 38.7\}\,\text{nm}$ and $\{1.8, 29.4\}\,\text{nm}$, respectively. Because each mode will not sample all of the disk’s physical irregularities, the roughness estimated by the statistical analysis is slightly greater than the roughness calculated from the doublet splittings. Hence, the doublet splitting places an upper bound and more accurate value for $Q_{ss}$ [Fig. 3.2(c)]. The statistical analysis gives a lower bound, although neither bound is strict in the theoretical sense.

Through Eq. (1.24), $Q_{ss}$ and $Q_{rad}$ are removed from the measured $Q_i$ to obtain limits on $Q_a$ [Fig. 3.2(d)]. To relate cavity losses to material properties, the material absorption rate ($\gamma_{a,p}$) for the $p$th mode is given by $\gamma_{a,p} = 2\pi c/\lambda_o Q_a$. We weight each measured doublet equally and average over all families in a band to determine an average $\overline{\gamma}_a$. Table 3.1 compiles the average absorption rates for both GaAs and Al$_{0.18}$Ga$_{0.82}$As across the three wavelength ranges. The average rates are 540 percent larger at 980 nm than at 1600 nm and 80 percent greater in AlGaAs than in GaAs.

The measured absorption may be due to a number of sources. Although nonlinear-absorption-induced optical bistability was measured for internal cavity energies as low as
Table 3.1: Summary of GaAs and Al$_x$Ga$_{1-x}$As material absorption rates.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Method</th>
<th>Splitting</th>
<th>Statistical</th>
</tr>
</thead>
<tbody>
<tr>
<td>GaAs ($\lambda \approx 980$ nm)</td>
<td>all modes</td>
<td>3.47 ± 0.59</td>
<td>3.29 ± 0.75</td>
</tr>
<tr>
<td></td>
<td>TE</td>
<td>2.94 ± 1.24</td>
<td>2.61 ± 1.28</td>
</tr>
<tr>
<td></td>
<td>TM</td>
<td>4.08 ± 0.91</td>
<td>4.06 ± 0.89</td>
</tr>
<tr>
<td>AlGaAs ($\lambda \approx 980$ nm)</td>
<td>all modes</td>
<td>5.84 ± 0.13</td>
<td>4.00 ± 0.91</td>
</tr>
<tr>
<td></td>
<td>TE</td>
<td>5.77 ± 1.61</td>
<td>3.44 ± 2.31</td>
</tr>
<tr>
<td></td>
<td>TM</td>
<td>6.03 ± 1.56</td>
<td>5.39 ± 1.73</td>
</tr>
<tr>
<td>GaAs ($\lambda \approx 1460$ nm)</td>
<td>all modes</td>
<td>0.94 ± 0.696</td>
<td>0.888 ± 0.692</td>
</tr>
<tr>
<td></td>
<td>TE$_{p=1}$</td>
<td>0.514 ± 0.085</td>
<td>0.444 ± 0.108</td>
</tr>
<tr>
<td></td>
<td>TM$_{p=1}$</td>
<td>0.495 ± 0.089</td>
<td>0.467 ± 0.095</td>
</tr>
<tr>
<td>AlGaAs ($\lambda \approx 1460$ nm)</td>
<td>all modes</td>
<td>1.73 ± 0.50</td>
<td>1.43 ± 0.76</td>
</tr>
<tr>
<td></td>
<td>TE$_{p=1}$</td>
<td>1.32 ± 0.22</td>
<td>0.882 ± 0.380</td>
</tr>
<tr>
<td></td>
<td>TM$_{p=1}$</td>
<td>2.30 ± 0.49</td>
<td>2.39 ± 0.40</td>
</tr>
<tr>
<td>GaAs ($\lambda \approx 1600$ nm) – TE</td>
<td>0.507 ± 0.186</td>
<td>0.460 ± 0.185</td>
<td></td>
</tr>
<tr>
<td>AlGaAs ($\lambda \approx 1600$ nm) – TE</td>
<td>0.968 ± 0.179</td>
<td>0.629 ± 0.173</td>
<td></td>
</tr>
</tbody>
</table>

106 aJ, the losses reported in Table 3.1 were all taken at input powers well below the non-linear absorption threshold. Free carrier absorption can also be neglected given the nominally undoped material and relatively short wavelengths studied here [132]. The Urbach tail makes a small contribution in the 980-nm band ($\leq 15$ percent) and is negligible otherwise [132]. This leaves deep electron (hole) traps as the major source contributing to bulk material absorption in the observed resonances. Similar wavelength dependent absorption has been observed in photocurrent measurements of MBE-grown AlGaAs waveguides [133] and attributed to sub-bandgap trap levels associated with vacancy complexes and oxygen incorporation during growth [134]. Given the high surface-volume ratio of the microdisks, another possible source of loss is surface-state absorption. The sensitivity to absorption from surface states can be quantified by the $p$th mode’s energy overlap with the disk’s surface, $\Gamma'_p$; TM modes are more surface-sensitive than TE modes [56] whereas both polarizations are almost equally sensitive to the bulk. The calculated surface overlap ratio is $\Gamma'_TM/\Gamma'_TE \approx 2.65$ for $p = 1$ modes in the 1460-nm band, where all surfaces of the disk (top, bottom, and etched edge) are treated equally. For these modes the measured absorption ratio is $\gamma_{a,TM}/\gamma_{a,TE} = 1.74 \pm 0.47 \ (0.96 \pm 0.23)$ in the AlGaAs (GaAs) microdisks, which indicates the presence of significant surface-state absorption in the AlGaAs resonators and
dominant bulk absorption in the GaAs disks. In the 980-nm band, the data is consistent with bulk absorption $[\gamma_{a,\text{TM}}/\gamma_{a,\text{TE}} = 1.05\pm0.40 (1.39\pm0.66)$ for the AlGaAs (GaAs) devices] although the results are less conclusive due to the larger scatter in the data.

In summary, after accounting for radiation and surface scattering losses, we measure greater sub-bandgap absorption in Al$_{0.18}$Ga$_{0.82}$As microdisks than in similar GaAs resonators, and the absorption in both materials decreases toward longer wavelengths. This finding is at least partially reflected in recent cQED experiments with a single QD near or in the strong-coupling regime [121–123, 135–142] that show the longest cavity photon lifetimes ($\tau_{\text{ph}} \approx 70$ ps) are achieved in devices at longer wavelengths ($\lambda_o = 1294.5$ nm) [138].

From the polarization dependence of the measured optical loss, we infer that both surface states and bulk states contribute to the residual absorption in these structures. Our results imply that reductions in the optical loss of AlGaAs-based microphotonics, especially at the shorter wavelengths of <1 $\mu$m and in high Al content alloys, will require further study and reduction of deep level traps, and that surface passivation techniques [143, 144] will also likely be important.

---

$^3$Longer cavity lifetimes ($\tau_{\text{ph}} \approx 80$ ps at $\lambda_o = 940.5$ nm) have been demonstrated in large AlAs/GaAs micropillars ($d \approx 5$ $\mu$m) with thick DBR mirrors (>32 layer pairs on each side of the defect) and low eccentricity, but Rabi splitting has only been observed in smaller pillars with higher loss ($d = 3$ $\mu$m with $\tau_{\text{ph}} = 20$ ps) [139].
Chapter 4

Optical Processes in Epitaxial c-Er$_2$O$_3$ on Silicon

Significant progress in the last decade has been made developing passive and active silicon optical components; however, efficient generation of light within a Si platform remains a technical and commercial challenge [79]. Efforts to incorporate Er$^{3+}$ into the Si material system, with erbium’s emission in the 1550-nm telecommunications band, have met with limited success. Amorphous Er$^{3+}$-doped glass waveguides on Si provide insufficient gain (<4 dB/cm [145, 146]) for dense photonic integration, while doped silicon allotropes are limited by other effects such as Auger recombination [147] and free-carrier absorption [148].

Here we describe the characterization of stoichiometric single-crystal Er$_2$O$_3$ grown on Si by atomic layer epitaxy (ALE).\(^1\) We measure a peak resonant absorption of 364 dB/cm at 1535 nm, negligible background absorption (<3 dB/cm), and strong cavity-polariton effects. The observed radiative efficiency from 1520 to 1650 nm is 0.09% with cooperative upconversion producing strong green and red emission for Er$^{3+}$ excitation levels as low as 2%.

Spurred by the growing power consumption of high-speed electrical interconnects for multicore processors [149,150], optical networks have become an attractive option to achieve Tb/s on-chip bandwidth [120,151,152]. Following the initial demonstration of silicon waveguide devices [153], there has been significant development in adding optical functionality to silicon microelectronics and, similarly, applying the efficiency and infrastructure of modern complementary metal-oxide-semiconductor (CMOS) processing to optical telecommunication components. While silicon exhibits low loss across the 1300-nm and 1550-nm

\(^1\)This work was done in collaboration with Translucent, Inc., where these epitaxial rare-earth oxide films were developed and characterized (XRD, RHEED, TEM, etc.).
telecommunication windows, unstrained silicon lacks any significant Pockels coefficient and produces little emission from its 1.1 eV indirect bandgap [79]. Free-carrier dispersion and four-wave mixing provide some inherent active functionality, such as modulation with rates exceeding 1 GHz [154] and wavelength conversion [114], but considerable research, especially concerning light emission and detection, has focused on integrating silicon with other optical materials such as SiGe [155, 156] and the III-Vs [157, 158]. In this chapter we describe the growth, processing, and optical properties of c-Er$_2$O$_3$ films on Si. Similar to stoichiometric polycrystalline Er$^{3+}$ materials [159–164], Er$_2$O$_3$ allows for a 100-fold increase in Er$^{3+}$ concentration over conventional Er-doped glasses [165], making it an attractive material for on-chip emission and amplification in the 1550-nm wavelength band. Developed simultaneously for optoelectronic [166] and high-$\kappa$ dielectric [167] applications, epitaxially grown Er$_2$O$_3$ films can be incorporated into precisely controlled heterostructures and superlattices, which may also allow for efficient electrical injection. Oxides incorporating multiple cation species provide additional flexibility in designing the emission spectrum and dynamics as a number of rare-earth ions may be interchangeable in the R$_2$O$_3$ lattice [168]. Beyond application to chip-based optical networks, the strong cooperative upconversion within these films may also be used for visible light generation in solid-state lighting and displays and infrared-to-visible energy conversion in photovoltaics.

4.1 Rare-Earth Ions in the R$_2$O$_3$ Bixbyite Lattice

The electronic configuration for elements in the lanthanide series is [Kr]$4d^{10}4f^x$ where $x$ is the number of additional electrons filling the 4f shell. Since the valence shell is constant, these atoms are chemically similar and approximately the same size. Crystalline R$_2$X$_3$ compounds containing trivalent rare-earth ions (e$^-$ configuration: [Kr]$4d^{10}4f^{x-1}$) typically assume two crystal structures: the hexagonal La$_2$O$_3$ structure and the cubic (Fe,Mn)$_2$O$_3$ bixbyite structure. The hexagonal arrangement is more common among rare-earth oxysulfides and oxyselenides while the bixbyite crystal is prevalent in sesquioxides when $0.60 \lesssim r(R)/r(O) \lesssim 0.87$, where $r$ is the ionic radius. Bixbyite’s large 16-molecule unit cell in Fig. 4.1(a) contains 32 metal ions with 24 distributed on low-symmetry sites ($C_2$) and 8 on sites with inversion symmetry ($C_{3i}$). The bixbyite cell also contains 48 oxygen atoms arranged in an incomplete cubic close packing [169].
Figure 4.1: (a) The bixbyite unit cell for $R_2O_3$ with the ionic radii approximately correct for Er$_2$O$_3$. A few extra $C_2$ ions are included to help illustrate the crystal’s symmetry even though they are not in the unit cell. Detailed views of the oxygen coordination at the (b) $C_2$ and (c) $C_{3i}$ sites demonstrate inversion symmetry is only present at eight lattice positions. In each diagram, the cube faces identify the $\{100\}$ planes. (d) Crystal field levels studied on $C_2$ sites in $R_2O_3$ and lanthanide-doped $Y_2O_3$ crystals [170–174]. The emission wavelength is given for transitions back to the ground state, although no all transitions may fluoresce. The La$^{3+}$ and Lu$^{3+}$ ions are omitted because intra-$4f$ optical transitions are not possible when the $4f$ shell is empty and full, respectively. More complete spectra for ions in LaCl$_3$ are available in Ref. [175].
Since the 4f orbitals have less radial extent than the 5s and 5p, the 4f wavefunctions are only weakly perturbed by surrounding ligands which gives rise to atomlike spectra for intra-4f transitions in solid or liquid hosts. The collective state for the 4f electrons is described using the \((2S+1)L_J\) Russell-Saunders (RS) terms, where \(S\), \(L\), and \(J\) are the sum of the spin (\(S = |\sum_i m_{s,i}|\)), orbital (\(L = |\sum_i m_{\ell,i}|\)), and total angular momentum quantum numbers (\(J = |\sum_i (m_{\ell,i}+m_{s,i})|\)), respectively, for all electrons in the shell. Mutual Coulomb interaction between these electrons and spin-orbit coupling lift the degeneracy of these states, but the other filled shells do not contribute to first order because they are spherically symmetric. In Russell-Saunders coupling, the spin-orbit interaction is small compared to the Coulomb repulsion, and the system can be treated perturbatively. However, rare-earth ions are in the intermediate coupling regime where the Coulomb and spin-orbit terms are approximately equal. The resulting eigenstates are not pure RS wavefunctions but rather superpositions that are typically calculated by parametrizing the Hamiltonian in terms of tensor operators along with a few adjustable coefficients, diagonalizing, and then fitting to experimental spectra for free ions. Conventionally, the superpositions are still labeled as RS terms according to the state they reduce to with zero spin-orbit coupling—these states are generally the largest component of the superposition.

In crystals, the inhomogeneous electric field from neighboring ions (i.e., the crystal field) splits the \((2S+1)L_J\) terms of the free ion by breaking their degeneracy in \(M_J\), but the field does not appreciably shift the levels from the free-ion case. The spread of these Stark-shifted levels depends on the ions present and their coordination but is generally less than the spacing between RS terms (several hundred \(\text{cm}^{-1}\) compared to several thousand \(\text{cm}^{-1}\)). For Er\(^{3+}\) and other ions with an odd number of electrons, each RS term splits into \((J + \frac{1}{2})\) sublevels that are doubly degenerate (i.e., the Kramer’s degeneracy); the sublevels of the first three Er\(^{3+}\) terms \(\{^4I_{15/2}, ^4I_{13/2}, ^4I_{11/2}\}\) are indexed with \(\{Z_{1−8}, Y_{1−7}, X_{1−6}\}\), respectively. First principles calculation of the crystal-field splitting magnitude are ordinarily not possible; a point-charge model for the crystal along with the free-ion wavefunctions underestimates the splitting by \(10\times\) because the orbitals’ spatial extent and overlap must be considered. Again the free-ion Hamiltonian and crystal-field interaction are parametrized with tensor operators and fit to experimental data for transitions up to \(\sim30,000\text{ cm}^{-1}\); increasingly complex and subtle refinement reduces the error in the fits. Since transitions within a shell conserve orbital momentum, electric-dopole emission or
absorption can only occur when the $4f^x$ states mix with $4f^{x-1}n\ell$ states with opposite parity through odd components of lattice vibrations or of the crystal field. The odd-parity crystal fields are most significant to radiative transitions. Coupling to the phonon bath produces a small temperature-dependent line shift but is more important for line broadening and nonradiative processes. In the bixbyite crystal, ions at the $C_2$ sites experience a crystal field without inversion symmetry [Fig. 4.1(b)] and have a small induced electric dipole moment giving a radiative lifetime on the order of milliseconds. Without these odd interaction terms, only magnetic dipole and electric quadrupole transitions are possible in ions at the high symmetry $C_{3i}$ sites [Fig. 4.1(c)]. The optical spectra and crystal field splitting in Fig. 4.1(d) have been thoroughly analyzed for all rare-earth ions on $C_2$ sites in $Y_2O_3$ and many pure sesquioxides [173,174], while their intensities were investigated within the Judd-Ofelt formalism [176]. Some data has also been published for select transitions and ions on the $C_{3i}$ sites [177,178]. Thorough reviews of the optical processes in crystals containing rare-earth ions are available in Refs. [175,179].

4.1.1 Growth

As shown in Fig. 4.2, single-crystal stoichiometric $Er_2O_3$ can be grown via ALE on Si(111) or Si(100) on-axis wafers without a buffer layer. Like most lanthanide sesquioxides [180], the film’s bixbyite crystal structure is oriented along $Er_2O_3$((111)) [Er$_2$O$_3$(110)] on Si(111)$^2$ [Si(100)] and reasonably lattice matched to the Si substrate; see Table 4.1 [168]. To minimize both erbium silicide formation and native SiO$_x$ growth, the O$_2$:Er ratio during deposition is 1:5. High growth temperatures between 650 and 900°C result in more homogeneous films [measured by transmission electron microscopy (TEM) and X-ray diffraction (XRD)], smoother surfaces [~1 nm roughness by atomic force microscopy (AFM), TEM, and reflection high-energy electron diffraction (RHEED)], and stronger C-band$^3$ photoluminescence with a narrower linewidth at 1536 nm [Fig. 4.5(a)]. Films on Si(111) are consistently higher quality and have been grown up to 200 nm thick. We observe no evidence of erbium clustering using TEM or visible upconversion [35,181].

$^2$The $Er_2O_3$(111) orientation is rotated 180° about the Si(111) surface normal.

$^3$The short wavelengths (S), conventional (C), and long wavelengths (L) telecommunications windows (bands) are relative to the region of lowest optical loss in silica fiber ($\lambda \approx 1550$ nm) and occur at 1460–1530 nm, 1530–1565 nm, and 1565–1625 nm, respectively. These designations are not strictly applied in this report as that the absorption extends into the E-band (extended, 1360–1460 nm) and the emission in Fig. 4.9(b) continues through the U-band (ultralong wavelengths, 1625–1675 nm).
Figure 4.2: Er$_2$O$_3$ grown via ALE on Si(111). (a) TEM image of an Er$_2$O$_3$ layer on Si(111). (b) High-resolution TEM image of Er$_2$O$_3$-Si interface showing a sharp boundary between the two materials. (c) XRD spectrum for Er$_2$O$_3$ on Si(111) and a reference Si(111) sample; the (⋆) peaks designate strained layers as described in the text.
Table 4.1: Lattice constant ($a_o$) and mismatch for $R_2O_3$ bixbyite crystals on (111)Si

<table>
<thead>
<tr>
<th>Crystal</th>
<th>$a_o$ (Å)</th>
<th>Mismatch*</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_2O_3$</td>
<td>10.604</td>
<td>-0.0238</td>
</tr>
<tr>
<td>$La_2O_3$</td>
<td>11.36</td>
<td>0.0458</td>
</tr>
<tr>
<td>$Ce_2O_3$†</td>
<td>11.16</td>
<td>0.0274</td>
</tr>
<tr>
<td>$Pr_2O_3$</td>
<td>11.152</td>
<td>0.0267</td>
</tr>
<tr>
<td>$Nd_2O_3$</td>
<td>11.077</td>
<td>0.0198</td>
</tr>
<tr>
<td>$Pm_2O_3$‡</td>
<td>10.99</td>
<td>0.0118</td>
</tr>
<tr>
<td>$Sm_2O_3$</td>
<td>10.932</td>
<td>0.0064</td>
</tr>
<tr>
<td>$Eu_2O_3$</td>
<td>10.866</td>
<td>0.0004</td>
</tr>
<tr>
<td>$Gd_2O_3$</td>
<td>10.813</td>
<td>-0.0045</td>
</tr>
<tr>
<td>$Tb_2O_3$</td>
<td>10.730</td>
<td>-0.0122</td>
</tr>
<tr>
<td>$Dy_2O_3$</td>
<td>10.667</td>
<td>-0.0180</td>
</tr>
<tr>
<td>$Ho_2O_3$</td>
<td>10.607</td>
<td>-0.0235</td>
</tr>
<tr>
<td>$Er_2O_3$</td>
<td>10.547</td>
<td>-0.0290</td>
</tr>
<tr>
<td>$Tm_2O_3$</td>
<td>10.488</td>
<td>-0.0344</td>
</tr>
<tr>
<td>$Yb_2O_3$</td>
<td>10.439</td>
<td>-0.0389</td>
</tr>
<tr>
<td>$Lu_2O_3$</td>
<td>10.391</td>
<td>-0.0434</td>
</tr>
<tr>
<td>$Si$</td>
<td>5.431</td>
<td>—</td>
</tr>
</tbody>
</table>

* The lattice mismatch is relative to twice the Si unit cell: ($a_{ROx}/2a_{Si} - 1$).
† Ce$_2$O$_3$ is readily oxidized to form CeO$_2$ under ambient conditions.
‡ Promethium is radioactive and has no isotopes with half-lives longer than 17.7 years.

In post-growth XRD analysis, the dominant peaks in Fig. 4.2(c) are due to the substrate’s Si\{111\} and film’s Er$_2$O$_3$\{111\} planes. We employ two configurations for XRD analysis using a standard Cu K$_\alpha$ source: (1) a low angular resolution configuration that is sensitive to more material phases and (2) a high angular resolution configuration. The low-resolution setup utilizes a mirror to provide a wider angular divergence and higher intensity (shorter integration times) for the X-rays. With the (110) wafer flat of the Si(111) substrate 30° misaligned from the X-ray beam, this measurement is sensitive to more secondary material phases as in Fig. 4.2(b). Lesser peaks are associated with additional lines from the X-ray source (K$_\beta$, W) and with diffraction from minority Er$_2$O$_3$ phases; there is no XRD evidence of erbium silicates or silicides at the Er$_2$O$_3$-Si interface [167, 182, 183]. The \{211\}, \{442\}, and \{822\} families of peaks correspond to slightly strained volumes with
surfaces nearly parallel to the dominant (111) surface. Given the Er$_2$O$_3$ films’ homogeneous TEM cross sections and the relative intensities of the secondary XRD peaks, the minority phases compose a small fraction of the material. The high resolution setup aligns the wafer flat parallel to the X-ray beam and uses a monochromator to narrow the beam’s angular divergence. High-resolution XRD spectra only include the Si\{111\} and Er$_2$O$_3$\{111\} peaks.

In addition to growth of Er$_2$O$_3$ on Si substrates, we have also demonstrated growth on the top Si device layer of (100)-oriented silicon-on-insulator substrates. Curiously, regions exposed to prolonged XRD analysis exhibit lower photoluminescence yield.

### 4.1.2 Processing

Chip-based photonics require a number of micron- and sub-micron-scale elements, such as waveguides and resonators. To minimize scattering in these devices, processing techniques must be developed that produce little surface roughness. At room temperature, we have found Er$_2$O$_3$ is nearly impervious to HF (49% by wt.), KOH (30% by wt.), Piranha (3:1 concentrated H$_2$SO$_4$ to 30% H$_2$O$_2$), HCl (37% by wt.), and reactive ion etching (RIE) using C$_4$F$_8$:SF$_6$. At 70°C, concentrated HCl, H$_2$SO$_4$, and KOH etch Er$_2$O$_3$, but they leave ragged edges and appear to be slightly anisotropic. Reactive ion dry etching (RIE) with a Cl$_2$-catalyzed plasma will etch the material but not remove it from the surface—droplets of involatile etch by-products form on the wafer. Because these deposits do not sublime at process temperatures ($\leq205^\circ$C) but are water soluble, we believe these droplets are erbium chlorides, which are known to have a low vapor pressure [184]. Rather than chemically etching the Er$_2$O$_3$ device layer, we use ion milling with high Ar$^+$ flux and high DC bias.

To fabricate Er$_2$O$_3$ microdisk resonators, a 400-nm low-stress SiN$_x$ hard mask is grown on the Er$_2$O$_3$-Si wafer by PECVD. Since the PECVD layer is conformal, any surface imperfections in the Er$_2$O$_3$ film will be reproduced in the SiN$_x$; early oxide growth was polycrystalline and had significant roughness [Fig. 4.3(a)]. The microdisk patterns are defined using electron-beam lithography and resist reflow with ZEP520A resist [42] and then are transferred to the SiN$_x$ using a C$_4$F$_8$:SF$_6$ RIE. We then mill the Er$_2$O$_3$ with an Ar$^+$ plasma. The Ar$^+$ mill exhibits approximately 1:1 selectivity with the SiN$_x$ hard mask, and the resulting side walls [Fig. 4.3(b)] feature mild striations due to magnification of residual roughness in the Er$_2$O$_3$ film. While these striations should be minimized, they are less detrimental than in Si or Al$_x$Ga$_{1-x}$As because the refractive index contrast is lower ($n_d \approx 2$). Finally, an
Figure 4.3: Processing Er$_2$O$_3$ microdisk cavities. (a) A reflowed resist mask on an early sample develops small pseudopodia as the ZEP clings to features in the SiN$_x$ that are transferred up from an irregular Er$_2$O$_3$ layer. (b) Er$_2$O$_3$ microdisk edge prior to the SF$_6$ undercut. (c) Hybrid Er$_2$O$_3$-Si microdisk (78 nm Er$_2$O$_3$, 188 nm Si, 1 µm SiO$_x$) after the final HF undercut of the buried oxide. (d) SEM image of the final microdisk cavities showing the effects of the relaxed strain and the SF$_6$ undercut. These devices are isolated for use with a straight taper.
isotropic SF$_6$ dry etch simultaneously partially undercuts the silicon substrate and removes the remaining SiN$_x$ mask. Once the film is released from the substrate in Fig. 4.3(d), the cavities “flower” as the compressively strained film partially relaxes. Since residual organics cannot be removed from this material using Piranha, initial tests indicate a zero-bias O$_2$ plasma will clean the surface and not damage the epitaxy. Given its low reactivity, the Er$_2$O$_3$ layer provides a convenient mask for subsequent processing of hybrid structures. Processing hybrid Er$_2$O$_3$-Si resonators from Er$_2$O$_3$-SOI wafers, as in Fig. 4.3(c), employs the same SiN$_x$ hard mask and Ar$^+$ mill. With the Er$_2$O$_3$ layer acting as a mask, we etch the silicon device layer with an anisotropic C$_4$F$_8$:SF$_6$ RIE and undercut the buried oxide with concentrated HF. Many samples are prepared to optimize the process parameters at each step in order to minimize the microdisk side wall roughness. The measurements described here concern \(~150\) nm thick Er$_2$O$_3$ on Si(111) and Er$_2$O$_3$ microdisks with a radius of \(~20 \mu m\).

### 4.2 Optical Properties

Once the SiN$_x$ is removed and the disks are undercut, the emission and absorption properties of Er$_2$O$_3$ are again investigated using dimpled fiber-taper waveguides to characterize the disks’ WGMs and excite the Er$^{3+}$ optical transitions. The fiber-taper also offers high photoluminescence collection efficiency from both the WGMs (\(~10\%\) [70] and unpatterned film (\(~0.1\%\) [185]. Once fiber coupled, the pump and emission can be easily demultiplexed for sensitive pump-probe and pulsed measurements. The cavity dimensions (\(R \approx 20 \mu m\) and \(h = 150\) nm) are chosen to give minimal bending loss across the S/C/L-bands and well separated TE first- and second-order radial modes, as in Fig. 4.4. Conveniently, these cavity modes and the taper waveguide are almost phase matched (\(n_{\text{eff}}^{\text{cav}} \approx 1.17\) and \(n_{\text{eff}}^{\text{wg}} \approx 1.26\)), which gives near critical coupling for a wide range (1450–1560 nm) when the taper is placed in contact with the disk edge. For a fixed taper position, WGMs across the S/C/L bands exhibit different coupling depths because of the Er$^{3+}$ absorption spectrum. On the edges of the absorption band (near 1420 and 1620 nm), \(\gamma_i \ll \gamma_e\), and the resonances are significantly over coupled. In the center of the band, \(\gamma_i \approx \gamma_e\), and the modes are near critical coupling. The shallower transmission features near 1535 and 1549 nm are attributed to cavity-polariton and single-pass absorption effects.
4.2.1 Absorption and Emission Spectra for the $^4I_{13/2} \rightarrow ^4I_{15/2}$ Transition

To observe the crystal-field Stark splitting for the $^4I_{13/2} \rightarrow ^4I_{15/2}$ transition in this crystalline host, we measure both the room temperature photoluminescence (PL) and absorption spectra. We obtain the emission spectrum by placing the fiber-taper in contact with an undercut part of the film and pumping at $\lambda = 981.4$ nm (exciting the $X_1$ sublevel in the $^4I_{11/2}$ term). The fiber-collected luminescence [Fig. 4.5(a)] displays little inhomogeneous broadening of the Er$^{3+}$ transitions compared to amorphous hosts, and is qualitatively similar to the spectrum reported for polycrystalline Er$_2$O$_3$ deposited by pulsed-laser ablation [163]. In comparison to large crystals of Er$_2$O$_3$ and dilute Er$^{3+}$-doped Y$_2$O$_3$ produced by flame fusion [172, 178], the low temperature PL spectrum of these Er$_2$O$_3$ films exhibit more inhomogeneous broadening, also shown in Fig. 4.5(a). The 8 K PL also indicates some of these lines may be homogeneously broadened at room temperature—e.g., the linewidth for the 1549 nm peak is 56.6 GHz at 8 K and 134.7 GHz at 300 K.

The absorption spectrum [Fig. 4.5(b)] can be determined by fitting the low-power intrinsic loss rate of many TE$_{p=1}$ WGMs. While this WGM-linewidth method gives a digital rather than continuous spectrum, it involves measuring spectral widths rather than intensity changes, and coupling losses are insignificant and easily calculated from cavity-waveguide loading. Since there are several loss mechanisms that contribute to $\gamma_i$, radiation and surface scattering losses must be accounted for when using cavity modes to determine the absorption spectrum (similar to §3.2). By choosing appropriate cavity dimensions, the calculated $\gamma_{\text{rad}}$ is negligible for the fundamental modes: $\gamma_{\text{rad}} < 82$ MHz $= 0.026$ cm$^{-1}$. Since Rayleigh scattering increases toward longer wavelengths [41, 69], scattering losses can be
Figure 4.5: Emission and absorption spectra. (a) Thin-film PL spectrum at 300 and 8 K while pumping at 981 nm. The dominant peaks at 8 K are presented off the scale to make smaller features more visible. Emission is observed for ions on both $C_2$ and $C_{3i}$ sites; the peaks at 1535.8 and 1548.6 nm correspond to the transition between the lowest Stark levels ($Y_1 \leftrightarrow Z_1$) of the $^4I_{13/2}$ and $^4I_{15/2}$ manifolds on the $C_2$ and $C_{3i}$ sites, respectively [172, 177, 178]. (b) Composite absorption spectrum at room temperature. Different color $\times$’s correspond to the intrinsic linewidths for modes of different microdisks; the $\star$’s correspond to absorption peaks inferred from non-Lorentzian cavity resonances. The gray region is the resulting spectrum for the McCumber reciprocity relation.

bounded as less than the minimum intrinsic loss rate near 1625 nm: $\gamma_{ss} \lesssim \min(\gamma_i) = 2.5 \text{ GHz} = 0.79 \text{ cm}^{-1}$ [Fig. 4.6(a)]. This 2.5 GHz has a small absorption component as $^4I_{13/2} \rightarrow ^4I_{15/2}$ emission into WGMs is observed out to $\sim 1660$ nm. Because absorption at these wavelengths ($\lambda_0 \gtrsim 1600$ nm) excites electrons from the upper sublevels of the $^4I_{15/2}$ manifold, the absorption decreases significantly at $T \approx 8$ K when only the lowest sublevel ($Z_1$) is populated. Figure 4.6 shows a WGM in a similar device at cryogenic temperatures with $\gamma_i = 1.02 \text{ GHz}$, which is a more reasonable estimate for $\gamma_{ss}$ given the etch quality in Fig. 4.3(b). Since the sample used to reconstruct the absorption spectrum was not tested at low temperature, we will use the 2.5 GHz in Fig. 4.6 as an upper bound on $\gamma_{ss}$ which implies scattering accounts for less than 1% of the peaks, and $\gamma_i \approx \gamma_a$ except in the spectral tails. The second-order microdisk resonances exhibit slightly greater $\gamma_i$’s due to mode coupling between the $\text{Er}_2\text{O}_3$ WGMs and the lossy WGMs of the Si pedestal; $\gamma_{rad}$ is also significant for the TE$_{p=2}$ modes in the L-band. Because we are unable to bound the additional loss rate associated with the pedestal, the second-order modes are not used to establish the $^4I_{15/2} \rightarrow ^4I_{13/2}$ absorption spectrum.
Figure 4.6: Sample fits to cavity transmission resonances. (a) Singlet fit for a sample at $T = 300$ K according to Eq. (1.18). (b) Doublet fit for a sample at $T = 9$ K. (c) Sample non-Lorentzian fit (solid line) of an asymmetric cavity resonance. The inferred absorption peak and a Lorentzian fit (dashed line) to the same data are also included. (d) Triplet fit (solid line) for modes straddling the absorption peak at 1537 nm. The underlying Lorentzian resonances are displayed as dashed lines.

Due to the small inhomogeneous broadening of the Er$^{3+}$ transitions, it is possible for an underlying absorption peak to change appreciably across a cavity linewidth. In this case, the resulting non-Lorentzian cavity resonance can be fit [Fig. 4.6(c)] using Eq. (1.18) and including a Lorentzian absorption profile,

$$\gamma_i \approx \gamma_a(\omega) = \frac{\gamma_0(\delta \omega_a/2)^2}{(\omega - \omega_a)^2 + (\delta \omega_a/2)^2},$$

where $\omega_a$, $\delta \omega_a$, and $\gamma_0$ are the center, full width at half maximum, and amplitude of the absorption peak. For the observed resonances, it has been unnecessary to include the absorption peak's effect on the real part of the refractive index through the Kramers-Kronig relation—this shifts the resonance center and would be significant for precise dispersion studies. The inferred absorption peaks from fitting these asymmetric resonances around 1457.3, 1462.8, 1478.8, 1545.6, and 1556.9 nm agree well with the data from nearby Lorentzian resonances. Cavity modes near the $Y_1 \leftrightarrow Z_1$ transitions at 1537 and 1549 nm hybridize with the electronic states and become cavity-polaritons (discussed further in §4.2.5). To accurately determine the loss rates near these wavelengths, it is often necessary to fit a triplet of
modes (two polariton modes and a neighboring broad WGM). This fit can be accomplished by modifying Eq. (1.17a) to include two additional modes:

$$t(\omega) = s + i\sqrt{\gamma_e,1}a_{cw,1}(\omega, \omega_1) + i\sqrt{\gamma_e,2}a_{cw,2}(\omega, \omega_2) + i\sqrt{\gamma_e,3}a_{cw,3}(\omega, \omega_3).$$  \hspace{1cm} (4.2)

This approximation holds for weakly coupled WGMs as long as the phase shift of $t$ is small. When the modes are more deeply coupled, Eq. (4.2) does not accurately account for the interference between the modes, and the resulting fits describe a system that does not conserve energy. After extracting the three absorption rates from data like Fig. 4.6(d), the rate for the $p = 1$ mode is identified and used in Fig. 4.5(b). When the TE$_{p=1}$ mode is resonant with the absorption peak, the cavity and electronic linewidths average, and the linewidth for either polariton branch can be used.

The absorption spectrum can also be obtained using the McCumber reciprocity relation between the absorption ($\sigma_a$) and emission ($\sigma_e$) cross sections [165]. Since the crystal field splitting of a RS term is approximately equal to or greater than the thermal energy, the occupation probabilities for the sublevels are not equal, and the relationship between the total absorption and emission rates between two levels (arising from the Einstein A and B coefficients) no longer holds. Within a single manifold, the electrons will thermalize on a time scale of $\sim$100 ps via interactions with the phonon bath. On the millisecond timescale for radiative transitions, the thermal distribution within the sublevels biases the emission (absorption) toward red (blue) wavelengths. This situation is identical to phonon-terminated lasers [186], and the cross sections are related by

$$\sigma_a(\omega) = \left( \frac{Z_2}{Z_1} \right) \sigma_e(\omega) \exp \left( \frac{\hbar \omega - E_{21}}{k_B T} \right),$$  \hspace{1cm} (4.3)

where $E_{21}$ is the energy difference between the two manifolds, $Z_2$ and $Z_1$ are the partition functions for sublevels within the upper and lower manifolds, $k_B$ is the Boltzmann constant, and $T$ is the temperature. Unlike in glass, spectroscopy of the Stark levels is available for ions in many crystals [e.g., Fig. 4.1(d)], and Eq. (4.3) can be evaluated exactly. Using the energies for the sublevels of $^4I_{13/2}$ and $^4I_{15/2}$ on $C_2$ from Ref. [172], the results from Eq. (4.3) in Fig. 4.5(b) are in reasonable agreement with the linewidth data. The McCumber spectrum is scaled to the maximum absorption peak at 1535 nm because the magnitude
of the PL data is not calibrated. The shortcoming of this approach is that it assumes the dipoles are identical and independent. These assumptions are valid in many systems such as lightly doped glasses, but neither are strictly true in Er$_2$O$_3$. Differences between the McCumber theory and the data occur primarily at the edges of the spectrum and may reflect different dynamics for ions on the $C_2$ and $C_{3i}$ sites. As will be discussed in the following sections, emission from the magnetic dipole transitions on the $C_{3i}$ sites is disproportionately large for the observed absorption, and the excitation migration rate (§4.2.3) may be asymmetric for transfers between the $C_2$ and $C_{3i}$ sublattices.

The resulting spectrum [Fig. 4.5(b)] again reflects the weak inhomogeneous broadening of the Er$^{3+}$ transitions and qualitatively agrees with the peaks’ positions and widths in the PL spectrum. Accounting for material and modal dispersion with scattering losses bound at <1% of the peak values, the absorption maxima of 267.0 GHz at 1535 nm and 175.4 GHz at 1479 nm correspond to 83.9 cm$^{-1}$ ($\delta \omega_{\text{fsr}}/2\pi \approx 1.59$ THz) and 56.2 cm$^{-1}$ ($\delta \omega_{\text{fsr}}/2\pi \approx 1.56$ THz), respectively. The narrow peaks and low losses in the tails further indicate there is little background absorption, and all 83.9 cm$^{-1}$ (364 dB/cm) can be attributed to resonant $^4I_{15/2} \rightarrow ^4I_{13/2}$ excitation, which sets a high maximum value for potential gain.

### 4.2.2 Upconversion

While the 1450–1650 nm band is most useful for hybrid Er$^{3+}$-silicon optical networks, upconversion into visible transitions is also present and has been partly investigated for polycrystalline Er$_2$O$_3$ on Si [159, 162]. We determine the upconversion spectrum by transferring a small piece ($\sim 2 \mu m^2 \times 150$ nm) of Er$_2$O$_3$ onto the fiber-taper and pumping with <3 mW at 1536.7 nm. The taper-collected PL is then measured in a Czerny-Turner spectrometer with a silicon CCD camera. The visible PL contains emission from many levels with significant emission near 550 and 670 nm and exhibits little inhomogeneous broadening, as in Fig. 4.5(a). The relative intensities in Fig.4.7(a,b) may not reflect the actual strength of each transition because we are unable to correct for the unknown taper collection efficiency across the visible range; we have corrected for the grating and CCD efficiencies. Spectroscopic data over a large range must also be corrected for second-order diffraction from the spectrometer—i.e., the grating equation for a given diffraction angle ($\theta_d$) is simultaneously satisfied for wavelengths ($\lambda_j$) in two different orders ($m_j$): $a(\sin \theta_d - \sin \theta_i) = m_1 \lambda_1 = m_2 \lambda_2$ where $a$ is the grating pitch and $\theta_i$ is the incident angle. Usually gratings are blazed to
operate at a given order and wavelength while all other orders have significantly lower efficiencies.\(^4\) In this measurement, the second-order grating response is only observed near 800 nm where the sharp \(^4G_{11/2}\) line at 391 nm appears as a small peak at the edge of the \(^4I_{9/2}\) emission—the second-order \(^4G_{11/2}\) intensity is approximately 10% of the \(^4I_{9/2}\) peak at 805 nm. To remove the second-order response associated with the \(^4G_{11/2}\) and \(^2H_{9/2}\) levels, we scale the spectrum at 385–425 nm to the second-order \(^4G_{11/2}\) peak at \(\sim782\) nm and subtract. This method is acceptable for broad surveys such as this, but more accurate studies should filter out the shorter wavelengths to avoid producing negative counts near 815–840 nm.

The pump-power dependence [Fig. 4.7(c)] of the three primary upconversion bands provides insight into the specific upconversion mechanism [Fig. 4.7(d)]; we use a microdisk cavity for these curves because they offer a greater signal at low pump power. The nearly quadratic dependence of the 800 nm emission suggests pairwise upconversion out of the \(^4I_{13/2}\) multiplet followed by excited-state absorption or a second upconversion event (\(^4I_{9/2} + ^4I_{13/2} \rightarrow ^2H_{11/2}\)) to produce the nearly cubic dependence around 550 and 670 nm. Subsequent absorption or energy transfer then connects \(^2H_{11/2} \rightarrow ^4S_{3/2}\) to even higher levels with energies in the near UV. Emission of multiple phonons or a single mid-IR photon connects this ladder of states to levels that are not directly excited [dashed relaxations in Fig. 4.7(d)]. Other paths may also contribute to the upconversion process, but this pathway is consistent with the pump dependence and involves the least energy mismatch between excitations—the phonon bath must provide or absorb any excess so that energy is conserved.

Since there are currently no efficient solid-state emitters at green wavelengths (most green lasers are frequency-doubled emission from Nd\(^3+\):YAG lasers), there is considerable interest in developing efficient rare-earth upconversion sources and lasers for displays, lighting, and compact projectors [187–190]. Initial investigations of upconversion in these Er\(_2\)O\(_3\) cavities at high power are inconclusive. We use a pulsed C-band excitation because the poor thermal conductivity for microdisk WGMs gives a damage threshold of \(P_d \approx 30\) mW. Because of the high Er\(^3+\) density, long pulses with large peak powers are necessary to provide enough energy to reach transparency for visible transitions; these 20-\(\mu\)m disks contain \(\sim10^{11}\) ions and require \(>18\) nJ for transparency at 550 nm (or equivalently a 20-ns square

\(^4\)The grating used in these measurements is blazed for operation near 500 nm in the first order with \(a = 0.83\) \(\mu\)m (1200 lines/mm); the antireflection coating on the CCD is optimized for 550 nm.
Figure 4.7: Upconversion behavior. (a,b) Fiber-taper collected Er$_2$O$_3$ upconversion spectrum while pumping at 1536.7 nm; spectroscopic identifications are made by comparison to known Er$^{3+}$ transitions back to $^4$I$_{15/2}$ [172]. NB: The scale is varied across the spectrum to make weaker transitions more visible. (c) Pump-power dependence for the integrated PL in the 550, 670, and 800 nm bands, which are in reasonable agreement with Ref. [187]. The data sets are offset vertically for clarity. (d) Proposed upconversion path. (e) Setup for measuring visible upconversion using a fiber-taper as the collection optic. (f) Setup for high-power pulsed C-band excitation. Two polarization controllers are needed to independently match the polarizations of the electro-optic modulator (EOM) and the cavity resonance. To optimize the modulator bias voltage and input polarization, we constantly monitor our extinction ratio (>35 dB) and pulse shape (square pulse, 100 ps rise time, 120 ps fall time) with an optical spectrum analyzer (OSA) and a digital communications analyzer (DCA)—cf. §A.2. The DCA is also used to record the transmitted waveform to check for any dynamic cavity detuning during the pulse width, similar to Ref. [63]. The second VOA and the hot mirror (average 85% transmission over 450–675 nm, 90% reflection over 750–1600 nm) protect the detectors from the high-power pulses at ~1550 nm.
pulse with a peak power of 900 mW). To generate these long high-energy pulses in the C-band, we modulate the input to an erbium-doped fiber amplifier (EDFA). Pulses with a high-extinction ratio and low duty cycle can fully extract the energy stored in the amplifier and produce peak powers well above the saturation output, which is usually specified for CW operation. With 10 mW from the tunable diode laser, the arrangement in Fig. 4.7(f) can produce pulses with energy up to 300 nJ (input pulses: (~33 dB extinction, 20 ns wide, 100 kHz repetition), output peak power: 15.8 W, EDFA saturation for CW operation: 315 mW). The upconversion bands around 550 and 670 nm do not saturate smoothly at high power and exhibit a weak kink (at 46 nJ/pulse) just prior to rolling off. Unfortunately, the poor thermal stability of the thin disks is inadequate to consistently reproduce these results and untangle the various experimental complexities (drifting waveguide coupling, thermo-optic shift of the pump mode, further upconversion into the blue and near UV, nonradiative decay, spontaneous and stimulated emission, etc.).

4.2.3 Effective $^4I_{13/2}$ Lifetime

Depopulation of $^4I_{13/2}$ by cooperative upconversion adds an additional complication to measuring the $^4I_{13/2} \rightarrow ^4I_{15/2}$ lifetime. To mitigate the upconversion effects, fluorescence decay measurements are performed by uniformly exciting a fundamental WGM at 1473.4 nm with 10-ns square pulses and a peak absorbed power of 21.7 $\mu$W. These lifetime measurements are typically performed by pumping an unpatterned film with a chopped laser at 980 nm (or at shorter wavelengths), but it is difficult to determine the density and distribution of excited ions, which is important in high Er$^{3+}$-density samples. Because of the weak PL signal, we apply a pulse-delay technique and a single photon counter to sample the C/L-band fluorescence decay curve [Fig. 4.8]. To reduce dark counts, InGaAs/InP avalanche photodiodes (APDs) are only gated above their breakdown voltage for a short time (~50 ns), which is not suitable for decay curves with $10^{-6}$–$10^{-2}$ s lifetimes. To circumvent the APD’s narrow gate width, we use the 50-ns window to discretely sample the decay curve. Centering the $^4I_{13/2}$ PL peak in the 50-ns gate, we observe the lowest noise and highest signal-to-noise ratio.

$^5$NB: These high-power pulses eventually damaged the 980-nm pump diodes in the EDFA. During its repair, the 980/1550 demultiplexing filters were doubled to further protect the diodes during these experiments. An optical isolator should also be placed at the EDFA output to prevent reflections from returning to the amplifier.

$^6$Because of poor impedance matching to the photodiode in the id Quantique id201 APD, there is significant oscillation in the gate voltage (and thus in the photon detection probability and dark counts) during the first ~20 ns of each detection window. To avoid these effects as in Ref. [191], the delay is set so the pulse arrives in the second half of a 100-ns window; the first 50 ns are then discarded. The lowest noise was
Figure 4.8: Measurement of $^4I_{13/2}$ lifetime. (a) Experimental setup to measure the $^4I_{13/2} \rightarrow ^4I_{15/2}$ lifetime by resonantly pumping a WGM at 1473 nm. Abbreviations: short-pass filter (SPF, pass 1460–1500 nm, reflect 1527–1610 nm), long-pass filter (LPF, pass 1527–1610 nm, reflect 1455–1500 nm), and InGaAs/InP avalanche photodiode (APD). A variable electronic delay is necessary to synchronize the APD gate with the arrival of the PL pulse. The OSA and DCA are again used to monitor the pulse extinction and waveform. Black lines represent optical fiber; blue lines designate coaxial cable. Dashed lines correspond to alternative connections. (b) Pulse period measurement of the lifetime for C/L-band emission. Fitting the data in the shaded area gives a lifetime of $5.7 \pm 0.9 \mu s$; this uncertainty and the dashed curves mark $\pm 2\sigma$ confidence for the fit. Inset: sample histogram for a pulse period of $9.89 \mu s$ (101.1 kHz repetition rate).
arrival of a pump pulse in the detection window simultaneously acquires the PL’s rise and initial decay due to the \(N\)th pulse along with the decay associated with the \((N-1)\)th pulse. The appearance of the \(N\)th pulse serves as a marker for sampling the \((N-1)\)th decay curve at a fixed delay—i.e., the pulse period separates the \((N-1)\)th peak from its tail just before the \(N\)th pulse. Several histograms (128 ps/bin resolution) with varying delays are used to construct the fluorescence decay in Fig. 4.8. Background histograms are taken at each point without the pulses to remove dark counts and PL counts from when the pump is nominally “off”—with \(-35\) dB extinction, the CW excitation is 6.9 nW. As the pulse period approaches the PL lifetime, the data deviates from a single exponential curve because decay from multiple pulses contributes to the PL tail prior to the \(N\)th pulse’s arrival. Data at longer periods is limited by a constant noise floor linked to the small portion of pump laser spontaneous emission that is not blocked by the filters. For \(\tau_{\text{eff}} \approx 10\) µs, probabilistic simulation suggests that a fitting region between 10 and 20 µs gives the greatest confidence unbiased estimate of the decay lifetime. Fitting the points in this range gives an effective lifetime of \(5.7\pm0.9\) µs for a peak excitation of 21.7 µW and \(\pm2\sigma\) uncertainty. With the effective lifetime much lower than the measured 8±0.5 ms (7.8±2.2 ms calculated) radiative lifetime in lightly Er\(^{3+}\)-doped bulk Y\(_2\)O\(_3\) [192], nonradiative relaxation is a major concern.

The nonradiative decay of rare-earth ions was extensively investigated during the early development of inorganic gain crystals for solid-state lasers, and accurate phenomenological models have been established to describe the two principal mechanisms: multiphonon relaxation and cooperative relaxation (also known as concentration quenching) [193]. First considering multiphonon emission, the high yield of the \(^4I_{13/2} \rightarrow ^4I_{15/2}\) transition in low Er\(^{3+}\) density samples is because the 0.8 eV (6500 cm\(^{-1}\)) energy gap is too large for fast depopulation. Using parameters from Er\(^{3+}\) relaxation in Y\(_2\)O\(_3\) at low temperature [192, 194] and the close similarity of the Er\(_2\)O\(_3\) and Y\(_2\)O\(_3\) vibrational spectra [195], we estimate an effective lifetime of 4.2 s for relaxation at 300 K via emission of 12 phonons (\(\sim550\) cm\(^{-1}\)). Cooperative relaxation encompasses several decay and sensitizing mechanisms where the excitation is nonradiatively transferred between ions through multipole or exchange interactions—we will limit the discussion to processes involving a single ion species. Self-quenching, where a donor ion decays to an intermediate level by exciting a low-level transition in a neighbor, is significant for higher levels in Er\(^{3+}\) (e.g., \(^4S_{3/2}\)) [196], but it is inactive for the first ex-
cited level ($^4I_{13/2}$). The most probable relaxation path in Er$_2$O$_3$ is the loss of energy to quenching (acceptor) sites either by direct transfer or from resonant excitation migration through multiple ions [197]. Acceptors are usually nonluminescent impurities (1–10 ppm of Fe, W, and Re and >10 ppm of Ta are present in the erbium ALE source, impurity levels in the final Er$_2$O$_3$ films are unknown) and/or perturbed electronic states near surfaces, grain boundaries, or dislocations. At timescales shorter than the radiative lifetime in high purity crystals at 300 K, diffusion via electric dipole-dipole interactions becomes extremely rapid and >10$^5$ transfers are possible before reaching an acceptor [198, 199]; with a mean Er$^{3+}$–Er$^{3+}$ separation of $\sim$3.3 Å, excitations in a 150-nm Er$_2$O$_3$ film are never more than $\sim$230 steps from a surface. Although transfer through (Y$_{1-x}$Tb$_x$)$_3$Al$_5$O$_{12}$ ($0.1 \leq x \leq 1.0$, minimum mean Tb$^{3+}$–Tb$^{3+}$ spacing of $\sim$4.2 Å at $x = 1.0$) was consistent with dipole-dipole coupling [200], high donor concentrations, as in Er$_2$O$_3$, may further increase the migration rate through short-range exchange and/or electric quadrupole-quadrupole interactions [201]. However even for dipole-dipole interactions, excitation migration at 300 K to a small concentration of acceptors will effectively quench the $^4I_{13/2}$ emission in these Er$_2$O$_3$ films. Since it seems unlikely that these interactions can be screened (perhaps by introducing various dopants), increasing the mean Er–Er separation appears to be the only way to slow this transfer rate and the nonradiative relaxation. Assuming the decay is diffusion limited due to a low density of acceptors and conservatively assuming dipole-dipole interactions, the nonradiative decay rate is proportional to the Er$^{3+}$ concentration [199, 202]. Ion-ion coupling resulting in cooperative upconversion is governed by the same multipole interactions and is also enhanced by the excitation diffusion.

4.2.4 Power-Dependent Radiative Efficiency

Because the fiber-taper offers adjustable waveguide coupling and a low-loss method for PL excitation and collection [70], the efficiency for emission into the cavity modes can be empirically determined. Utilizing wide band-pass filters, the S-band transmission and C/L-band emission can be observed simultaneously [Fig. 4.9(a)]—all cavity-based PL measurements are pumped on resonance with a WGM. The long-pass and short-pass filters separate their pass bands ($P \leftrightarrow CM$ ports) and reflection bands ($R \leftrightarrow CM$ ports) with high directivity (>55 dB) and low insertion loss (<1 dB). By collecting and filtering PL in the cavity’s “reflection” channel, we achieve >100 dB isolation at the pump wavelength. Some residual
spontaneous emission from the laser diode does bleed through the filters and produces a ~10 pW signal at 1519 nm with our pump near maximum power. Since the radiative efficiency measurements only concern WGMs at $\lambda > 1520$ nm, this peak is ignored; in other cases it can be reduced below 1 pW by inserting an additional short-pass filter (pass 1450–1495 nm) after the laser. We employ an optical spectrum analyzer to measure the PL at peak powers $\gtrsim 10$ pW and a spectrometer with a liquid-N$_2$ cooled InGaAs detector array at lower powers (down to peak powers of ~0.1 fW).

The cavity-coupled radiative efficiency ($\eta_{\text{obs}}$) is given by the ratio of the total power emitted into the cavity modes divided by the absorbed pump power. This measurement requires careful characterization of the pump mode along with all modes in the desired emission band. For these c-Er$_2$O$_3$ microdisks, surface scattering and radiation losses are negligible compared to the Er$^{3+}$ absorption, and the absorbed power ($P_a$) is nearly equal to the dropped power ($P_d$),

$$P_a = \frac{\gamma_i}{\gamma_i} P_d \approx P_d = (1 - T_{\text{cav}}) \frac{P_{\text{in}}}{\sqrt{T_t}},$$

(4.4)

where $T_{\text{cav}}$ is the cavity-waveguide transmission at the pump wavelength, $T_t$ is the end-to-end fiber-taper transmission, and $P_{\text{in}}$ is the pump power measured at the fiber-taper input. Equation 4.4 assumes the taper’s loss is symmetric about the taper-device coupling region. While transmission loss in fiber-tapers is usually dominated by bending loss in the taper mount which is symmetric about the coupling region, small bits of dust on the taper will scatter light from the fundamental mode and produce asymmetric loss. In this case, $\sqrt{T_t}$ is replaced by the one-sided waveguide transmission ($T_1$), which can be found from $T_t$ and the ratio of another quantity that depends on $P_d$ (e.g., thermo-optic wavelength shift or peak PL yield) when using either end of the waveguide as the input. For these measurements, $T_t = 0.68$ with symmetric loss. Using the collected emission [$P(\lambda)$] spectrum as in Fig. 4.9(b), the total power emitted into the cavity modes ($P_{\text{cav}}$) is given by

$$P_{\text{cav}} = \frac{2}{\delta \lambda_{\text{rbw}} \sqrt{T_t}} \left[ \sum_n \left( \frac{\gamma_i,n + \gamma_e,n}{\gamma_e,n} \right) \int \frac{P_n(\lambda)}{T_1(\lambda)} d\lambda \right],$$

(4.5)

where the summation is over the cavity modes at $\lambda > 1520$ nm and $\delta \lambda_{\text{rbw}}$ is the spectral resolution bandwidth of the grating and detector. The factor of 2 compensates for equal
Figure 4.9: (a) Setup for measuring C/L-band PL and \( \eta_{\text{obs}} \). (b) Observed cavity-coupled PL for the device in Fig. 4.4; the TE\(_{p=1}\) modes are highlighted in gray. (c) Power-dependent radiative efficiency of the \( ^4I_{13/2} \rightarrow ^4I_{15/2} \) transition. The marker color indicates the pump-mode wavelength from blue (1460.9 nm) to dark red (1494.5 nm) while (o) and (+) designate first- and second-order radial pump modes, respectively. Black markers represent data from devices on another wafer processed with wet chemical etching and using first- and second-order pump modes (spanning 1437.6–1490.9 nm). The inset shows green upconverted luminescence from a TE\(_{p=1}\) mode [35].
emission into the degenerate clockwise and counterclockwise traveling-wave modes of the disk and hence into the forward- and backward-propagating modes of the waveguide. While the C/L-band transmission $|T_{\text{f}}(\lambda)|$ is fairly flat for the filters used to (de)multiplex the PL and the pump beam, the transmission does slowly decrease at longer wavelengths. In the emission band, $T_{\text{f}} = 0.66$ with symmetric loss. To establish and correct for the fraction of the PL in the WGMs that is collected by the taper, we find the total cavity loss rate for each cavity mode in the emission band by measuring the $\gamma_{i,n}$ under weak loading and $\gamma_{e,n}$ at the fixed taper position used during PL collection. Then the fraction of $P_{\text{cav}}$ coupled into the forward propagating mode is the ratio of the loss rate into the waveguide ($\gamma_{e,n}$) over the loss rate into all channels ($\gamma_{i,n} + \gamma_{e,n}$), which gives a correction factor of $(\gamma_{i,n} + \gamma_{e,n})/\gamma_{e,n}$. There is no evidence for significant parasitic loading ($\gamma_p \approx 0$) during this measurement.

For example, the WGMs near the emission peaks at 1535 and 1549 nm are nearly critically coupled in Fig. 4.4 ($\gamma_{i,n} \approx \gamma_{e,n}$), and the waveguide collection factor for these modes is $(\gamma_{i,n} + \gamma_{e,n})/\gamma_{e,n} \approx 2$. Since the ratio of the intensity of individual WGMs over the total emitted power is constant at $P_d \lesssim 4 \text{ mW}$, the total integrated PL is proportional to the intensity of the strongest emission line. Using this proportionality, the radiative efficiency measurement can be extended to excitation levels low enough that the largest PL peak is just above the noise floor.

We analyze the power dependence of $\eta_{\text{obs}}$ for several pump modes across the S-band on two different samples. In Fig. 4.9, $\eta_{\text{obs}}$ decreases as upconversion becomes the dominant path for depopulating $^4I_{13/2}$. Fitting the data to

$$\eta_{\text{obs}}(P_a) = \frac{\eta_0}{1 + (P_a/P_{\text{up}})^x} \quad (4.6)$$

gives $x = 0.61 \pm 0.07$ and $P_{\text{up}} = 204 \pm 47 \mu\text{W}$, at which point the efficiency has dropped by half and there are equal rates for nonradiative relaxation and cooperative upconversion. Expressed in terms of an effective lifetime ($\tau_{\text{eff}}$) and the state’s lifetime at zero pump power ($\tau_0$)

$$\tau_{\text{eff}}(P_a) = \tau_0 \left(1 + (P_a/P_{\text{up}})^x\right), \quad (4.7)$$

residual upconversion can be removed from the previous lifetime estimate of $5.7 \pm 0.9 \mu\text{s}$ at $21.7 \mu\text{W}$ to give a zero-power effective $^4I_{13/2}$ lifetime of $\tau_0 = 7.2 \pm 1.2 \mu\text{s}$. Since the radiative lifetime ($\tau_{\text{rad}}$) in these Er$_2$O$_3$ films should be very similar to the 8 ms lifetime in Y$_2$O$_3$ [192],
\( \tau_o \) can be related to the overall emission efficiency across the S/C/L-bands \((\eta_{\text{scI}})\) and the observed spontaneous emission coupling factor \((\beta_{\text{obs}})\) for the chosen cavity resonances:

\[
\frac{\tau_o}{\tau_{\text{rad}}} = \eta_{\text{scI}} = \frac{\eta_o}{\beta_{\text{obs}}}. \tag{4.8}
\]

With \( \tau_{\text{rad}} = 8 \text{ ms} \) and \( \eta_o = 3.4 \times 10^{-5} \), we find the \( ^4I_{13/2} \rightarrow ^4I_{15/2} \) transition is almost completely nonradiative \((\tau_o \approx \tau_{nr})\) with \( \eta_{\text{scI}} = 9.0 \times 10^{-4} \) and \( \beta_{\text{obs}} = 0.038 \), which is consistent with estimates of \( \beta_{\text{obs}} \) based on the cavity mode spectrum and negligible Purcell enhancement of \( \tau_{\text{rad}} \) (see §B). Finally, at \( 204 \mu \text{W} \), upconversion reduces the lifetime to \( \tau_{\text{eff}} = \frac{1}{2} \tau_o \) and gives \( \sim 3 \times 10^{20} \) ions/cm\(^3\) in \( ^4I_{13/2} \). Based on the more rigorous analysis of Nikonorov et al. [203], we can equate the nonradiative decay \((1/\tau_o)\) and upconversion rates \([1/(C_{\text{up}} N_{13/2})]\) at this power to estimate the cooperative upconversion coefficient \((C_{\text{up}})\):

\[
C_{\text{up}} = \frac{2hcV}{\lambda_p \tau_o^2 P_{\text{up}}} = (5.1 \pm 2.1) \times 10^{-16} \text{ cm}^3/\text{s}, \tag{4.9}
\]

for a \( ^4I_{13/2} \) rate equation of the form \( \partial_t N_{13/2} = \ldots - C_{\text{up}} N_{13/2}^2 \) and where \( V = 20.1 \mu \text{m}^3 \) is the volume of \( \text{Er}_2\text{O}_3 \) excited by the cavity mode, \( \lambda_p \approx 1480 \text{ nm} \) is the pump wavelength, and \( P_{\text{up}} = 204 \pm 47 \mu \text{W} \) is the power when the nonradiative and upconversion rates are equal. This upconversion coefficient is extremely large and similar to that found in cosputtered \( \text{Er}_2\text{O}_3/\text{Al}_2\text{O}_3 \) [204].\(^7\)

### 4.2.5 Cavity Polaritons

Following the observation of normal mode splitting in a semiconductor microcavity [11], vacuum-Rabi splitting has been observed over a growing range of materials and temperatures, and the exploitation of cQED and strong coherent light-matter coupling has become a promising route to create more efficient light emitters. Normal mode splitting at room temperature has been verified in planar cavities containing various active layers, including InGaAs quantum wells (QWs) [205], II-VI QWs [12], group III-nitrides [206], and with intersubband transitions in a quantum cascade structure [207]. In exciton-polariton systems, there has been considerable effort to understand the low-temperature thermodynamics of

---

\(^7\)The definition of \( C_{\text{up}} \) is somewhat arbitrary depending on the conventions used in the rate equations. Kik and Polman use \( \partial_t N_{13/2} = \ldots - 2C_{\text{up}}' N_{13/2}^2 \) and find \( C_{\text{up}}' = \frac{2}{C_{\text{up}} = 3.5 \times 10^{-16} \text{ cm}^3/\text{s} \), which is within the uncertainty in Eq. (4.9).
the boson quasiparticles and to utilize stimulated scattering toward the lower polariton branch [208, 209]. Optoelectronic functionalities based on efficient polariton emission from single excitons, such as parametric gain [210] and electroluminescence (EL) [211, 212], have now also progressed above $T \gtrsim 200\, \text{K}$.

Rabi splitting and the associated cavity-polariton modes can be described as a consequence of linear dispersion [213] or as non-perturbative coupling between the dipole(s) and an optical cavity, while perturbative coupling to a single ion produces Purcell-enhanced emission (without Rabi oscillation) [214]. Beyond quantum wells and single dipoles with large oscillator strengths, the collective oscillation of atomic and molecular ensembles is another viable method to obtaining light from polariton states. Due to their large dipole moment, planar microcavities incorporating organic semiconductor emitters exhibit massive room-temperature Rabi splitting [215] and EL [216]. While in a complementary approach with high density ($\rho = 2.7 \times 10^{22}\, \text{cm}^{-3}$), the narrow inhomogeneous distribution of optical transitions in stoichiometric erbium crystals can also intrinsically produce large dispersive resonances in the refractive index and splitting of the cavity modes. Cavity polaritons have been observed around a single Er$^{3+}$ transition in polycrystalline Er$_2$O$_3$ produced by sputtering or oxidation of an erbium film [217], but the effect was quenched at $T \gtrsim 40\, \text{K}$ [218].

In this section, we analyze the properties of low temperature ($T \approx 8\, \text{K}$) and high temperature ($T > 361\, \text{K}$) cavity-polaritons formed between the microdisk WGMs and the Er$^{3+}$ $^4I_{13/2} \leftrightarrow ^4I_{15/2}$ transitions in epitaxial Er$_2$O$_3$-on-Si.\(^8\)

Analyzing the polariton response involves continuously tuning a cavity mode across the Er$^{3+}$ transitions. As a mode is shifted through an optical transition, the resonances anticross (i.e., the vacuum-Rabi splitting) and produce symmetric hybrid modes (i.e., the cavity-polaritons), which appear in both the cavity transmission and PL. In terms of cQED, the fundamental quantity for cavity-polaritons is the Rabi frequency that describes the coherent coupling between the two-level system(s) and a near-resonant optical mode. For a uniform distribution of emitters, the many-dipole Rabi frequency ($\bar{g}_N$) is given by

$$|\bar{g}_N|^2 = N|\bar{g}_1|^2 \approx \frac{\bar{\mu}^2 \omega_c \rho \Gamma}{2 \hbar m^2 \epsilon_0}, \quad \text{(4.10)}$$

where $\bar{g}_1$ is the mean coupling rate experienced by a single dipole, $N$ is the number of emitters, $\omega_c$ is the cavity frequency, $\rho$ is the density of emitters, $\Gamma$ is the decay rate, $\hbar$ is the reduced Planck constant, $m$ is the mass of the emitters, and $\epsilon_0$ is the vacuum permittivity.

dipoles in the cavity, \( \vec{\mu} \) is the electric dipole moment of the optical transition, \( \omega_c \) (\( \lambda_c \)) is frequency (wavelength) of the cavity resonance, \( n \) is the refractive index of the dipole medium, \( \rho \approx N/(V_{\text{eff}} \Gamma) \) is the dipole density in \( \text{Er}_2\text{O}_3 \), \( V_{\text{eff}} \approx 70(\lambda_c/n)^3 \) is the optical mode volume from Eq. (1.28), and \( \Gamma \) is the optical mode’s electric-field energy overlap with the dipole ensemble from Eq. (1.29) \[6, 219\]—see also §A.4. In \( \text{Er}_2\text{O}_3 \) resonators, a large \( \text{Er}^{3+} \) density overcomes the small \( 4^1I_{13/2} \rightarrow 4^1I_{15/2} \) dipole moment to produce resolved polariton modes.

### 4.2.5.1 Low-Temperature Polaritons

Measurements at cryogenic temperatures are conducted in a continuous flow \(^4\text{He} \) cryostat with the fiber-taper waveguide and sample mounted on piezoelectric stages with a base temperature of 8 K. At cryogenic temperatures, the homogeneous linewidths of the \( \text{Er}^{3+} \) transitions narrow significantly, and 99.8% of the ions reside in the \( Z_1 \) sublevel of \( 4^1I_{15/2} \) (compared to 24.3% at 300 K). As in Fig. 4.5(a), the observed linewidths are due to inhomogeneous broadening from ions in strained volumes or near dislocations. To study the polariton modes, the repeatable deposition and sublimation of thick \( \text{N}_2 \) films (\( n \approx 2 \)) can tune the \( \text{Er}_2\text{O}_3 \) cavity modes more than 25 nm. Compared to GaAs disks which only tune up to 4 nm \[220\], the \( \text{Er}_2\text{O}_3 \) resonators are thinner with lower refractive index. The additional scattering loss due to the roughness of the \( \text{N}_2 \) ice also remains negligible compared to the resonant \( \text{Er}^{3+} \) absorption. Away from the \( \text{Er}^{3+} \) transitions, the dominant sources of loss are surface scattering at the milled sidewalls and parasitic coupling from the fiber-taper.\(^9\) Testing the cavity transmission over a range from 1423 to 1625 nm, we measure nine anticrossings between 1455.5 and 1548.6 nm (Fig. 4.10). Considering Rabi splitting within the context of linear absorption \[213\], we expect that polaritons will form at 8 K around transitions originating from the lowest sublevel (\( Z_1 \)) of \( 4^1I_{15/2} \), and we observe splitting about excitations from \( Z_1 \) to all seven Stark sublevels (\( Y_1\ldots7 \)) of \( 4^1I_{13/2} \) on the \( C_2 \) lattice sites—the inhomogeneous distributions for the \( Y_3 \) and \( Y_4 \) sublevels overlap around 1516 nm. Rabi splitting also occurs about three transitions (\( Z'_1 \rightarrow Y'_{1\ldots3} \)) on the \( C_{3i} \) sites. Table 4.2 summarizes the average measured Rabi splittings (\( \delta \omega_{\text{RVS}} \)) and the transition linewidth (\( \Delta \omega \)).

---

\(^9\)These measurements were conducted during a period when many tapers were breaking in the cryostat. This problem was likely due to large thermal gradients when the fiber was placed in contact with the sample. One solution was to use larger, more robust tapers, but the final resolution was to wait longer for the temperatures to equilibrate.
Figure 4.10: Cavity transmission spectra in the (a) S-band and (b) C-band at 8.6 K tuned using N₂ deposition. The shallow background ripple is residual-etalon noise from the tunable lasers or ringing from the notch filters used to filter out that noise. (c) Taper-collected PL spectrum as the cavity is tuned. Equal spaced contours (Δ = 0.002) are used to display low-intensity features. The Er³⁺ ions are excited by resonantly pumping a WGM near 1480 nm. (d) Identified transitions in the bulk PL spectrum at 8.6 K.
Table 4.2: Summary of observed cavity-polaritons at $T = 8.6$ K

<table>
<thead>
<tr>
<th>Wavelength (nm)</th>
<th>Transition (site)</th>
<th>$\delta \omega_{\text{vrs}} / 2\pi$ (GHz)</th>
<th>$\Delta \omega / 2\pi$ (GHz)</th>
<th>$\gamma_{\text{inh}} / 2\pi$ (GHz)</th>
<th>$\alpha$ (cm$^{-1}$)</th>
<th>$n_g$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1455.5</td>
<td>$Z_1\rightarrow Y_7$ ($C_2$)</td>
<td>320</td>
<td>235</td>
<td>230</td>
<td>387</td>
<td>-182</td>
</tr>
<tr>
<td>1461.1</td>
<td>$Z_1\rightarrow Y_6$ ($C_2$)</td>
<td>372</td>
<td>254</td>
<td>248</td>
<td>488</td>
<td>-230</td>
</tr>
<tr>
<td>1495.0</td>
<td>$Z_1\rightarrow Y_5$ ($C_2$)</td>
<td>365</td>
<td>296</td>
<td>290</td>
<td>410</td>
<td>-193</td>
</tr>
<tr>
<td>1516.4</td>
<td>$Z_1\rightarrow Y_3, Y_4$ ($C_2$)</td>
<td>377</td>
<td>387</td>
<td>382</td>
<td>338</td>
<td>-159</td>
</tr>
<tr>
<td>1528.0</td>
<td>$Z_1\rightarrow Y_2$ ($C_2$)</td>
<td>247</td>
<td>164</td>
<td>158</td>
<td>340</td>
<td>-160</td>
</tr>
<tr>
<td>1530.4</td>
<td>$Z'_1\rightarrow Y'<em>3$ ($C</em>{3i}$)</td>
<td>146</td>
<td>74</td>
<td>68</td>
<td>259</td>
<td>-119</td>
</tr>
<tr>
<td>1535.6</td>
<td>$Z_1\rightarrow Y'_1$ ($C_2$)</td>
<td>399</td>
<td>136</td>
<td>130</td>
<td>1075</td>
<td>-510</td>
</tr>
<tr>
<td>1545.2</td>
<td>$Z'_1\rightarrow Y'<em>2$ ($C</em>{3i}$)</td>
<td>68</td>
<td>39</td>
<td>32</td>
<td>104</td>
<td>-47</td>
</tr>
<tr>
<td>1548.6</td>
<td>$Z'_1\rightarrow Y'<em>1$ ($C</em>{3i}$)</td>
<td>69</td>
<td>65</td>
<td>59</td>
<td>65</td>
<td>-29</td>
</tr>
</tbody>
</table>

of the corresponding $\text{Er}^{3+}$ transitions, which are easily identified by comparison with crystal field studies [172, 178]. At higher temperatures, the transmission spectrum exhibits polaritons around additional dispersive transitions along with a concomitant decrease in splitting around the polaritons in Fig. 4.10 as the ions are thermally excited from $Z_1$ to higher Stark levels of $^4 I_{15/2}$.

Unlike the canonical Rabi-split spectrum of two coupled oscillators, the eigenstates for a cavity mode coupled to an inhomogeneously broadened ensemble of dipoles consist of two polariton modes and $(N - 1)$ states with little photonic character and energies determined by the original inhomogeneous distribution [221, 222]. These additional states appear in Fig. 4.10(a,b) as an additional transmission dip as the optical field in the waveguide evanescently senses the large absorption peaks in the $\text{Er}_2\text{O}_3$. The single-pass absorption features occur at the center of the polariton resonances and do not tune with the cavity. They persist when far detuned from any mode, and the response is independent of the mean intracavity photon number during degenerate pump-probe measurements (from $\sim 0.00021$ to at least $\sim 210$ photons). When the taper is touching an undercut but unpatterned $\text{Er}_2\text{O}_3$ film, the slab transmission spectrum contains all the absorption resonances with the same wavelengths and linewidths as in Table 4.2.

While resonantly a pumping a WGM tuned near 1480 nm, the 1525–1555 nm PL spectrum [Fig. 4.10(c)] displays the same anticrossings. However, there are no significant peaks between the cavity-polariton resonances. Weak peaks are occasionally visible between sharp

\[10\] The difference between the observed and reported wavelengths is due to drift in our laser’s wavelength calibration.
>10× stronger cavity-polariton lines, but these peaks are due to spontaneous emission from a small number of excited ions that is collected by the nearby taper. The tuned PL spectra also reveal transitions that do not terminate in the lowest sublevels. At 1553.9 nm, the $Y_1 \rightarrow Z_3$ ($C_2$ site) emission is barely detectable above the noise floor unless resonant with a WGM. The data in Fig. 4.10 reflect the trend in Fig. 4.5 and 4.10(d) where emission from ions on $C_{3i}$ sites is disproportionately large relative to the $C_{3i}$ absorption and the $C_2$ emission. To illustrate this point without rigorous calculations, the polariton splitting [per ion, using the $\sqrt{N}$ dependence in Eq. (4.10)] at the $Z_1 \rightarrow Y_1$ line on both sites roughly follows the absorption cross sections ($\sigma_a$) calculated for erbium-doped $Y_2O_3$ [178]

$$\frac{\delta \omega_{C_2}}{\sqrt{N_{C_2}}} = 3.3 \left( \frac{\delta \omega_{C_{3i}}}{\sqrt{N_{C_{3i}}}} \right) \quad \text{and} \quad \sigma_{C_2}^a = 4.8\sigma_{C_{3i}}^a,$$

as expected under weak pumping. However, the $Y_1 \rightarrow Z_1$ emission (per ion) in Fig. 4.10(d) at 8 K deviates significantly from the emission cross sections ($\sigma_e$) obtained for both sites:

$$\frac{P_{C_2}}{N_{C_2}} = 0.02 \left( \frac{P_{C_{3i}}}{N_{C_{3i}}} \right) \quad \text{and} \quad \sigma_{C_2}^e = 1.04\sigma_{C_{3i}}^e.$$

As discussed previously, a possible explanation for the emission behavior is asymmetric excitation transfer between the $^4I_{13/2}$ levels on the $C_2$ and $C_{3i}$ sites. The interion diffusion rate slows at low temperature because the thermal occupancy of the Stark levels limits the paths that can participate (any nonresonant transfers require phonons to conserve energy), but resonant transfer through the lowest sublevels is always active. One mechanism that may contribute to this asymmetry is the misalignment of the $Y_1$ and $Y_1'$ sublevels. The different crystal fields at the $C_2$ and $C_{3i}$ sites can split identical sublevels between 3 and 240 cm$^{-1}$ (mean absolute shift of 65 cm$^{-1} = 8.1$ meV) [177], and if the $Y_1$ level is higher, excitations transferred to the $C_{3i}$ sublattice may encounter a sizable energy barrier for returning to the $C_2$ ions ($k_BT = 0.74$ meV at 8.6 K). The PL peak at 1545.1 nm in Fig. 4.10(d) complicates this argument because it contains contributions from two unresolved transitions: $Y_1 \rightarrow Z_2$ ($C_2$, nominally 1544.7 nm) and $Y_2' \rightarrow Z_1'$ ($C_{3i}$, nominally 1545.4 nm). If the $Y_1/Y_1'$ misalignment is significant, then the emission from the $C_2$ site must dominate at this wavelength because it has a larger emission cross section in dilute $Y_2O_3$. However, the $Y_1 \rightarrow Z_2$

---

[11] By the nature of the crystal field calculations, the energy levels are found relative to the ground state and without reference to an absolute scale.
cross section is not large enough to account for the disparity with the $Y_1 \rightarrow Z_1$ ($C_2$) peak at 1535.5 nm, which is markedly weaker and should have similar amplitude. If emission from the $C_{3i}$ ions is more intense, another mechanism must be present because it is unlikely that the $Y_2'$ level is also shifted significantly below $Y_1$. Temperature-dependent lifetime measurements for the individual transitions would help clarify this matter (as in Ref. [223]); further information could also be obtained from tertiary alloys such as $(Y_1-xEr_x)2O_3$ similar to the thorough studies of La$_{1-x}Er_xF_3$ [202] and $(Y_1-xTb_x)3Al_5O_{12}$ [200].

The polariton resonances provide useful spectra features for studying the optical transitions, but the homogeneous ($\gamma_h$) and inhomogeneous linewidths $\gamma_{\text{inh}}$ must first be considered. The homogeneous linewidth (and the dephasing time $T_2$: $\gamma_h/2\pi = 1/(\pi T_2)$) is extremely sensitive to the temperature, transition, and host matrix. For example with the $^4F_{9/2} \rightarrow ^4I_{15/2}$ transition in Er$^{3+}$:YLiF$_4$ at 4.2 K, photon echos by MacFarlane et al. show $T_2$ decreases from $>4\mu s$ to 0.4 $\mu s$ when the Er$^{3+}$ concentration increases from 0.1% to 1% [224]. Without previous studies of dephasing in Er$_2$O$_3$, we conservatively use $\gamma_h/2\pi = 10$ GHz at 8 K, which is similar to erbium-doped glass [225,226], and we deconvolve $\{\gamma_h, \gamma_{\text{inh}}\}$ from $\Delta \omega$ in Table 4.2 using the Voigt profile. Inhomogeneous broadening of quantum well excitons has been shown to damp the coherent oscillations with the cavity field [227, 228] without significantly reducing the spectral Rabi splitting [221, 222]. In the limit $\bar{g}_N \gg \gamma_{\text{inh}} > \gamma_h$, the linewidth of the polariton modes becomes $\frac{1}{2}(\gamma_\ell + \gamma_h)$, where $\gamma_\ell$ is again the loaded linewidth of the cavity mode. However, this averaging does not occur when $\bar{g}_N$ and $\gamma_{\text{inh}}$ are similar because the cavity mode experiences more absorption and broadens as it approaches resonance with the optical transition [229]. The magnitudes of $\{\gamma_{\text{inh}}, \bar{g}_N\}$ in these Er$_2$O$_3$ films is in a regime that gives $\delta\omega_{\text{trs}} \approx 2\bar{g}_N$ and only modest narrowing of the polariton modes around the most dispersive transitions; we observe line narrowing of $\sim 20\%$ for the polaritons at 1535 nm.

Using the observed splitting and Eq. (4.10), we find an effective dipole moment $|\langle \vec{\mu} \rangle|$ for each cavity polariton and estimate the absorption coefficient ($\alpha$) and the group index ($n_g$) under weak excitation. For an ensemble of dipoles interacting with an optical field, the electromagnetic susceptibility ($\chi$) can be obtained from a density matrix approach by assuming the dipoles consist of only two levels ($\Delta E = h\omega_o$) and including a phenomenological
dephasing rate:

\[ \chi(\omega) \equiv \chi' - i\chi'' = \frac{\pi|\vec{\mu}|^2(\omega_o - \omega)T_2}{\epsilon_o \hbar} \Delta N \mathbb{L}(\omega) - i\frac{\pi|\vec{\mu}|^2}{\epsilon_o \hbar} \Delta N \mathbb{L}(\omega), \quad (4.11) \]

where \( \Delta N \) is the population difference between the two states (per volume) and the homogeneously broadened lineshape \( \mathbb{L}(\omega) = \mathbb{L}(\omega, \omega_o, \gamma_h) \) is a Lorentzian with width \( \gamma_h \) and normalization such that \( \int_{-\infty}^{\infty} \mathbb{L}(\omega)d\omega = 1 \) [230]. Assuming that the fields are weak and \( \hbar \omega_o \gg k_B T \), the dipoles remain in their ground states, and the frequency- and power-dependent \( \Delta N \) equals the dipole density (\( \rho \)). Then the propagation constant (\( \beta' \)) for a monochromatic plane wave \( \vec{E}e^{i(\omega t - \beta'z)} \) becomes

\[ \beta' \equiv \omega \sqrt{\mu_o \epsilon'} \approx \beta \left( 1 + \frac{\epsilon_o}{2\epsilon} \chi(\omega) \right) \quad \text{... for } |\chi| \ll 1, \quad (4.12) \]

where \( \epsilon' = \epsilon + \epsilon_o \chi \) is the complex permittivity, \( \epsilon \) is the permittivity that accounts for nonresonant contributions to the medium’s polarization, and \( \beta = \omega \sqrt{\mu_o \epsilon} \) is the propagation constant in the absence of the two level transition. The change in the refractive index (\( \Delta n \)) and the intensity absorption coefficient (\( \alpha \)) can be found from the real and imaginary parts of Eq. (4.12):

\[ \frac{\Delta n}{n} = \frac{\Re(\beta') - \beta}{\beta} = \frac{\chi'(\omega)}{2n^2}, \quad (4.13) \]

\[ \alpha = -2\Im(\beta') = \frac{\beta \chi''(\omega)}{n^2}, \quad (4.14) \]

with the refractive index far from the transition is given by \( n = \sqrt{\epsilon/\epsilon_o} \). Strong transitions with narrow linewidths will also affect the group index:

\[ n_g = n + \omega \left( \frac{dn}{d\omega} \right) \approx n + \frac{\omega}{2n} \left( \frac{d\chi'}{d\omega} \right) \quad \text{... for } |\chi| \ll 1. \quad (4.15) \]

Equation 4.10 can be combined with the formulas above to simply express the peak absorption coefficient and group index for a homogeneously broadened transition in terms of polariton parameters:

\[ \alpha(\omega_o) = \frac{4n|\vec{g}_N|^2}{c\gamma_h \Gamma} \quad \text{and} \quad n_g(\omega_o) = n \left( 1 - \frac{4|\vec{g}_N|^2}{\gamma_h^2 \Gamma} \right). \quad (4.16) \]
A common approach to modeling inhomogeneous broadening is superimpose the dielectric response of many independent transitions that are evenly distributed in frequency space [221, 222]. The amplitude of the dipole moment $\{\vec{\mu}_j\}$ at each frequency is given a Gaussian weight $\{f_i\}$ according to the inhomogeneous distribution such that $|\vec{\mu}_j| = \sqrt{f_j} \langle \vec{\mu} \rangle$ and $\sum f_j = 1$. If the frequency separation between the transitions is less than their homogeneous, their combined susceptibility is a smooth function, from which $\alpha$ and $n_g$ can be obtained using Eqs. (4.14) and (4.15).

The results of this model for the polariton resonances at 8.6 K are displayed in Fig. 4.11; the peak absorption coefficient and group index at each transition are summarized in Table 4.2. Compared to the absorption spectrum at $T = 300$ K in Fig. 4.5(b), the larger low-temperature peaks reflect the trends in Eq. (4.16) regarding narrower homogeneous linewidths and increased population of the lowest sublevel—recall the collective coupling rate depends on the population: $|\bar{g}_N| = \sqrt{N}|\bar{g}_1|$. Rare-earth ions commonly used in slow-light experiments because of their long radiative lifetimes. For example, coherent population oscillations have been used to achieve $n_g \approx 10^8$ by opening a narrow hole at 1536 nm in inhomogeneously broadened Er$^{3+}$:Y$_2$SiO$_5$ [231], but Fig. 4.11(b) also illustrates the poten-
tial for high-bandwidth fast-light experiments in Er$_2$O$_3$ without coherent control fields. At the absorption peaks, pulse advancement of $t_{\text{adv}} \approx 36$ ps (for $-10$ dB signal attenuation) is possible at all the observed $Z_1 \rightarrow Y_x$ lines, which is expected based on Eq. (4.16) for a homogeneously broadened transition:

$$ t_{\text{adv}} = (1 - n_g) \frac{L}{c} = 0.230(1 - n_g) \frac{A}{\alpha c} \approx 0.230 \frac{A}{\gamma_h}, $$

(4.17)

where $A$ is the desired signal attenuation in dB and by assuming $n_g \ll 0$. Despite the small advancement, the large $\gamma_{\text{inh}}$ value allows the advancement-bandwidth product ($t_{\text{adv}} \Delta \omega/2\pi = 4.9$ for $Z_1 \rightarrow Y_1$) to considerably exceed those in other fast-light systems, such as Cr$^{3+}$:BeAl$_2$O$_4$ ($t_{\text{adv}} \Delta \nu = 0.061$) [232]. The prospects for collective coherent emission are discussed at the end of the next section.

4.2.5.2 High-Temperature Polaritons

At 300 K, the transmission spectra of Er$_2$O$_3$ microdisks contain split resonances near 1537 and 1549 nm similar to those at 8 K. Unfortunately, the N$_2$ adsorption technique used to tune the microdisk resonances at cryogenic temperatures has no convenient analogue near 300 K, and digitally etching [233] the devices is not ideal due to the chemical resistance of Er$_2$O$_3$. The temperature dependence of the refractive index offers a convenient and reversible alternative, but the low tuning rate for these modes (9.7 pm/K, corresponding to $dn/dT = 1.9 \times 10^{-5}$ K$^{-1}$) limits the wavelength range for simply heating the entire sample (maximum $\Delta \lambda \approx 0.6$ nm for $\Delta T \approx 60$ K). Instead, we use optical absorption and the low thermal conductivity associated with undercut microcavities to locally heat the optical mode volume of a single device, similar to Ref. [234] but more efficient. Pumping conventional microdisks, as in the previous sections, with $\sim 3$ mW of absorbed power only achieves $\Delta \lambda \approx 2.0$ nm, which corresponds to $\Delta T \approx 200^\circ$C. To obtain a wider tuning range, we further decrease the resonator’s thermal conductivity by fabricating microdisks with three thin spokes supporting a 4-$\mu$m wide ring where the mode resides [Fig. 4.12(a)]. The spokes, nominally 2-$\mu$m wide, incur some additional scattering [Fig. 4.12(b)] but only marginally degrade the $Q$ of the $\text{TE}_{p=1}$ WGMs—the second-order WGMs experience higher scattering losses due to greater overlap with the spokes [Fig. 4.12(d)]. With 3.6 mW of pump power in Fig. 4.12(e), we are able to red-shift a WGM from 1489.8 to 1495.5 nm. Assuming the
Figure 4.12: (a) Optical image of a spoked Er\textsubscript{2}O\textsubscript{3} microdisk. (b) Visible upconversion luminescence shows the spatial profile of a fundamental mode and additional scattering where the spokes attach to the ring. (c) Scanning electron micrograph showing the smooth edge of an undercut microdisk. (d) Electric field profile (|E|^2) of the first- and second-order radial modes calculated using FEM. (e) Thermo-optic tuning of a single WGM in a spoked microdisk. The pump power increases \(\sim 200 \mu W\) between traces from blue (146 \(\mu W\)) to red (3.58 mW).

To tune the resonator and observe the anticrossing, we perform a counterpropagating pump-probe measurement that simultaneously heats the cavity, monitors the cavity transmission, and collects the Er\textsuperscript{3+} luminescence [Fig. 4.13]. The counterpropagating configuration and edge-pass filters combine/split the beams with <1 dB loss and provide >100 dB pump-probe isolation at the probe detector and OSA. Unlike many cQED experiments that

\begin{itemize}
  \item thermo-optic coefficient of Er\textsubscript{2}O\textsubscript{3} rate remains constant, the 5.7 nm change corresponds to \(\Delta T = 590 K\) and a final temperature of \(T = 890 K\); currently we have no method to verify temperatures above \(\sim 375 K\).
\end{itemize}
must minimize $V_{\text{eff}}$, the uniform dipole density for planar microdisks [see Eq. (4.10)] makes $\bar{g}_N$ independent of $V_{\text{eff}}$ so $\Gamma$ and $\rho$ become more important design parameters. For these experiments, a microdisk radius of $\sim 20 \mu m$ is again chosen so the quasi-TE mode spectrum consists of evenly spaced first- and second-order radial modes—$\Gamma = 0.67$ and has a weak dependence on radius. This radius is also chosen because the FSR between modes of a single family is approximately equal to the separation of the polaritons ($\Delta \lambda \approx 12 \text{ nm}$) so the cavity can be tuned onto resonance simultaneously with both transitions. The cavities’ radii are also varied slightly to acquire a device with first-order WGMs slightly blue detuned from the polariton features at 1537 and 1549 nm. With the fiber-taper input/collection waveguide in contact with the disk to stabilize the coupling against mechanical noise, we pump a fundamental WGM near 1490 nm with 0.5 to 2.5 mW in the clockwise direction to heat the disk and drive the $^4I_{13/2} \leftrightarrow ^4I_{15/2}$ transition. While the 1490-nm pump laser is tuned to maximize the dropped pump power, a second $\sim 200 \text{ nW}$ tunable laser probes the cavity’s transmission spectrum in the counterclockwise direction. The C-band PL is also measured by simply blocking this probe beam.

As the cavity is heated, the split resonances clearly anticross in both transmission [Fig. 4.14(a)] and PL [Fig. 4.14(b)] while the bare cavity modes tune as a quadratic polynomial of the pump power—the nonlinearity stems from Er$^{3+}$ cooperative upconversion
which produces a cascade of multiphonon transitions. Parts of these anticrossings are also observed in separate measurements when we heat the sample stage from 300 to 361 K, but, as discussed previously, this temperature range is insufficient to tune a cavity mode completely across the Er$^{3+}$ transitions. In both measurements when the cavity modes are detuned from the Er$^{3+}$ transitions as in Fig. 4.14(d), weak transmission features remain near the polariton resonance wavelengths due to single-pass loss in the taper waveguide as it evanescently senses the Er$_2$O$_3$ absorption spectrum.

In addition to the spectral anticrossing, the linewidths of the polariton branches average as the cavity modes and Er$^{3+}$ lines become resonant, further indicating the hybridization of the photonic and atomic degrees of freedom. For each trace in Fig. 4.14(a), we fit a series of Lorentzian valleys [Fig. 4.14(d)] to extract the resonances’ loaded linewidths and demonstrate they become equal near the anticrossing for both sets of polaritons [Fig. 4.14(e,f)]. At higher pump powers (higher cavity temperature), the widths of the Er$^{3+}$ lines broaden, which is likely due to added phonon-related dephasing [225, 226].

To quantitatively describe the polaritons and obtain $\bar{g}_N$, we diagonalize and fit a coupled-oscillator Hamiltonian to the Rabi-split eigenfrequencies in Fig. 4.14(a). Assuming the quasi-TE cavity modes are linearly polarized, $\bar{g}_N$ becomes real, and we use five parameters for the tuning of each cavity-polariton pair: $\delta \omega_{\text{vrs}}$, the Er$^{3+}$ transition wavelength ($\lambda_o$), and the three coefficients that describe the quadratic tuning of the bare cavity mode. For $\delta \omega_{\text{vrs}} \approx 2|\bar{g}_N|$, the fit in Fig. 4.14(c) gives $\bar{g}_N/2\pi$ of 99 and 57 GHz for the Er$^{3+}$ transitions at 1537.2 and 1549.2 nm, respectively. The weak coupling to each ion ($\bar{g}_1/2\pi \approx 0.3$ and 0.1 MHz for $N \approx 10^{11}$) reflects the necessity of high ion densities to achieve resolved vacuum-Rabi splitting while compensating for erbium’s small dipole moment. Based on the spectroscopy of Er$^{3+}$ in Y$_2$O$_3$ [178] and the data in Fig. 4.10 at 8 K, we associate the 1537.2 and 1549.2 nm splittings with $^4I_{13/2} \leftrightarrow ^4I_{15/2}$ transitions between their lowest Stark levels on the $C_2$ (Y$_1 \leftrightarrow Z_1$) and $C_{3i}$ (Y$_1' \leftrightarrow Z_1'$) lattice sites, respectively. However, neighboring resonances between higher sublevels have significant dipole moments and may also contribute to the polariton behaviour—e.g., the Y$_2 \leftrightarrow Z_2$ ($C_2$) and Y$_3 \leftrightarrow Z_3$ ($C_{3i}$) transitions near 1537 and 1549 nm, respectively [178].

The calculation in the previous section can be reversed using the absorption data in Fig. 4.5(b) to estimate effective dipole matrix elements that include contributions from multiple transitions. However, there is some ambiguity in deconvolving $\{\gamma_h, \gamma_{\text{inh}}\}$ from the
Figure 4.14: (a) Normalized probe transmission and (b) PL as a function of wavelength and pump power. To compensate for upconversion effects, the PL spectra are normalized to give a constant value for the peak near 1525 nm. (c) Comparison of measured resonance wavelengths, the coupled-oscillator eigenvalues (solid lines), and the uncoupled transition and cavity responses (dashed lines). The device temperature is inferred from the bare cavity mode detuning and a constant tuning rate of 9.7 pm/K. (d) Transmission for two devices tuned on and off resonance with the Er\textsuperscript{3+} transitions; a sample fit is also included. (e,f) Linewidth averaging of the polariton modes.
measured high-temperature linewidths ($\Delta \omega$). We use the $\gamma_{\text{inh}}$ values for $Y_1 \leftrightarrow Z_1$ and $Y_1' \leftrightarrow Z_1'$ in Table 4.2 which assumed $\gamma_h/2\pi = 10$ GHz at 8.6 K; errors in $\gamma_{\text{inh}}$ due to this assumption (error < 6 GHz) are less significant at 300 K because the total linewidths are larger. Greater uncertainty arises if nearly resonant transitions between other Stark levels further inhomogeneously broaden the absorption lines. For simplicity, we assume this increase is marginal compared to $\gamma_h$. Finally, the measured values for $\{\Delta \omega\}$ differ in the absorption spectrum [obtained from Fig. 4.14(e,f) when the WGMs are not resonant] and in the PL spectrum [from the full width at half maximum in Fig. 4.5(a)]. The results for $\{\delta \omega_{\text{vrs}}, \Delta \omega, \gamma_h\}$ are summarized in Table 4.3 along with the calculated $\bar{g}_N$ values for each transition. For comparison, the Rabi splitting for a homogeneously broadened atomic ensemble is

$$\delta \omega_{\text{vrs}} = 2 \sqrt{\bar{g}_N^2 - \frac{1}{4} \left(\kappa - \frac{\gamma_h}{2}\right)^2}, \quad (4.18)$$

where $\kappa$ is the cavity field decay rate in the absence of the atomic transition [214]. Neglecting the inhomogeneous broadening, the fitted splitting around 1537.2 nm and Eq. (4.18) give $|\bar{g}_N|/2\pi = 105$ GHz with $\{\delta \omega_{\text{vrs}}, \kappa, \gamma_h\}/2\pi = \{198, 113, 361\}$ GHz. Given that $\gamma_{\text{inh}}$ can also slightly reduce the observed splitting, the method in the previous section predicts 113 GHz with the absorption spectrum linewidth (which is preferred in Ref. [222]). Following the generally observed behavior for the magnetic-dipole transitions on the $C_{3i}$ sites, the Rabi splitting around $Y_1' \leftrightarrow Z_1'$ at 1549.2 nm is larger than expected from the absorption. Treating this polariton resonance as a single homogeneous transition shows $\delta \omega_{\text{vrs}} \approx 2|\bar{g}_N|$ and $\gamma_h$ is not significant—in Eq. (4.18), $|\bar{g}_N|/2\pi = 58$ GHz with $\{\delta \omega_{\text{vrs}}, \kappa, \gamma_h\}/2\pi = \{114, 50, 57\}$ GHz.

We are also working to understand why high-temperature polaritons only form around $Y_1 \leftrightarrow Z_1$ transitions and not around transitions involving higher Stark levels with similar

---

12The inhomogenous linewidths obtained at $T = 8$ K are repeated for completeness.
absorption peaks.

The robust collective coupling we have observed between Er\(^{3+}\) transitions and the optical modes in Er\(_2\)O\(_3\) resonators is an appealing route to improving the material’s efficiency or gain through cQED effects, like cavity-enhanced superradiance [235] and the Purcell Effect [4]. Unfortunately to achieve superradiant emission, with yield increasing as \(N^2\), the ensemble must be first excited into a Dicke state that quickly decoheres due to fast dephasing above cryogenic temperatures and excitation diffusion [236]. Because collective relaxation of the polariton modes is difficult to initiate, Purcell enhanced emission from individual Er\(^{3+}\) ions is a more promising direction and can be realized by decreasing \(V_{\text{eff}}\) to increase the coupling rate per ion [6, 214]. From the perspective of a single excited ion in the present devices, the other \((N-1)\) ions create a dispersive resonance in the refractive index that generates the paired polariton modes [213]. The weak coupling between each ion and the large WGM at rate \(\bar{g}_1\) can then be treated perturbatively, which causes negligible enhanced for these system parameters—the expected Purcell “enhanced” lifetime \(\kappa/2\bar{g}_1^2 > 150\text{ms} [214]\) is much longer than the \(7.2\mu s\) \((8\text{ms})\) bulk nonradiative (radiative) lifetime. Quantum master equation models produce similar behaviour using a single Er\(^{3+}\) ion and a “privileged” ion coupled to an optical cavity mode at rates of \(\bar{g}_1\) and \(\bar{g}_N\), respectively. By integrating proposed ultrasmall mode volumes formed using nanoscale slots [18] with recent low-loss photonic crystals in Si [237], a hybrid Si cavity containing a Er\(_2\)O\(_3\) slot \([\Delta n \approx 1.5, V_{\text{eff}} \approx 0.18(\lambda/n)^3, Q \approx 10^6, \kappa/2\pi \approx 100\text{MHz}]\) could exhibit an enhanced emission lifetime of \(\kappa/2\bar{g}_1^2 \approx 0.23\mu s\).

### 4.2.6 Rate Equation Estimates

In more conventional cavities where the Purcell Effect is negligible, the large Er\(_2\)O\(_3\) upconversion rate at low pump power alters the population dynamics for the Er\(^{3+}\) levels when compared to the quasi-two-level system present in fiber amplifiers [165]. To accurately describe an Er\(_2\)O\(_3\) cavity at the pump rates used during the radiative efficiency measurement (Fig. 4.9), a rate equation model must include the populations of several excited states (likely all levels up through \(^4S_{3/2}\) and \(^2H_{11/2}\)) and the accompanying coupling terms and coefficients (e.g., radiative and nonradiative decay into other levels or upconversion paths). However with only the \(\{\tau_{\text{rad}}, \tau_{\text{nr}}, C_{\text{up}}\}\) parameters for \(^4I_{13/2}\), we can roughly estimate the power necessary to reach transparency in the C/L-bands by solving a simplified three-state
model for pumping at $\sim 1480$ nm:

$$N_1 = N_{\text{Er}} - N_2 - N_3,$$

$$\frac{dN_2}{dt} = -\frac{N_2}{\tau_{\text{nr}}} - C_{\text{up}}N_2^2 + s\Phi(N_1 - rN_2),$$

$$\frac{dN_3}{dt} = -\frac{N_3}{\tau_{\text{nr}}} + \frac{1}{2}C_{\text{up}}N_2^2.$$ (4.21)

The populations $\{N_1, N_2, N_3\}$ represent densities for ions in the $\{^4I_{15/2}, ^4I_{13/2}, ^4I_{9/2}\}$ states, respectively; higher states are neglected because we lack reasonable estimates for the appropriate upconversion and/or excited-state absorption coefficients. The $^4I_{11/2}$ level is omitted because it quickly decays to $^4I_{13/2}$ through multiphonon emission—this assumption produces a lower effective upconversion rate into $^4I_{9/2}$. For simplicity, the crystal-field splitting for all levels is ignored except that the emitted photons from $^4I_{13/2}$ are red-shifted from pump photons. The total ion density on $C_2$ lattice sites ($N_{\text{Er}}$) is $2.05 \times 10^{22}$ cm$^{-3}$, and we use our estimated value for the cooperative upconversion coefficient $C_{\text{up}} = 5.1 \times 10^{-16}$ cm$^{-3}$/s.

We assume the $^4I_{13/2}$ and $^4I_{9/2}$ lifetimes are approximately equal ($\tau_{\text{nr}} = 7.2$ $\mu$s) because both transitions to the ground state are dipole forbidden and likely subject to similar excitation diffusion rates. This analysis excludes spontaneous emission (low radiative efficiency) and stimulated emission (leaky cavity at the emission wavelength). The pump photon flux within the cavity is given by $\Phi$, and $r$ is the ratio of the emission and absorption cross sections at the pump wavelength. For Er$^{3+}$-doped silica, $r \approx 1/3$ at $1480$ nm [165] and accounts for emission stimulated by the pump beam. The adjustable parameter $s = 2.5 \times 10^{-12}$ encompasses a number of factors including the value of the absorption cross section; it is set to give $N_2 \approx 2.7 \times 10^{20}$ ions/cm$^3$ at $P_d = 204$ $\mu$W—corresponding to the inferred values from the radiative efficiency results in Fig. 4.9. This model for the Er$^{3+}$ transitions is then applied to a microdisk cavity (20 $\mu$m radius, material volume $V = \Gamma V_{\text{eff}} = 20.1$ $\mu$m$^3$).

Steady-state solutions to these rate equations give approximate pump powers at which $N_2 \approx N_3$ and at which these transitions near transparency. By $P_d \approx 0.6$ mW, there are $\sim 5 \times 10^{20}$ ions/cm$^3$ in both the $^4I_{13/2}$ and $^4I_{9/2}$ levels. These microdisks approach transparency for $^4I_{9/2} \rightarrow ^4I_{15/2}$ ($\lambda \approx 800$ nm) and $^4I_{13/2} \rightarrow ^4I_{15/2}$ ($\lambda \approx 1480$ nm) with pump powers of $P_d \approx 18$ mW and $P_d \approx 130$ mW, respectively. To improve the accuracy, this model should include $^4S_{3/2}$, which would encompass the minimum number of levels to account for the green emission. Since the upconverted luminescence is most intense around $\lambda \approx 550$ nm,
the combined upconversion-emission path for $^4I_{9/2} \rightarrow ^4S_{3/2} \rightarrow ^4I_{15/2}$ may provide a fast route back to the ground state. A fast green relaxation would increase the transparency thresholds and may make the $^4I_{9/2} \rightarrow ^4I_{15/2}$ relaxation a secondary process. Exploratory pulsed measurements using a Tm$^{3+}$-doped fiber amplifier at 1480 nm also suggest some of the upper levels may be longer lived (on the order of 100 µs). In summary, the upconversion processes close to transparency are quick enough to produce substantial populations in every level up to and including $^4S_{3/2}$, and the $^4I_{13/2}$ level is likely not the first level to reach transparency when pumping at $\sim 1480$ nm.

### 4.3 Conclusions

The rate-equation estimates suggest a prohibitive amount of power is required to invert the $^4I_{15/2}$ manifold in microphotonic Er$_2$O$_3$ cavities due to the fast excitation diffusion to nonradiative sites and the large cooperative upconversion coefficient. Since the rapid upconversion prevents transparency in the C-band until most of the ions are sequestered in higher states, individual upper levels (e.g., $^4S_{3/2}$) may be the first to invert relative to the ground state making Er$_2$O$_3$ upconversion green lasers a possibility [187–190]. Epitaxial Er$_2$O$_3$-on-Si might also be developed into an incoherent visible emitter—rough estimates based on the camera’s sensitivity in Fig. 4.9 give a green radiative efficiency on the order of 5% for an absorbed power density of 0.15 mW/µm$^3$. Additionally, there is potential for using Er$_2$O$_3$ to shift infrared radiation into the visible spectrum (i.e., above the bandgap of silicon) as part of multijunction silicon solar cells.

To achieve technological maturity for the original application of Er$_2$O$_3$ to waveguide amplifiers and lasers for on-chip optical networks, it is possible, but improbable, that the performance of pure Er$_2$O$_3$ films can be improved by identifying and eliminating the quenching impurities and finding suitable dopants to screen the ion-ion interactions. More likely, future progress will be due to the precision and flexibility of the ALE growth. Ternary oxides with Y and Gd, such as $(Y_{1-x}Er_x)_2O_3$, would be better lattice matched to silicon, and their lack of infrared and visible transitions would slow the upconversion and non-radiative processes by increasing the inter-ion separation for the optically active species. Detailed relaxation studies (as in Ref. [202]) of these alloys are necessary to characterize the excitation diffusion and to optimize the emitter density for the desired gain. Electrolu-
minescence may be achieved through R$_2$O$_3$-Si heterostructures and superlattices [238-240], but efficient excitation must be explored first including the alignment of the R$_2$O$_3$ 4$f$ levels to the Si bands. Finally, the spectral response of rare-earth oxides can be designed for specific applications. For example, (Gd$_{1-x}$Pr$_x$)$_2$O$_3$ may be more suitable for silicon photonics because the $^1G_4\rightarrow^3H_5$ transition in Pr$^{3+}$ at 1310 nm does not terminate in the ground state ($^3H_4$) and operates at the same wavelength as SiGe electro-absorption modulators [156] and APDs [241], and the phase space for engineering sensitization and cross-relaxation processes increases significantly in quaternary oxides. As in III-V systems, high-quality epitaxy will become crucial in controlling the material’s structure and optical properties while moving toward CMOS-compatible rare-earth laser diodes and photodetectors.
Appendix A

Experimental Supplement

A.1 Mach-Zehnder Interferometer

The accurate determination of spectral linewidths is essential to much of this research. The typical testing arrangement in Fig. 2.4 employs an external-cavity tunable diode laser to sweep across the desired wavelengths and a computer to synchronize the laser’s reference output (or start trigger) with the analog transmission/reflection signal acquired by the photodetector. The method and accuracy of this synchronization depends on the laser model and is particular to each unit. For example, the piezo-tuning mode for the New Focus Velocity lasers applies a sawtooth waveform to a piezoelectric element attached to the tuning mirror; the computer also receives this waveform and applies a set calibration to its amplitude to determine the wavelength. Unfortunately, the wavelength span can vary up to 10% from day to day for a given voltage, and the forward and backward scans are hysteretic (see Ref. [60] for more information). Another common problem in older Velocity lasers is that the position sensor on their tuning arm (used for coarse motor scans) can begin to malfunction and produce an offset between the laser’s true wavelength and the wavelength indicated by the laser head—this offset can be removed during post-processing using the OSA. In general, the repeatability issues with the Velocity lasers do not impede daily testing, but precise measurements over several days (e.g., Ref. [56]) require calibration. In contrast, the discontinued Vidia lasers only send a single pulse at the beginning of the scan (after their acceleration phase), and the computer simply assumes the laser is sweeping at the set rate after that time. To date, there has been no need to repeatedly calibrate narrow Vidia sweeps.

To calibrate a laser’s sweep range, we use a fiber-based Mach-Zehnder interferome-
ter consisting of two 50:50 splitters connected by two SMF-28e patch cords with different lengths. The entire assembly is placed inside a insulated cooler with fiber thru-ports to prevent thermal drift. The interference fringe spacing is given by \( \delta_{fs} = \frac{\lambda^2}{\bar{n}_{eff} L} \), where \( \bar{\lambda} \) is the average wavelength between the two fringes, \( \bar{n} \) is the fiber’s effective index at \( \bar{\lambda} \), and \( L \) is the path length difference between the two arms. For \( L \approx 1 \) m, the reference spacing and wavelength are \( \delta_o = 1.57 \pm 0.03 \) pm at \( \lambda_o = 1460.5 \) nm. The fringe spacing can be measured by performing laser scans over a small range without the customary acceleration phase, finding the start and stop wavelengths on the OSA, and dividing by the number of fringes. The resulting transmission interferogram consists of a chirped signal as a function of wavelength:

\[
T(\lambda) = T_o \sin \left[ \frac{2\pi \lambda_o^2}{\delta_o} \left( \frac{1}{\lambda} - \frac{1}{\lambda_1} \right) \right] + T_{dc},
\]

where \( \{T_o, T_{dc}\} \) are amplitudes to account for incomplete interference contrast and \( \lambda_1 \) is the wavelength of a fringe maximum [to set the phase of \( T(\lambda) \)]. At \( N \) fringes from the maximum at \( \lambda_1 \), the wavelength (\( \lambda \)) and the uncertainty (\( \sigma_\Delta \)) in the separation (\( \Delta = \lambda - \lambda_1 \)) are

\[
\lambda = \frac{\lambda_o^2}{\lambda_o^2/\lambda_1 - N\delta_o} \quad \text{and} \quad \sigma_\Delta = \left| \frac{N\lambda_o^2}{(\lambda_o^2/\lambda_1 - N\delta_o)^2} \right| \sigma_\delta,
\]

where \( +N \) corresponds to a red shift (\( \lambda > \lambda_1 \)), \( \sigma_\delta \) is the uncertainty in the measured \( \delta_o \) fringe spacing, and \( \sigma_\Delta \) assumes \( \delta_o \) is the only significant source of error. For practical measurements, the fractional uncertainty in the calibrated linewidth (\( \sigma_\Delta/\Delta \)) is the same as the fractional uncertainty in the fringe spacing (\( \sigma_\delta/\delta_o \)).

The calibration procedure for spectral measurements depends on the desired accuracy. For general exploratory testing, the piezo-tuning ranges should be calibrated when the laser arrives from the factory and then periodically verified; the DC motor accuracy can be assessed by comparing the wavelength set point and the laser output on the OSA. Much of the early data in §3.2 and §2.2 was taken by checking the piezo calibration daily or even more frequently. A more accurate procedure is to acquire a real-time calibration signal on a second detector, as in Fig. 2.4. The most expedient use of this data is to find the true wavelength span of the data by counting fringes and then scale any raw linewidths by the ratio of the true and raw spans. This method was used throughout §4.2, but it assumes the scan rate is constant. The most accurate method is to simply count fringes starting at the beginning of the scan and to ignore the wavelength data from the laser. Time used for
further improvements would likely be better spent building another MZI with smaller $\delta_o$ and $\sigma_\delta$.

The pure fringe counting method has been implemented to calibrate large wavelength ranges while studying dispersion-engineered microdisks for degenerate four-wave mixing. Designed and fabricated by Q. Lin, the disks consist of a Si core and SiO$_2$ cladding to control the GVD, similar to the waveguides in Ref. [117]. The disks are repeated oxidized to tune the ZDWL for the TM$_{p=1}$ modes toward 1550 nm in order to achieve momentum and energy conservation for four-wave mixing between three consecutive modes. Because future experiments call for a coarse wavelength division multiplexer (20 nm channel spacing), the FSRs separating the signal and idler modes from the pump mode must be approximately equal to the channel spacing (2530 GHz) with mismatched comparable to the modes’ linewidths ($\sim$0.6 GHz). For comparing two wavelength spans, the measurement is less sensitive to uncertainty in $\delta_o$, and the uncertainty in the difference $(\Delta_1 - \Delta_2)$ between the two spans is

$$\sigma(\Delta_1 - \Delta_2) = \left| \frac{N_1 \lambda_o^2}{(\lambda_o^2/\lambda_1 - \delta_o N_1)^2} - \frac{N_2 \lambda_o^2}{(\lambda_o^2/\lambda_2 - \delta_o N_2)^2} \right| \sigma_\delta.$$  

(A.3)

Using the MZI and sampling every $\sim$0.1 pm, Fig. A.1 shows a sample calibrated data set for a device with signal and idler modes well matched about the pump mode at 1528.8 nm. The difference between the calibrated and raw wavelengths is at most 0.25 nm, and the scan rate is clearly not linear. The pump-signal and pump-idler separation is $\sim$2676.7 GHz with a difference of 0.2$\pm$2.8 GHz, which is a relative error of 0.1%. To achieve the necessary accuracy of $\leq$0.02%, the MZI fringe spacing must be increased to $\delta_o \approx 7$ pm if the uncertainty in the spacing cannot be reduced ($\sigma_\delta = \pm 0.03$ pm).

In the current implementation, the MZI reference signal is loaded into an $(N \times 1)$ matrix and analyzed as a function of the matrix index. First, the mean is subtracted, and a zero-finding routine determines the “fractional” indices where the fringes cross the axis. These positions in the matrix are separated in wavelength by half the fringe spacing, and they provide a piecewise function for interpolating the indices of the original vector onto calibrated wavelengths. Since the reference and transmission signals are sampled simultaneously, the same calibrated wavelength vector applies to the transmission data. The dependence of the fringe spacing on wavelength is included by quadratically scaling $\delta_{fs}$ at each “zero” crossing.
Figure A.1: (a) Raw TE- and TM-polarized transmission spectra for a Si-SiO$_2$ microdisk after calibrating the wavelength. The MZI reference signal was acquired concurrently with the TM spectrum. Nearly equal FSRs between the highlighted TM$_{p=1}$ modes shows the ZDWL is near 1530 nm. The inset contains a 10-pm section of the calibration signal showing measured data points (∗) and the interpolated “zero” crossings (◦). (b) Difference between the calibrated and raw wavelengths for a long Vidia laser scan.
to \( \delta'_{fs} = \delta_o (\lambda'/\lambda_o)^2 \) where the reference splitting again is \( \delta_o = 1.57 \text{ pm} \) at \( \lambda_o = 1460.5 \text{ nm} \). It assumes the starting wavelength is accurate; the lasers’ ±0.05 nm repeatability is sufficient for these purposes. This method also ignores dispersion in the optical fiber because it introduces a correction an order of magnitude smaller than the present uncertainty in \( \delta_{fs,o} \). For completeness, the dispersion can be calculated using parameters obtained from the manufacturer (Corning for SMF-28e) and is presented below for \( 1200 \text{ nm} \leq \lambda \leq 1625 \text{ nm} \):

\[
D \equiv \frac{-2\pi c}{\lambda^2} \frac{\partial^2 k}{\partial \omega^2} = \frac{\pi c S_o}{2} \left[ \frac{1}{\omega} - \frac{\omega^3}{\omega_0^4} \right],
\]

\[
n_g = \pi^2 c^3 S_o \left[ \frac{\omega^2}{2\omega_0^6} + \frac{1}{2\omega^2} \right] + C_1,
\]

\[
n_{eff} = \pi^2 c^3 S_o \left[ \frac{\omega^2}{6\omega_0^6} - \frac{1}{2\omega^2} \right] + C_1 + \frac{C_2}{\omega},
\]
1) Connect the laser source and a FPC\textsuperscript{1} to the modulator, and affix the modulator and patch cords to the optical table. Maximizing the extinction ratio requires great stability in the polarization and source wavelength. Let the laser warm up and adjust any fiber and electrical connections prior to this procedure.

2) Attach the modulator’s output pigtail to the table. The output does not need to be stabilized to the same extent as the input, but movement of the output fiber should not cause the modulator’s input to shift.

3) With the output connected to an OSA or power meter, maximize the transmission polarization and modulator’s DC bias to calibrate the modulator’s insertion loss. This step needs to be done only once.

4) Optimize the polarization and DC bias to minimize the transmission. The optimal polarization should be close to that found in the previous step.

5) Wait. The modulator takes a few hours to fully settle. Periodically readjust the polarization (should not significantly change) and DC bias to minimize transmission. The bias must be monitored frequently at first while the transmission is changing rapidly. At $>30 \text{ dB}$ extinction, the transmission will be very sensitive to any small changes.

6) Once the EOM has settled, it requires small tweaks to maintain the maximum extinction.

7) Connect the EOM output to an oscilloscope. Be aware of the maximum peak power for the high-speed detectors, and attenuate the optical pulses if needed.

8) Set the amplitude of the RF signal/pulses to obtain the desired optical waveform and measure the peak power. Prior to acquiring data, it is prudent to examine the optical signal produced by the modulator and a given voltage source. For example, the trailing edge of pulses from the Avtech pulse generator (AVPP–1–C) overshoots zero and produces a small secondary pulse. These after pulses are $\sim15 \text{ dB}$ attenuated from the main pulse and occur $\sim15 \text{ ns}$ after the trailing edge; the exact waveform is

\textsuperscript{1}The FPC may not be necessary when using polarization-maintaining (PM) fiber, but this procedure assumes the polarization of a standard single-mode fiber must be matched to the modulator’s PM input fiber.
dependent on the pulse width and amplitude. In addition, the optical transmission between the pulses is increased 1.2 dB relative to the transmission without the RF signal. For small amplitudes and widths, there are also small prepulses.

9) Throughout the experiment, adjust the signal and bias amplitudes to maintain the pulse shape and extinction ratio because the modulator response will continue to slowly drift. It is convenient to incorporate a few splitters into the optical path so the pulses can be easily monitored without constantly breaking fiber connections.

The high extinction modulator (>40 dB) utilizes polarization filters at the input and output and an additional port for fine tuning the DC bias, and the above procedure can be used to generate pulses with an extinction ratio >62 dB with one modification. Once the bias and polarization are optimized for maximum transmission, the input polarization should not be changed at any point. Otherwise the laser will be partially attenuated by the polarization filters and not the LiNbO$_3$ MZI within the modulator.

**A.3 Effective Index for the HE$_{11}$ Mode**

Because of the ubiquitous use of fiber-tapers throughout these measurements, it is often convenient to have a fast, accurate function to calculate the effective index for the fundamental HE$_{11}$ taper mode. For the cylindrical step-index waveguide with infinite cladding, the eigenvalues require the solutions of a transcendental equation similar to Eq. (1.6), although without the radiating fields (see Ref. [49]). Unfortunately, the complexity of this equation makes numerical solvers prone to bugs. While benchmarking various versions of these solvers, it was noticed that the graphical solutions presented by Barnoski [242] and later Yariv and Yeh [49] contained slight errors when compared to code implementing their analytic solutions and independent FEM models. These errors are approximately 1%–3% and plotted in Fig. A.2.

**A.4 Coherent Coupling to a Uniform Density of Dipoles**

Because the coherent coupling rate between a dipole and an optical mode in a cavity depends on the local electric field at the dipole, significant effort is given to localizing the emitter at an antinode of the cavity as in laser-cooled Cs atoms within a Fabry-Pérot cavity [243]
Figure A.2: Comparison of $n_{\text{eff}}$ values for the HE$_{11}$ mode of a fiber-taper or step-index optical fiber with $n_1 = 1.45$ (silica) and $n_2 = 1.0$ (air). Data points for the published curves by Barnoski [242] and Yariv and Yeh [49] were scanned and digitally extracted from the original sources. The analytic solver was implemented by P. Barclay. The FEM model calculated $n_{\text{eff}}$ at a fixed wavelength (1550 nm) while varying the core radius from 0.2 to 1.5 µm.

and precisely aligning photonic crystal cavities to buried quantum dots [244]. For many-body systems, the collective interaction must account for the field seen at every dipole. Cooperative interactions delocalize the excitations in Er$_2$O$_3$ and highly doped solids, and the dipoles can be treated approximately as a gas. Following Ref. [219], the collective coupling rate ($\bar{g}_N$) is proportional to the mean rate experienced by each dipole ($\bar{g}_1$):

$$|\bar{g}_N|^2 = N|\bar{g}_1|^2 = \int \rho(\vec{r})|g_1(\vec{r})|^2 dV,$$  \hspace{1cm} (A.5)

where $\rho(\vec{r})$ and $N$ are the distribution and number of dipoles, respectively. The coupling rate for a single emitter within the cavity mode is

$$g_1(\vec{r}) = \sqrt{\frac{\mu^2 \omega_c}{2\hbar c(\vec{r})V_{\text{eff}}}} U(\vec{r}),$$  \hspace{1cm} (A.6)

where $U(\vec{r})$ describes the cavity field and is normalized so $\int |U(\vec{r})|^2 dV = V_{\text{eff}}$ [6]. If the dipole density ($\rho_o$) within the solid is uniform, Eq. (A.5) simplifies to

$$\bar{g}_N = \frac{\mu^2}{2\hbar^2 c_o V_{\text{eff}}} \int_{\delta V} |U(\vec{r})|^2 dV,$$  \hspace{1cm} (A.7)
where $n$ is the refractive index of the dipole medium. The integral is taken over the volume containing the dipoles ($\delta V$) where $\rho(\vec{r}) \neq 0$. By inspecting Eqs. (1.28) and (1.29), 
\[ \int_{\delta V} |U(\vec{r})|^2 dV = \Gamma V_{\text{eff}}, \]
and the collective coupling rate is independent of the cavity size,
\[ |\bar{g}_N|^2 = N|\bar{g}_1|^2 \approx \left| \frac{\bar{\mu}}{\bar{\mu}} \sqrt{\omega_c \rho \Gamma} \right|^2 \frac{\omega_c \rho \Gamma}{2\hbar n^2 \epsilon_0}, \] (A.8)
which is consistent with the classical approach of treating many-atom Rabi splitting as a consequence of absorption [213].
Appendix B

Calculation of the $\beta$-Factor for Er$_2$O$_3$ Microdisks

In laser physics, the $\beta$-factor is the ratio near the lasing threshold of the spontaneous emission into the lasing mode over the emission into all modes, and it can range from $10^{-6}$ to $10^{-5}$ for large gas lasers to nearly $10^0$ for few quantum-dot microcavity lasers [245]. In this work, we modify this definition so $\beta$ is the ratio of emitted power into a chosen subset of modes ($P'_{c,T}$) over the total power emitted into free-space ($P_{fs,T}$) and cavity modes ($P_{c,T}$):

$$\beta = \frac{P'_{c,T}}{P_{fs,T} + P_{c,T}}. \quad (B.1)$$

From Fermi’s Golden rule, the total emitted power ($P_T$) in a given spectral range is approximately

$$P_T = \frac{2\pi}{\hbar} \int \rho_f |\langle \psi_f | \hat{H}_{int} | \psi_i \rangle|^2 N_{Er} \hbar \omega_i d\omega_i, \quad (B.2)$$

where $\rho_f$ is the density of final states and $N_{Er}$ is the number of excited Er$^{3+}$ ions. We express the final density of states as a product of the density of electronic states ($\rho_e$) per energy per ion and the density of emission modes ($\rho_m$) per unit frequency. Using a semiclassical electric-dipole interaction $\hat{H}_{int} = -q\vec{E} \cdot \vec{r}$ and averaging over all polarizations and wave vectors for a fixed dipole orientation, $P_T$ can be expressed as

$$P_T = \frac{2\pi}{3\hbar} \int \rho_f |\langle \vec{E}(\omega_i) \rangle|^2 |\vec{\mu}(\omega_i)|^2 N_{Er} \hbar \omega_i d\omega_i, \quad (B.3)$$

where $\langle \vec{E}(\omega_i) \rangle$ is the time-averaged electric-field strength per emitted photon and the $|\psi_i\rangle \rightarrow |\psi_f\rangle$ transition’s dipole moment is $\vec{\mu}(\omega_i) = -q \langle \psi_f | \vec{x} | \psi_i \rangle$. For emission into free-space
modes, $\rho_m$ and $\langle \vec{E}(\omega_i) \rangle$ are given by

$$\rho_m = \rho_{fs}(\omega_i) = \frac{V_b \omega_i^2 n^3}{\pi^2 \epsilon_0^3}, \quad (B.4)$$

$$|\langle \vec{E}(\omega_i) \rangle| = \sqrt{\frac{\hbar \omega_i}{2n^2 \epsilon_0 V_b}}, \quad (B.5)$$

where $V_b$ is the volume of the “box” containing the free-space modes and $n \approx 2.0$ is the refractive index of Er$_2$O$_3$. Since Eq. (B.3) for the free-space modes is proportional to the measured photoluminescence spectrum in Fig. 4.5 from the unpatterned film, the spectral dependence of $\rho_e(\omega_i)|\vec{p}(\omega_i)|^2$ is known without obtaining its explicit form. Utilizing the free-space PL data also accounts for the thermal distribution of electrons within the $^4I_{13/2}$ and $^4I_{15/2}$ manifolds (width $\Delta E \approx 2k_B T$ at $T = 300 K$).

For emission into the microdisk cavity modes, $\rho_m$ is

$$\rho_m = \rho_c(\omega_i) = \sum_j 2L(\omega_i, \omega_j, \delta \omega_j), \quad (B.6)$$

where $L(\omega_i, \omega_j, \delta \omega_j)$ is a Lorentzian with a center at the $j$th cavity mode ($\omega_j$) and full width at half maximum ($\delta \omega_j$) given by the mode’s loaded linewidth ($\gamma_e + \gamma_s$); $L(\omega_i, \omega_j, \delta \omega_j)$ is normalized such that $\int_{-\infty}^{+\infty} L(\omega_i) d\omega_i = 1$. The factor of 2 accounts for the degenerate clockwise and counterclockwise traveling-wave modes. Since the cavity field is not spatially uniform, the average field strength per photon in the $j$th cavity mode experienced by the ions is

$$|\langle \vec{E}(\omega_i) \rangle_j| = \sqrt{\frac{\hbar \omega_i \vartheta_j}{2n^2 \epsilon_0 V_{c,j}}}, \quad (B.7)$$

where $V_{c,j}$ is the cavity mode volume and $\vartheta_j$ accounts for the overlap between the $j$th emission mode and the distribution of excited ions. Because the excited ion distribution depends on the intensity of the pump mode and Eq. (B.3) includes $|\langle \vec{E}(\omega_i) \rangle_j|^2$, $\vartheta_j$ is a scalar integral over the cavity volume

$$\vartheta_j = \frac{\int |\vec{E}_j|^2 |\vec{E}_p|^2 dV}{\max(|\vec{E}_j|^2) \int |\vec{E}_p|^2 dV}, \quad (B.8)$$

with the field components of the pump mode ($\vec{E}_p$) and $j$th emission mode ($\vec{E}_j$) computed with finite-element models. Figure B.1 shows this method is able to reasonably predict the
Figure B.1: Comparison of calculated and measured spontaneous emission into $\text{Er}_2\text{O}_3$ WGMs. The PL intensity is calculated using Eqs. (B.3,6–8), and the measured power into each mode is integrated and corrected for collection losses according to Eq. (4.5). Since the exact values for $\rho_e(\omega_i)|\vec{\mu}(\omega_i)|^2$ and the absolute collection efficiency are unknown, the theoretical and empirical values for $P_{c,j}$ are normalized to their maxima—i.e., this graph compares the PL intensity in each mode relative to modes near the emission peak. The PL spectrum was excited by pumping a $\text{TE}_{p=1}$ mode at 1493.3 nm [repeated from Fig. 4.9(b)]; the $\text{TE}_{p=1}$ emission modes are again highlighted in gray. Beyond our lasers’ ranges, the fractional PL in the observed peaks at $\lambda > 1620$ nm is overestimated because the WGMs are extremely over coupled and we can only approximate $\{\gamma_i, \gamma_e\}$.

fractional emission into each WGM.

Using the thin film PL data and the cavity mode parameters, the $\beta$-factor can be calculated for any individual or collection of modes. While $P_{c,T}$ in Eq. (B.1) includes a summation over all cavity modes (both observed and unobserved$^1$), the $\theta_j/V_{c,j}$ factor heavily weights the contribution of the modes with low radial order. In this analysis we include the quasi-TE modes of the first 8 radial families; the quasi-TM modes are poorly confined and have little overlap with the $\text{Er}^{3+}$ ions. For first and second radial-order emission modes at $\lambda > 1520$ nm as in the radiative efficiency measurement, we estimate $\beta_{\text{obs}} = 0.091$ which is in reasonable agreement with the experimental value of 0.038. Increasing the sum to include all observed modes across the S/C/L-bands (Fig. 4.4) gives $\beta_{12} = 0.127$. By including all the cavity modes in $P'_{c,T}$, $\beta_T = 0.227$ is the fraction of the total $^4I_{13/2} - ^4I_{15/2}$ photoluminescence that is emitted into the microdisk WGMs.

$^1$The higher-order quasi-TE modes have large bending losses ($Q < 100$) and are poorly phase matched to the taper waveguide. Since they cannot be observed in transmission or taper-collected PL, all necessary parameters are obtained through finite-element simulations.
Bibliography


109


[153] R. Soref and J. Larenzo, “All-silicon active and passive guided-wave components for \( \lambda = 1.3 \) and 1.6 \( \mu m \),” IEEE J. Quantum Electron. 22(6), 873–879 (1986).


111


