# Generating Pictures from Waves:
# Aspects of Image Formation

by

## Anthony Accardi

S.B. Mathematics
Massachusetts Institute of Technology, 1998
S.B., M.Eng. Electrical Engineering and Computer Science
Massachusetts Institute of Technology, 1998

Submitted to the Department of
Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2010

Author . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Department of Electrical Engineering and Computer Science
April 27, 2010

Certified by . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Gregory W. Wornell
Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Accepted by . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Terry P. Orlando
Professor of Electrical Engineering and Computer Science
Chairman, Committee for Graduate Students

# Generating Pictures from Waves:
# Aspects of Image Formation
by
Anthony Accardi

Submitted to the Department of Electrical Engineering and Computer Science
on April 27, 2010, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Electrical Engineering and Computer Science

## Abstract

The research communities, technologies, and tools for image formation are diverse. On the one hand, computer vision and graphics researchers analyze incoherent light using coarse geometric approximations from optics. On the other hand, array signal processing and acoustics researchers analyze coherent sound waves using stochastic estimation theory and diffraction formulas from physics. The ability to inexpensively fabricate analog circuitry and digital logic for millimeter-wave radar and ultrasound creates opportunities in comparing diverse perspectives on image formation, and presents challenges in implementing imaging systems that scale in size. We present algorithms, architectures, and abstractions for image formation that relate the different communities, technologies, and tools. We address practical technical challenges in operating millimeter-wave radar and ultrasound systems in the presence of phase noise and scattering.

We model a broad class of physical phenomena with isotropic point sources. We show that the optimal source location estimator for coherent waves reduces to processing an image produced by a conventional camera, provided the sources are well-separated relative to the system resolution, and in the limit of small wavelength and globally incoherent light. We introduce quasi light fields to generalize the incoherent image formation process to coherent waves, offering resolution tradeoffs that surpass the traditional Fourier uncertainty principle by leveraging time-frequency distributions. We show that the number of sensors in a coherent imaging array defines a stable operating point relative to the phase noise. We introduce a digital phase tightening algorithm to reduce phase noise. We present a system identification framework for multiple-input multiple-output (MIMO) ultrasound imaging that generalizes existing approaches with time-varying filters. Our theoretical results enable the application of traditional techniques in incoherent imaging to coherent imaging, and vice versa. Our practical results suggest a methodology for designing millimeter-wave imaging systems. Our conclusions reinforce architectural principles governing transmitter and receiver design, the role of analog and digital circuitry, and the tradeoff between data rate and data precision.

Thesis Supervisor: Gregory W. Wornell
Title: Professor of Electrical Engineering and Computer Science

# Acknowledgments

How do you gain a unique insight into how things work? I came back to this question many times throughout my graduate program, and was treated to a deeper answer than I anticipated. Though the story acknowledges a single person, it describes my gratitude for everyone else equally well.

I was meeting with a professor I hadn't seen for many years, and was describing the many different perspectives on a particular problem that appeared in the literature, at conferences, and in hallway conversations. "I'd like to think that one would first start with the physics, and then build algorithms on top of the physical principles." he ventured, then quickly followed with "But I suppose there are other approaches." No one who knew this professor or his background would be surprised by his remark. I was instead struck by the *way* he said it, as a devout believer proclaiming his truth yet respectfully acknowledging that there are people of other faiths in the world. Moreover, I recognized that his perspective was part of my own religion, but it had not always been so. It was 15 years ago, as a sophomore in his class, that I was first exposed to this way of thinking, always asking why, digging for the next theory that addresses the holes in the first, reaching a fundamental understanding before deciding how to engineer.

Not only is your unique perspective, born of your community, the key to your unique insight, but it is in this way that I have unconsciously grown and learned from the many people I have met on my journey. I am grateful to several generations of students, many professors and colleagues, my advisers, my friends, my family, and Samantha, for all they have given, for what I have learned, and for the wisdom they have imparted. Thank you.

# Contents

# List of Figures

# List of Symbols and Notation

| | |
|---|---|
| $t$ | time |
| $\mathbf{r}$ | position |
| $\mathbf{s}$ | unit direction |
| $\mathbf{r}_\perp, \mathbf{s}_\perp$ | projected versions of $\mathbf{r}, \mathbf{s}$ |
| $x_i(t)$ | signal emitted from $i^{\text{th}}$ source |
| $y_i(t)$ | signal measured at $i^{\text{th}}$ sensor |
| $z_i(t)$ | signal transmitted by $i^{\text{th}}$ transducer |
| $v_i(t)$ | additive white Gaussian noise at $i^{\text{th}}$ sensor |
| $\phi_i(t)$ | ergodic phase noise at $i^{\text{th}}$ sensor |
| $\nu$ | temporal frequency |
| $\omega_{\text{c}}$ | carrier frequency |
| $\lambda$ | wavelength |
| $k$ | wave number |
| $U(\mathbf{r})$ | scalar field |
| $a(\mathbf{s})$ | plane wave component |
| $\mathbf{a}(\mathbf{r})$ | array steering vector |
| $B_\theta$ | beam pattern at angle $\theta$ |
| $\tau$ | tilt |
| $S^{\text{SL}}$ | sidelobe suppression |
| $\theta_{\text{BW}}$ | angular half-power beam width |
| $\mathbf{F}(\mathbf{r})$ | energy flux density vector |
| $I(\mathbf{s})$ | radiant intensity |
| $L(\mathbf{r}, \mathbf{s})$ | light field, radiance |
| $\mathfrak{W}(\mathbf{r}, \mathbf{s}_\perp)$ | Wigner distribution |
| $K(\mathbf{r}_1, \mathbf{r}_2)$ | kernel indicating choice of quasi light field |
| $W(\mathbf{r}_1, \mathbf{r}_2, \nu)$ | cross-spectral density |
| $\lambda_k$ | Markov chain transition probability |
| $\bar{\lambda}_k$ | $1 - \lambda_k$ |
| $\pi_k$ | steady-state probability vector component |
| $S_P$ | generalized binomial sum |
| $\mathbb{Z}$ | set of integers |
| $\delta(\cdot)$ | Dirac delta function |
| $\text{sgn}(\cdot)$ | signum function |
| $u(\cdot)$ | unit step function |
| $\text{rect}(\cdot)$ | unit rectangle function |
| $(x)_n$ | falling factorial function of length $n$ |
| $\mathbf{A}^{\text{t}}$ | matrix transpose of $\mathbf{A}$ |
| $\mathbf{A}^*$ | conjugate transpose of $\mathbf{A}$ |
| $a * b$ | convolution of $a$ and $b$ |
| $a \circledast b$ | circular convolution of $a$ and $b$ |
| $|\cdot|$ | $\ell^2$-norm |
| $\langle \cdot \rangle$ | average over time or ensemble |

# Chapter 1

# Two imaging communities meet

Mankind has manipulated electromagnetic radiation to form images for millennia, yet the definition of an image today is broad and context-dependent. Images were originally physically-based visual representations that reflect how the human visual system perceives the world. The camera obscura, a dark room with a small hole in a wall to admit outside light, was the first man-made image formation device [1]. By looking at the illuminated wall opposite the hole, one could view an inverted image of the outside scene. Much later, Leonardo da Vinci drew an analogy between the operation of the camera obscura and the eye, which we now know functions in a similar way by using a lens to focus the scene onto the retina. Physical images, such as those produced by a conventional camera or the eye, attempt to accurately reflect how we physically see the world.

Despite the physical roots of image formation, today an image can refer to any type of two-dimensional representation, such as an abstract painting or the visualization of quantitative data. Sometimes different image types are overlaid, for example when a color flow map depicting blood velocity is superimposed on a physical ultrasound image of the human body [2]. These images are designed to satisfy application requirements, such as those for diagnosing specific medical conditions. An ultrasound machine can process pressure field measurements to estimate specific parameters of interest, such as blood velocity, rather than to create what the human eye would see if it could perceive pressure waves as it does visible light. Unlike physical images, such task-oriented images incorporate estimated parameters in a manner that aids the completion of a specific task.

We distinguish between two broad communities that research image formation. The coherent community studies coherent waves, such as the electromagnetic waves in radar and the pressure waves in ultrasound, that can be directly produced, manipulated, and measured by today's technology using electronic circuits. The incoherent community studies incoherent waves, such as the light from the sun, that cannot yet be directly produced, manipulated, or measured. The distinction between these communities disappears as technology evolves; we can produce coherent light with lasers [3], build nanostructures to manipulate light with precision [4], and infer wave phase from holograms [5]. Nevertheless, the different roots of the coherent and incoherent communities have led to different terminology, tools, and insights, even as the

communities now overlap to some degree.

Two technological innovations are responsible for the overlap in the applications that the coherent and incoherent communities study (Figure 1-1). First, we can now fabricate integrated circuits that operate at high frequencies corresponding to coherent electromagnetic radiation of millimeter wavelength [6, 7, 8]. Previously, radar was primarily used to estimate the parameters of a scene, such as the position of an aircraft [9]. Now, due to the small wavelength, millimeter-wave radar is capable of forming pictures at a high enough resolution to visualize a scene, representing what the human eye would perceive if it could process millimeter-wave radiation as it does light in the visible spectrum. Applications range from automotive collision avoidance [10] to concealed weapons detection [11]. Second, we now understand how to infer directional information from incoherent light without having to directly measure the phase of the underlying electric and magnetic fields [12, 13]. Previously, a camera was primarily used to produce a physical image of a scene. Now, a plenoptic camera is capable of estimating depth information and other scene parameters using an array of lenses, and taking pictures that can be digitally refocused [14]. Evidently, the coherent and incoherent communities are addressing similar challenges in parameter estimation and visualization, but have approached the application space in different ways.
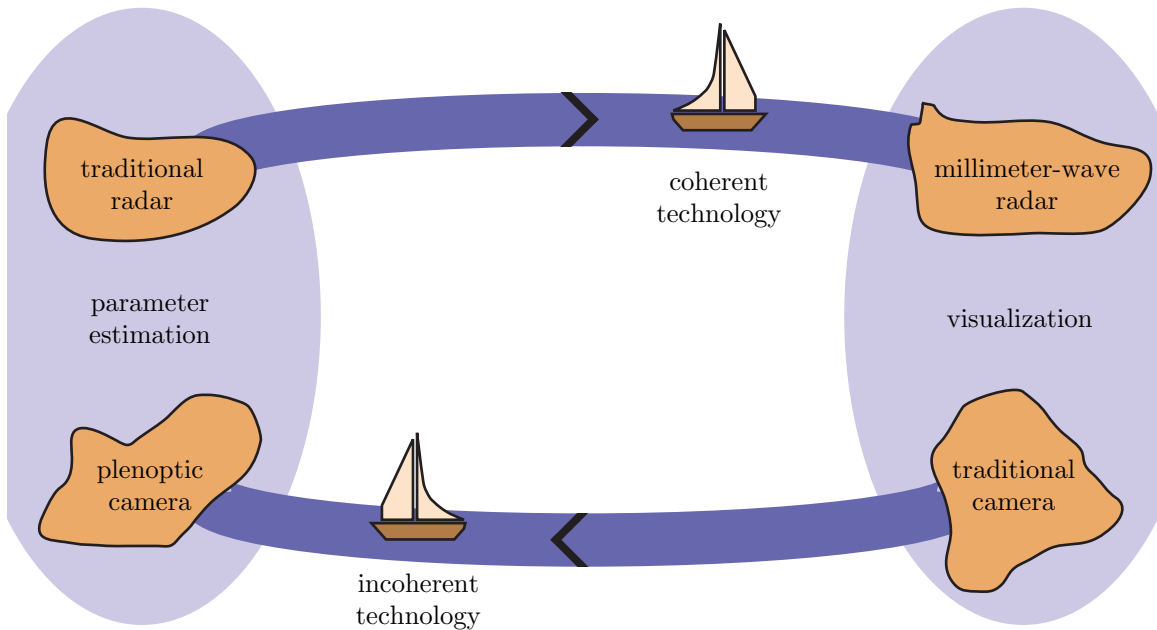


Figure 1-1: The evolution of coherent and incoherent wave processing technologies has spanned the same application space but in opposing directions, so that there is limited overlap between the corresponding tools and research communities. Coherent technology has evolved from traditional radar for estimating parameters such as aircraft position to millimeter-wave radar for producing pictures. Incoherent technology has evolved from the traditional camera for imaging a scene to the plenoptic camera for estimating parameters such as object depth.

Our goal is to unify the coherent and incoherent perspectives of image formation, so we can apply the insights and techniques developed in one community to the other. Specifically, our primary theoretical objective is to demonstrate how and under what conditions the problem of optimally estimating point source locations, using coherent waves, is mathematically and physically identical to finding the brightest spots on a conventional photograph, using incoherent light. First, we show how classical coherent parameter estimation problems can be visualized in Chapter 2. Second, we show how to process incoherent light to estimate parameters of objects in a scene in Chapter 3. Third, we introduce a construct called *quasi light fields* to bridge the visualization of coherent and incoherent waves in Chapter 4. We achieve our objective by assembling these pieces.

Once we have established the precise connection between coherent and incoherent imaging, we address the implementation challenges and practical design considerations that arise in the latest coherent imaging systems, focusing on millimeter-wave radar and ultrasound. Managing phase noise in millimeter-wave radar circuitry is challenging and essential for accurate imaging. We explore to what extent large radar arrays are robust to phase noise in Chapter 5, and propose a digital phase tightening algorithm to reduce unacceptably high levels of phase noise in Chapter 6. Fast ultrasound systems that produce images at a high frame rate are desirable for many medical applications, but are challenging to implement due to the slow speed of sound. We outline new opportunities for fast ultrasound based on ideas from multiple-input multiple-output (MIMO) radar [15] and system identification in Chapter 7.

Both radar and ultrasound millimeter-wave imaging systems are uniquely positioned to benefit from a deeper understanding of the relationship between coherent and incoherent image formation, because they bridge their roots in coherent processing with an application space that has traditionally been the domain of incoherent processing. We can therefore translate well-studied problems and solutions from incoherent imaging applications to a coherent imaging context, where we typically have more options and tradeoffs available due to the coherent structure.

# Chapter 2

# Processing coherent waves to form images

From an engineer's perspective, images are used as aids to complete a task. To make the role of image formation concrete, we explore how images naturally arise as tools to solve optimization problems using coherent wave measurements as raw data. Specifically, we introduce source localization as a simple model that illustrates how low-level image processing can solve higher-level parameter estimation problems, provided that the imaging system is of sufficiently high quality. In source localization, $M$ sensors at known positions measure waveforms emitted by $P$ sources at unknown positions, and we wish to estimate the source locations given the sensor measurements. The number of sensors indicates the quality of the imaging system, and the number of sources indicates the complexity of the parameter estimation task to complete. We can model a broad class of physical phenomena with such point sources. Provided that the imaging system is of sufficiently high quality relative to the desired resolution, the optimal source localization algorithm is well-approximated by first forming an image from the sensor measurements with the conventional beamformer, and then choosing the $P$ brightest pixels as estimates of the source locations.

We model signal propagation in a way that encompasses both the electromagnetic waves used in radar [9] as well as the pressure waves used in sonar [16] and ultrasound [17]. Depending on their type, the waves are produced and measured by either antennas or transducers. The sensors measure the amplitude and phase of the waves at different points in space, and electronic circuits process the measured waveforms. Early coherent wave processing applications were limited to locating airplanes and sea vessels. However, circuit technology has been continuously advancing to process higher frequency waveforms, so that today we can manufacture millimeter-wave ultrasound systems [17] and complete millimeter-wave radar systems using inexpensive silicon fabrication technology [7, 8]. The importance of millimeter wavelength is that it is small enough to achieve a useful resolution for many imaging applications. The images produced as aids to solve the source localization problem in conventional radar are themselves of primary interest in millimeter-wave imaging.

We first relate source localization to scene visualization by formulating the source localization problem and employing various beamforming algorithms as potential so-

lutions (Section 2.1). If there are enough sensors relative to the desired resolution between sources, then processing the image produced by the conventional beamformer is approximately optimal. Many alternative coherent wave processing techniques address a different scenario, where one or more sources are close together. Some alternative techniques, such as Capon's method, process different types of images that achieve better resolution than the conventional beamformer. However, unlike these alternative techniques, the conventional beamformer forms an image in a manner that is physically identical to the operation of a conventional camera and the human eye. Because of its important physical connection to incoherent image formation, we focus on the conventional beamformer to define performance metrics and address design tradeoffs in the presence of phase noise and timing inaccuracies (Section 2.2).

We will show how to compute physically-based images with quasi light fields in Chapter 4, where the choice of quasi light field corresponds to employing a particular estimate of the correlation structure of the radiation when implementing the conventional beamformer. A similar choice must be made to implement Capon's method, where the resulting tradeoffs have been extensively studied. Therefore, our primary motivation for introducing alternative techniques such as Capon's method is not to catalogue the already well-established strategies for resolving closely-spaced sources, but rather to emphasize the potential of realizing new tradeoffs in producing physically-based images using quasi light fields.

## 2.1  Visualizing source locations with arrays

We relate image formation to parameter estimation through the source localization problem. We first formulate source localization using an array of sensors. We then derive the deterministic maximum likelihood location estimator (Section 2.1.1). We show that when the sources are far apart, the maximum likelihood estimate is equivalent to forming an image using the conventional beamformer and finding the brightest pixels (Section 2.1.2). It is our thesis that the image formed by the conventional beamformer has a physical basis related to conventional photography. The maximum likelihood location estimator corresponds to a quasi light field that produces one type of physically-based image, and other choices of quasi light field lead to other types of physically-based images. The choice of quasi light field corresponds to a choice of how to estimate the correlation structure of the radiation. Alternative beamformers, such as Capon's method, attempt to overcome the conventional beamformer's limited resolution, but do not produce images that have the same physical interpretation. We describe how the same choice of correlation structure estimate impacts the robustness of Capon's method, motivating the importance of quasi light field choice (Section 2.1.3).

We formulate the source localization problem by considering $P$ isotropic point sources and $M$ isotropic point sensors. Point sources and sensors are simple to analyze, yet serve as building blocks for more complicated scenarios. For example, we can formulate the higher-level edge detection problem by determining the boundary of a large number of point sources. More generally, the number of sources $P$ indicates

the complexity of the parameter estimation task, and the number of sensors $M$ indicates the amount of data available, which affects the quality of the estimates. We can model complicated scenarios that exhibit anisotropic behavior and directional radiation sources in both the near and far zone by assembling an appropriate collection of isotropic point sources. We limit our discussion here to perfectly coherent radiation, although we will show how to generalize our results to any state of partial coherence in Chapter 4.

The sources emit waveforms $x_1(t), \ldots, x_P(t)$ in free space, and the sensors measure the responses $y_1(t), \ldots, y_M(t)$ (Figure 2-1). We may think of the sources as airplanes and the sensors as a set of isotropic antennas organized in an array. For simplicity, we assume that the sensors are equally spaced every distance $d$ along a linear array, at positions $\mathbf{r}_1^M, \ldots, \mathbf{r}_M^M$. We assume that the sources emit their waveforms isotropically from positions $\mathbf{r}_1^P, \ldots, \mathbf{r}_P^P$, but it will be convenient to also consider the far-zone case where the $j^{th}$ source lies far away with respect to the array size, along angle $\theta_j^P$ with respect to the array. In the far-zone case, the spherical wavefronts of the source waveforms are approximately planar at the array.
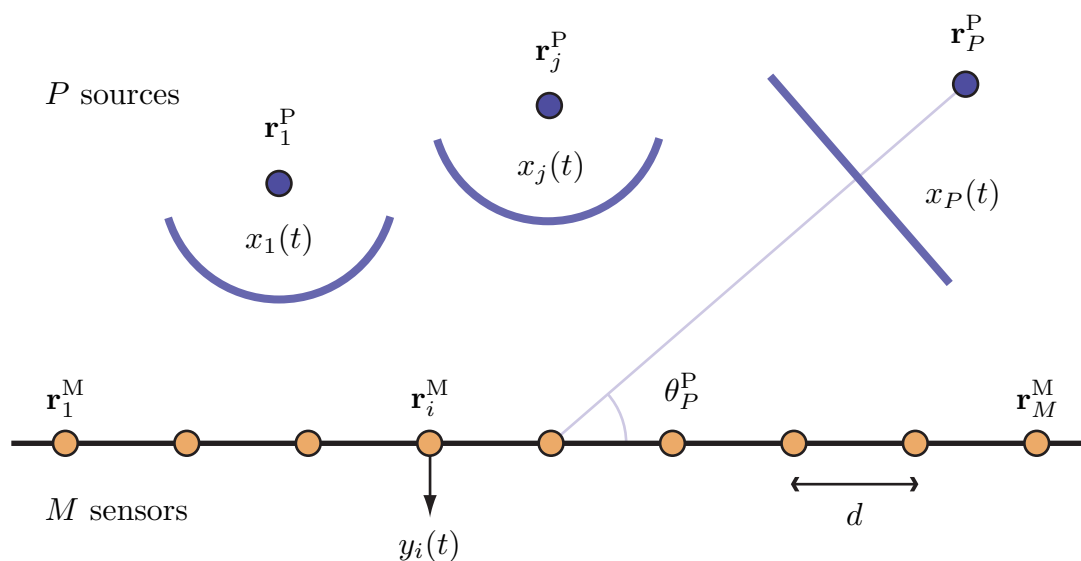


Figure 2-1: In the source localization problem, we estimate the position $\mathbf{r}_j^P$ or direction of arrival $\theta_j^P$ of each of $P$ sources emitting waveforms $x_j(t)$, using a linear array of $M$ sensors equally spaced every distance $d$ at positions $\mathbf{r}_i^M$ that measure and record waveforms $y_i(t)$. In this example, the $P^{th}$ source is in the far zone.

In general, the emitted waveforms may have arbitrary spectral content. Without loss of generality, we make a narrowband assumption to simplify the physics and model wave propagation as a linear operation on the waveform envelopes. Specifically, we assume that the emitted waveforms are narrowband about some carrier frequency $\omega_c$ so that $x_j(t)$ represents the envelope of the complex analytic signal that is physically transmitted (Appendix A.1). That is, the $j^{th}$ source gives rise to the

scalar field $U$, which at position $\mathbf{r}$ and time $t$ is

$$U(\mathbf{r}, t) = \mathrm{Re}\left\{ x_j\left( t - |\mathbf{r} - \mathbf{r}_j^{\mathrm{P}}|\frac{k}{\omega_{\mathrm{c}}} \right) \frac{\exp\left[ i(k|\mathbf{r} - \mathbf{r}_j^{\mathrm{P}}| - \omega_{\mathrm{c}}t) \right]}{|\mathbf{r} - \mathbf{r}_j^{\mathrm{P}}|} \right\}. \tag{2.1}$$

The scalar field $U$ solves the scalar wave equation, either as a component of the electric field according to Maxwell's equations [18], or as a pressure field according to the laws of acoustics [19]. In (2.1), $k = 2\pi/\lambda$ is the wave number and $\lambda$ is the wavelength, so that the wave travels at speed $c = \omega_{\mathrm{c}}/k$. The narrowband assumption implies that the envelope $x_j(t)$ remains constant over the time it takes the wave to propagate, so that we may approximate

$$x_j\left( t - |\mathbf{r} - \mathbf{r}_j^{\mathrm{P}}|\frac{k}{\omega_{\mathrm{c}}} \right) \approx x_j(t) \tag{2.2}$$

in (2.1). Consequently, the waveform emitted by a source is simply attenuated and phase-shifted when measured at a remote sensor, and we may ignore the time dependence $\exp(-i\omega_{\mathrm{c}}t)$. The narrowband assumption therefore allows us to work exclusively with the complex envelopes $x_j(t)$ and model propagation by multiplication with a complex constant.

We apply the above propagation model to complete our formulation. For a single source, each sensor records the propagated waveforms corrupted by additive white complex Gaussian noise $v_i(t)$ of zero mean and variance $\sigma_v^2$. The noise at different sensors is uncorrelated. For a single source at $\mathbf{r}_j^{\mathrm{P}}$, we express the output at all sensors in vector form by

$$\mathbf{y}(t) = \mathbf{a}(\mathbf{r}_j^{\mathrm{P}})x_j(t) + \mathbf{v}(t), \tag{2.3}$$

where $\mathbf{y}(t) = [y_1(t) \cdots y_M(t)]^{\mathrm{t}}$, $\mathbf{v}(t) = [v_1(t) \cdots v_M(t)]^{\mathrm{t}}$, and $\mathbf{a}(\mathbf{r}_j^{\mathrm{P}})$ is the steering vector that reflects the propagation of $x_j(t)$ to each sensor in the array. Specifically, the $n^{\mathrm{th}}$ component of $\mathbf{a}(\mathbf{r}_j^{\mathrm{P}})$ is

$$a_n(\mathbf{r}_j^{\mathrm{P}}) = \frac{\exp\left( ik|\mathbf{r}_n^{\mathrm{M}} - \mathbf{r}_j^{\mathrm{P}}| \right)}{|\mathbf{r}_n^{\mathrm{M}} - \mathbf{r}_j^{\mathrm{P}}|}. \tag{2.4}$$

We will frequently ignore the attenuation when the signal power drop across the sensor array is negligible, and instead use:

$$a_n(\mathbf{r}_j^{\mathrm{P}}) = \exp\left( ik|\mathbf{r}_n^{\mathrm{M}} - \mathbf{r}_j^{\mathrm{P}}| \right). \tag{2.5}$$

We may at times further make a far-zone approximation and ignore a constant phase offset when the source is far away, so that

$$a_n(\mathbf{r}_j^{\mathrm{P}}) = a_n(\theta_j^{\mathrm{P}}) = \exp(-iknd\cos\theta_j^{\mathrm{P}}). \tag{2.6}$$

We add the contributions from all $P$ sources to obtain

$$\mathbf{y}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{v}(t), \tag{2.7}$$

where $\mathbf{A} = [\mathbf{a}(\mathbf{r}_1^{\mathrm{P}}) \cdots \mathbf{a}(\mathbf{r}_P^{\mathrm{P}})]$ and $\mathbf{x}(t) = [x_1(t) \cdots x_P(t)]^{\mathrm{t}}$. The source localization problem is to determine the source positions $\mathbf{r}_j^{\mathrm{P}}$ given the sensor measurements in (2.7).

Our formulation assumes perfectly coherent radiation, as the scalar field $U(\mathbf{r}, t)$ is a deterministic function in both space and time. The scalar field is more accurately modeled by a random process, especially at optical frequencies, where random fluctuations due to the spontaneous emission of radiation from atoms or the mechanical vibrations of the mirrors at the end of a laser cavity result in an uncertain field [91]. In order to relate coherent and incoherent image formation within this stochastic framework, we use the tools of optical coherence theory to make the intuitive notion of coherence mathematically precise.

When one says that a radiation source is temporally or spatially coherent, the implication is that the field at the source possesses a statistical similarity in either time or space that will result in interference if the field is allowed to recombine after propagation. There are many details left unspecified, such as whether there is a statistical similarity at other points in space and time away from the source, and the precise measure of the degree of statistical similarity. The mutual coherence function serves as a single precise specification of the coherence structure of the scalar field, and is defined as the ensemble average of the product of the field at different points in space and time (Appendix A.4):

$$\Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau) = \langle U(\mathbf{r}_1, t + \tau) U^*(\mathbf{r}_2, t) \rangle. \tag{2.8}$$

Classical interference effects are expressed in terms of time-averaged intensity measurements, which can be expressed in terms of $\Gamma$ for ergodic fields [91]. Although the mutual coherence function ultimately allows us to treat coherent and incoherent image formation with a common framework, the key concepts are more transparent without the additional stochastic baggage that comes with $\Gamma$. Therefore, when we consider coherent image formation, we assume that the field is deterministic everywhere. And when we consider incoherent image formation, we assume that the sources are spatially incoherent so that their contributions at a remote sensor add in intensity. Throughout our treatment, the mutual coherence function serves as the underlying formalism that allows us to rigorously relate these two types of image formation (Appendix B.2).

## 2.1.1   Deterministic maximum likelihood estimation

We derive the well-known deterministic maximum likelihood estimator for source location, so named because we assume that the source signals $\mathbf{x}(t)$ are deterministic yet unknown [20]. We assume we know the sensor positions $\mathbf{r}_i^{\mathrm{M}}$, but not the noise variance $\sigma_v^2$. We seek the maximum likelihood estimator of the unknown source

positions $R^{\mathrm{P}} = \{\mathbf{r}_1^{\mathrm{P}}, \ldots, \mathbf{r}_P^{\mathrm{P}}\}$ given $N$ time samples of all sensor measurements $\mathbf{y}(t)$. The likelihood function is

$$L(R^{\mathrm{P}}, \mathbf{x}(t), \sigma_v^2) = \prod_{t=1}^{N} \left(\pi\sigma_v^2\right)^{-M} \exp\left(-\frac{|\mathbf{y}(t) - \mathbf{A}\mathbf{x}(t)|^2}{\sigma_v^2}\right), \tag{2.9}$$

so that the log-likelihood function is

$$l(R^{\mathrm{P}}, \mathbf{x}(t), \sigma_v^2) = -NM\log(\pi\sigma_v^2) - \frac{1}{\sigma_v^2}\sum_{t=1}^{N}|\mathbf{y}(t) - \mathbf{A}\mathbf{x}(t)|^2. \tag{2.10}$$

The desired position estimates $\hat{R}^{\mathrm{P}}$ maximize (2.10).

First, for fixed $\sigma_v^2$ and $R^{\mathrm{P}}$, the source signals $\mathbf{x}(t)$ that maximize (2.10) are those that minimize $|\mathbf{y}(t) - \mathbf{A}\mathbf{x}(t)|^2$ at each $t$. This minimization is a classic linear least-squares problem from linear algebra, whose solution follows from the normal equations:

$$\hat{\mathbf{x}}(t) = (\mathbf{A}^*\mathbf{A})^{-1}\mathbf{A}^*\mathbf{y}(t). \tag{2.11}$$

Next, substituting the optimal source signals $\mathbf{x}(t) = \hat{\mathbf{x}}(t)$ into (2.10) yet still keeping $\sigma_v^2$ fixed, we note that the source positions $R^{\mathrm{P}}$ that maximize (2.10) are those that minimize

$$\sum_{t=1}^{N}\left|\mathbf{y}(t) - \mathbf{A}(\mathbf{A}^*\mathbf{A})^{-1}\mathbf{A}^*\mathbf{y}(t)\right|^2. \tag{2.12}$$

We recognize

$$\Pi_{\mathbf{A}} = \mathbf{A}(\mathbf{A}^*\mathbf{A})^{-1}\mathbf{A}^* \tag{2.13}$$

in (2.12) as a projection onto the columns of $\mathbf{A}$. Therefore, minimizing (2.12) is equivalent to maximizing the energy in the projected subspace, so that

$$\hat{R}^{\mathrm{P}} = \arg\max_{R^{\mathrm{P}}} \sum_{t=1}^{N}|\Pi_{\mathbf{A}}\mathbf{y}(t)|^2. \tag{2.14}$$

Intuitively, the maximum likelihood estimator in (2.14) identifies the steering vectors that define a signal subspace that best contains the measurements $\mathbf{y}(t)$. While the maximum likelihood estimator provides optimal source location estimates, it is generally computationally complex to calculate, requiring a multidimensional search.

## 2.1.2  Conventional beamforming

We show that the maximum likelihood estimator reduces to an efficient search of an image produced by the conventional Bartlett beamformer when the sources are far apart. We first consider the scenario for a single source, to understand how the conventional beamformer focuses at a point and maximizes the expected energy emitted from a source at the point of focus. Additionally, we show that the conventional beamformer corresponds to a particular way of estimating the correlation structure of the radiation, which we will later connect to a specific choice of quasi light field in

Chapter 4. In the far zone, the image produced by the conventional beamformer is a visual representation of the spatial spectrum of the scene, so that image resolution is limited by the Fourier uncertainty principle. When the sources are far apart compared with beamformer resolution, the conventional beamformer solves the source localization problem for multiple sources.

To interpret source localization as an image search, we first rewrite (2.14) to express the maximum likelihood estimator as a bilinear function of the measured signals:

$$
\begin{aligned}
\hat{R}^{\mathrm{P}} &= \underset{R^{\mathrm{P}}}{\arg\max} \sum_{t=1}^{N} [\Pi_{\mathbf{A}} \mathbf{y}(t)]^* \, \Pi_{\mathbf{A}} \mathbf{y}(t) \\
&= \underset{R^{\mathrm{P}}}{\arg\max} \sum_{t=1}^{N} \mathbf{y}^*(t) \Pi_{\mathbf{A}} \mathbf{y}(t) \\
&= \underset{R^{\mathrm{P}}}{\arg\max} \sum_{t=1}^{N} \mathbf{y}^*(t) \mathbf{A} \left(\mathbf{A}^* \mathbf{A}\right)^{-1} \mathbf{A}^* \mathbf{y}(t).
\end{aligned}
\tag{2.15}
$$

When there is only a single source, then $\mathbf{A} = \mathbf{a}(\mathbf{r})$ is a single steering vector, so that

$$
\hat{\mathbf{r}}_1^{\mathrm{P}} = \underset{\mathbf{r}}{\arg\max} \sum_{t=1}^{N} \frac{|\mathbf{a}^*(\mathbf{r})\mathbf{y}(t)|^2}{\mathbf{a}^*(\mathbf{r})\mathbf{a}(\mathbf{r})}.
\tag{2.16}
$$

Furthermore, if we ignore signal attenuation across the sensor array, $\mathbf{a}^*(\mathbf{r})\mathbf{a}(\mathbf{r}) = M$ is constant per (2.5), so that

$$
\hat{\mathbf{r}}_1^{\mathrm{P}} = \underset{\mathbf{r}}{\arg\max} \frac{1}{N} \sum_{t=1}^{N} \left| \frac{\mathbf{a}^*(\mathbf{r})\mathbf{y}(t)}{M} \right|^2.
\tag{2.17}
$$

Equation (2.17) links parameter estimation to image formation. The maximum likelihood estimate of the position of a single source is found by computing an image with pixel values of

$$
P^{\mathrm{CBF}}(\mathbf{r}) = \frac{1}{N} \sum_{t=1}^{N} \left| \frac{\mathbf{a}^*(\mathbf{r})\mathbf{y}(t)}{M} \right|^2
\tag{2.18}
$$

at each position $\mathbf{r}$, and then choosing the brightest pixel. The image is the objective function, parameterized by the potential locations of the source.

The quantity $(1/M)\mathbf{a}^*(\mathbf{r})\mathbf{y}(t)$ is the conventional Bartlett beamformer focused at point $\mathbf{r}$, which we now generalize to show that (2.18) implicitly assumes a particular estimate of the correlation structure of $\mathbf{y}(t)$. The conventional beamformer maximizes the power emitted from a source at $\mathbf{r}$ [21], by choosing weights $\mathbf{w}$ that optimize:

$$
\max_{\mathbf{w}} E\left[|\mathbf{w}^*\mathbf{y}(t)|^2\right] \text{ such that } |\mathbf{w}|^2 = 1.
\tag{2.19}
$$

We compute

$$
E\left[|\mathbf{w}^*\mathbf{y}(t)|^2\right] = E\left[|x_1(t)|^2\right] |\mathbf{w}^*\mathbf{a}(\mathbf{r}_1^{\mathrm{P}})|^2 + \sigma_v^2 |\mathbf{w}|^2,
\tag{2.20}
$$

so that by the Cauchy-Schwarz inequality, the optimal weights are

$$\mathbf{w} = \frac{\mathbf{a}(\mathbf{r}_1^{\mathrm{P}})}{\sqrt{\mathbf{a}^*(\mathbf{r}_1^{\mathrm{P}})\mathbf{a}(\mathbf{r}_1^{\mathrm{P}})}}, \tag{2.21}$$

which substituting into $E\left[|\mathbf{w}^*\mathbf{y}(t)|^2\right]$ and scaling by $1/M$ yields the spectrum

$$P^{\mathrm{GBF}}(\mathbf{r}) = E\left[\left|\frac{\mathbf{a}^*(\mathbf{r})\mathbf{y}(t)}{M}\right|^2\right]. \tag{2.22}$$

If we estimate the correlation matrix $E[\mathbf{y}(t)\mathbf{y}^*(t)]$ in (2.22) by averaging $N$ samples of $\mathbf{y}(t)\mathbf{y}^*(t)$, we obtain the image formed by the conventional beamformer in (2.18). We will show how to compute other physically-based coherent images with other estimates for the correlation matrix in Chapter 4.

We now illustrate how the conventional beamformer $(1/M)\mathbf{a}^*(\mathbf{r})\mathbf{y}(t)$ focuses on the point $\mathbf{r}$. The measured signals for a single source are

$$\mathbf{y}(t) = \mathbf{a}(\mathbf{r}_1^{\mathrm{P}})x_1(t) + \mathbf{v}(t). \tag{2.23}$$

When the beamformer is focused on the source, $\mathbf{r} = \mathbf{r}_1^{\mathrm{P}}$, so $\mathbf{a}^*(\mathbf{r})\mathbf{a}(\mathbf{r}_1^{\mathrm{P}}) = M$ is maximized. For values of $\mathbf{r}$ further away from $\mathbf{r}_1^{\mathrm{P}}$, the beamformer is focused on empty space, so $\mathbf{a}^*(\mathbf{r})\mathbf{a}(\mathbf{r}_1^{\mathrm{P}})$ is diminished. The beamformer performs generalized spatial Fourier analysis on the measured signals, as $\mathbf{a}^*(\mathbf{r})\mathbf{y}(t)$ reduces to the space-domain version of the discrete-time short-time Fourier transform in the far-zone per (2.6). The resulting image in (2.18) is consequently the time average of spatial periodograms of the measured signals [22]. The spatial resolution of the periodograms is directly related to the resolution of the image produced by the beamformer. The performance of the beamformer is therefore limited by the Fourier uncertainty principle. Although the resolution may be improved by increasing the size of the sensor array to widen the window, no amount of time averaging will improve the resolution.

We now consider the maximum likelihood estimator for multiple sources. The matrix product $\mathbf{A}^*\mathbf{A}$ in (2.15) has a constant diagonal of maximal value $M$, while the off-diagonal entries have the form $\mathbf{a}^*(\mathbf{r}_i)\mathbf{a}(\mathbf{r}_j)$. Each off-diagonal entry indicates the response of a conventional beamformer focused at one position $\mathbf{r}_i$ when the source is at another position $\mathbf{r}_j$. Therefore, when the maximum likelihood objective function is evaluated near the true source locations, every pair of sources that are far apart from each other relative to the beamformer's resolution contributes two zero entries in $\mathbf{A}^*\mathbf{A}$. The sources can be reindexed and clustered into groups that are far apart from each other, so that $\mathbf{A}^*\mathbf{A}$ is block-diagonal. If all the sources are far apart from each other, then

$$\mathbf{A}^*\mathbf{A} \approx M\mathbf{I} \tag{2.24}$$

when evaluated near the true source locations. We substitute (2.24) into (2.15) to obtain an approximation to the maximum likelihood estimator for well-separated

sources:

$$\hat{R}^{\mathrm{P}} = \arg\max_{R^{\mathrm{P}}} \frac{1}{P}\sum_{j=1}^{P}\frac{1}{N}\sum_{t=1}^{N}\left|\frac{\mathbf{a}^{*}(\mathbf{r}_{j}^{\mathrm{P}})\mathbf{y}(t)}{M}\right|^{2}. \tag{2.25}$$

Equation (2.25) estimates the source locations by the $P$ brightest pixels of the image formed by the conventional beamformer in (2.18). We conclude that the conventional beamformer produces an image that is suitable to locate well-separated point sources.

### 2.1.3   Alternative beamformers and images

The conventional beamformer cannot resolve closely-spaced sources as well as the maximum likelihood estimator. However, the maximum likelihood estimator is computationally complex to implement directly. Researchers have therefore developed alternative beamformers that, like the conventional beamformer, efficiently process an image, but unlike the conventional beamformer, achieve resolution that is not limited by the Fourier uncertainty principle. We describe two alternative beamformers: Capon's method that actively blocks interference from those sources not at the point of focus, and the multiple signal classification (MUSIC) algorithm that leverages assumptions about the correlation structure of the sources. Neither Capon's method nor the MUSIC algorithm produces a physically-based image described by quasi light fields, which offer different tradeoffs in resolving the position and direction of the radiation that are not limited by the Fourier uncertainty principle. However, implementing both Capon's method and the MUSIC algorithm requires the selection of an estimator for the correlation structure of $\mathbf{y}(t)$, which corresponds to selecting a particular quasi light field when implementing the conventional beamformer in (2.22). Evidently, at least one aspect of the choice of quasi light field has been previously studied in alternative beamforming contexts, where the choice has a significant practical impact.

   In contrast with the conventional beamformer that maximizes filter output power, Capon's method [23] minimizes output power, subject to the constraint that any signal from the point of focus $\mathbf{r}$ is undistorted:

$$\min_{\mathbf{w}} E\left[|\mathbf{w}^{*}\mathbf{y}(t)|^{2}\right] \text{ such that } \mathbf{w}^{*}\mathbf{a}(\mathbf{r}) = 1. \tag{2.26}$$

With this constraint, minimizing output power reduces noise and interference. The optimal weights are found using Lagrange multipliers:

$$\mathbf{w} = \frac{\left[E\left[\mathbf{y}(t)\mathbf{y}^{*}(t)\right]\right]^{-1}\mathbf{a}(\mathbf{r})}{\mathbf{a}^{*}(\mathbf{r})\left[E\left[\mathbf{y}(t)\mathbf{y}^{*}(t)\right]\right]^{-1}\mathbf{a}(\mathbf{r})}. \tag{2.27}$$

We use the optimal weights to compute the filter output power $E\left[|\mathbf{w}^{*}\mathbf{y}(t)|^{2}\right]$ to obtain an image of the scene:

$$P^{\mathrm{CAP}}(\mathbf{r}) = \frac{1}{\mathbf{a}^{*}(\mathbf{r})\left[E\left[\mathbf{y}(t)\mathbf{y}^{*}(t)\right]\right]^{-1}\mathbf{a}(\mathbf{r})}. \tag{2.28}$$

Unlike the image produced by the conventional beamformer in (2.18), the image produced by Capon's method is not a true frequency spectrum. Capon's method is able to achieve better resolution than the conventional beamformer by adaptively blocking interfering signals arriving from locations other than the point of focus [21, 20]. However, Capon's method is sensitive to errors in estimating the correlation matrix $E\left[\mathbf{y}(t)\mathbf{y}^*(t)\right]$ in (2.28), and performs poorly when the estimate is inaccurate. Capon's method can be made more robust by applying tapers to the correlation matrix estimate, resulting in design features such as diagonal loading [24]. In Chapter 4, we show that applying these same tapers to the correlation matrix estimate in (2.22) corresponds to choosing alternative quasi light fields.

In order to resolve closely-spaced sources, Capon's method blocks interfering sources, while the MUSIC algorithm instead leverages assumptions on the correlation structure of the sources. Intuitively, the optimal maximum likelihood source location estimator identifies the steering vectors that span a subspace that best contains the measured signals. In fact, the noise subspace is orthogonal to the steering vectors at the true source locations. This observation leads to a number of signal subspace techniques for source localization, including the MUSIC algorithm [21, 20]. The MUSIC algorithm estimates the noise subspace, and then identifies source locations that yield steering vectors orthogonal to the noise subspace. Provided that the source signals are not coherent and that there are more sensors than sources, the MUSIC source location estimates converge to the true values. The MUSIC algorithm thereby performs well when the source assumptions are valid, and poorly otherwise. The MUSIC algorithm chooses the brightest pixels in the image

$$P^{\mathrm{MUS}}(\mathbf{r}) = \frac{\mathbf{a}^*(\mathbf{r})\mathbf{a}(\mathbf{r})}{\mathbf{a}^*(\mathbf{r})\hat{\Pi}^{\perp}\mathbf{a}(\mathbf{r})}, \tag{2.29}$$

where $\hat{\Pi}^{\perp}$ is an estimate of the projection onto the noise subspace. While estimating the noise subspace to compute $\hat{\Pi}^{\perp}$, the MUSIC algorithm implicitly estimates the correlation matrix $E\left[\mathbf{y}(t)\mathbf{y}^*(t)\right]$, just like the conventional beamformer and Capon's method. The MUSIC algorithm also illustrates the opportunities available when the sources are uncorrelated. While it may seem unreasonable to assume that the sources are uncorrelated in a coherent imaging context, we have some control over the source signals in active imaging, which we address in Chapter 7.

## 2.2   Beamformer performance and tradeoffs

We can visualize any objective function and call it an image, and Section 2.1.3 revealed a taste of the many possible formulations. We focus our attention on the conventional beamformer, because of all the coherent images we have so far discussed, those produced by the conventional beamformer have a deep physical connection to incoherent image formation. By focusing on one beamformer, we can precisely define performance metrics and evaluate design tradeoffs in the presence of phase noise and timing inaccuracies. First, we characterize beamformer performance by describing

the tilt, sidelobe suppression, and width of the beamformer's beam pattern (Section 2.2.1). Second, we model phase noise and begin to explore its potential impact on performance (Section 2.2.2). Third, we discuss implementing the beamformer with time delays at both the carrier and intermediate frequencies (Section 2.2.3).

## 2.2.1   The beam pattern characterizes tradeoffs

The conventional beamformer maximizes the output power when a single source is present at the point of focus. However, to characterize the beamformer's performance, we must understand the output when a waveform is emitted from all other points in space. Due to constructive interference at the carrier frequency, the magnitude of the beamformer response is greatest when the source lies at the point of focus, and drops off as the source moves further away, tracing out a multilobed response. The beamformer response magnitude is therefore shaped like a beam when viewed as a function of source position. The response amplitude, called the beam pattern, characterizes the beamformer's performance.

The beam pattern is a function of three spatial variables and can be shaped through the design of the sensor array geometry and additional amplitude weights. In general, the beam pattern must be computed numerically. For simplicity, we restrict our attention to sources far away from a regularly-spaced linear array and do not apply amplitude weights. With these restrictions, the beam pattern is a function of just one variable, the direction of arrival of an incoming waveform, and the major tradeoffs are governed by coarse design parameters such as array size and wavelength.

We now examine the constructive interference at work in beamformer operation. We use a linear array of equally-spaced sensors to estimate the envelope of a narrowband plane wave arriving from a particular angle $\theta_T$ at an angular frequency $\omega_c$ (Figure 2-2). Say there is a single source in the far zone in direction $\theta_1^P$ with envelope $x_1(t)$. We ignore additive receiver noise, setting $\sigma_v^2 = 0$. The beamformer computes

$$g(t) = \frac{1}{M}\mathbf{a}^*(\theta_T)\mathbf{y}(t) = \frac{1}{M}\mathbf{a}^*(\theta_T)\mathbf{a}(\theta_1^P)x_1(t). \tag{2.30}$$

The beamformer delays the signals recorded at each sensor so that the wavefronts from a hypothetical plane wave propagating along the target angle $\theta_T$ are aligned, and then averages the results. If the only observed signal is in fact a single plane wave propagating along the steering angle so that $\theta_1^P = \theta_T$, then $\mathbf{a}^*(\theta_T)\mathbf{a}(\theta_1^P) = M$ and the beamformer will recover the source signal $x_1(t)$. However, other waves arriving from different angles will be attenuated by harmonic interference. This interference at the carrier frequency is how the beamformer achieves directional selectivity and is able to focus at a specific point or in a specific direction.

To quantify the beamformer's directional selectivity, we substitute the steering vector (2.6) into (2.30), and recognize that the beamformer output $g(t)$ is a geometric series:

$$g(t) \;=\; \frac{1}{M}\sum_{n=0}^{M-1} \exp\left[iknd(\cos\theta_T - \cos\theta_1^P)\right] x_1(t) \tag{2.31}$$
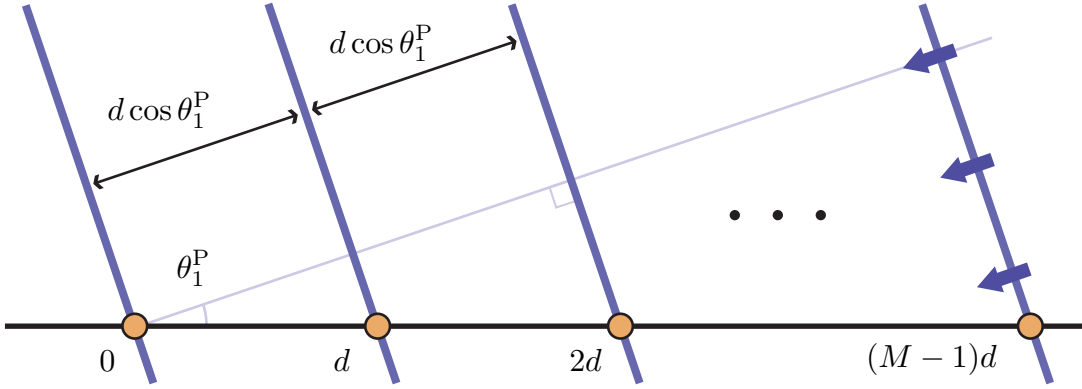
Figure 2-2: When steered in the direction of an incoming plane wave along angle $\theta_1^{\mathrm{P}}$, the conventional beamformer estimates the wave's envelope by delaying the signal measurements made at each sensor to align their wavefronts, and then averaging the results.

$$
= \frac{1}{M} \frac{1 - \exp\left[iMkd\left(\cos\theta_{\mathrm{T}} - \cos\theta_1^{\mathrm{P}}\right)\right]}{1 - \exp\left[ikd\left(\cos\theta_{\mathrm{T}} - \cos\theta_1^{\mathrm{P}}\right)\right]} x_1(t)
$$

$$
= \frac{1}{M} \exp\left[i(M-1)\frac{kd}{2}\left(\cos\theta_{\mathrm{T}} - \cos\theta_1^{\mathrm{P}}\right)\right] \frac{\sin\left[\frac{Mkd}{2}\left(\cos\theta_{\mathrm{T}} - \cos\theta_1^{\mathrm{P}}\right)\right]}{\sin\left[\frac{kd}{2}\left(\cos\theta_{\mathrm{T}} - \cos\theta_1^{\mathrm{P}}\right)\right]} x_1(t).
$$

Thus, in the absence of noise, the amplitude of the beamformer output is proportional to the plane wave signal, by a factor that depends on both the steering direction $\theta_{\mathrm{T}}$ and the angle of the incoming plane wave $\theta$:

$$
B_\theta = \frac{1}{M} \frac{\sin\left[\frac{M\pi d}{\lambda}\left(\cos\theta_{\mathrm{T}} - \cos\theta\right)\right]}{\sin\left[\frac{\pi d}{\lambda}\left(\cos\theta_{\mathrm{T}} - \cos\theta\right)\right]}, \tag{2.32}
$$

where we have used the relation $k = 2\pi/\lambda$. The factor $B_\theta$ is called the beam pattern for the array.

The beam pattern exhibits a multilobed response. From the geometry, $\theta$ and $\theta_{\mathrm{T}}$ take on values between 0 and $\pi$, so $|\cos\theta_{\mathrm{T}} - \cos\theta|$ can be as large as 2 in an extreme steering situation. The beam pattern achieves a maximum when the wave propagates along the direction the beamformer is steered in, when $\theta = \theta_{\mathrm{T}}$. As $\theta$ moves away from $\theta_{\mathrm{T}}$, the beam pattern decreases, defining a main lobe of sensitivity pointing in the steering direction. However, as $\theta$ moves even further, the beam pattern increases again, tracing out a sidelobe. In fact, if $d$ is larger than $\lambda/2$, the peaks of these sidelobes can reach the maximum. These sidelobes are then called grating lobes, and they represent ambiguities in direction that we are unable to resolve with the array. We eliminate grating lobes by using a dense array and setting $d = \lambda/2$. Although the beam pattern conveys a lot of information, we focus on three key performance

metrics: tilt, sidelobe suppression, and beam width (Figure 2-3).
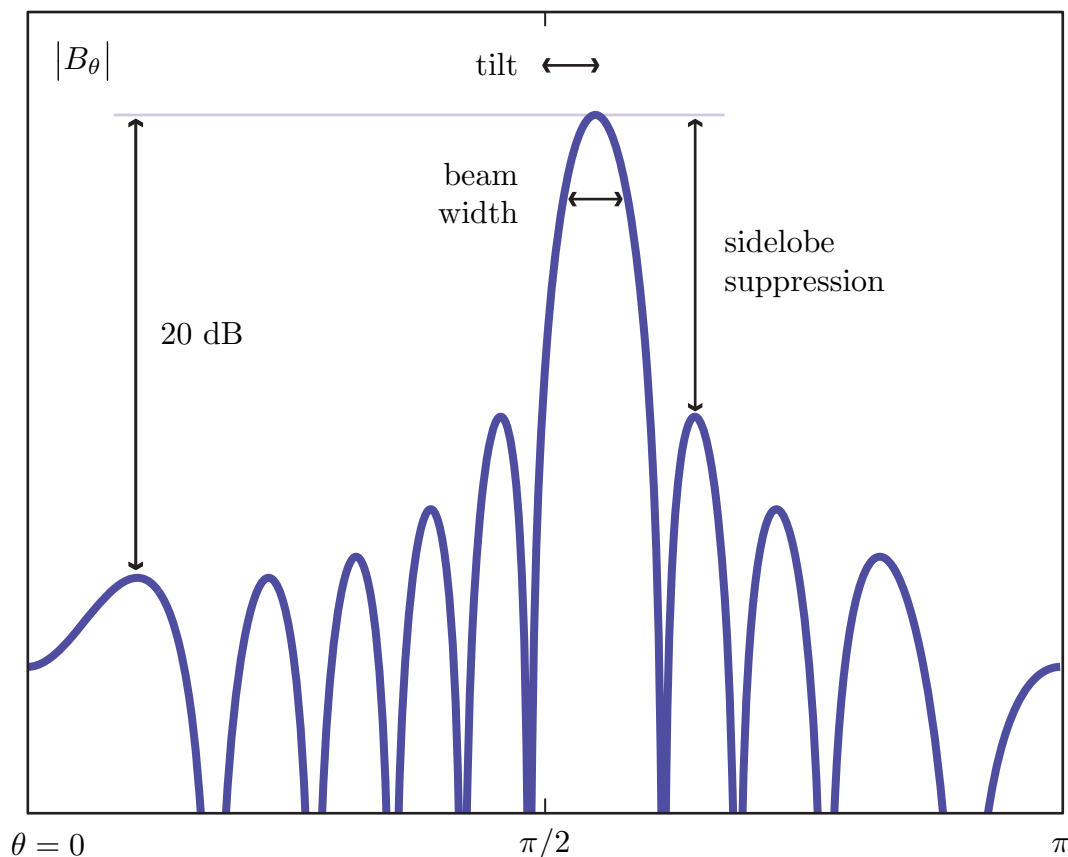


Figure 2-3: We characterize the performance of a beamformer by the tilt, sidelobe suppression, and beam width of its beam pattern $|B_\theta|$. This beamformer has $M = 10$ sensors and is intended to be steered towards $\theta_\mathrm{T} = \pi/2$.

The beam pattern can be shaped using amplitude weights and by modifying the array geometry. Since the beam pattern calculation in (2.31) is a discrete-time Fourier transform of a rectangular pulse, we can use standard digital filter design techniques to trade off sidelobe suppression and main lobe width [25]. We can also use irregularly-spaced arrays to eliminate grating lobes with fewer sensors, but at the expense of sidelobe suppression [20].

**Tilt**

Tilt is the offset between the intended steering direction $\theta_\mathrm{T}$ and the actual steering direction where the beam pattern reaches its maximum value. The tilt can only be nonzero in the presence of noise. We define tilt as the random variable

$$\tau = \arg\max_\theta |B_\theta(t)| - \theta_\mathrm{T}. \tag{2.33}$$

The tilt may vary rapidly over successive moments in time as noise causes the main lobe to wobble. But as the beam pattern is averaged over time, this wobbling will result in a broadened averaged main lobe with a smaller overall steering error. Tilt is robust to noise for moderate to large array sizes $M$.

**Sidelobe suppression**

Sidelobe suppression $S^{\mathrm{SL}}$ is the height of the first sidelobe relative to the main lobe, and indicates the susceptibility of the beamformer to interference from plane waves arriving from directions other than the steering direction $\theta_{\mathrm{T}}$. In the absence of noise, sidelobe suppression depends on the array size $M$, but becomes relatively constant at $M$ increases. To compute the sidelobe suppression of the beam pattern for the unweighted uniform linear array in (2.32), we note that the main lobe maximum is attained when $\cos\theta_{\mathrm{T}} - \cos\theta = 0$ and that the first sidelobe peak approximately occurs when the numerator is $-1$, when $\cos\theta_{\mathrm{T}} - \cos\theta = 3\lambda/2Md$. For large values of $M$, we make a small-angle approximation in the denominator of (2.32), so that the first sidelobe peak is

$$\left| \frac{-1}{M\frac{\pi d}{\lambda}(\cos\theta_{\mathrm{T}} - \cos\theta)} \right| = \frac{2}{3\pi}, \tag{2.34}$$

which is 13.46 dB down from the maximum. Engineers apply amplitude weights to further reduce the sidelobe suppression in beamformer implementations. Sidelobe suppression is only a useful performance metric when the beam pattern has not been corrupted by too much noise. At excessive noise levels, the shape of the beam pattern may not be clear enough to discern the first sidelobe, let alone its peak.

**Beam width**

Beam width is the width of the main lobe, and indicates the resolution at which the beamformer can discern two waves propagating at different angles: a narrow beam can resolve small differences in the direction of propagation, while a broader beam will blur a wider range of directions together. Deciding where the main lobe ends is a matter of convention; we use the half-power beam width, defined as the width of the interval over the main lobe where $|B_\theta|^2 > 1/2$. In the absence of noise, the beam width across a target is approximately

$$\text{beam width} \approx \frac{\rho}{\sin\theta_{\mathrm{T}}}\,\theta_{\mathrm{BW}}, \tag{2.35}$$

where $\rho$ is the range to the target and $\theta_{\mathrm{BW}}$ is the angular beam width

$$\theta_{\mathrm{BW}} \approx \lambda/Md, \tag{2.36}$$

which is the inverse of the width of the sensor array measured in wavelengths (Appendix B.1). Equation (2.35) demonstrates that the beam widens as the target moves further away and as the beamformer is steered towards the ends of the array [20]. Like sidelobe suppression, the beam width is only a useful metric when the beam

pattern has not been corrupted by too much noise, so that the main lobe structure of the beam pattern is clearly evident. The application requirements for beam width and tilt are usually related; we typically wish to be able to steer the beam at the same resolution that the beam can resolve.

## 2.2.2   Modeling phase noise

Phase noise is introduced by timing jitter in the receiver circuitry at each sensor and interferes with the coherent combination responsible for the beamformer's directional selectivity. Large sensor arrays are robust to phase noise and exhibit a threshold effect, so that exceeding a certain noise level negatively impacts beamformer performance. We will quantitatively explore the impact of phase noise in Chapter 5. For now, we incorporate phase noise into our source localization model and indicate the complications that arise.

To incorporate phase noise into our source localization model, we assume that the underlying timing jitter is small enough that the envelope is unchanged. Consequently, we model phase noise with a diagonal matrix

$$\Phi = \mathrm{diag}\left(e^{i\phi_1(t)}, \ldots, e^{i\phi_M(t)}\right) \tag{2.37}$$

so that

$$\mathbf{y}(t) = \Phi\mathbf{A}\mathbf{x}(t). \tag{2.38}$$

We have omitted the additive noise term $\mathbf{v}(t)$ in (2.7) for simplicity, and we assume that each $\phi_i(t)$ is an ergodic noise process with zero mean and variance $\sigma_\phi^2$.

We generalize (2.32) to compute the beam pattern with phase noise. Setting $d = \lambda/2$, the beam pattern

$$|B_\theta(t)| = \frac{1}{M}\left|\sum_{n=0}^{M-1} \exp\left\{i\left[n\pi\left(\cos\theta_{\mathrm{T}} - \cos\theta\right) + \phi_n(t)\right]\right\}\right| \tag{2.39}$$

is a random process indexed by the parameters $\theta$ and $\theta_T$. For $M = 2$ sensors, the beam pattern can be expressed in closed form:

$$\begin{aligned}|B_\theta(t)| &= \frac{1}{2}\left|\exp(i\phi_0(t)) + \exp\left\{i\left[\pi(\cos\theta_{\mathrm{T}} - \cos\theta) + \phi_1(t)\right]\right\}\right| \\ &= \left|\cos\left[\frac{\phi_0(t) - \phi_1(t)}{2} - \frac{\pi}{2}(\cos\theta_{\mathrm{T}} - \cos\theta)\right]\right|. \end{aligned} \tag{2.40}$$

For $M = 2$, the phase noise evidently preserves the shape of the beam pattern but steers it in an undesirable direction. Although the impact of phase noise on the beam pattern is more subtle for larger arrays, we can observe a key mechanism at work in (2.40): the phase noise is averaged, thereby reducing its effective variance.

There are different sources of averaging inherent in beamformer operation that mitigate the impact of phase noise on performance. First, beamforming across multiple sensors at a single instant in time spatially averages the phase noise, as in (2.40).

Second, the beamformer output is averaged over multiple ($N$) time samples that define a beamforming interval, temporally averaging the phase noise, as in (2.18). To distinguish between these two averaging effects, we distinguish between the *instantaneous beam pattern* computed at a particular time and the *time-averaged beam pattern* computed over the beamforming interval. According to our model, the beam pattern is ergodic, so that the time-averaged beam pattern converges to the *expected beam pattern $E[\|B_\theta(t)\|]$*.

### 2.2.3   Timing accuracy requirements

Depending on the implementation, it may be desirable to implement the beamforming algorithm by delaying the measured signals in time. The timing accuracy required depends on the frequency at which the delay is applied, and timing inaccuracies can be modeled as additive phase noise. Specifically, according to (2.1) a time delay of $\phi/\omega_c$ is required to introduce a phase shift of $\phi$ radians. The carrier frequency may be so high that this time delay is too small to be accurately implemented. In this case, we can modulate the measured signal by $\exp\left[i(\omega_0 t + \zeta(t))\right]$ to reduce the carrier frequency down to the intermediate frequency $\omega_I = \omega_c - \omega_0$. We can now introduce the phase shift of $\phi$ radians with less timing precision, by applying the larger time shift $\phi/\omega_I$. However, we have introduced additional phase noise $\zeta(t)$ in the process. Phase noise accumulates at the time scale that the processing is performed, and any phase noise introduced via modulation is additive once converted to radians. The optimal beamforming implementation will depend on the phase noise characteristics of the various components.

## 2.3   Summary and future challenges

We have described how to process coherent waves to locate sources and visualize a scene. We form images by visualizing objective functions for source location. The maximum likelihood estimator performs a multidimensional search in general, but reduces to searching for the brightest pixels in an image produced by the conventional beamformer when the sources are well-spaced. When the sources are closer together than the conventional beamformer's resolution, alternative techniques such as Capon's method and the MUSIC algorithm may resolve them. To circumvent the Fourier uncertainty principle that constrains the conventional beamformer's resolution, Capon's method attempts to block interfering sources while the MUSIC algorithm leverages assumptions about the correlation structure of the sources. We have explored the performance tradeoffs inherent in beamformer design by examining the tilt, sidelobe suppression, and beam width of the beam pattern of the conventional beamformer. Our model incorporates phase noise and timing inaccuracies.

The source localization problem provides enough structure to relate the low-level imaging of points to high-level visualization tasks. The number of sensors indicates the quality of the imaging system by specifying the beam width and thereby the resolution of the conventional beamformer. The number of sources indicates the

complexity of the visualization task. We can discretize a continuous scene according to the beamformer resolution, and represent objects with point sources of varying amplitudes. We interpret a high-level visualization task, such as the detection of an edge, as a source localization problem of locating the boundary of a collection of point sources. We solve the source localization problem by forming images composed of pixels computed by a beamformer.

The images produced by the conventional beamformer are physically related to conventional photography. In the thesis, we generalize image formation using quasi light fields, allowing us to view the conventional beamformer as a specific type of time-frequency distribution that makes a specific tradeoff between localizing energy in position and direction. We can therefore implement alternatives to the conventional beamformer to make different tradeoffs when forming images to solve high-level visualization problems. The choice of quasi light field corresponds to the choice of estimator of the correlation matrix $E[\mathbf{y}(t)\mathbf{y}^*(t)]$. Although both Capon's method and the MUSIC algorithm do not produce physically-based images from quasi light fields, implementing them involves making the same choice of correlation matrix estimate, which has a significant practical impact on performance.

# Chapter 3

# Manipulating incoherent light to form images

Having described coherent image formation in Chapter 2, we next describe how to manipulate incoherent light to form images, allowing us to compare coherent and incoherent image formation. Image formation from visible incoherent light differs from the coherent wave processing used in radar and ultrasound. A major difference is that the wavelength of the visible spectrum varies from 380 nm to 750 nm, corresponding to frequencies around 500 terahertz, which is too high to directly measure and manipulate with electronic circuits. Consequently, instead of measuring the amplitude and phase of the electric field with antennas, lenses focus light from a scene onto a plane, where photodetectors measure the energy distribution and compute the pixel values of the resulting image. Due to the physical limitations of lenses, there are approximations and tradeoffs inherent in this classical perspective on image formation that are absent in coherent wave processing.

The consumer market demand for digital cameras has fueled the emerging field of computational photography, which extends the classical optics perspective on image formation to enable technologies such as refocusing and dynamic viewpoint generation [26]. These technologies leverage the directional information of traveling waves, and were previously the exclusive domain of coherent wave processing. In computational photography, directional information is extracted from visible light without having to measure the phase, for example by using arrays of lenses [14]. An abstraction called the light field represents the directional power flow of incoherent light. From the light field, we can simulate different traditional cameras and estimate object parameters such as depth [27]. The light field thereby enables for incoherent light many applications familiar in coherent wave processing.

The light field is a common currency for camera comparison that invites us to rethink the conventional beamformer's role in image formation. Specifically, the light field decouples information capture and image formation: a generalized camera measures the light field, and a computer subsequently generates the desired image. We can therefore compare cameras by their ability to capture the light field. Although the traditional light field is only valid for incoherent light, we can interpret the beamformer as a device that captures a coherent light field. The light field captured by the

beamformer motivates the search for a universal light field that can relate coherent wave processing to conventional computational photography.

We first introduce the classical optics perspective on incoherent image formation (Section 3.1). Next, we discuss how computational photography generalizes incoherent image formation using the light field (Section 3.2). We then derive a coherent light field captured by the conventional beamformer, from which we construct the image for an ideal single-lens camera (Section 3.3). We explain why the typical tools from optics are inadequate for deriving a universal light field (Section 3.4), in preparation for unifying image formation in Chapter 4.

# 3.1   Classical optics maps a scene to an image plane

Classical optical image formation involves using an optical system, such as a lens, to direct the light radiating from a portion of a scene onto a surface, where the intensities corresponding to individual pixels are measured. The optical system establishes a connection between object points comprising the scene and image points on the surface [18]. Forming an image of a complicated scene thereby reduces to the problem of forming an image of a point. After describing how to perfectly image a point (Section 3.1.1), we make several conventional approximations to define the ideal single-lens camera (Section 3.1.2).

## 3.1.1   A Cartesian oval perfectly images a point

The simplest type of optical system is a single refracting surface, which is a discontinuity between media of different refractive indices $n_1$ and $n_2$. Only one type of surface, the Cartesian oval, can form a perfect image of a point, for the following reason. To create a perfect image of an object point $\mathbf{r}^P$, we must translate any cone of rays diverging from $\mathbf{r}^P$ into a cone of rays converging to an image point $\mathbf{r}^M$; that is, wavefronts leaving $\mathbf{r}^P$ must converge at $\mathbf{r}^M$ (Figure 3-1). Any ray from $\mathbf{r}^P$ to $\mathbf{r}^M$ must traverse the same optical path length

$$n_1|\mathbf{a} - \mathbf{r}^P| + n_2|\mathbf{r}^M - \mathbf{a}| = \text{constant} \tag{3.1}$$

for an arbitrary point $\mathbf{a}$ on the surface [28]. Equation (3.1) defines a Cartesian oval in $\mathbf{a}$.

## 3.1.2   The ideal camera perfectly images multiple points

An ideal camera perfectly images multiple points in a scene simultaneously; real cameras are approximations that deviate from this ideal. The difficulty in realizing an ideal camera is that a surface must be a Cartesian oval to perfectly image a single point, yet this surface will not perfectly image neighboring points, as each neighboring point would require its own incompatible Cartesian oval. Because we seek a single refracting surface to image a collection of points in a scene, we must settle for an approximate, rather than a perfect, image. Spherical surfaces both create
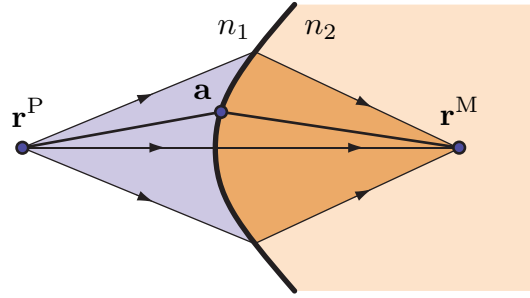
Figure 3-1: A refracting surface in the shape of a Cartesian oval perfectly images a point $\mathbf{r}^P$ onto a point $\mathbf{r}^M$. The transition to an optically denser material with index of refraction $n_2 > n_1$ bends the light so that any ray from $\mathbf{r}^P$ to $\mathbf{r}^M$ traverses the same optical path length regardless of the point of intersection $\mathbf{a}$ with the surface.

a reasonable image of flat scene elements and are easy to manufacture, for which they are widely used in optical systems. The spherical surface approximately images each object point $\mathbf{r}_i^P$ onto a corresponding image point $\mathbf{r}_i^M$, as the surface is similar to that of a Cartesian oval for rays that intersect at a small angle with respect to the surface normal (Figure 3-2). In this manner, the spherical refracting surface approximately images a small spherical patch $\sigma_o$ onto a small spherical patch $\sigma_i$. If the points on $\sigma_o$ are close to each other, then the surfaces $\sigma_o$ and $\sigma_i$ are approximately flat. We then say that $\sigma_o$ lies in the object plane and $\sigma_i$ in the image plane [28].
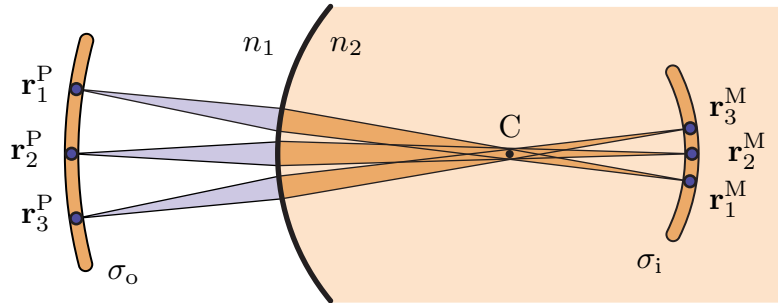


Figure 3-2: A refracting surface in the shape of a sphere with center $C$ approximately images points $\mathbf{r}_1^P$, $\mathbf{r}_2^P$, and $\mathbf{r}_3^P$ lying on a spherical patch $\sigma_o$ onto points $\mathbf{r}_1^M$, $\mathbf{r}_2^M$, and $\mathbf{r}_3^M$ lying on a spherical patch $\sigma_i$. The point images are not perfect because the transition to a higher index of refraction $n_2 > n_1$ is only similar to that of a Cartesian oval for rays that intersect the surface at a small angle with respect to the surface normal. When all points on $\sigma_o$ are close to each other, the spherical patches are approximately flat. We then say that $\sigma_o$ lies in the object plane and $\sigma_i$ in the image plane.

The spherical surface construction defines an ideal camera that forms an image of the entire object plane at a specified distance away, called the focus range. The pixel value indicating the power radiating from $\sigma_o$ towards the optical system is given

by the flux through $\sigma_i$, as might be measured by a photodetector. We introduce approximation errors by limiting the optical system to a single physical refracting surface. The approximation errors, called aberrations, are non-negligible and the subject of much analysis and practical concern [18]. Multiple lenses are typically used to build a larger optical system to manage the tradeoffs between different types of aberrations for a specific application, all to compensate for the simple physical limitation of only being able to place one lens at each location.

Practically, we do not place a photodetector inside the medium of greater refractive index $n_2$, which is usually a solid material such as glass. Instead, we shape the material at both ends, so that the light returns to the original medium, which is typically air. The resulting lens can be thought of as two refracting surfaces. If the first surface is designed to form a virtual image and the second surface a real image, then the lens will mimic the behavior of our original single refracting surface, to which our analysis above is applicable [28].

## 3.2   The light field in computational photography

The computational photography perspective on image formation is a simplified abstraction of classical optics; we work with ideal camera models using the light field, but the resulting power calculations are limited in applicability and greatly simplified compared with the underlying physics. The light field grew out of a desire to index all the potential visual stimuli presented to the human visual system. Light field calculations are based on radiometry, a phenomenological theory of energy transport that models optical power flow along rays in space. By integrating appropriate bundles of light field rays, one can simulate the image formation process of many different conventional cameras. The light field therefore acts as an intermediary between two stages of image formation: first a camera captures the light field by manipulating incoherent light and measuring its intensity distribution, and then a computer processes the captured light field to compute the pixel values of the desired image.

We first describe how the light field emerged from an image indexing problem (Section 3.2.1). Next, we formulate the traditional light field for incoherent light and review the laws of radiometry as applied to light fields (Section 3.2.2). Then we process the light field to form various images (Section 3.2.3).

### 3.2.1   The light field represents potential images

Motivated by a desire to understand human vision, Adelson and Bergen introduced the plenoptic function to represent all potential input to the visual system [29]. Specifically, the plenoptic function describes the two-dimensional image viewed by an observer placed at a specific position, and is a function of 7 parameters:

$$L(\theta, \phi, \lambda, t, V_x, V_y, V_z). \tag{3.2}$$

The observer is located at the three-dimensional position $(V_x, V_y, V_z)$. The spherical coordinates $\theta$ and $\phi$ specify a direction from the observer's position that indicates a pixel. The wavelength $\lambda$ and time $t$ further parameterize the visual stimuli.

We obtain the light field from the plenoptic function by sampling or integrating over two parameters, time and wavelength [12]. The light field $L(\mathbf{r}, \mathbf{s})$ specifies radiance for every position $\mathbf{r}$ and every unit direction $\mathbf{s}$; that is, for every ray in three-dimensional space. Radiance is power per unit projected area per unit solid angle, and has SI unit $\mathrm{W}/(\mathrm{m}^2\text{-sr})$, watts per meter square per steradian. The light field has a physical interpretation: $L(\mathbf{r}, \mathbf{s})$ is the nonnegative output of a photodetector at position $\mathbf{r}$ looking back along direction $\mathbf{s}$. Formally, spectral radiance is used when considering emissions at a particular wavelength, and radiance is spectral radiance integrated over all wavelengths, but we will not make this distinction. Radiometry is similar to geometric optics in that they both model light as traveling along rays, but radiometry further models power transport and only makes accurate predictions when the light is incoherent and the wavelength is small. Under such conditions, radiance is constant along rays in a lossless medium [30], which allows us to eliminate a redundant light field parameter. We thereby obtain a 4-dimensional light field from the 7-dimensional plenoptic function.

## 3.2.2   Pixel values are integrals of traditional light field rays

The light field is a useful tool for incoherent imaging because it acts as an intermediary between the camera and the picture, decoupling information capture and image production: the camera measures the light field, from which many different traditional pictures can be computed. We define a pixel in the image of a scene by a surface patch $\sigma$ and a virtual aperture (Figure 3-3). Specifically, we define the pixel value as the power $P$ radiated by $\sigma$ towards the aperture, just as an ideal single-lens camera would measure. According to radiometry, $P$ is an integral over a bundle of light field rays [30]:

$$P = \int_\sigma \int_{\Omega_\mathbf{r}} L(\mathbf{r}, \mathbf{s}) \cos \psi \, \mathrm{d}^2 s \, \mathrm{d}^2 r, \tag{3.3}$$

where $L(\mathbf{r}, \mathbf{s})$ is the radiance at position $\mathbf{r}$ and in unit direction $\mathbf{s}$, $\psi$ is the angle that $\mathbf{s}$ makes with the surface normal at $\mathbf{r}$, and $\Omega_\mathbf{r}$ is the solid angle subtended by the virtual aperture at $\mathbf{r}$. The images produced by many different conventional cameras can be computed from the light field using (3.3) [31].

Since the light field is constant along rays in a lossless medium, we can measure the light field remotely. To measure the light field on the surface of a scene, we follow the rays for the images we are interested in, and intercept those rays with our camera hardware (Figure 3-3). However, our hardware must be capable of measuring the radiance at a point and in a specific direction; a conventional camera that simply measures the irradiance at a point is insufficient. We can discern directional power flow using a lens array, as is done in a plenoptic camera [14].
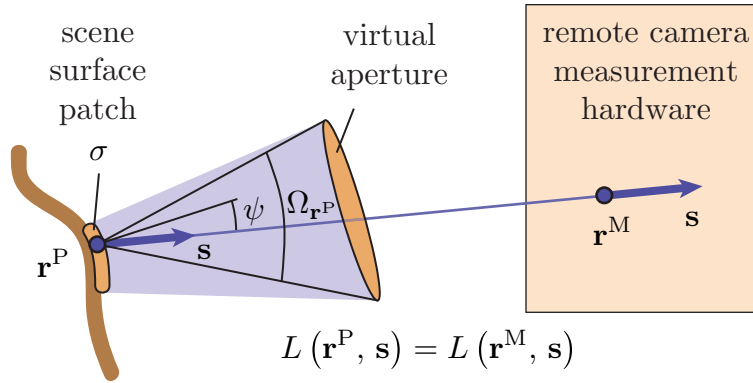
Figure 3-3: We can compute the value of each pixel in an image produced by an arbitrary virtual camera, defined as the power emitted from a scene surface patch towards a virtual aperture, by integrating an appropriate bundle of light field rays that have been previously captured with remote hardware.

### 3.2.3 Simulating cameras from the light field

We generalize the notion of a camera as a device that measures the light field, and we generalize the notion of a picture as an estimate of the light field. From this perspective, any image produced by a conventional camera can be computed on demand from a captured light field, well after the generalized picture is taken and the data is collected [12, 13]. For example, the plenoptic camera estimates the light field, from which one can compute images that different single-lens cameras focused at different depths could have produced [14]. We simulate the ideal single-lens camera by computing image pixel values from the light field and relating the camera's parameters to the light field geometry.

We simulate the ideal single-lens camera in two dimensions with three parameters: the camera is focused at range $R$, has aperture width $A$, and resolution $V$ (Figure 3-4). We select coordinates so that the center of the aperture, which coincides with the center of the camera hardware, is on the $z$-axis at $(0, R)$. Let $B$ denote the bundle of rays corresponding to the pixel centered at $\mathbf{r}^{\mathrm{P}} = (y_0, 0)$, and let $L(y, \psi)$ denote the light field on the sensor array, at position $(y, R)$ and angle $\psi$. The power corresponding to the pixel value at $\mathbf{r}^{\mathrm{P}}$ is

$$P = \int_{-A/2}^{A/2} \int_{-\pi/2}^{\pi/2} B(y, \psi) L(y, \psi) \cos \psi \, \mathrm{d}\psi \, \mathrm{d}y, \tag{3.4}$$

where $B(y, \psi)$ is an indicator function for the region of integration. The pixel values are therefore integrals of bundles of light field rays, and different camera parameters correspond to different ray bundles.

We relate the camera parameters to the ray bundle $B$ using the two-plane parameterization of the light field [12]. The ideal single-lens camera maps the object plane onto the image plane (Figure 3-5a). Pixels correspond to scene surface patches in
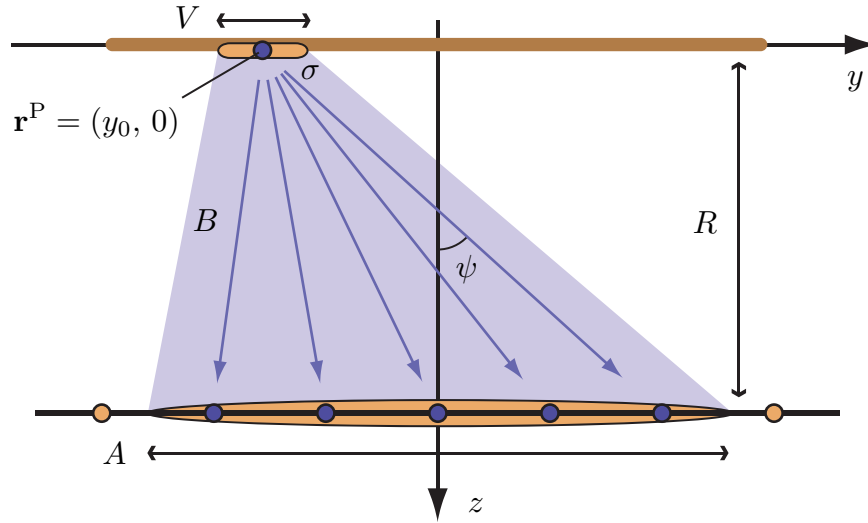
Figure 3-4: We define the ideal single-lens camera, which associates a pixel value with a flat scene surface patch of interest $\sigma$ that is parallel to the aperture. The camera integrates the light field over the bundle of rays $B$ to determine the pixel value associated with $\sigma$. We show representative rays from $\mathbf{r}^P$, the center of $\sigma$, to each sensor in a hypothetical array. We represent a ray pointing to $(y, R)$ that makes an angle $\psi$ with respect to the $z$-axis by the coordinate pair $(y, \psi)$. The camera is parameterized by the aperture diameter $A$, focus range $R$, and resolution $V$.

the object plane that are imaged onto photodetectors in the image plane. Each pixel value is the energy measured at the corresponding photodetector, which is equivalent to the energy emitted from the corresponding surface patch towards the lens. The bundle of rays $B$ consists of those rays that pass through both the scene surface patch and the lens.

We represent each ray as a two-dimensional point using a two-plane parameterization. We define the first plane $a$ as the plane in which the lens lies. We define the second plane $b$ parallel to plane $a$, a fixed distance $R_0$ away. A given ray then intersects the two planes, say at points $a_0$ and $b_0$. We plot each ray using $(a, b)$ coordinates (Figure 3-5b). The bundle of rays corresponding to each pixel is given by a different parallelogram in this two-plane parameterization. We can compute the pixel values by integrating the light field over the parallelograms. A ray intersects the planes at points related by

$$b = \frac{R_0}{R}v + \frac{R - R_0}{R}a. \tag{3.5}$$

From (3.5), we see that the parallelograms in the two-plane parameterization have height $A$, width $(R_0/R)V$, and slope $R/(R - R_0)$. Therefore, by integrating over parallelograms of different shapes, we can effectively change the simulated camera's focus range, aperture width, and resolution.
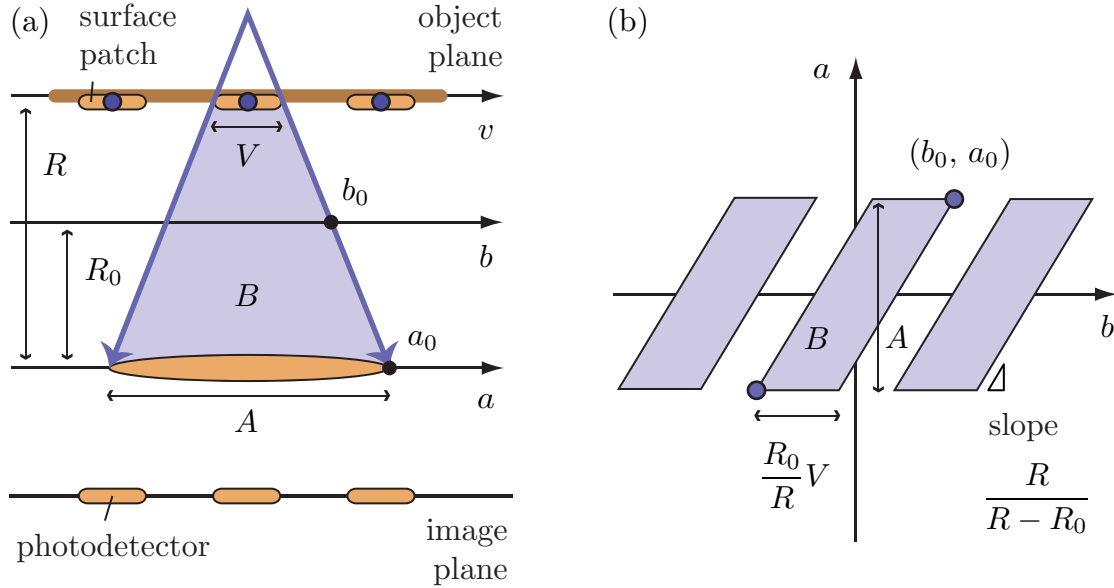
Figure 3-5: The pixel values for an ideal single-lens camera are each determined by integrating the light field over a bundle $B$ of rays that (a), in the scene, are directed from a surface patch to the lens, and (b), in the two-plane parameterization, correspond to a parallelogram. The camera focus range $R$, aperture width $A$, and resolution $V$ determine the shape of the parallelogram. Each pixel corresponds to a different scene surface patch and therefore a different parallelogram.

## 3.3  Beamformer versus incoherent technology

The classical optics and computational photography perspectives on image formation invite comparisons with the coherent wave processing discussed in Chapter 2. Compared with classical optics, the coherent array processor is not physically limited to using a single lens system to image an entire scene. Rather, the processor can simulate the propagation of light through a different lens for each pixel in the image. We demonstrate that the conventional beamformer is a special case that simulates a Cartesian oval that perfectly images the point of focus (Section 3.3.1).

Compared with computational photography, coherent image formation can be interpreted as light field processing. The light field serves as a common currency for camera comparison: while it is difficult to compare a camera with a telephoto lens with one with a wide-angle lens, it is natural to compare how two different cameras estimate the light field, from which both telephoto and wide-angle images can be computed [27]. We view the beamformer as a device that captures a coherent light field, and use this light field to compute an ideal single-lens camera image to compare with the image produced by the conventional beamformer (Section 3.3.2).

### 3.3.1 The beamformer simulates a Cartesian oval

In both classical optics and coherent wave processing, forming an image of a scene reduces to the problem of forming an image of a point. Although it may appear that the similarities between classical optics and coherent wave processing end here, the operation of the conventional beamformer mimics the mathematics that a refracting surface implicitly performs on electromagnetic waves. In general, sampling the scalar field with a sensor array offers many possibilities: we can simulate the effect of arbitrarily complicated optical systems on the waves we measure, swapping lenses in the middle of a calculation, all on a computer and without having to face any of the physical layout and manufacturing limitations in classical optics. In particular, we can perfectly image each point $\mathbf{r}$ in a scene by simulating a Cartesian oval focused at $\mathbf{r}$ (Figure 3-6). This would be physically impossible with classical optics, as we could not simultaneously pick a different lens for each point of interest in the scene.
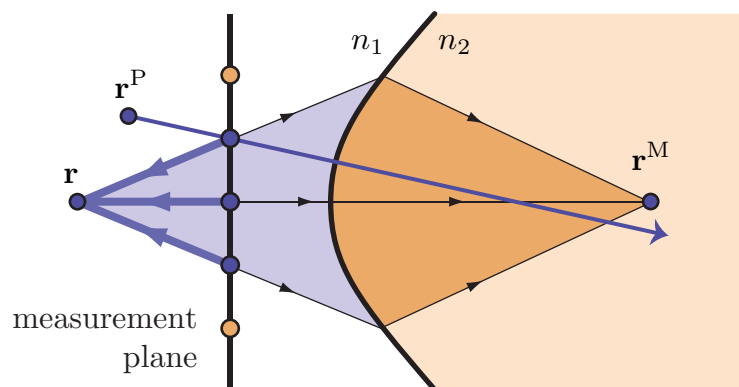


Figure 3-6: The conventional beamformer simulates a Cartesian oval by regressing signals measured at each sensor back in time so that the wavefronts align at the point of focus $\mathbf{r}$, and then adding the resulting waveforms. A physical Cartesian oval would similarly process the signals by aligning their wavefronts and combining at $\mathbf{r}^{\mathrm{M}}$ instead.

Provided that the scalar field is coherent, the beamforming algorithm is a shortcut for simulating a Cartesian oval. Specifically, by computing the inner product $\mathbf{a}^{*}(\mathbf{r})\mathbf{y}(t)$ in (2.18) using the phase shifts in (2.5), the conventional beamformer regresses the measured signals back in time so that the wavefronts align at the point of focus $\mathbf{r}$, and then adds the resulting waveforms. Apart from a delay, this processing is equivalent to what a physical Cartesian oval would do to light emitted from $\mathbf{r}$, by aligning the wavefronts along each direction and combining them at $\mathbf{r}^{\mathrm{M}}$.

We now compare the beamformer with a Cartesian oval when the source is at $\mathbf{r}^{\mathrm{P}}$, away from the point of focus. On the one hand, a physical Cartesian oval would steer a ray from $\mathbf{r}^{\mathrm{P}}$ away from $\mathbf{r}^{\mathrm{M}}$ so it would not contribute to the image there, at least in the geometric optics limit as $\lambda \to 0$. On the other hand, even though an individual sensor cannot distinguish between a ray arriving from $\mathbf{r}$ and one arriving from $\mathbf{r}^{\mathrm{P}}$, the beamformer will attenuate the signal from $\mathbf{r}^{\mathrm{P}}$ through destructive interference.

Intuitively, the end result is similar because the beamformer enforces the constant optical path length condition in (3.1), from which the shape of the Cartesian oval is determined.

It is convenient to associate the sensors used by a beamformer with a virtual lens (Figure 3-7). If we use the entire sensor array as a beamformer, differences in viewing angle are averaged away, as if a single large lens were imaging the point. If instead we use a smaller subarray as a beamformer, we can select specific viewing angles, but at the cost of worse resolution due to a wider beam width per (2.36). A correspondingly smaller lens would match the viewing angle of the smaller beamformer, but would collect less light and produce a dimmer image than a large lens.
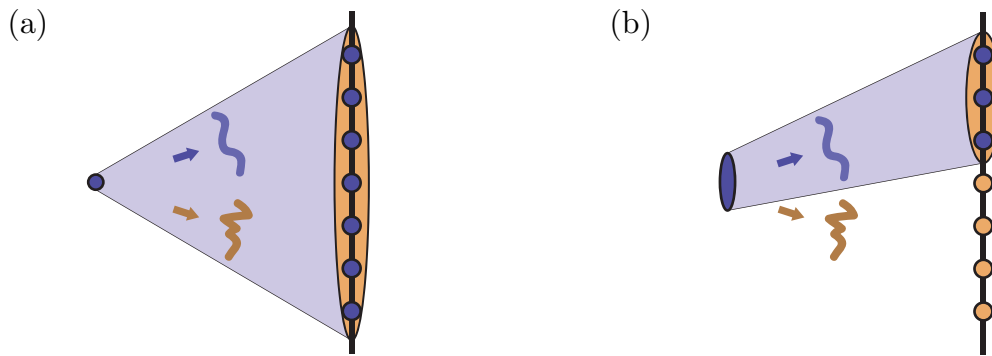


Figure 3-7: We associate a virtual lens with the sensors used by a beamformer. The beamformer (a) averages over different viewing angles when using a large array but (b) can select a specific viewing angle at the cost of worse resolution when using a smaller subarray.

### 3.3.2 The beamformer acts as a light-field camera

Although the traditional light field and radiometry only apply to incoherent radiation, we can interpret the beamformer as a device that measures a coherent light field. A beamformer estimates a spherical wave diverging from the point of focus, which we can decompose into the directional power flow represented by the light field using the laws of radiometry. The radial symmetry ensures that we avoid the difficulties we would otherwise encounter if we attempted a similar construction with an arbitrary coherent field, as we shall see in Chapter 4. After we derive the light field captured by the beamformer, we use it to simulate an ideal single-lens camera. We call the result the beamformer camera with pixel values $P^{\mathrm{BFC}}$.

In two dimensions, the beamformer estimates a cylindrical wave $U$ radiating from a point source $\mathbf{r}$ (Figure 3-8). We choose coordinates so that the origin is at $\mathbf{r}$, and we assume that the beamformer's sensor array is located a distance $\rho$ away in direction $\mathbf{s}$. The beamformer is focused at $\mathbf{r}$ and has a beam width of $\rho\theta_{\mathrm{BW}}\sec\psi$ at $\mathbf{r}$ per (2.35).
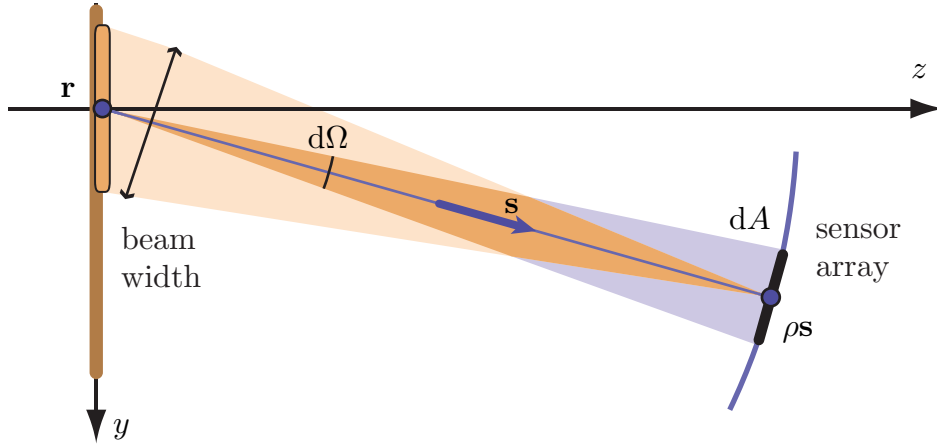
Figure 3-8: In two dimensions, a beamformer estimates a cylindrical wave radiating from a point source $\mathbf{r}$, from which we infer a coherent light field $L(\mathbf{r}, \mathbf{s})$. The sensor array is located a distance $\rho$ away from $\mathbf{r}$ in direction $\mathbf{s}$. We first compute the flux through a small portion of a sphere with differential area $\mathrm{d}A$ that subtends a differential solid angle $\mathrm{d}\Omega$ at $\mathbf{r}$. $I(\mathbf{s})$ is then flux per unit solid angle. The beamformer views the point source through a beam of width $\rho\theta_{\mathrm{BW}} \sec \psi$ at $\mathbf{r}$, so that the light field is simply $I(\mathbf{s})$ divided by $\rho\theta_{\mathrm{BW}} \sec \psi$.

At frequency $\nu$, wave number $k$, and for sufficiently large $\rho$ the scalar field is

$$U(\rho\mathbf{s}, t) = \frac{C}{\sqrt{\rho}} \exp\left[i(k\rho - 2\pi\nu t)\right], \qquad (3.6)$$

for some complex constant $C$. Neglecting variations in distance between $\mathbf{r}$ and each of the individual sensors in the array, the magnitude of the beamformer output, $|g|$, provides an estimate of $|C|/\sqrt{\rho}$, the magnitude of $U$.

To compute the light field for the cylindrical wave, we first use scalar field theory to calculate the radiant intensity, which we then distribute over the source plane (the $y$-axis) to deduce the radiance. The energy flux density vector for scalar fields is

$$\mathbf{F}(\rho\mathbf{s}, t) = -\frac{1}{4\pi k\nu}\left[\frac{\partial U^*}{\partial t}\nabla U + \frac{\partial U}{\partial t}\nabla U^*\right]. \qquad (3.7)$$

We substitute (3.6) into (3.7) to obtain the energy flux density vector

$$\mathbf{F}(\rho\mathbf{s}) = \frac{|C|^2}{\rho}\mathbf{s}. \qquad (3.8)$$

The differential flux $\mathrm{d}\Phi$ at $\rho\mathbf{s}$ through a spherical surface of differential area $\mathrm{d}A$ is therefore

$$\mathrm{d}\Phi = \frac{|C|^2}{\rho}\mathrm{d}A. \qquad (3.9)$$

The differential solid angle $d\Omega$ subtended by the surface is

$$d\Omega = \frac{dA}{\rho}. \tag{3.10}$$

Thus, as we vary $\mathbf{s}$ and move the sensor array along the circular arc of radius $\rho$, we measure a constant flux per unit solid angle, which is known in radiometry as radiant intensity $I(\mathbf{s})$ [30]:

$$I(\mathbf{s}) = \frac{d\Phi}{d\Omega} = |C|^2. \tag{3.11}$$

The point source at $\mathbf{r}$ is an idealized abstraction that is useful for understanding beamformer behavior. Radiometry requires a more realistic model, where the source is distributed over a plane instead of consolidated at a point. Since the beamformer incorporates contributions from all points within its beam, it is natural to replace the point source with an equivalent planar source that has projected area equal to the beam width at $\mathbf{r}$. We distribute $I(\mathbf{s})$ uniformly over this beam width $\rho\theta_{\mathrm{BW}} \sec\psi = \rho\theta_{\mathrm{BW}}/s_z$, so that the radiance is

$$L(\mathbf{r}, \mathbf{s}) = \frac{1}{\theta_{\mathrm{BW}}} s_z \frac{|C|^2}{\rho}. \tag{3.12}$$

Having derived the light field for a cylindrical wave, we now express the light field in terms of the beamformer output and integrate the light field per (3.4) to compute $P^{\mathrm{BFC}}$ for the ideal single-lens camera. If the resolution $V$ is small enough, we may assume that over the ray bundle $B$, $|C|^2/\rho$ is a constant given by the value at the point of focus $\mathbf{r}^{\mathrm{P}}$. Recalling that $|g|^2$ is the beamformer's estimate of $|C|^2/\rho$, and switching to $(y, \psi)$ coordinates, we conclude that the light field on $B$ for the beamformer camera is

$$L^{\mathrm{BFC}}(y, \psi) = \frac{\cos\psi}{\theta_{\mathrm{BW}}} |g|^2. \tag{3.13}$$

We substitute (3.13) into (3.4) to evaluate the pixel values of the beamformer camera,

$$P^{\mathrm{BFC}} = K|g|^2 \tag{3.14}$$

for the constant

$$K = \frac{1}{\theta_{\mathrm{BW}}} \int_{-A/2}^{A/2} \int_{-\pi/2}^{\pi/2} B(y, \psi) \cos^2\psi \, d\psi \, dy. \tag{3.15}$$

Evidently, each beamformer camera pixel value is a scaled version of the energy of the beamformer output, consistent with how beamformers are conventionally used to form images.

## 3.4  In search of a universal light field

We have used the beamformer to capture a light field for coherent waves, allowing us to relate and compare cameras for incoherent light and coherent waves through

their light field estimates. We seek a universal light field that can be computed from physical quantities such as the electric field that (i) incorporates the information from sensor measurements without presuming a beamforming algorithm and (ii) describes both coherent and incoherent image formation using the same framework. The quasi light fields, which we present in Chapter 4, are the proper universal generalization of the light field. In the remainder of this chapter, we discuss how traditional tools in optics are inadequate for describing a universal light field. We consider both geometric optics [18] and Fourier optics [32].

## 3.4.1   Geometric optics explains rays

The idea that light travels along rays dates back to Euclid [33]. Although geometric optics accounts for light traveling along rays in the limit $\lambda \to 0$, it does not account for the directional energy distribution at each point in space that the light field represents. Geometric optics is rigorously grounded in electromagnetic field theory (Appendix A.2). We define a scalar optical path function $S(\mathbf{r})$ and assume that the electric and magnetic fields at a specific frequency $\nu$ have the form

$$
\begin{aligned}
\mathbf{E}(\mathbf{r}, t) &= \mathbf{e}(\mathbf{r}) \exp\left[i 2\pi\nu(S(\mathbf{r}) - t)\right], \\
\mathbf{H}(\mathbf{r}, t) &= \mathbf{h}(\mathbf{r}) \exp\left[i 2\pi\nu(S(\mathbf{r}) - t)\right].
\end{aligned}
\tag{3.16}
$$

In the limit as the wavelength $\lambda \to 0$, the optical path function $S$ obeys the eikonal equation

$$
\nabla^2 S = n^2,
\tag{3.17}
$$

so that light bends when the index of refraction $n$ changes. The level sets of $S$ represent wavefronts. The time-averaged Poynting vector points in the direction of the gradient $\nabla S$ of the optical path function, which defines oriented curves called geometrical light rays (Figure 3-9). The eikonal equation (3.17) predicts how wavefronts refract and the light rays bend as the index $n$ varies. In a homogeneous medium, $n$ is constant so that the light rays are straight lines. Thus while electromagnetic energy does not in general flow along a straight line according to the Poynting vector, it approximately does in a macroscopic sense in a homogeneous medium, for small wavelengths when averaged over time.

We can construct a degenerate light field from the optical path function, where energy is concentrated in a single direction $\nabla S(\mathbf{r})$ at each point $\mathbf{r}$:

$$
L(\mathbf{r}, \mathbf{s}) = |\langle \mathbf{S}(\mathbf{r}, t) \rangle| \delta\left(\mathbf{s} - \frac{\nabla S(\mathbf{r})}{|\nabla S(\mathbf{r})|}\right),
\tag{3.18}
$$

where $\langle \cdot \rangle$ denotes the time average and $\delta$ is the Dirac delta function. Geometric optics allows for many wavefronts to overlap, each pointing in a different direction, but it is the coherence structure of the radiation that determines how overlapping waves are superimposed to yield a directional energy distribution. Thus without a means to model coherence, geometric optics does not account for the light field.
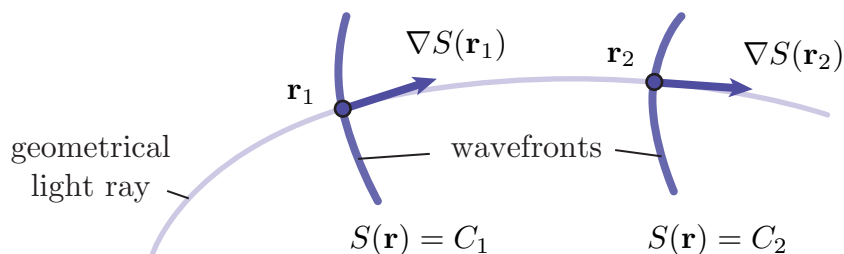
Figure 3-9: Geometric optics predicts that light travels along curved rays along the gradient of the optical path function $S$ and orthogonal to the wavefronts.

### 3.4.2 Fourier optics extracts direction from phase

We noted that the conventional beamformer extracts directional information from coherent waves when focused in the far zone, by performing a Fourier transform on the scalar field (Section 2.1.2). More generally, the two-dimensional spatial Fourier transform of the scalar field $U(\mathbf{r})$ over any plane $z = z_1$ is

$$a(\mathbf{s}) = \left(\frac{k}{2\pi}\right)^2 \int U(\mathbf{r}) \exp(-ik\mathbf{s} \cdot \mathbf{r}) \mathrm{d}^2 r, \tag{3.19}$$

so that the inverse transform

$$U(\mathbf{r}) = \int a(\mathbf{s}) \exp(ik\mathbf{s} \cdot \mathbf{r}) \mathrm{d}^2 s \tag{3.20}$$

decomposes the scalar field into the sum of plane waves with amplitude $a(\mathbf{s})$ propagating into the half-space $z > z_1$. Equation (3.20) thereby provides an angular decomposition of the field in terms of plane waves [32].

Although the Fourier transform connects two different representations of the scalar field, one in terms of position and the other in terms of direction, it does not account for the light field, which simultaneously decomposes energy in both position and direction. However, modern time-frequency representations in signal processing [34, 35] do provide this accurate description of the light field.

## 3.5  Summary and future challenges

We have described the classical optics perspective on image formation and how the field of computational photography has enabled new applications for incoherent light by introducing the light field. The light field decouples data collection from image production, and serves as a common currency for camera comparison. Specifically, the computational photography perspective invites us to recast the conventional beamformer as an estimator of a coherent light field, and to search for a universal light

field that can describe both beamformer operation as well as incoherent imaging. Geometric optics and Fourier optics do not explain light field behavior, which requires more sophisticated tools to unify the coherent and incoherent perspectives on image formation. We present the relevant research in optics, quantum mechanics, and signal processing to extend the light field to coherent radiation with quasi light fields in Chapter 4.

# Chapter 4

# Unified theory of image formation based on quasi light fields

We have described both coherent and incoherent image formation, and have made several preliminary comparisons in Chapter 3. We now present a unified theory of image formation that incorporates both the coherent and incoherent perspectives. Imaging technologies such as dynamic viewpoint generation are engineered for incoherent light using the traditional light field, and for coherent waves using electromagnetic field theory. We present a model of coherent image formation that strikes a balance between the utility of the light field and the comprehensive predictive power of Maxwell's equations. We synthesize research in optics and signal processing to formulate, capture, and form images from *quasi light fields*, which extend the light field from incoherent to coherent radiation [36]. Our coherent cameras generalize the conventional beamforming algorithm in sensor array processing, and invite further research on alternative notions of image formation.

We motivate the application of quasi light fields to sensor array processing by using the conventional beamformer in (2.18) to operate directly on the scalar field $U$ at a single time instant ($N = 1$), resulting in pixel values

$$P\left(\mathbf{r}^{\mathrm{P}}\right) = \left| \frac{1}{M} \sum_n a_n^*\left(\mathbf{r}^{\mathrm{P}}\right) U\left(\mathbf{r}_n^{\mathrm{M}}\right) \right|^2. \qquad (4.1)$$

We have previously noted that the summation in (4.1) is the space-domain version of the discrete-time short-time Fourier transform in the far zone (Section 2.1.2), and that the spatial Fourier transform decomposes the scalar field into plane wave components (Section 3.4.2). We therefore interpret (4.1) as the value of a coherent light field $L(\mathbf{r}, \mathbf{s})$, where the position $\mathbf{r}$ is the location of the sensor array, and the direction $\mathbf{s}$ is the direction from the point of focus $\mathbf{r}^{\mathrm{P}}$ towards the sensor array. In fact, we will show that the conventional beamformer in (4.1) computes the spectrogram quasi light field. By processing the scalar field measurements in different ways, we obtain other quasi light fields that exhibit different properties. In this chapter, we express the quasi light fields as functions of the scalar field samples, which can readily be computed to form images using sensor arrays, as with the conventional beamformer

in (4.1).

## 4.1   Previous work and historical context

The light field represents radiance as a function of position and direction, thereby decomposing optical power flow along rays. The light field is an important tool used in many imaging applications in different disciplines, but is traditionally limited to incoherent light. In computer graphics, a rendering pipeline can compute new views at arbitrary camera positions from the light field [12]. In computational photography, a camera can measure the light field and later generate images focused at different depths, after the picture is taken [14]. In electronic displays, an array of projectors can present multiple viewpoints encoded in the light field, enabling 3D television [37]. Many recent incoherent imaging innovations have been made possible by expressing image pixel values as appropriate integrals over light field rays.

For coherent imaging applications, the value of decomposing power by position and direction has long been recognized without the aid of a light field, since the complex-valued scalar field encodes direction in its phase. A hologram encodes multiple viewpoints, but in a different way than the light field [38]. An ultrasound machine generates images focused at different depths, but from air pressure instead of light field measurements [17]. A Wigner distribution function models the operation of optical systems in simple ways, by conveniently inferring direction from the scalar field instead of computing nonnegative light field values [39]. Comparing these applications, coherent imaging uses the scalar field to achieve results similar to those that incoherent imaging obtains with the light field.

Our goal is to provide a model of coherent image formation that combines the utility of the light field with the comprehensive predictive power of the scalar field. The similarities between coherent and incoherent imaging motivate exploring how the scalar field and light field are related, which we address by synthesizing research across three different communities. Each community is concerned with a particular Fourier transform pair and has its own name for the light field. In optics, the pair is position and direction, and Walther discovered the first generalized radiance function by matching power predictions made with radiometry and scalar field theory [40]. In quantum physics, the pair is position and momentum, and Wigner discovered the first quasi-probability distribution, or phase-space distribution, as an aid to computing the expectation value of a quantum operator [41]. In signal processing, the pair is time and frequency, and while instantaneous spectra were used as early as 1890 by Sommerfeld, Ville is generally credited with discovering the first nontrivial quadratic time-frequency distribution by considering how to distribute the energy of a signal over time and frequency [42]. Walther, Wigner, and Ville independently arrived at essentially the same function, which is one of the ways to express a light field for coherent radiation in terms of the scalar field.

The light field has its roots in radiometry, a phenomenological theory of radiative power transport that began with Herschel's observations of the sun [43], developed through the work of astrophysicists such as Chandrasekhar [44], and culminated with

its grounding in electromagnetic field theory by Friberg et al. [45]. The light field represents radiance, which is the fundamental quantity in radiometry, defined as power per unit projected area per unit solid angle. Illuminating engineers would integrate radiance to compute power quantities, although no one could validate these calculations with the electromagnetic field theory formulated by Maxwell. Gershun was one of many physicists who attempted to physically justify radiometry, and who introduced the phrase *light field* to represent a three-dimensional vector field analogous to the electric and magnetic fields [46]. Gershun's light field is a degenerate version of the one we discuss, and more closely resembles the time-averaged Poynting vector that appears in a rigorous derivation of geometric optics [18]. Subsequently, Walther generalized radiometry to coherent radiation in two different ways [40, 47], and Wolf connected Walther's work to quantum physics [48], ultimately leading to the discovery of many more generalized radiance functions [49] and a firm foundation for radiometry [45].

Meanwhile, machine vision researchers desired a representation for all the possible pictures a pinhole camera might take in space-time, which led to the current formulation of the light field. Inspired by Leonardo da Vinci, Adelson and Bergen defined a plenoptic function to describe "everything that can be seen" as the intensity recorded by a pinhole camera, parametrized by position, direction, time, and wavelength [29]. Levoy and Hanrahan tied the plenoptic function more firmly to radiometry, by redefining Gershun's phrase *light field* to mean radiance parametrized by position and direction [12]. Gortler et al. introduced the same construct, but instead called it the lumigraph [13]. The *light field* is now the dominant terminology used in incoherent imaging contexts.

Our contribution is to describe and characterize all the ways to extend the light field to coherent radiation, and to interpret coherent image formation using the resulting extended light fields. We call our extended light fields *quasi light fields*, which are analogous to the generalized radiance functions of optics, the quasi-probability and phase-space distributions of quantum physics, and the quadratic class of time-frequency distributions of signal processing. Agarwal et al. have already extended the light field to coherent radiation [49], and the signal processing community has already classified all of the ways to distribute power over time and frequency [35]. Both have traced their roots to quantum physics. But to our knowledge, no one has connected the research to show (i) that the quasi light fields represent *all* the ways to extend the light field to coherent radiation, and (ii) that the signal processing classification informs which quasi light field to use for a specific application. We further contextualize the references, making any unfamiliar literature more accessible to specialists in other areas.

In order to generate coherent images using the framework for incoherent light described in Chapter 3, we must overcome three challenges. First, we must determine how to measure power flow by position and direction to formulate a coherent light field. Second, we must capture the coherent light field remotely and be able to infer behavior at the scene surface. Third, we must be able to use equation (3.3) to produce correct power values, so that we can form images by integrating over the coherent light field. We address each challenge in a subsequent section.

We formulate the quasi light fields in Section 4.2, by reviewing and relating the relevant research in optics, quantum physics, and signal processing. In Section 4.3, we describe how to capture quasi light fields, discuss practical sampling issues, and illustrate the impact of light field choice on energy localization. In Section 4.4, we describe how to form images with quasi light fields. We derive a light field camera, demonstrate and compensate for diffraction limitations in the near zone, and generalize the conventional beamforming algorithm in sensor array processing. We remark on the utility of quasi light fields and future perspectives on image formation in Section 4.5.

## 4.2   Formulating quasi light fields

We motivate, systematically generate, and characterize the quasi light fields by relating existing research. We begin in Section 4.2.1 with research in optics that frames the challenge of extending the light field to coherent radiation in terms of satisfying a power constraint required for radiometry to make power predictions consistent with scalar field theory. While useful in developing an intuition for quasi light fields, the power constraint does not allow us to easily determine the quasi light fields. We therefore proceed in Section 4.2.2 to describe research in quantum physics that systematically generates quasi light fields satisfying the power constraint, and that shows how the quasi light fields are true extensions that reduce to the traditional light field under certain conditions. While useful for generating the quasi light fields, the quantum physics approach does not allow us to easily characterize them. Therefore, in Section 4.2.3 we map the generated quasi light fields to the quadratic class of time-frequency distributions, which has been extensively characterized and classified by the signal processing community. By relating research in optics, quantum physics, and signal processing, we express all the ways to extend the light field to coherent radiation, and provide insight on how to select an appropriate quasi light field for a particular application.

   We assume a perfectly coherent complex scalar field $U(\mathbf{r})$ at a fixed frequency $\nu$ for simplicity, although we comment in Section 4.5 on how to extend the results to broadband, partially coherent radiation. The radiometric theory we discuss assumes a planar source at $z = 0$. Consequently, although the light field is defined in three-dimensional space, much of our analysis is confined to planes $z = z_0$ parallel to the source. Therefore, for convenience, we use $\mathbf{r} = (x, y, z)$ and $\mathbf{s} = (s_x, s_y, s_z)$ to indicate three-dimensional vectors and $\mathbf{r}_\perp = (x, y)$ and $\mathbf{s}_\perp = (s_x, s_y)$ to indicate two-dimensional, projected versions.

### 4.2.1   Intuition from optics

An extended light field must produce accurate power transport predictions consistent with rigorous theory; thus the power computed from the scalar field using wave optics determines the allowable light fields via the laws of radiometry. One way to find extended light fields is to guess a light field equation that satisfies this power

constraint, which is how Walther identified the first extended light field [40]. The scenario involves a planar source at $z = 0$ described by $U(\mathbf{r})$, and a sphere of large radius $\rho$ centered at the origin. We use scalar field theory to compute the flux through part of the sphere, and then use the definition of radiance to determine the light field from the flux.

According to scalar field theory, the differential flux $d\Phi$ through a portion of the sphere subtending differential solid angle $d\Omega$ is given by integrating the radial component of the energy flux density vector $\mathbf{F}$. From diffraction theory (Appendix A.3), the scalar field in the far zone is

$$U^{\infty}(\rho\mathbf{s}) = -\frac{2\pi i}{k} s_z \frac{\exp(ik\rho)}{\rho} a(\mathbf{s}), \tag{4.2}$$

where $k = 2\pi/\lambda$ is the wave number, $\lambda$ is the wavelength, and

$$a(\mathbf{s}) = \left(\frac{k}{2\pi}\right)^2 \int U(\mathbf{r}) \exp(-ik\mathbf{s} \cdot \mathbf{r}) \, d^2r \tag{4.3}$$

is the plane wave component in direction $\mathbf{s}$ [32]. Now

$$\mathbf{F}^{\infty}(\rho\mathbf{s}) = \left(\frac{2\pi}{k}\right)^2 a(\mathbf{s})a^*(\mathbf{s})\frac{s_z^2}{\rho^2}\mathbf{s}, \tag{4.4}$$

so that

$$d\Phi = \left(\frac{2\pi}{k}\right)^2 s_z^2 a(\mathbf{s})a^*(\mathbf{s}) \, d\Omega. \tag{4.5}$$

According to radiometry, radiant intensity is flux per unit solid angle

$$I(\mathbf{s}) = \frac{d\Phi}{d\Omega} = \left(\frac{2\pi}{k}\right)^2 s_z^2 a(\mathbf{s})a^*(\mathbf{s}). \tag{4.6}$$

Radiance is $I(\mathbf{s})$ per unit projected area [30], and this is where the guessing happens: there are many ways to distribute (4.6) over projected area by factoring out $s_z$ and an outer integral over the source plane, but none yield light fields that satisfy all the traditional properties of radiance [50]. One way to factor (4.6) is to substitute the expression for $a(\mathbf{s})$ from (4.3) into (4.6) and change variables:

$$I(\mathbf{s}) = s_z \int \left[ \left(\frac{k}{2\pi}\right)^2 s_z \int U\left(\mathbf{r} + \frac{1}{2}\mathbf{r}'\right) U^*\left(\mathbf{r} - \frac{1}{2}\mathbf{r}'\right) \exp\left(-ik\mathbf{s}_\perp \cdot \mathbf{r}'_\perp\right) d^2r' \right] d^2r. \tag{4.7}$$

The bracketed expression is Walther's first extended light field

$$L^{\mathrm{W}}(\mathbf{r}, \mathbf{s}) = \left(\frac{k}{2\pi}\right)^2 s_z \mathfrak{W}(\mathbf{r}, \mathbf{s}_\perp/\lambda), \tag{4.8}$$

where

$$\mathfrak{W}(\mathbf{r}, \mathbf{s}_\perp) = \int U\left(\mathbf{r} + \frac{1}{2}\mathbf{r}'\right) U^*\left(\mathbf{r} - \frac{1}{2}\mathbf{r}'\right) \exp\left(-i2\pi\mathbf{s}_\perp \cdot \mathbf{r}'_\perp\right) \mathrm{d}^2 r' \qquad (4.9)$$

is the Wigner distribution [34]. We may manually factor (4.6) differently to obtain other extended light fields in an *ad hoc* manner, but it is hard to find and verify the properties of all extended light fields this way, and we would have to individually analyze each light field that we do manage to find. So instead, we pursue a systematic approach to exhaustively identify and characterize the extended light fields that guarantee the correct radiant intensity in (4.6).

Expressions for quasi light fields, such as that for the Wigner in (4.8), can be used directly by sensor arrays to form images. For example, instead of computing pixel values using the conventional beamformer as in (4.1), we can use the Wigner quasi light field to compute

$$P\left(\mathbf{r}^{\mathrm{P}}\right) = \sum_n U\left(\mathbf{r}_n^{\mathrm{M}}\right) U^*\left(\mathbf{r}_{-n}^{\mathrm{M}}\right) a_{2n}\left(\mathbf{r}^{\mathrm{P}}\right). \qquad (4.10)$$

All quasi light fields are similarly easy to implement with sensor arrays. However, different quasi light fields exhibit different computational structure and properties. For example, the pixel values in (4.10) can be computed using a discrete Fourier transform, yet may be negative-valued. We will explore quasi light field properties and tradeoffs after we formulate the remaining quasi light fields.

Our image formation approach will generally assign complex numbers to image pixel values. While traditional pixel values are nonnegative and have a physical interpretation in terms of power, negative and complex pixel values do not. Nonetheless, allowing complex-valued pixels offers other advantages. For coherent radiation, we discussed how images naturally arise as an intermediate quantity when solving the source localization problem. The resulting nonnegative pixel values in (2.18) happen to have a physical power interpretation. For incoherent radiation, we made a similar arbitrary choice by defining a pixel value in terms of a power quantity in (3.3). If an image with complex-valued pixels enables a better solution to a problem than an image with nonnegative pixels, then we are justified in allowing complex-valued pixels, even if doing so precludes a physical interpretation. With this perspective, nonnegativity is just one of many quasi light field properties, whose importance depends on the particular application.

## 4.2.2  Generating explicit extensions from quantum physics

The mathematics of quantum physics provides us with a systematic extended light field generator that factors the radiant intensity in (4.6) in a structured way. Walther's extended light field in (4.8) provides the hint for this connection between radiometry and quantum physics. Specifically, Wolf recognized the similarity between Walther's light field and the Wigner phase-space distribution [41] from quantum physics [48]. Subsequently, Agarwal, Foley, and Wolf repurposed the mathematics behind phase-space representation theory to generate new light fields instead of distributions [49]. We summarize their approach, define the class of quasi light fields, describe how

quasi light fields extend traditional radiometry, and show how quasi light fields can be conveniently expressed as filtered Wigner distributions.

Agarwal et al.'s key insight was to introduce a position operator $\hat{\mathbf{r}}_\perp$ and a direction operator $\hat{\mathbf{s}}_\perp$ that obey the commutation relations (Appendix A.5)

$$[\hat{x}, \hat{s}_x] = i\lambda/2\pi, \quad [\hat{y}, \hat{s}_y] = i\lambda/2\pi, \tag{4.11}$$

and to map the different ways of ordering the operators to different extended light fields. This formulation is valuable for two reasons. First, (4.11) is analogous to the quantum-mechanical relations for position and momentum, allowing us to exploit the phase-space distribution generator from quantum physics for our own purposes, thereby providing an explicit formula for extended light fields. Second, in the geometric optics limit as $\lambda \to 0$, the operators commute per (4.11), so that all of the extended light fields collapse to the same function that can be related to the traditional light field. Therefore, Agarwal et al.'s formulation not only provides us with different ways of expressing the light field for coherent radiation, but also explains how these differences arise as the wavelength becomes non-negligible.

We now summarize the phase-space representation calculus that Agarwal and Wolf invented [51] to map operator orderings to functions, which Agarwal et al. later applied to radiometry [49], culminating in a formula for extended light fields. The phase-space representation theory generates a function $\tilde{L}^\Omega$ from any operator $\hat{L}$ for each distinct way $\Omega$ of ordering collections of $\hat{\mathbf{r}}_\perp$ and $\hat{\mathbf{s}}_\perp$. So by choosing a specific $\hat{L}$ defined by its matrix elements using the Dirac notation (Appendix A.5)

$$\left\langle \mathbf{r}_\perp^{\mathrm{R}} \middle| \hat{L} \middle| \mathbf{r}_\perp^{\mathrm{C}} \right\rangle = U\left(\mathbf{r}^{\mathrm{R}}\right) U^*\left(\mathbf{r}^{\mathrm{C}}\right), \tag{4.12}$$

and supplying $\hat{L}$ as input, we obtain the extended light fields

$$L^\Omega(\mathbf{r}, \mathbf{s}) = \left(\frac{k}{2\pi}\right)^2 s_z \tilde{L}^\Omega\left(\mathbf{r}_\perp, \mathbf{s}_\perp\right) \tag{4.13}$$

as outputs. The power constraint from Section 4.2.1 translates to a minor constraint on the allowed orderings $\Omega$, so that $L^\Omega$ can be factored from (4.6). Finally, there is an explicit formula for $L^\Omega$ [51], which in Friberg et al.'s form [45] reads

$$\begin{aligned} L^\Omega(\mathbf{r}, \mathbf{s}) &= \frac{k^2}{(2\pi)^4} s_z \iiint \tilde{\Omega}\left(\mathbf{u}, k\mathbf{r}_\perp''\right) \exp\left[-i\mathbf{u} \cdot (\mathbf{r}_\perp - \mathbf{r}_\perp')\right] \exp\left(-ik\mathbf{s}_\perp \cdot \mathbf{r}_\perp''\right) \\ &\quad \times U\left(\mathbf{r}' + \frac{1}{2}\mathbf{r}''\right) U^*\left(\mathbf{r}' - \frac{1}{2}\mathbf{r}''\right) \mathrm{d}^2u\, \mathrm{d}^2r'\, \mathrm{d}^2r'', \end{aligned} \tag{4.14}$$

where $\tilde{\Omega}$ is a functional representation of the ordering $\Omega$.

Previous research has related the extended light fields $L^\Omega$ to the traditional light field, by examining how the $L^\Omega$ behave for globally incoherent light of a small wavelength, an environment technically modeled by a quasi-homogeneous source in the geometric optics limit where $\lambda \to 0$. As $\lambda \to 0$, $\hat{\mathbf{r}}_\perp$ and $\hat{\mathbf{s}}_\perp$ commute per (4.11), so

that all orderings $\Omega$ are equivalent and all of the extended light fields $L^\Omega$ collapse to the same function. Since, in the source plane, Foley and Wolf showed that one of those light fields behaves like traditional radiance [52] for globally incoherent light of a small wavelength (Appendix B.2), all of the $L^\Omega$ behave like traditional radiance for globally incoherent light of a small wavelength. Furthermore, Friberg et al. showed that many of the $L^\Omega$ are constant along rays, for globally incoherent light of a small wavelength [45]. The $L^\Omega$ thereby subsume the traditional light field, and *globally incoherent light of a small wavelength* is the environment in which traditional radiometry holds.

To more easily relate $L^\Omega$ to the signal processing literature, we conveniently express $L^\Omega$ as a filtered Wigner distribution. We introduce a function $\Pi$ and substitute

$$\tilde{\Omega}(\mathbf{u}, \mathbf{v}) = \iint \Pi(-\mathbf{a}, -\mathbf{b}) \exp\left[-i(\mathbf{a} \cdot \mathbf{u} + \mathbf{b} \cdot \mathbf{v})\right] \mathrm{d}^2 a \, \mathrm{d}^2 b \qquad (4.15)$$

into (4.14), integrate first over $\mathbf{u}$, then over $\mathbf{a}$, and finally substitute $\mathbf{b} = \mathbf{s}'_\perp - \mathbf{s}_\perp$:

$$
\begin{aligned}
L^\Omega(\mathbf{r}, \mathbf{s}) &= \left(\frac{k}{2\pi}\right)^2 s_z \iint \Pi(\mathbf{r}_\perp - \mathbf{r}'_\perp, \mathbf{s}_\perp - \mathbf{s}'_\perp) \mathfrak{W}(\mathbf{r}', \mathbf{s}'_\perp/\lambda) \, \mathrm{d}^2 r' \, \mathrm{d}^2 s' \\
&= \left(\frac{k}{2\pi}\right)^2 s_z \, \Pi(\mathbf{r}_\perp, \mathbf{s}_\perp) \otimes \mathfrak{W}(\mathbf{r}, \mathbf{s}_\perp/\lambda). \qquad (4.16)
\end{aligned}
$$

The symbol $\otimes$ in (4.16) denotes convolution in both $\mathbf{r}_\perp$ and $\mathbf{s}_\perp$. Each filter kernel $\Pi$ yields a different light field. There are only minor restrictions on $\Pi$, or equivalently on $\tilde{\Omega}$. Specifically, Agarwal and Wolf's calculus requires that [51]

$1/\tilde{\Omega}$ be an entire analytic function with no zeros on the real component axes.

$$(4.17)$$

Agarwal et al.'s derivation additionally requires that

$$\tilde{\Omega}(\mathbf{0}, \mathbf{v}) = 1 \text{ for all } \mathbf{v}, \qquad (4.18)$$

so that $L^\Omega$ satisfies the laws of radiometry and is consistent with (4.6) [49].

We call the functions $L^\Omega$, the restricted class of extended light fields that we have systematically generated, *quasi light fields*, in recognition of their connection with quasi-probability distributions in quantum physics.

### 4.2.3  Characterization from signal processing

Although we have identified the quasi light fields and justified how they extend the traditional light field, we must still show that we have found all possible ways to extend the light field to coherent radiation, and we must indicate how to select a quasi light field for a specific application. We address both concerns by relating quasi light fields to bilinear forms of $U$ and $U^*$ that are parameterized by position and direction. First, such bilinear forms reflect all the different ways to represent the energy distribution of a complex signal in signal processing, and therefore contain all

possible extended light fields, allowing us to identify any unaccounted for by quasi light fields. Second, we may use the signal processing classification of bilinear forms to characterize quasi light fields and guide the selection of one for an application.

To relate quasi light fields to bilinear forms, we must express the filtered Wigner distribution in (4.16) as a bilinear form. Towards this end, we first express the filter kernel $\Pi$ in terms of another function $K$:

$$\Pi(\mathbf{a}, \mathbf{b}) = \int K\left(-\mathbf{a} + \frac{\lambda}{2}\mathbf{v}, -\mathbf{a} - \frac{\lambda}{2}\mathbf{v}\right) \exp(-i2\pi\mathbf{b} \cdot \mathbf{v}) \, \mathrm{d}^2 v. \qquad (4.19)$$

We substitute (4.19) into (4.16), integrate first over $\mathbf{s}'_\perp$, then over $\mathbf{v}$, and finally substitute

$$\mathbf{r}^{\mathrm{R}} = \mathbf{r}' + \frac{1}{2}\mathbf{r}'', \quad \mathbf{r}^{\mathrm{C}} = \mathbf{r}' - \frac{1}{2}\mathbf{r}'' \qquad (4.20)$$

to express the quasi light field as

$$\begin{aligned}
L(\mathbf{r}, \mathbf{s}) &= \left(\frac{k}{2\pi}\right)^2 s_z \iint U\left(\mathbf{r}^{\mathrm{R}}\right) \left\{ K\left(\mathbf{r}^{\mathrm{R}}_\perp - \mathbf{r}_\perp, \mathbf{r}^{\mathrm{C}}_\perp - \mathbf{r}_\perp\right) \right. \\
&\quad \left. \times \exp\left[-ik\mathbf{s}_\perp \cdot \left(\mathbf{r}^{\mathrm{R}}_\perp - \mathbf{r}^{\mathrm{C}}_\perp\right)\right] \right\} U^*\left(\mathbf{r}^{\mathrm{C}}\right) \mathrm{d}^2 r^{\mathrm{R}} \, \mathrm{d}^2 r^{\mathrm{C}}.
\end{aligned} \qquad (4.21)$$

We recognize that (4.21) is a bilinear form of $U$ and $U^*$, with kernel indicated by the braces.

The structure of the kernel of the bilinear form in (4.21) limits $L$ to a shift-invariant energy distribution. Specifically, translating the scalar field in (4.21) in position and direction orthogonal to the $z$-axis according to

$$U(\mathbf{r}) \to U\left(\mathbf{r} - \mathbf{r}^0\right) \exp\left(ik\mathbf{s}^0_\perp \cdot \mathbf{r}_\perp\right) \qquad (4.22)$$

results in a corresponding translation in position and direction in the light field, after rearranging terms:

$$L(\mathbf{r}, \mathbf{s}) \to L\left(\mathbf{r} - \mathbf{r}^0, \mathbf{s} - \mathbf{s}^0\right). \qquad (4.23)$$

Such shift-invariant bilinear forms comprise the *quadratic class* of time-frequency distributions, which is sometimes misleadingly referred to as Cohen's class [35].

The quasi light fields represent *all* possible ways of extending the light field to coherent radiation. This is because any reasonably defined extended light field must be shift-invariant in position and direction, as translating and rotating coordinates should modify the scalar field and light field representations in corresponding ways. Thus, on the one hand, an extended light field must be a quadratic time-frequency distribution. On the other hand, (4.21) implies that quasi light fields span the entire class of quadratic time-frequency distributions, apart from the constraints on $\Pi$ described at the end of Section 4.2.2. The constraint in (4.18) is necessary to satisfy the power constraint in (4.6), which any extended light field must satisfy. The remaining constraints in (4.17) are technical details concerning analyticity and the location of zeros; extended light fields strictly need not satisfy these mild constraints, but the

light fields that are ruled out are well-approximated by light fields that satisfy them.

We obtain a concrete sensor array processing interpretation of quasi light fields by grouping the exponentials in (4.21) with $U$ instead of $K$:

$$
L(\mathbf{r}, \mathbf{s}) = \left( \frac{k}{2\pi} \right)^2 s_z \iint \left\{ U\left(\mathbf{r}^{\mathrm{R}}\right) \exp\left[ ik\mathbf{s} \cdot \left(\mathbf{r} - \mathbf{r}^{\mathrm{R}}\right) \right] \right\} K\left( \mathbf{r}_\perp^{\mathrm{R}} - \mathbf{r}_\perp, \mathbf{r}_\perp^{\mathrm{C}} - \mathbf{r}_\perp \right)
$$
$$
\times \left\{ U\left(\mathbf{r}^{\mathrm{C}}\right) \exp\left[ ik\mathbf{s} \cdot \left(\mathbf{r} - \mathbf{r}^{\mathrm{C}}\right) \right] \right\}^* \mathrm{d}^2 r^{\mathrm{R}} \, \mathrm{d}^2 r^{\mathrm{C}}. \tag{4.24}
$$

The integral in (4.24) is the expected value of the energy of the output of a spatial filter with impulse response $\exp(ik\mathbf{s} \cdot \mathbf{r})$ applied to the scalar field, when using $K$ to estimate the correlation $E[U(\mathbf{r}^{\mathrm{R}})U^*(\mathbf{r}^{\mathrm{C}})]$ by

$$
U\left(\mathbf{r}^{\mathrm{R}}\right) K\left( \mathbf{r}_\perp^{\mathrm{R}} - \mathbf{r}_\perp, \mathbf{r}_\perp^{\mathrm{C}} - \mathbf{r}_\perp \right) U^*\left(\mathbf{r}^{\mathrm{C}}\right). \tag{4.25}
$$

That is, the choice of quasi light field corresponds to a choice of how to infer coherence structure from scalar field measurements. The application of the spatial filter $\exp(ik\mathbf{s} \cdot \mathbf{r})$ is the continuous-time equivalent of the operation of the conventional beamformer in the far zone in Section 2.1.2. $K$ serves the role of a covariance matrix taper by inferring the coherence structure, analogous to Guerci's application of covariance matrix tapers to Capon's method [24]. But for our immediate purposes, the sensor array processing interpretation in (4.24) allows us to cleanly separate the choice of quasi light field in $K$ from the plane wave focusing in the exponentials.

Several signal processing books meticulously classify the quadratic class of time-frequency distributions by their properties, and discuss distribution design and use for various applications [34, 35]. We can use these resources to design quasi light fields for specific applications. For example, if we desire a light field with fine directional localization, we may first try the Wigner quasi light field in (4.8), which is a popular starting choice. We may then discover that we have too many artifacts from interfering spatial frequencies, called *cross terms*, and therefore wish to consider a reduced interference quasi light field. We might try the modified B-distribution, which is a particular reduced interference quasi light field that has a tunable parameter to suppress interference. Or, we may decide to design our own quasi light field in a transformed domain using ambiguity functions. The resulting tradeoffs can be tailoring to specific application requirements.

## 4.3  Capturing quasi light fields

To capture an arbitrary quasi light field, we sample and process the scalar field. In incoherent imaging, the traditional light field is typically captured by instead ignoring phase and making intensity measurements at a discrete set of positions and directions, as is done in the plenoptic camera [14]. While it is possible to apply the same technique to coherent imaging, only a small subset of quasi light fields can be captured this way, limiting the available tradeoffs. In comparison, all quasi light

fields can be computed from the scalar field, as in (4.16). We therefore sample the scalar field with a discrete set of sensors placed at different positions in space, and subsequently process the scalar field measurements to compute the desired quasi light field. We describe the capture process for three specific quasi light fields in Section 4.3.1, and demonstrate the different localization properties of these quasi light fields via simulation in Section 4.3.2.

## 4.3.1   Sampling the scalar field

To make the capture process concrete, we capture three different quasi light fields. Capturing quasi light fields is very similar to the sensor array processing calculations in (4.1) and (4.10), which we now justify. For simplicity, we consider a two-dimensional scene and sample an arbitrary scalar field with a linear array of sensors regularly spaced along the $y$-axis (Figure 4-1). Our model can account for general coherent phenomena, but assumes a particular sensor geometry to enable us to express the captured quasi light fields in a simple form. With this geometry, the scalar field $U$ is parameterized by a single position variable $y$, and the discrete light field $\ell$ is parameterized by $y$ and the direction component $s_y$. The sensor spacing is $d/2$, which we assume is fine enough to ignore aliasing effects. This assumption is practical for long-wavelength applications such as millimeter-wave radar. For other applications, aliasing can be avoided by applying an appropriate pre-filter. From the sensor measurements, we compute three different quasi light fields, including the spectrogram and the Wigner.
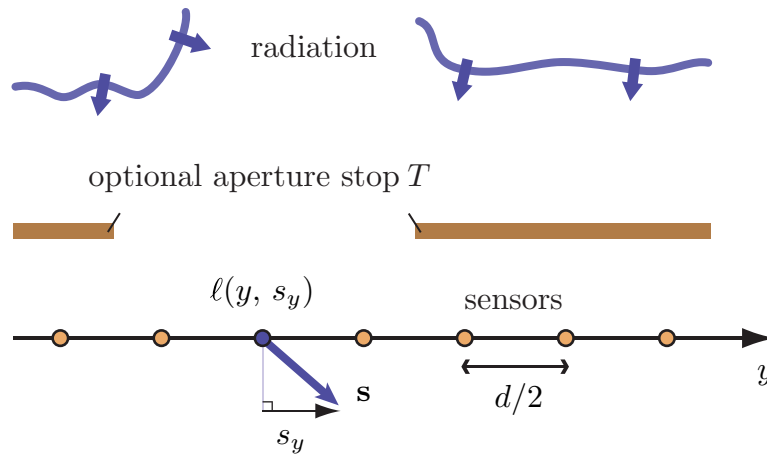


Figure 4-1: We capture a discrete quasi light field $\ell$ by sampling the scalar field at regularly-spaced sensors and processing the resulting measurements. We may optionally apply an aperture stop $T$ to mimic traditional light field capture, but this restricts us to capturing a subset of quasi light fields.

Although the spectrogram quasi light field is attractive because it can be captured like the traditional light field by making intensity measurements, it is inferior to other

quasi light fields in jointly localizing energy in both position and direction. Zhang and Levoy explain [53] how to capture the spectrogram by placing an aperture stop specified by a transmission function $T$ over the desired position $y$ before computing a Fourier transform to extract the plane wave component in the desired direction $s_y$, and previously Ziegler et al. used the spectrogram as a coherent light field to represent a hologram [38]. The spectrogram is an important quasi light field because it is the building block for the quasi light fields that can be directly captured by making intensity measurements, since all nonnegative quadratic time-frequency distributions, and therefore all nonnegative quasi light fields, are sums of spectrograms [35]. Ignoring constants and $s_z$, we compute the discrete spectrogram from the scalar field samples by

$$\ell^{\mathrm{S}}(y, s_y) = \left| \sum_n T(nd) U(y + nd) \exp\left(-iknds_y\right) \right|^2. \tag{4.26}$$

The Wigner quasi light field is a popular choice that exhibits good energy localization in position and direction [35]. We already identified the Wigner quasi light field in (4.8); the discrete version is

$$\ell^{\mathrm{W}}(y, s_y) = \sum_n U(y + nd/2) U^*(y - nd/2) \exp\left(-iknds_y\right). \tag{4.27}$$

Evidently, the spectrogram and Wigner distribute energy over position and direction in very different ways. Per (4.26), the spectrogram first uses a Fourier transform to extract directional information and then computes a quadratic energy quantity, while the Wigner does the reverse, per (4.27). On the one hand, this reversal allows the Wigner to better localize energy in position and direction, since the Wigner is not bound by the classical Fourier uncertainty principle as the spectrogram is. On the other hand, the Wigner's nonlinearities introduce cross-term artifacts by coupling energy in different directions, thereby replacing the simple uncertainty principle with a more complicated set of tradeoffs [35].

Although the spectrogram is similar to the traditional light field in having nonnegative values, no quasi light field exhibits all the properties of the traditional light field [50]. We introduce a third quasi light field to capture, that exhibits a different traditional light field property. Specifically, the traditional light field is zero where the scalar field is zero, while the support of both the spectrogram and Wigner spills over into regions where the scalar field is zero. In contrast, the conjugate Rihaczek quasi light field, which can be obtained by substituting (4.3) for $a^*(\mathbf{s})$ in (4.6) and factoring, is identically zero at all positions where the scalar field is zero and for all directions in which the plane wave component is zero:

$$L^{\mathrm{R}}(\mathbf{r}, \mathbf{s}) = s_z U^*(\mathbf{r}) \exp(ik\mathbf{s} \cdot \mathbf{r}) a(\mathbf{s}). \tag{4.28}$$

However, unlike the nonnegative spectrogram and the real Wigner, the Rihaczek is complex-valued, as each of its discoverers independently observed: Walther in optics [47], Kirkwood in quantum physics [54], and Rihaczek in signal processing [55]. The

discrete conjugate Rihaczek quasi light field is

$$\ell^{\mathrm{R}}(y, s_y) = U^*(y) \exp\left(iky s_y\right) \sum_n U(nd) \exp\left(-iknd s_y\right). \tag{4.29}$$

## 4.3.2  Localization tradeoffs

Different quasi light fields localize energy in position and direction in different ways, so that the choice of quasi light field affects the potential resolution achieved in an imaging application. We illustrate the diversity of behavior by simulating a plane wave propagating past a screen edge and computing the spectrogram, Wigner, and Rihaczek quasi light fields from scalar field samples (Figure 4-2). This simple scenario stresses the main tension between localization in position and direction: each quasi light field must encode the position of the screen edge as well as the downward direction of the plane wave. The quasi light fields serve as intermediate representations used to jointly estimate the position of the screen edge and the orientation of the plane wave. Although simple, our simulation illustrates the tradeoffs in implementing different quasi light fields with a sensor array, as in (4.1) and (4.10).
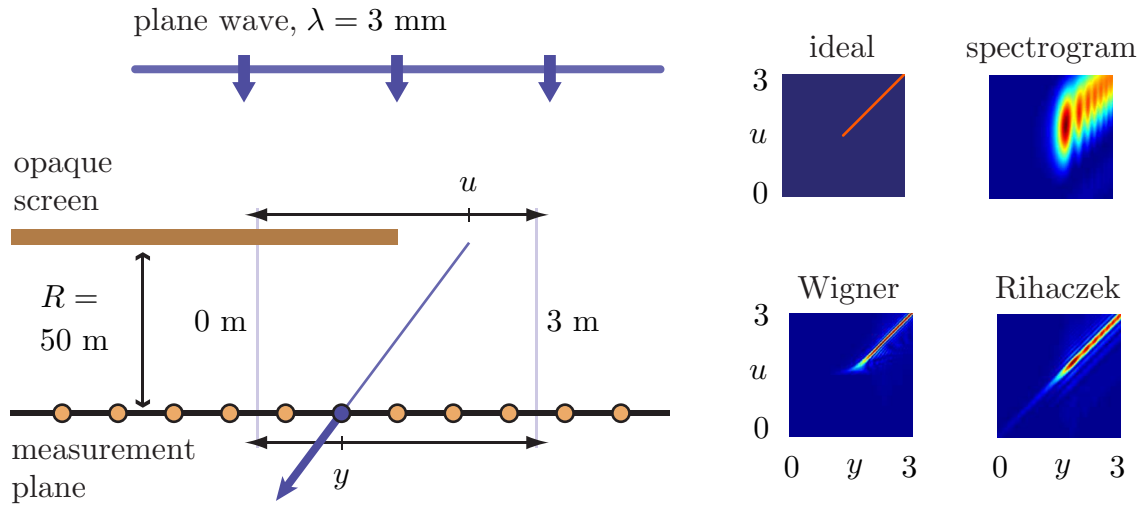


Figure 4-2: The spectrogram does not resolve a plane wave propagating past the edge of an opaque screen as well as other quasi light fields, such as the Wigner and Rihaczek. We capture all three quasi light fields by sampling the scalar field with sensors and processing the measurements according to (4.26), (4.27), and (4.29). The ringing and blurring in the light field plots indicate the diffraction fringes and energy localization limitations.

Our simulation accurately models diffraction using our implementation of the plane-wave propagation method (Appendix C.1), which is the same technique used in commercial optics software to accurately simulate wave propagation [56]. We propagate a plane wave with wavelength $\lambda = 3$ mm a distance $R = 50$ m past the screen edge, where we measure the scalar field and compute the three discrete light

fields using (4.26), (4.27), and (4.29). To compute the light fields, we set $d = \lambda/10$, run the summations over $|n| \leq 10/\lambda$, and use a rectangular window function of width 10 cm for $T$. We plot $\ell^{S}$, $|\ell^{W}|$, and $|\ell^{R}|$ in terms of the two-plane parameterization of the light field [12], so that each ray is directed from a point $u$ in the plane of the screen towards a point $y$ in the measurement plane, and so that $s_y = (y - u)/\left[R^2 + (y - u)^2\right]^{1/2}$.

We compare each light field's ability to estimate the position of the screen edge and the orientation of the plane wave (Figure 4-2). Geometric optics provides an ideal estimate: we should ideally only see rays pointing straight down ($u = y$) past the screen edge, corresponding to a diagonal line in the upper-right quadrant of the light field plots. Instead, we see blurred lines with ringing. The ringing is physically accurate and indicates the diffraction fringes formed on the measurement plane. The blurring indicates localization limitations. While the spectrogram's window $T$ can be chosen to narrowly localize energy in either position or direction, the Wigner narrowly localizes energy in both, depicting instantaneous frequency without being limited by the classical Fourier uncertainty principle [35].

It may seem that the Wigner light field is preferable to the others and the clear choice for all applications. While the Wigner light field possesses excellent localization properties, it exhibits cross-term artifacts due to interference from different plane wave components. An alternative quasi light field such as the Rihaczek can strike a balance between localization and cross-term artifacts, and therefore may be a more appropriate choice, as discussed at the end of Section 4.2.3. If our goal were to only estimate the position of the screen edge, we might prefer the spectrogram; to jointly estimate both position and plane wave orientation, we prefer the Wigner; and if there were two plane waves instead of one, we might prefer the Rihaczek. One thing is certain, however: we must abandon nonnegative quasi light fields to achieve better joint localization tradeoffs, as all nonnegative quadratic time-frequency distributions are sums of spectrograms and hence are constrained by the Fourier uncertainty principle [35].

## 4.4  Image formation from quasi light fields

We wish to form images from quasi light fields for coherent applications similarly to how we form images from the traditional light field for incoherent applications, by using (3.3) to integrate bundles of light field rays to compute pixel values (Figure 3-3). However, simply selecting a particular captured quasi light field $L$ and evaluating (3.3) raises three questions about the validity of the resulting image. First, is it meaningful to distribute coherent energy over surface area by factoring radiant intensity in (4.6)? Second, does the far-zone assumption implicit in radiometry and formalized in (4.2) limit the applicability of quasi field fields? And third, how do we capture quasi light field rays remotely if, unlike the traditional light field, quasi light fields need not be constant along rays?

The first question is a semantic one. For incoherent light of a small wavelength, we *define* an image in terms of the power radiating from a scene surface towards an

aperture, and physics tells us that this uniquely specifies the image (Section 4.2), which may be expressed in terms of the traditional light field. If we attempt to generalize the same definition of an image to partially coherent, broadband light, and specifically to coherent light at a non-zero wavelength, we must ask how to isolate the power from a surface patch towards the aperture, according to classical wave optics. But there is no unique answer; different isolation techniques correspond to different quasi light fields. Therefore, to be well-defined, we must extend the definition of an image for coherent light to include a particular choice of quasi light field, which corresponds to a particular factorization of radiant intensity.

The second and third questions speak of assumptions in the formulation of quasi light fields and in the image formation from quasi light fields that can lead to coherent imaging inaccuracies when these assumptions are not valid. Specifically, unless the scene surface and aperture are far apart, the far-zone assumption in (4.2) does not hold, so that quasi light fields are incapable of modeling near-zone behavior. Also, unless we choose a quasi light field that is constant along rays, such as an angle-impact Wigner function [57], remote measurements might not accurately reflect the light field at the scene surface [58], resulting in imaging inaccuracies. Therefore, in general, integrating bundles of remotely captured quasi light field rays produces an approximation of the image we have defined. We assess this approximation by building an accurate near-zone model in Section 4.4.1, simulating imaging performance of several coherent cameras in Section 4.4.2, and showing how our image formation procedure generalizes the classic beamforming algorithm in Section 4.4.3. We additionally make a more precise coherent camera comparison in Section 4.4.4.

## 4.4.1   Near-zone radiometry

We take a new approach to formulating light fields for coherent radiation that avoids making the assumptions that (i) the measurement plane is far from the scene surface and (ii) light fields are constant along rays. The resulting light fields are accurate in the near zone, and may be compared with quasi light fields to understand quasi light field limitations. The key idea is to express a near-zone light field $L(\mathbf{r}, \mathbf{s})$ on the measurement plane in terms of the infinitesimal flux at the point where the line containing the ray $(\mathbf{r}, \mathbf{s})$ intersects the scene surface (Figure 4-3). First we compute the scalar field at the scene surface, next we compute the infinitesimal flux, and then we identify a light field that predicts the same flux using the laws of radiometry. In contrast with Walther's approach (Section 4.2.1), (i) we do not make the far-zone approximation as in (4.2), and (ii) we formulate the light field in the measurement plane instead of in the source plane at the scene surface. Therefore, in forming an image from a near-zone light field, we are not limited to the far zone and we need not relate the light field at the measurement plane to the light field at the scene surface.

The first step in deriving a near-zone light field $L$ for the ray $(\mathbf{r}, \mathbf{s})$ is to use the scalar field on the measurement plane to compute the scalar field at the point $\mathbf{r}^{\mathrm{P}}$ where the line containing the ray intersects the scene surface. We choose coordinates so that the measurement plane is the $xy$-plane, the scene lies many wavelengths away in the negative $z < 0$ half-space, and $\mathbf{r}$ is at the origin. We denote the distance between the
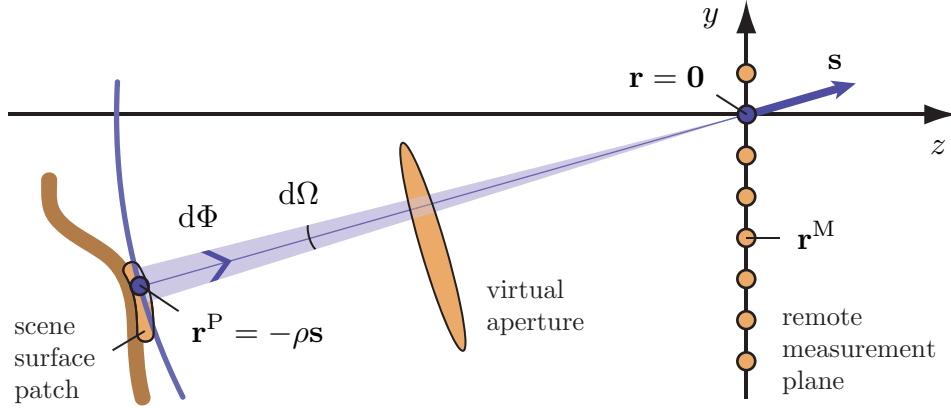
Figure 4-3: To ensure that integrating bundles of remote light field rays in the near zone results in an accurate image, we derive a light field $L_\rho^{\mathrm{R}}(\mathbf{r}, \mathbf{s})$ in the measurement plane from the infinitesimal flux $\mathrm{d}\Phi$ at the point $\mathbf{r}^{\mathrm{P}}$ where the ray originates from the scene surface patch. We thereby avoid making the assumptions that the measurement plane is far from the scene and that the light field is constant along rays.

source $\mathbf{r}^{\mathrm{P}}$ on the scene surface and the point of observation $\mathbf{r}$ by $\rho$. Under a reasonable bandwidth assumption, the inverse diffraction formula expresses the scalar field at $\mathbf{r}^{\mathrm{P}}$ in terms of the scalar field on the measurement plane [59]:

$$U(\mathbf{r}^{\mathrm{P}}) = \frac{ik}{2\pi} \int U(\mathbf{r}^{\mathrm{M}}) \frac{-z^{\mathrm{P}}}{|\mathbf{r}^{\mathrm{P}} - \mathbf{r}^{\mathrm{M}}|} \frac{\exp(-ik|\mathbf{r}^{\mathrm{P}} - \mathbf{r}^{\mathrm{M}}|)}{|\mathbf{r}^{\mathrm{P}} - \mathbf{r}^{\mathrm{M}}|} \, \mathrm{d}^2 r^{\mathrm{M}}. \tag{4.30}$$

Next, we compute the differential flux $\mathrm{d}\Phi$ through a portion of a sphere at $\mathbf{r}^{\mathrm{P}}$ subtending differential solid angle $\mathrm{d}\Omega$. We obtain $\mathrm{d}\Phi$ by integrating the radial component of the energy flux density vector

$$\mathbf{F}(\mathbf{r}^{\mathrm{P}}) = -\frac{1}{4\pi k \nu} \left[ \frac{\partial U^*}{\partial t} \nabla U + \frac{\partial U}{\partial t} \nabla U^* \right]. \tag{4.31}$$

To keep the calculation simple, we ignore amplitude decay across the measurement plane, approximating

$$|\mathbf{r}^{\mathrm{P}} - \mathbf{r}^{\mathrm{M}}| \approx |\mathbf{r}^{\mathrm{P}}| \tag{4.32}$$

outside the exponential in (4.30), and

$$\frac{\partial}{\partial |\mathbf{r}^{\mathrm{P}}|} |\mathbf{r}^{\mathrm{P}} - \mathbf{r}^{\mathrm{M}}| \approx 1, \tag{4.33}$$

when evaluating (4.31), resulting in

$$\mathbf{F}(-\rho \mathbf{s}) = \left( \frac{2\pi}{k} \right)^2 \tilde{a}(-\rho \mathbf{s}) \tilde{a}^*(-\rho \mathbf{s}) \frac{s_z^2}{\rho^2} \mathbf{s}, \tag{4.34}$$

where

$$\tilde{a}(-\rho\mathbf{s}) = \left(\frac{k}{2\pi}\right)^2 \int U(\mathbf{r}^{\mathrm{M}}) \exp(-ik|-\rho\mathbf{s} - \mathbf{r}^{\mathrm{M}}|) \, \mathrm{d}^2 r^{\mathrm{M}}. \tag{4.35}$$

Thus,

$$\mathrm{d}\Phi = \left(\frac{2\pi}{k}\right)^2 s_z^2 \tilde{a}(-\rho\mathbf{s}) \tilde{a}^*(-\rho\mathbf{s}) \, \mathrm{d}\Omega. \tag{4.36}$$

Finally, we factor out $s_z$ and an outer integral over surface area from $\mathrm{d}\Phi/\mathrm{d}\Omega$ to determine a near-zone light field. Unlike in Section 4.2.1, the nonlinear exponential argument in $\tilde{a}$ complicates the factoring. Nonetheless, we obtain a near-zone light field that generalizes the Rihaczek by substituting (4.35) for $\tilde{a}^*$ in (4.36). After factoring and freeing $\mathbf{r}$ from the origin by substituting $\mathbf{r} - \rho\mathbf{s}$ for $-\rho\mathbf{s}$, we obtain

$$\begin{aligned}
L_\rho^{\mathrm{R}}(\mathbf{r}, \mathbf{s}) &= s_z U^*(\mathbf{r}) \exp(ik\rho)\tilde{a}(\mathbf{r} - \rho\mathbf{s}) \\
&= \left(\frac{k}{2\pi}\right)^2 s_z U^*(\mathbf{r}) \exp(ik\rho) \\
&\quad \times \int U\left(\mathbf{r}^{\mathrm{M}}\right) \exp\left(-ik|\mathbf{r} - \rho\mathbf{s} - \mathbf{r}^{\mathrm{M}}|\right) \, \mathrm{d}^2 r^{\mathrm{M}},
\end{aligned} \tag{4.37}$$

where the subscript $\rho$ reminds us of this near-zone light field's dependence on distance.

$L_\rho^{\mathrm{R}}$ is evidently neither the traditional light field nor a quasi light field, as it depends directly on the scene geometry through an additional distance parameter. This distance parameter $\rho$ is a function of $\mathbf{r}$, $\mathbf{s}$, and the geometry of the scene; it is the distance along $\mathbf{s}$ between the scene surface and $\mathbf{r}$. We may integrate $L_\rho^{\mathrm{R}}$ over a bundle of rays to compute the image pixel values just like any other light field, as long as we supply the right value of $\rho$ for each ray. In contrast, quasi light fields are incapable of modeling optical propagation in the near zone, as it is insufficient to specify power flow along rays. We must also know the distance between the source and point of measurement along each ray.

We can obtain near-zone generalizations of all quasi light fields through the sensor array processing interpretation in Section 4.2.3. Recall that each quasi light field corresponds to a particular choice of the function $K$ in (4.24). For example, setting $K(\mathbf{a}, \mathbf{b}) = \delta(\mathbf{b})$, where $\delta$ is the Dirac delta function, yields the Rihaczek quasi light field $L^{\mathrm{R}}$ in (4.28). To generalize quasi light fields to the near zone, we focus at a point instead of a plane wave component by using a spatial filter with impulse response $\exp\left(-ik\,|\mathbf{r} - \rho\mathbf{s}|\right)$ instead of $\exp(ik\mathbf{s}\cdot\mathbf{r})$ in (4.24). The application of the spatial filter $\exp\left(-ik\,|\mathbf{r} - \rho\mathbf{s}|\right)$ is the continuous-time equivalent of the operation of the conventional beamformer in Section 2.1.2. Then, choosing $K(\mathbf{a}, \mathbf{b}) = \delta(\mathbf{b})$ yields $L_\rho^{\mathrm{R}}$, the near-zone generalization of the Rihaczek in (4.37), and choosing other functions $K$ yield near-zone generalizations of the other quasi light fields.

### 4.4.2  Near-zone diffraction limitations

We compute and compare image pixel values using the Rihaczek quasi light field $L^{\mathrm{R}}$ and its near-zone generalization $L_\rho^{\mathrm{R}}$, demonstrating how all quasi light fields implicitly

make the Fraunhofer diffraction approximation that limits accurate imaging to the far zone. First, we construct coherent cameras from $L^{\mathrm{R}}$ and $L_\rho^{\mathrm{R}}$. For simplicity, we consider a two-dimensional scene and sample the light fields, approximating the integral over a bundle of rays by the summation of discrete rays directed from the center $\mathbf{r}^{\mathrm{P}}$ of the scene surface patch to each sensor on a virtual aperture of diameter $A$, equally spaced every distance $d$ in the measurement plane (Figure 4-4a). Ignoring constants and $s_z$, we compute the pixel values for a far-zone camera from the Rihaczek quasi light field in (4.28),

$$P^{\mathrm{R}} = \sum_{|nd|<A/2} \left\{ \left[ U(nd) \exp\left( -iknds_y^n \right) \right]^* \sum_m U(md) \exp\left( -ikmds_y^n \right) \right\}, \qquad (4.38)$$

and for a near-zone camera from the near-zone generalization of the Rihaczek in (4.37),

$$P_\rho^{\mathrm{R}} = \left[ \sum_{|nd|<A/2} U(nd) \exp\left( -ik\Delta_n \right) \right]^* \left[ \sum_m U(md) \exp\left( -ik\Delta_m \right) \right]. \qquad (4.39)$$

In (4.38), $\mathbf{s}^n$ denotes the unit direction from $\mathbf{r}^{\mathrm{P}}$ to the $n^{\mathrm{th}}$ sensor, and in (4.39), $\Delta_n$ denotes the distance between $\mathbf{r}^{\mathrm{P}}$ and the $n^{\mathrm{th}}$ sensor.

By comparing the exponentials in (4.38) with those in (4.39), we see that the near-zone camera aligns the sensor measurements along spherical wavefronts diverging from the point of focus $\mathbf{r}^{\mathrm{P}}$, while the far-zone camera aligns measurements along plane wavefront approximations (Figure 4-4b). Spherical wavefront alignment makes physical sense in accordance with the Huygens-Fresnel principle of diffraction, while approximating spherical wavefronts with plane wavefronts is reminiscent of Fraunhofer diffraction. In fact, the far-zone approximation in (4.2) used to derive quasi light fields follows directly from the Rayleigh-Sommerfeld diffraction integral by linearizing the exponentials, which is precisely Fraunhofer diffraction. Therefore, all quasi light fields are only valid for small Fresnel numbers, when the source and point of measurement are sufficiently far away from each other.

We expect the near-zone camera to outperform the far-zone camera in near-zone imaging applications, which we demonstrate by comparing their ability to resolve small targets moving past their field of view. As a baseline, we introduce a third camera with nonnegative pixel values $P_\rho^{\mathrm{B}}$ by restricting the summation over $m$ in (4.39) to $|md| < A/2$, which is a scaled version of the beamformer camera $P^{\mathrm{BFC}}$ in (3.14). Alternatively, we could extend the summation over $n$ in (4.39) to the entire array, but this would average anisotropic responses over a wider aperture diameter, resulting in a different image. We simulate an opaque screen containing a pinhole that is backlit with a coherent plane wave (Figure 4-5). The sensor array is $D = 2$ m wide and just $R = 1$ m away from the screen. The virtual aperture is $A = 10$ cm wide and the camera is focused on a fixed 1 mm pixel straight ahead on the screen. The pinhole has width 1 mm, which is smaller than the wavelength $\lambda = 3$ mm, so the plane wavefronts bend into slightly spherical shapes via diffraction. We move the pinhole to
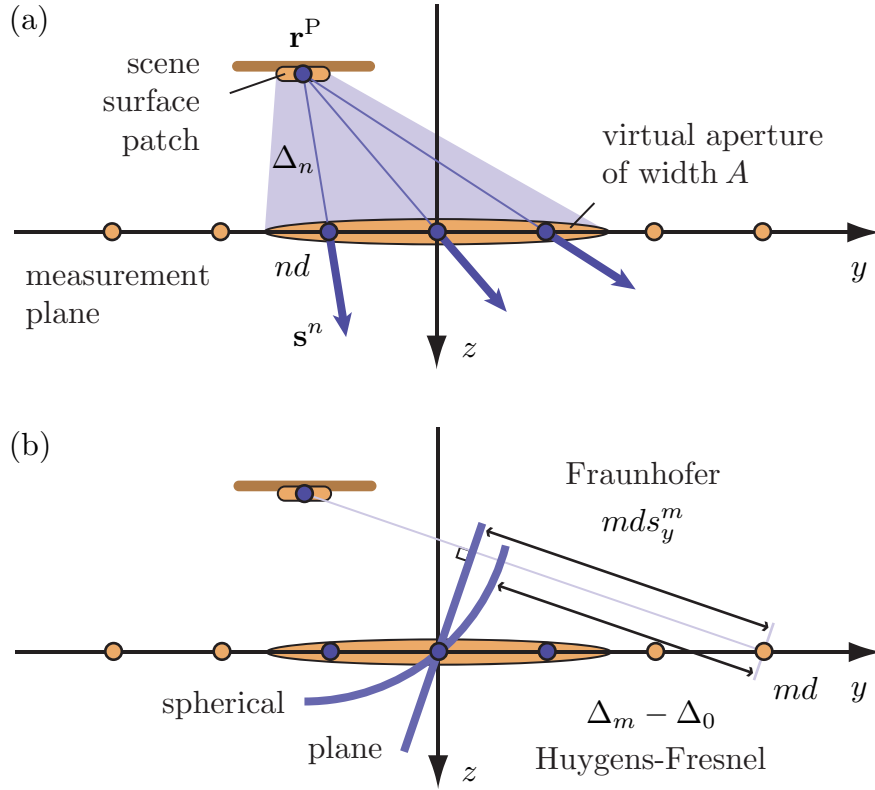
Figure 4-4: The near-zone light field results in a camera that aligns spherical wavefronts diverging from the point of focus $\mathbf{r}^{\mathrm{P}}$, in accordance with the Huygens-Fresnel principle of diffraction, while quasi light fields result in cameras that align plane wavefront approximations, in accordance with Fraunhofer diffraction. Quasi light fields are therefore only accurate in the far zone. We derive both cameras by approximating the integral over a bundle of rays by the summation of discrete light field rays (a), and we interpret the operation of each camera by how they align sensor measurements along wavefronts from $\mathbf{r}^{\mathrm{P}}$ (b).

the right, recording pixel values $|P^{\mathrm{R}}|$, $|P_\rho^{\mathrm{R}}|$, and $P_\rho^{\mathrm{B}}$ for each camera at each pinhole position. Due to the nature of the coherent combination of the sensor measurements that produces the pixel values, each camera records a multilobed response. The width of the main lobe indicates the near-zone resolution of the camera.

The near-zone camera is able to resolve the pinhole down to its actual size of 1 mm, greatly outperforming the far-zone camera which records a blur 66 cm wide, and even outperforming the beamformer camera. Neither comparison is surprising. First, with a Fresnel number of $D^2/R\lambda \approx 1333$, the Fraunhofer approximation implicitly made by quasi light fields does not hold for this scenario, so we expect the far-zone camera to exhibit poor resolution. Second, the near-zone camera uses the entire $D = 2$ m array instead of just the sensors on the virtual aperture that the beamformer camera is restricted to, and the extra sensors lead to improved resolution.
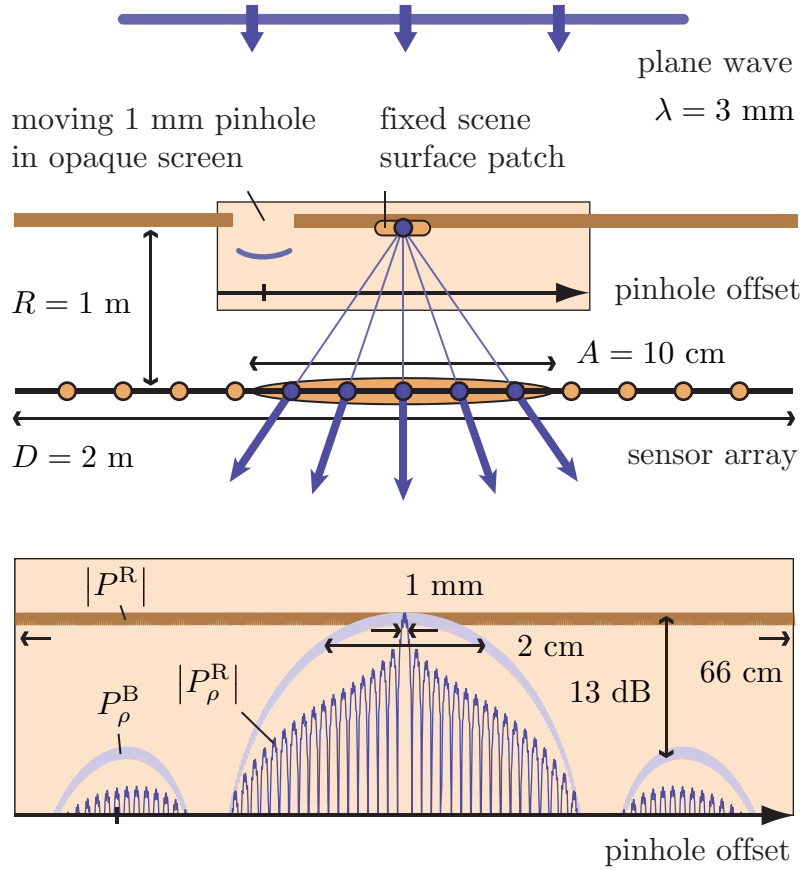
Figure 4-5: Images of nearby objects formed from pure quasi light fields are blurry. In the scene, a small backlit pinhole moves across the field of view of a sensor array that implements three cameras, each computing one pixel value for each pinhole position, corresponding to a fixed surface patch. As the pinhole crosses the fixed scene surface patch, the near-zone camera resolves the pinhole down to its actual size of 1 mm, while the far-zone camera records a blur 66 cm wide.

### 4.4.3   Generalized beamforming

We compare image formation from light fields with traditional perspectives on coherent image formation, by relating quasi light fields and our coherent cameras with the conventional beamforming algorithm from Chapter 2. The beamforming algorithm estimates a spherical wave diverging from a point of focus $\mathbf{r}^P$ by delaying and averaging sensor measurements. When the radiation is narrowband, the delays are approximated by phase shifts. With the sensor array geometry from Section 4.4.2, the beamformer output is

$$g = \sum_m T(md)U(md)\exp(-ik\Delta_m), \qquad (4.40)$$

where the $T(md)$ are amplitude weights used to adjust the beamformer's performance. As $\mathbf{r}^P$ moves into the far zone,

$$\Delta_m - \Delta_0 \to mds_y^m \to mds_y^0, \tag{4.41}$$

so that apart from a constant phase offset, (4.40) becomes a short-time Fourier transform

$$g^\infty = \sum_m T(md)U(md)\exp(-ikmds_y^0). \tag{4.42}$$

Evidently, $|g^\infty|^2$ is a spectrogram quasi light field, and we may select $T$ to be a narrow window about a point $\mathbf{r}$ to capture $L^S(\mathbf{r}, \mathbf{s}^0)$. We have already seen how quasi light fields generalize the spectrogram, and sensor array processors can readily compute alternative quasi light fields to achieve new tradeoffs for source localization and imaging performance.

Current beamformer applications instead typically select $T$ to be a wide window to match the desired virtual aperture, and assign the corresponding pixel value to the output power $|g|^2$. We can decompose the three cameras in Section 4.4.2 into such beamformers. First, we write $P_\rho^R$ in (4.39) in terms of two different beamformers,

$$P_\rho^R = g_1^* g_2, \tag{4.43}$$

where

$$g_1 = \sum_{|nd|<A/2} U(nd)\exp\left(-ik\Delta_n\right) \tag{4.44}$$

and

$$g_2 = \sum_m U(md)\exp\left(-ik\Delta_m\right), \tag{4.45}$$

so that the windows for $g_1$ and $g_2$ are rectangular with widths matching the aperture $A$ and sensor array $D$, respectively. Next, by construction

$$P_\rho^B = |g_1|^2. \tag{4.46}$$

Finally, in the far zone, $\mathbf{s}^n \to \mathbf{s}^0$ in (4.38) so that

$$P^R \to \left(g_1^\infty\right)^* g_2^\infty, \tag{4.47}$$

where $g_1^\infty$ and $g_2^\infty$ are given by (4.42) with the windows $T$ used in (4.44) and (4.45). In other words, the near-zone camera is the Hermitian product of two different beamformers, and is equivalent to the far-zone camera in the far zone.

We interpret the role of each component beamformer from the derivation of (4.39). Beamformer $g_1^*$ aggregates power contributions across the aperture using measurements of the conjugate field $U^*$ on the aperture, while beamformer $g_2$ isolates power from the point of focus using all available measurements of the field $U$. In this manner, the tasks of aggregating and isolating power contributions are cleanly divided between the two beamformers, and each beamformer uses the measurements from those sensors appropriate to its task. In contrast, the beamformer camera uses the same set of

sensors for both the power aggregation and isolation tasks, thereby limiting its ability to optimize over both tasks.

The near-zone camera achieves a new tradeoff between resolution and anisotropic sensitivity. We noted that the near-zone camera exhibits better resolution than the beamformer, for the same virtual aperture (Figure 4-5). This is not an entirely fair comparison because the near-zone camera is using sensor measurements outside the aperture, and indeed, a beamformer using the entire array would achieve comparable resolution. However, extending the aperture to the entire array results in a different image, as anisotropic responses are averaged over a wider aperture diameter. We interpret the near-zone camera's behavior by computing the magnitude

$$\left| P_\rho^{\mathrm{R}} \right| = \sqrt{|g_1|^2 |g_2|^2}. \tag{4.48}$$

Evidently, the pixel magnitude of the near-zone camera is the geometric mean of the two traditional beamformer output powers. $|P_\rho^{\mathrm{R}}|$ has better resolution than $|g_1|^2$ and better anisotropic sensitivity than $|g_2|^2$.

Image formation with alternative light fields uses the conjugate field and field measurements to aggregate and isolate power in different ways. In general, image pixel values do not neatly factor into the product of beamformers, as they do with the Rihaczek (Appendix B.3.1).

### 4.4.4  Explicit camera comparison

For clarity of presentation, we ignored the leading constants for the light fields in our derivation of the cameras in Section 4.4.2. We now restore these constants to demonstrate the minor differences they represent. We interpret the discretization of the bundle of rays in Section 4.4.2 by approximating the indicator function $B(y, \psi)$ in (3.4). For the ideal single-lens camera (Figure 3-4), the bundle of rays corresponds to the space bounded between the curves $y = y_0 - V/2 + R \tan \psi$ and $y = y_0 + V/2 + R \tan \psi$ (Figure 4-6). The subset of $B$ defined by the cone of rays with vertex $\mathbf{r}^{\mathrm{P}}$ corresponds to the curve $y = y_0 + R \tan \psi$. We therefore approximate $B(y, \psi)$ with a Dirac delta function along this curve. We weight the delta function with the width of the region of integration, $V \, \mathrm{d}\psi / \mathrm{d}y$. Since

$$\mathrm{d}\psi = \frac{\cos^2 \psi}{R} \mathrm{d}y, \tag{4.49}$$

the region of integration becomes

$$B(y, \psi) = \frac{V \cos^2 \psi}{R} \delta(y - y_0 - R \tan \psi). \tag{4.50}$$

We substitute (4.50) into (3.4) to obtain

$$P = \frac{V}{R} \int_{-A/2}^{A/2} \int_{-\pi/2}^{\pi/2} L(y, \psi) \delta(y - y_0 - R \tan \psi) \cos^3 \psi \, \mathrm{d}\psi \, \mathrm{d}y, \tag{4.51}$$

and then integrate out $\psi$:

$$P = \frac{V}{R} \int_{-A/2}^{A/2} L(y, \psi_y) \cos^3 \psi_y \, \mathrm{d}y, \tag{4.52}$$

where $\psi_y$ denotes the angle that a ray from $\mathbf{r}^P$ to $(y, R)$ makes with respect to the $z$-axis.
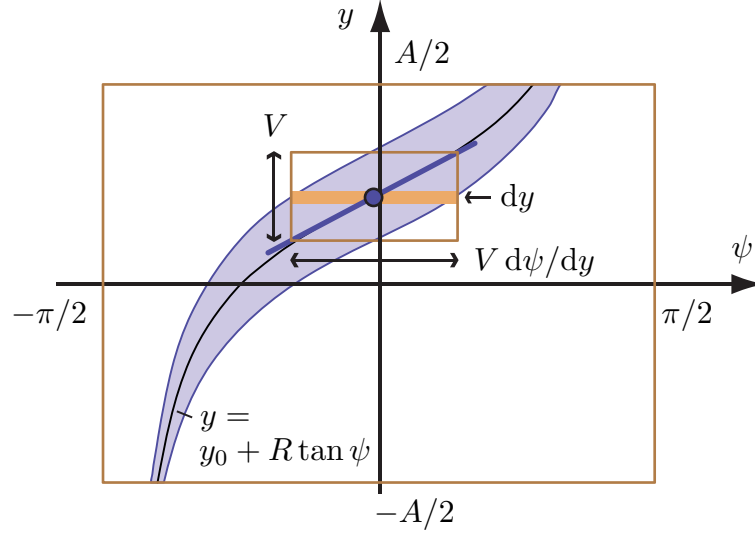


Figure 4-6: We approximate the region of integration of the light field over the bundle of rays $B$ for the ideal single-lens camera in Figure 3-4. The exact region $B(y, \psi)$ is bounded between the curves $y = y_0 - V/2 + R \tan \psi$ and $y = y_0 + V/2 + R \tan \psi$, separated by a vertical distance $V$. We collapse the ray bundle to a cone with vertex $\mathbf{r}^P$, corresponding to the curve $y = y_0 + R \tan \psi$. We approximate $B(y, \psi)$ by a Dirac delta function on this curve, weighted by the width of the rectangle, $V \, \mathrm{d}\psi/\mathrm{d}y$.

We now use (4.52) to recompute the pixel values for the near-zone camera and the beamformer camera. For the near-zone camera, we first express the near-zone generalization of the Rihaczek light field in $(y, \psi)$ coordinates as

$$L_\rho^R(y, \psi) = \frac{k}{2\pi} U^*(y) \exp(ik\Delta_y) \cos \psi \int U(y') \exp(-ik\Delta_{y'}) \, \mathrm{d}y', \tag{4.53}$$

where $\Delta_y$ is the distance between $\mathbf{r}^P$ and $(y, R)$. We substitute (4.53) into (4.52) and approximate the integrals with Riemann sums to obtain the near-zone camera pixel values

$$P_\rho^{R2} = \left[ \frac{1}{A/d} \sum_{|nd| < A/2} K_n U(nd) \exp(-ik\Delta_n) \right]^* \left[ \frac{1}{M} \sum_m U(md) \exp(-ik\Delta_m) \right], \tag{4.54}$$

where

$$K_n = \frac{VAD}{R\lambda}\cos^4\psi_{nd} \tag{4.55}$$

and $D = Md$ is the length of the sensor array.

For the beamformer camera, we substitute (3.13) into (4.52) and make a Riemann sum approximation to obtain $P^{\mathrm{BFC}} = K|g|^2$ just as in (3.14), but with a leading constant

$$K = \frac{VA}{R\theta_{\mathrm{BW}}} \frac{1}{A/d} \sum_{|nd|<A/2} \cos^4\psi_{nd}. \tag{4.56}$$

According to our approximation for the beam width $\theta_{\mathrm{BW}}$ in (2.36), $K$ is simply the average of the $K_n$:

$$K = \frac{1}{A/d} \sum_{|nd|<A/2} K_n. \tag{4.57}$$

This averaging is the only effect we have neglected by omitting the light field constants. We can obtain a deeper perspective on camera comparison by using principles from tomography (Appendix B.3.2).

## 4.5   Summary and future challenges

We enable the use of existing incoherent imaging tools for coherent imaging applications, by extending the light field to coherent radiation. We explain how to formulate, capture, and form images from quasi light fields. By synthesizing existing research in optics, quantum physics, and signal processing, we motivate quasi light fields, show how quasi light fields extend the traditional light field, and characterize the properties of different quasi light fields. We explain why capturing quasi light fields directly with intensity measurements is inherently limiting, and demonstrate via simulation how processing scalar field measurements in different ways leads to a rich set of energy localization tradeoffs. We show how coherent image formation using quasi light fields is complicated by an implicit far-zone (Fraunhofer) assumption and the fact that not all quasi light fields are constant along rays. We demonstrate via simulation that a pure light field representation is incapable of modeling near-zone diffraction effects, but that quasi light fields can be augmented with a distance parameter for greater near-zone imaging accuracy. We show how image formation using light fields generalizes the classic beamforming algorithm, allowing for new tradeoffs between resolution and anisotropic sensitivity. Quasi light fields play a role similar to the scalar field, as an abstract representation that describes how sources of radiation interact with the environment, and that can be processed to form images.

Although we have assumed perfectly coherent radiation, tools from partial coherence theory (i) allow us to generalize our results, and (ii) provide an alternative perspective on image formation. First, our results extend to broadband radiation of any state of partial coherence by replacing $U(\mathbf{r}^{\mathrm{R}})U^*(\mathbf{r}^{\mathrm{C}})$ with the cross-spectral density $W(\mathbf{r}^{\mathrm{R}}, \mathbf{r}^{\mathrm{C}}, \nu)$. $W$ provides a statistical description of the radiation, indicating how light at two different positions, $\mathbf{r}^{\mathrm{R}}$ and $\mathbf{r}^{\mathrm{C}}$, is correlated at each frequency $\nu$

(Appendix A.4). Second, $W$ itself may be propagated along rays in an approximate asymptotic sense [60, 61], which forms the basis of an entirely different framework for using rays for image formation, using the cross-spectral density instead of the light field as the core representation.

We present a model of coherent image formation that strikes a balance between utility and comprehensive predictive power. On the one hand, quasi light fields offer more options and tradeoffs than their traditional, incoherent counterpart. In this manner, the connection between quasi light fields and quasi-probability distributions in quantum physics reminds us of the potential benefits of forgoing a single familiar tool in favor of a multitude of useful yet less familiar ones. On the other hand, compared with Maxwell's equations, quasi light fields are less versatile. Therefore, quasi light fields are attractive to engineers who desire more versatility than traditional energy-based methods, yet a more specialized model of image formation than Maxwell's equations.

Unlike our treatment of additive noise in Chapter 2, we have neglected to model sensor measurement error when using arrays to capture quasi light fields. Our emphasis reflects the fact that a noise model was instrumental in deriving the conventional beamformer and its role in image formation, but primarily guides the implementation when forming images from quasi light fields. Noise analysis is discussed in the time-frequency distribution literature [35].

Quasi light fields illustrate the limitations of the simple definition of image formation ubiquitous in incoherent imaging. An image is the visualization of some underlying physical reality, and the energy emitted from a portion of a scene surface towards a virtual aperture is not a physically precise quantity when the radiation is coherent, according to classical electromagnetic wave theory. Perhaps a different image definition may prove more fundamental for coherent imaging, or perhaps a quantum optics viewpoint is required for precision. Although we have borrowed the mathematics from quantum physics, our entire discussion has been classical. Yet if we introduce quantum optics and the particle nature of light, we may unambiguously speak of the probability that a photon emitted from a portion of a scene surface is intercepted by a virtual aperture.

# Chapter 5

# Beamformer performance
# in the presence of phase noise

Having presented a unified theory of image formation in Chapter 4, we now address
the practical challenges in implementing millimeter-wave imaging systems, beginning
with the management of phase noise in radar. Specifically, using quasi light fields
for coherent imaging involves coherently combining scalar field measurements. Scalar
field processing extracts direction information from the phase, just as applying phase
shifts electronically steers the conventional beamformer. The addition of phase noise
therefore threatens our ability to extract direction information and form images.
Phase noise is particularly challenging in millimeter-wave radar, where the phase
shifts required to steer the beamformer correspond to timing delays on the order of a
picosecond at the carrier frequency, leading to aggressive noise margin requirements
on today's technology.

We explore the impact of phase noise on beamformer performance and outline a
methodology for translating application-level requirements to phase noise margins for
the underlying circuitry. For simplicity, we assume a conventional beamformer using
a linear sensor array focused in the far zone, with inter-sensor spacing $d = \lambda/2$. As in
Section 2.2.2, we model the phase noise at the $i^{\text{th}}$ sensor as an ergodic noise process
$\phi_i(t)$ with zero mean. The three key design parameters are the number of sensors $M$,
the phase noise standard deviation $\sigma_\phi$, and the beamforming interval $N$. We base our
performance metrics on the beam pattern magnitude when the beamformer is steered
perpendicular to the array at $\theta_{\text{T}} = \pi/2$:

$$|B_\theta(t)| = \left| \frac{1}{M} \mathbf{a}^*(\pi/2) \Phi \mathbf{a}(\theta) \right|. \tag{5.1}$$

Specifically, we consider the sidelobe suppression, tilt, and beam width of both the
expected beam pattern and, where applicable, the sample paths of the instantaneous
beam pattern. We estimate the expected beam pattern using Monte Carlo simulations
(Appendix C.2).

Our strategy is to explore the relationship between the design parameters and
performance metrics through simulations, simplify the results with analytic approxi-

mations, and support the conclusions with intuition. To simplify the tradeoff space, we separately analyze how to size the array relative to the phase noise and how many instantaneous beam patterns $|B_\theta(t)|$ to average over. The array size determines the steady state behavior of the beamformer in the presence of phase noise, which we coarsely classify into four different regions of operation. No amount of time averaging can impact the beamformer's steady state behavior, which we examine by first studying fixed array sizes of $M = 10$, 100, and 1000, and then generalizing for arbitrary size. After we size the array to achieve the desired region of operation, we analyze the second-order effects on performance due to time averaging. Depending on the application requirements, we may need to increase the array size to compensate for the lack of a sufficient number of instantaneous beam patterns to average. Therefore, although the array size, phase noise, and beamforming interval interact with each other to determine beamformer performance, it is conceptually simpler to postpone a discussion of time averaging and first study the steady state behavior.

We begin by describing the four stages of beamformer performance degradation that define four regions of operation (Section 5.1). There are certain noise thresholds, or breakpoints, at which the beamformer transitions to the next region. We define and characterize the breakpoints while analyzing the steady state behavior of sidelobe suppression, tilt, and beam width in the various regions of operation (Section 5.2). We then analyze how much time averaging is required to guarantee acceptable interference suppression and resolution within the stable region of operation (Section 5.3).

## 5.1   Regions of operation for different noise levels

We classify beamformer behavior into four regions of operation, based on both $M$ and $\sigma_\phi$ (Figure 5-1). For fixed array size $M$, the beamformer degrades over the four regions as the phase noise $\sigma_\phi$ increases from 0 to $\infty$ (Section 5.1.1). Larger values of $M$ push the points of transition between regions, called breakpoints, further out to higher noise levels. Beamformers thereby become more robust to phase noise as the number of sensors increases (Section 5.1.2).

### 5.1.1   Macroscopic beam pattern behavior

We fix the array size $M$ and describe the beamformer's behavior as $\sigma_\phi$ increases from 0. For small values of $\sigma_\phi$ below the sidelobe breakpoint $\sigma_\phi^{\mathrm{SB}}$, the expected beam pattern retains its shape even though the instantaneous beam pattern can be quite noisy, particularly around the sidelobes. In this stable region of operation, enough time averaging will virtually eliminate the effect of phase noise on performance. We therefore desire to operate in the stable region, below the sidelobe breakpoint, for applications that cannot tolerate any detrimental performance impact due to phase noise.

As the phase noise increases and surpasses the sidelobe breakpoint $\sigma_\phi^{\mathrm{SB}}$, the sidelobes begin to rise, resulting in greater interference from sources outside the steering direction. No amount of time averaging prevents the sidelobes from rising. How-
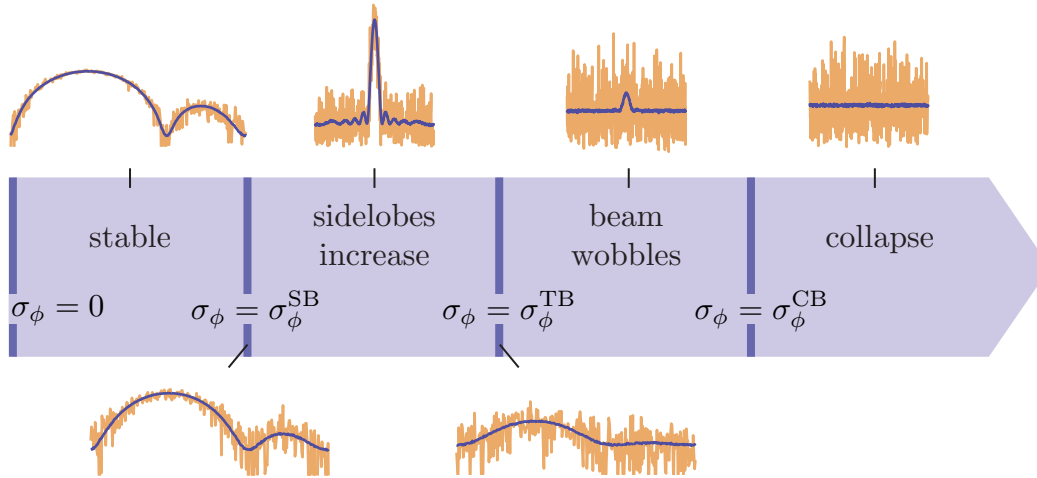
Figure 5-1: As phase noise increases, beamformer performance degrades over four different regions of operation. We illustrate the expected beam pattern along with a noisy sample path of the instantaneous beam pattern for various values of $\sigma_\phi$. At the sidelobe breakpoint $\sigma_\phi^{\mathrm{SB}}$, the sidelobes of the beam pattern begin to increase. At the tilt breakpoint $\sigma_\phi^{\mathrm{TB}}$, the sidelobes begin to compete with the main lobe, causing the beam to wobble. At the collapse breakpoint $\sigma_\phi^{\mathrm{CB}}$, the beam collapses.

ever, if the phase noise remains below the tilt breakpoint $\sigma_\phi^{\mathrm{TB}}$, the sidelobes do not compete with the main lobe. The maxima of the instantaneous beam patterns are clustered about the main lobe with high probability, so that the tilt exhibits a Gaussian distribution. We therefore desire to operate above the sidelobe breakpoint but below the tilt breakpoint for applications that do not have aggressive requirements for interference suppression.

As the phase noise increases further and surpasses the tilt breakpoint $\sigma_\phi^{\mathrm{TB}}$, the sidelobes become so large that they compete with the main lobe, which becomes difficult to identify in the instantaneous beam patterns. The beam wobbles about the intended steering direction and can point in an arbitrary direction at any moment in time, so that the tilt no longer exhibits a Gaussian distribution. If the phase noise remains below the collapse breakpoint $\sigma_\phi^{\mathrm{CB}}$, enough time averaging can still localize the beam so that the main lobe is evident, but the expected beam pattern otherwise degrades. We therefore desire to operate above the tilt breakpoint but below the collapse breakpoint for applications without aggressive performance requirements where large sensor arrays are prohibitively expensive.

When the phase noise surpasses the collapse breakpoint $\sigma_\phi^{\mathrm{CB}}$, the beam collapses to a random uniform distribution and loses all directional selectivity. Time averaging results in a blank, uniform image. It is useless to operate above the collapse breakpoint.

Larger arrays are more robust to phase noise. For example, when $M = 100$ and $\sigma_\phi = 0.2\pi$, corresponding to 1 ps at 100 GHz, the phase noise is below the sidelobe breakpoint and the expected beam pattern is stable. The sidelobes would rise for

the same array with twice the phase noise, when $\sigma_\phi = 0.4\pi$. However, a larger array with $M = 1000$ would have a stable beam pattern for phase noise up to $\sigma_\phi = 0.6\pi$, corresponding to 3 ps at 100 GHz.

## 5.1.2   Robustness of large arrays

We employ phasors to understand why a beamformer using a large sensor array is robust to phase noise. The beam pattern is the sum of $M$ complex numbers, each with a magnitude of $1/M$:

$$B_\theta(t) = \frac{1}{M}\mathbf{a}^*(\theta_\mathrm{T})\Phi\mathbf{a}(\theta). \tag{5.2}$$

As $M$ increases, we add additional sensors to grow the array, with the distance between adjacent sensors held constant and the beamformer output power normalized by $1/M$. We interpret beamforming as adding the corresponding $M$ phasors together, positioned head-to-tail, forming a random walk (Figure 5-2). There are $M$ steps of size $1/M$, so that $M$ determines the granularity of the walk. The steering vector $\mathbf{a}(\theta)$ is a sequence of phasors with linearly increasing phase, per (2.6). Adding the phasors in $\mathbf{a}(\theta)$ together forms a segmented arc with a curvature determined by $\theta$. The beamformer applies $\mathbf{a}^*(\theta_\mathrm{T})$ to unwrap the arc according to $\theta_\mathrm{T}$, so that when $\theta_\mathrm{T} = \theta$ the arc is completely flattened into a straight line. The phase noise $\Phi$ perturbs each segment of the arc, randomly rotating each phasor with a standard deviation of $\sigma_\phi$. The distance from the origin to the tip of the last phasor indicates the magnitude of the beam pattern. When phase noise is present, the arc is generally not regularly curved, so that steering cannot straighten the arc into a line to maximize the beam pattern. For low values of $M$, phase noise can dramatically change the curvature of the arc, while for high values of $M$, the granularity of the random walk is higher so that the same level of phase noise has less impact on the curvature. The curvature, and hence the beam pattern, is therefore better preserved when $M$ is large.

We employ vector spaces to understand how the impact of phase noise on beamformer steering error diminishes with array size. While the phase noise for an array with $M$ sensors resides in an $M$-dimensional space, the application of phase shifts to steer the array traverses a 1-dimensional subspace. Specifically, the steering subspace is generated by a sequence of $M$ regularly ascending phase shifts

$$\mathbf{u} = [1\,2\,\cdots\,M]^\mathrm{t}. \tag{5.3}$$

There is also a 1-dimensional invariant subspace given by

$$\mathbf{v} = [1\,1\,\cdots\,1]^\mathrm{t}, \tag{5.4}$$

as adding the same phase at every sensor does not change the beam pattern magnitude. For two sensors, $\mathbf{u}$ and $\mathbf{v}$ span the entire space, so that phase noise can always be interpreted as a change in steering, consistent with (2.40). For arbitrary $M$, we
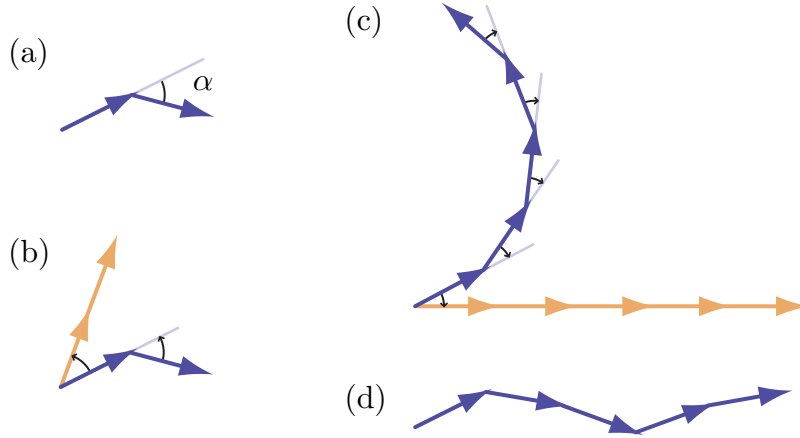
Figure 5-2: We interpret beamforming as unwrapping a segmented arc composed of phasors. The magnitude of the unwrapped arc corresponds to the magnitude of the beam pattern. The impact of phase noise is to randomly rotate each phasor. For two sensors (a), the addition of phase noise can always be interpreted as a change in steering direction (b). For five sensors, the phasors can generally only be perfectly straightened in the absence of phase noise (c), but the overall length is not substantially reduced even when significant phase noise is present (d).

can project the phase noise vector

$$\mathbf{n} = [\phi_1 \ \phi_2 \ \cdots \ \phi_M]^{\mathrm{t}} \tag{5.5}$$

onto $\mathbf{u}$ to obtain

$$\frac{\mathbf{n}^{\mathrm{t}}\mathbf{u}}{\mathbf{u}^{\mathrm{t}}\mathbf{u}}\mathbf{u}. \tag{5.6}$$

In (5.6), $\mathbf{u}$ is scaled by a random variable with zero mean and a variance that decays as $1/M^3$. Thus the impact of phase noise on steering diminishes as $M$ increases.

## 5.2 The steady state operating point

The size of the sensor array, relative to the phase noise, determines the steady state operating point. We analyze the behavior of the sidelobe suppression, tilt, and beam width in the presence of phase noise, and thereby determine the breakpoints at the boundaries of the regions of operation. We first describe how the sidelobes are stable for low phase noise but then grow linearly above the sidelobe breakpoint (Section 5.2.1). Next, we characterize the probability distribution of the tilt, which is Gaussian below the tilt breakpoint and uniform above the collapse breakpoint (Section 5.2.2). Then, we determine the robustness of the beam width to phase noise (Section 5.2.3). We postpone a discussion of the role of time averaging until Section 5.3.

## 5.2.1   Sidelobe suppression in presence of phase noise

Our simulations indicate that the sidelobes of the expected beam pattern are stable for low levels of phase noise, but then grow linearly until the sidelobe suppression is approximately $-6.2$ dB. We define the level of phase noise at which the sidelobes begin to grow as the sidelobe breakpoint. We estimate the sidelobe breakpoint and determine that it increases logarithmically in the array size.

**Sidelobes grow above sidelobe breakpoint**

To first order, the sidelobe suppression $S^{\mathrm{SL}}$ of the expected beam pattern is a piecewise linear function of the phase noise $\sigma_\phi$ (Figure 5-3). In the absence of phase noise, $S^{\mathrm{SL}}$ is a function of the array size $M$ that approaches $-13.5$ dB as $M \to \infty$ (Section 2.2.1). As $\sigma_\phi$ increases from zero, $S^{\mathrm{SL}}$ remains constant until $\sigma_\phi$ approaches the sidelobe breakpoint. The phase noise is small enough that the perturbations in the phasor sum in Section 5.1.2 are approximately linear, and therefore absent in the expected beam pattern used to compute $S^{\mathrm{SL}}$. Once the noise exceeds the sidelobe breakpoint, the phasor addition introduces a bias and the sidelobes grow linearly (in dB) with $\sigma_\phi$. The sidelobe breakpoint increases with $M$, so that the beam pattern remains stable in the presence of greater phase noise when using larger arrays. The slope of $S^{\mathrm{SL}}$ versus $\sigma_\phi$ also increases with array size, from 8.7 to 11.7 dB per radian as $M$ varies from 10 to 1000.
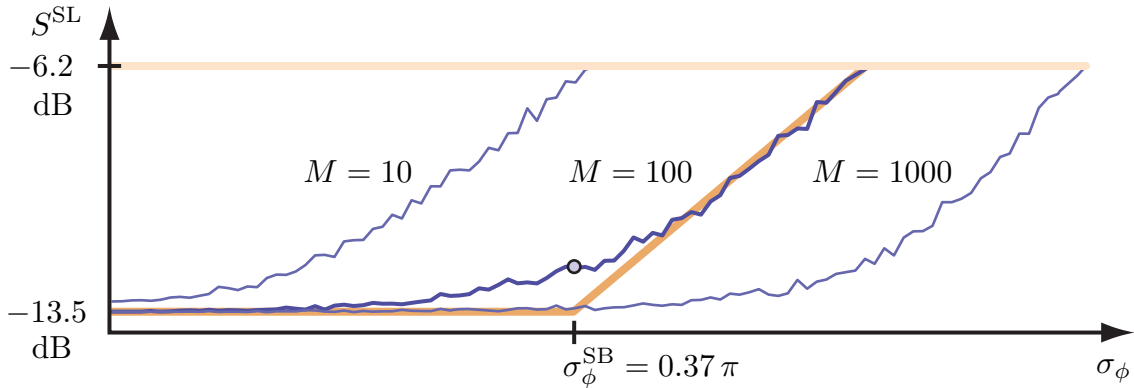


Figure 5-3: Sidelobe suppression is approximately a piecewise linear function of phase noise $\sigma_\phi$. The sidelobe suppression of the expected beam pattern remains constant until the noise approaches the sidelobe breakpoint $\sigma_\phi^{\mathrm{SB}}$, at which point the sidelobes grow approximately linearly with $\sigma_\phi$. The sidelobe breakpoint increases with the array size $M$, so that the beam pattern remains stable in the presence of greater phase noise when using larger arrays.

**Estimating the sidelobe breakpoint**

To estimate $\sigma_\phi^{\mathrm{SB}}$ for a particular array size $M$, we estimate the sidelobe suppression $S^{\mathrm{SL}}$ of the expected beam pattern for each $\sigma_\phi$ using a Monte Carlo simulation (Appendix C.2.2), and then fit a piecewise linear function to the resulting data. We then

estimate $\sigma_\phi^{\mathrm{SB}}$ to be the point at which the piecewise linear function changes slope. Empirically, for large values of $M$, our estimate of the sidelobe breakpoint occurs when $S^{\mathrm{SL}} \approx -12$ dB. When fitting the piecewise linear function we must take care to avoid excessively large values of $\sigma_\phi$, since as $\sigma_\phi \to \infty$ the sidelobe suppression $S^{\mathrm{SL}} \to 0$ dB, and so $S^{\mathrm{SL}}$ will eventually no longer grow linearly with $\sigma_\phi$ as the beam begins to wobble. However, $S^{\mathrm{SL}}$ grows linearly with $\sigma_\phi$ as long as the sidelobes do not compete with the main beam, which empirically holds when $S^{\mathrm{SL}} < -6.2$ dB. We therefore confine the data to $S^{\mathrm{SL}} < -6.2$ dB when fitting the piecewise linear function to estimate the sidelobe breakpoint.
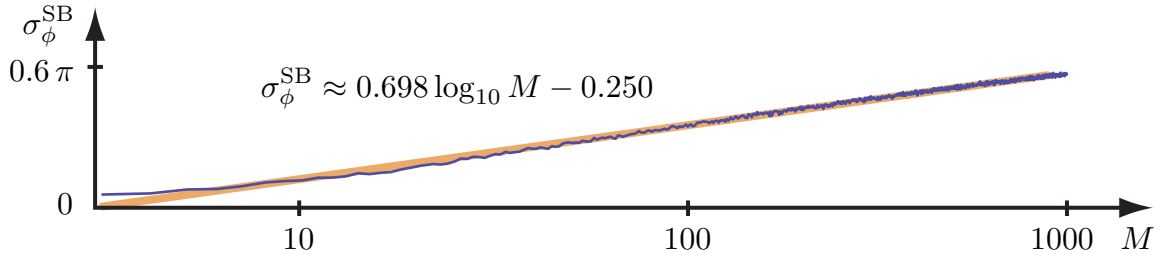


Figure 5-4: The sidelobe breakpoint grows approximately logarithmically in array size $M$, so that we may approximate the relationship with a line on a semi-log plot.

The sidelobe breakpoint grows logarithmically in array size $M$, especially for larger values of $M$ (Figure 5-4). We obtain an analytic approximation by performing a least-squares linear fit of $\sigma_\phi^{\mathrm{SB}}$ to the logarithm of $M$:

$$\sigma_\phi^{\mathrm{SB}} \approx 0.698 \log_{10} M - 0.250. \tag{5.7}$$

Given a model for phase noise, we can use (5.7) to size an array to guarantee that we operate in the stable operating region, so that the sidelobe suppression of the expected beam pattern is not significantly impacted by phase noise. Similarly, given an array of a certain size, we can use (5.7) to place requirements on the phase noise to guarantee stable operation.

## 5.2.2 Tilt in presence of phase noise

Our simulations indicate that the tilt exhibits a Gaussian distribution for low levels of phase noise and a uniform distribution for high levels of phase noise. We define the level of phase noise at which the tilt distribution ceases to be Gaussian as the tilt breakpoint, and we define the level of phase noise at which the tilt distribution becomes uniform as the collapse breakpoint. In between the tilt and collapse breakpoints, the tilt exhibits a mixed Gaussian/uniform distribution. We estimate the tilt breakpoint and determine that it increases logarithmically in the array size. We also provide intuition for how phase noise impacts tilt through the beam width.

**Tilt is Gaussian below tilt breakpoint**

As the phase noise increases so that $\sigma_\phi$ approaches the tilt breakpoint $\sigma_\phi^{\mathrm{TB}}$ and $S^{\mathrm{SL}}$ approaches $-6.2$ dB, the sidelobes begin to compete with the main lobe and the instantaneous beam pattern wobbles. We analyze this behavior and estimate the tilt breakpoint through the tilt random variable $\tau$ (Section 2.2.1). For fixed array size $M$, the probability distribution of $\tau$ changes from Gaussian to uniform as the phase noise increases (Figure 5-5).
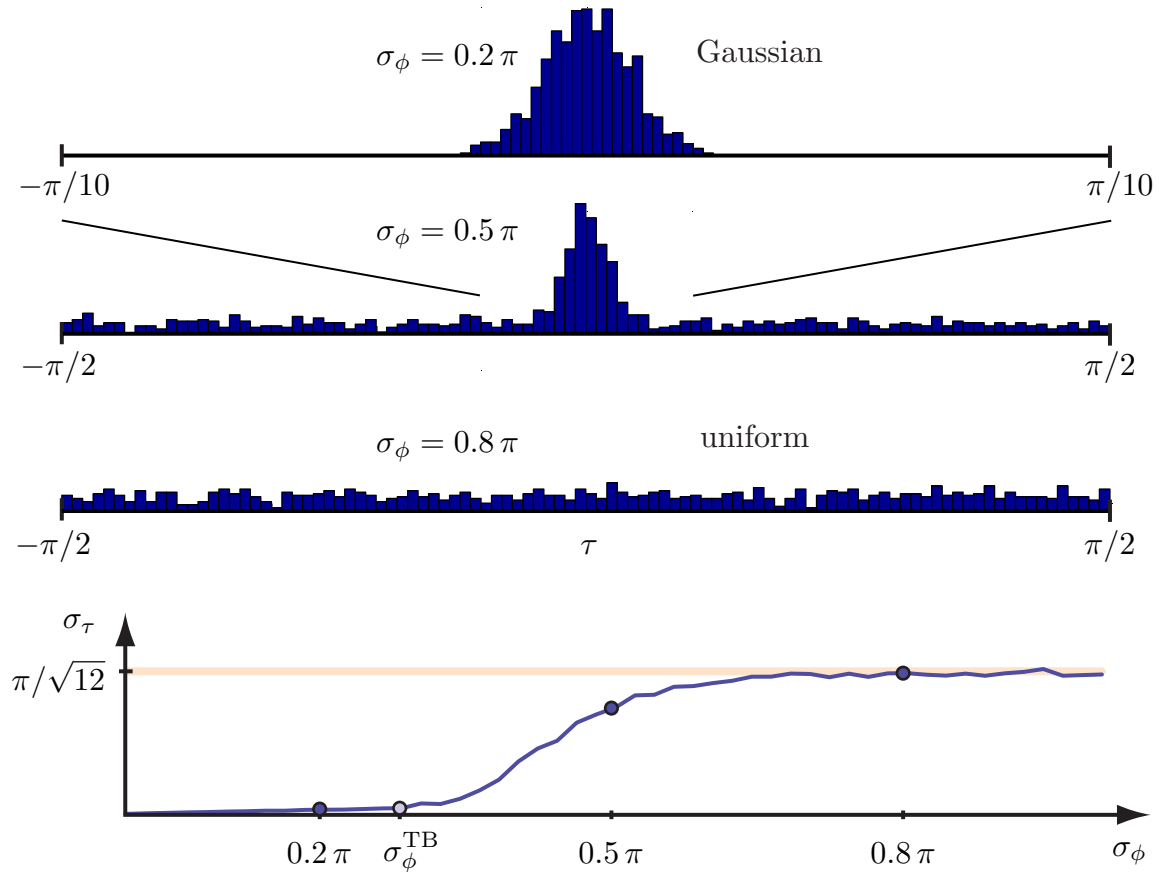


Figure 5-5: The distribution of beam pattern tilt $\tau$ changes from Gaussian to uniform as the phase noise $\sigma_\phi$ increases. When the phase noise is below the tilt breakpoint $\sigma_\phi^{\mathrm{TB}}$, the beam pattern maximum is locally concentrated about the main lobe. As $\sigma_\phi$ increases past $\sigma_\phi^{\mathrm{TB}}$, the sidelobes grow so that the maximum can occur elsewhere, resulting in a mixed Gaussian/uniform distribution. With enough phase noise, the beam collapses and the maximum is uniformly distributed over the entire steering range. The progression of the tilt standard deviation $\sigma_\tau$ mirrors the evolution of the distribution of $\tau$.

We define the tilt breakpoint $\sigma_\phi^{\mathrm{TB}}$ as the smallest value of $\sigma_\phi$ at which the distribution of $\tau$ is no longer Gaussian. Intuitively, the Gaussian distribution arises because for low noise levels, only points on the main lobe can compete for the maximum value of the beam pattern. Once $\sigma_\phi$ reaches the tilt breakpoint, the sidelobes compete with the main lobe and the beam pattern maximum can be achieved far away from the

main lobe. When there is enough phase noise to collapse the beam pattern, the maximum is equally likely to be achieved at any direction, so that the tilt has a uniform distribution. The relationship between the standard deviation of the tilt $\sigma_\tau$ and $\sigma_\phi$ mirrors the progression of the tilt distribution. Below the tilt breakpoint, $\sigma_\tau$ increases linearly with $\sigma_\phi$ as local competition for the maximum about the main lobe increases, and the Gaussian distribution widens. Above the tilt breakpoint, $\sigma_\tau$ increases at a significantly faster rate as the maximum can occur far away from the main lobe, so that the Gaussian distribution is mixed with a uniform distribution. Ultimately, $\sigma_\tau$ flattens out at the standard deviation $\pi/\sqrt{12}$ of a uniform random variable between $-\pi/2$ and $\pi/2$.

## Estimating the tilt breakpoint

To estimate the tilt breakpoint $\sigma_\phi^{\mathrm{TB}}$, we apply the Shapiro-Wilk hypothesis test [62] to determine the smallest value of $\sigma_\phi$ at which the distribution of $\tau$ is no longer Gaussian. Specifically, we run the Shapiro-Wilk test on samples of $\tau$ generated by a Monte Carlo simulation (Appendix C.2.1). We begin searching for the tilt breakpoint at a small value of $\sigma_\phi$ and then slowly increment $\sigma_\phi$. We reject the null hypothesis that $\tau$ has a Gaussian distribution at significance level 0.01. Once we obtain three such rejections in a row, we declare the corresponding value of $\sigma_\phi$ to be our estimate of the tilt breakpoint.
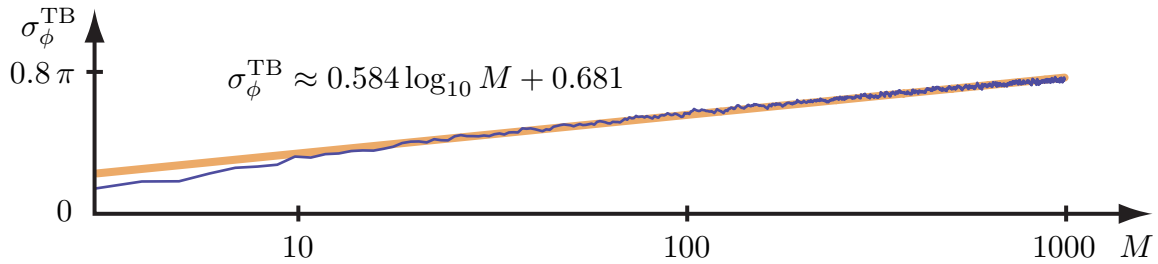


Figure 5-6: The tilt breakpoint grows approximately logarithmically in array size $M$, so that we may approximate the relationship with a line on a semi-log plot.

The tilt breakpoint grows logarithmically in the array size $M$, especially for larger values of $M$ (Figure 5-6). We obtain an analytic approximation for tilt breakpoint growth by performing a least-squares linear fit of $\sigma_\phi^{\mathrm{TB}}$ to the logarithm of $M$:

$$\sigma_\phi^{\mathrm{TB}} \approx 0.584 \log_{10} M + 0.681. \tag{5.8}$$

Given a model for the phase noise, we can use (5.8) to size an array to guarantee that we operate where the sidelobes do not interfere with the steering of the main beam.

**Intuition for tilt behavior**

We obtain a simple model of the effect of phase noise on tilt below the tilt breakpoint by analyzing the slope of $\sigma_\tau$. First, we estimate $\sigma_\tau$ for values of $\sigma_\phi < \sigma_\phi^{\mathrm{TB}}$ and for fixed $M$. We perform a least-squares linear fit to estimate the slope $\mathrm{d}\sigma_\tau/\mathrm{d}\sigma_\phi$, and repeat the process for each desired $M$. Then, we obtain an analytic approximation of the slope $\mathrm{d}\sigma_\tau/\mathrm{d}\sigma_\phi$ as a function of $M$ by performing a least-squares linear fit of the logarithm of $\mathrm{d}\sigma_\tau/\mathrm{d}\sigma_\phi$ to the logarithm of $M$ (Figure 5-7):

$$\log_{10} \frac{\mathrm{d}\sigma_\tau}{\mathrm{d}\sigma_\phi} \approx -1.216 \log_{10} M - 0.255. \tag{5.9}$$

We thereby obtain an expression for the standard deviation of the tilt in terms of both the phase noise $\sigma_\phi$ and the array size $M$:

$$\sigma_\tau \approx 0.556 \frac{\sigma_\phi}{M^{1.216}}. \tag{5.10}$$
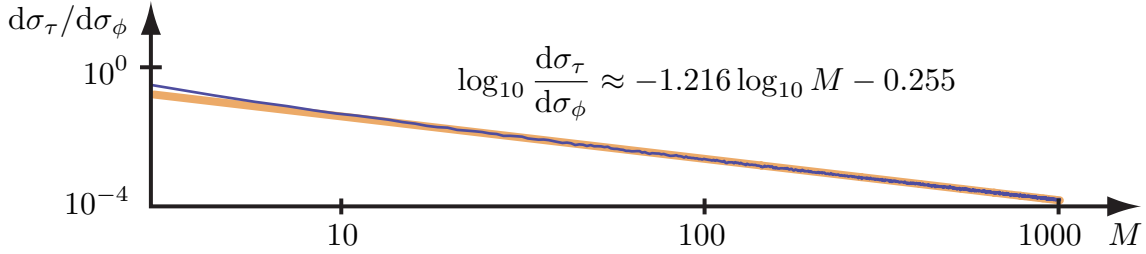


Figure 5-7: The tilt standard deviation slope decays with array size $M$, so that we may approximate the relationship with a line on a log-log plot.

We interpret the impact of phase noise on tilt by noting that $\sigma_\tau$ is approximately inversely proportional to $M$ in (5.10), and that the angular beam width $\theta_{\mathrm{BW}}$ is also inversely proportional to $M$ in (2.36). Therefore, we relate the tilt and phase noise by

$$\sigma_\tau \sim \sigma_\phi \, \theta_{\mathrm{BW}}. \tag{5.11}$$

Thus, for phase noise below the tilt breakpoint, the beam pattern maximum is only likely to be attained on the main lobe, and the array size impacts tilt by determining the width of that lobe.

## 5.2.3   Beam width in presence of phase noise

Phase noise does not substantially impact the beam width of the expected beam pattern when operating below the tilt breakpoint. We compare the half-power beam width when the phase noise is just below the tilt breakpoint at $\sigma_\phi = 0.9\sigma_\phi^{\mathrm{TB}}$ to the half-power beam width in the absence of phase noise. Our Monte Carlo simulations

(Appendix C.2.3) indicate that the percentage change

$$\frac{\left| \theta_{\mathrm{BW}}^{0.9\sigma_\phi^{\mathrm{TB}}} - \theta_{\mathrm{BW}}^0 \right|}{\theta_{\mathrm{BW}}^0} \tag{5.12}$$

is bound below 6% for all array sizes $M \leq 1000$. Therefore, the beam width of the expected beam pattern is essentially preserved below the tilt breakpoint.

## 5.3 The role of time averaging

Our analysis above allows us to size the sensor array relative to the phase noise to achieve the steady state behavior in the desired region of operation. We now assume that we are operating in the stable region, below the sidelobe breakpoint. The sidelobe suppression and beam width of the expected beam pattern are the same as if there were no phase noise. If the beamforming interval $N$ is large enough so that the time-averaged beam pattern converges to the expected beam pattern, then we can effectively average out the phase noise and expect as good performance as if the noise were absent. However, many applications restrict $N$. For example, in imaging applications, $N$ must be low enough so that the scene is stationary over the beamforming interval relative to the desired resolution, and a still lower $N$ enables a higher frame rate. To first order, deviations from noise-free behavior decay as $1/\sqrt{N}$ in the number $N$ of independent sensor measurements processed. We make the dependence on $N$ precise by defining appropriate metrics for sidelobe interference suppression (Section 5.3.1) and image resolution (Section 5.3.2).

### 5.3.1 Averaging to suppress sidelobe interference

We desire another metric for interference suppression besides sidelobe suppression, since $S^{\mathrm{SL}}$ of the expected beam pattern is only indicative of actual performance for large $N$, and the sidelobes of a noisy instantaneous beam pattern are not clearly identifiable. We therefore augment the sidelobe suppression metric to incorporate the variance of the beam pattern at the main lobe and sidelobe peaks. We define the sidelobe distortion $S^{\mathrm{D}}$, which we subsequently relate to the phase noise $\sigma_\phi$, array size $M$, and beamforming interval $N$.

To define the sidelobe distortion, we first express the familiar sidelobe suppression in convenient notation. In the absence of noise, the beam pattern achieves its maximum at $\theta = \pi/2$ and has a first sidelobe peak at $\theta = \theta_S$. The sidelobe suppression is

$$S^{\mathrm{SL}} = 20 \log_{10} \left( \frac{E\left[|B_{\theta_S}|\right]}{E\left[|B_{\pi/2}|\right]} \right). \tag{5.13}$$

We denote the expected values of the main and sidelobe peaks of the beam pattern by $\mu_M = E[|B_{\pi/2}|]$ and $\mu_S = E[|B_{\theta_S}|]$, and we denote the standard deviations of the main and sidelobe peaks of the time-averaged beam pattern by $\sigma_M$ and $\sigma_S$. With this

notation, the sidelobe suppression is

$$S^{\mathrm{SL}} = 20 \log_{10} \left( \frac{\mu_S}{\mu_M} \right). \tag{5.14}$$

In contrast, we define the sidelobe distortion as

$$
\begin{aligned}
S^{\mathrm{D}} &= 20 \log_{10} \left( \frac{\mu_S + \sigma_S}{\mu_M - \sigma_M} \right) - S^{\mathrm{SL}} \\
&= 20 \log_{10} \left( \frac{1 + \sigma_S/\mu_S}{1 - \sigma_M/\mu_M} \right).
\end{aligned}
\tag{5.15}
$$

Intuitively, $S^{\mathrm{D}}$ is an indication of how much worse sidelobe suppression can become if our estimates of the beam pattern at the main lobe and sidelobe peaks are off by one standard deviation. Note that as $N \to \infty$, both $\sigma_M \to 0$ and $\sigma_S \to 0$, so that $S^{\mathrm{D}} \to 0$. The sidelobe distortion is approximately linear in $\sigma_\phi$ when operating below the sidelobe breakpoint (Figure 5-8). We obtain an analytic approximation of the slope $\mathrm{d}S^{\mathrm{D}}/\mathrm{d}\sigma_\phi$ as a function of $M$ by performing a least-squares linear fit of the logarithm of the slope to the logarithm of $M$:

$$\log_{10} \frac{\mathrm{d}S^{\mathrm{D}}}{\mathrm{d}\sigma_\phi} \approx -0.235 \log_{10} M + 1.018. \tag{5.16}$$

We thereby obtain an expression for the sidelobe distortion in terms of both the phase noise $\sigma_\phi$ and the array size $M$:

$$S^{\mathrm{D}} \approx 10.43 \frac{\sigma_\phi}{M^{0.235}}. \tag{5.17}$$

We now relate $S^{\mathrm{D}}$ to the beamforming interval $N$ to understand the impact of time averaging. We make several approximations that are valid when $\sigma_\phi$ is small. First, we note that for low phase noise, most of the uncertainty in sidelobe suppression comes from the variability of the sidelobe peak, rather than the main lobe peak. We therefore assume that $\sigma_M \ll \mu_M$ in (5.15) so that

$$S^{\mathrm{D}} \approx 20 \log_{10} \left( 1 + \sigma_S/\mu_S \right). \tag{5.18}$$

Next, we make the first-order Taylor series approximation $\log(1 + x) \approx x$ to obtain

$$S^{\mathrm{D}} \approx \frac{20}{\log 10} \frac{\sigma_S}{\mu_S}. \tag{5.19}$$

Finally, we note that $\sigma_S$ decays by $1/\sqrt{N}$ when we average $N$ independent beam patterns, while $\mu_S$ remains constant. Therefore, the sidelobe distortion also decays by $1/\sqrt{N}$. Together with (5.17), we relate sidelobe distortion to the design parameters

Figure 5-8: The sidelobe distortion slope decays with array size $M$, so that we may approximate the relationship with a line on a log-log plot.

by

$$S^{\mathrm{D}} \sim \frac{\sigma_\phi}{M^{0.235}\sqrt{N}}. \tag{5.20}$$

### 5.3.2 Averaging for better resolution

We desire another metric for resolution besides the half-power beam width, since $\theta_{\mathrm{BW}}$ of the expected beam pattern is only indicative of actual performance for large $N$, and the width of the main lobe of a noisy instantaneous beam pattern can be difficult to identify. We therefore augment the beam width metric to incorporate the variance of the beam pattern at every direction $\theta$. We define the beam width distortion $\theta^{\mathrm{D}}$, which we relate to the phase noise $\sigma_\phi$, array size $M$, and beamforming interval $N$.
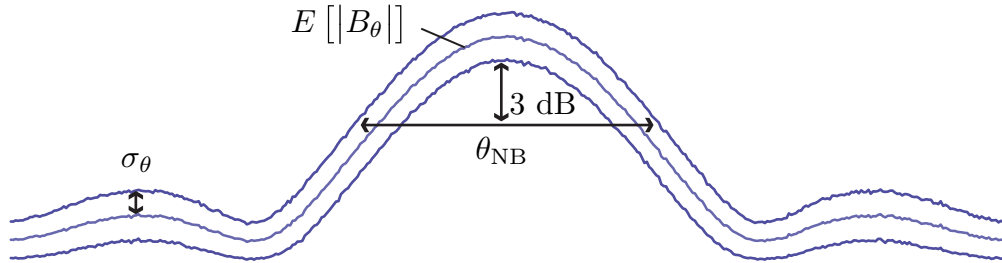


Figure 5-9: The noisy beam width corresponds to making estimation errors on the order of one standard deviation, and is defined as the main lobe of $E[|B_\theta|] + \sigma_\theta$ that is greater than $1/\sqrt{2}$ times the peak of $E[|B_\theta|] - \sigma_\theta$.

To define the beam width distortion, we offset the beam pattern magnitude $|B_\theta|$ by the standard deviation $\sigma_\theta$ of the time-averaged beam pattern at each angle $\theta$

(Figure 5-9). Recall that the half-power beam width $\theta_{\mathrm{BW}}$ is the width of the main lobe of $E[\|B_\theta\|]$ that is greater than $1/\sqrt{2}$ times the peak of $E[\|B_\theta\|]$. In contrast, we define the noisy beam width $\theta_{\mathrm{NB}}$ as the width of the main lobe of $E[\|B_\theta\|] + \sigma_\theta$ that is greater than $1/\sqrt{2}$ times the peak of $E[\|B_\theta\|] - \sigma_\theta$. The noisy beam width corresponds to making estimation errors of one standard deviation in one direction at the peak and one standard deviation in the opposite direction everywhere else.

We define the beam width distortion as the percentage change in the beam width due to phase noise:

$$\theta^{\mathrm{D}} = \frac{|\theta_{\mathrm{NB}} - \theta_{\mathrm{BW}}|}{\theta_{\mathrm{BW}}}. \tag{5.21}$$

Note that as $N \to \infty$, we have $\sigma_\theta \to 0$ so that $\theta^{\mathrm{D}} \to 0$. The beam width distortion is approximately linear in $\sigma_\phi$ when operating below the sidelobe breakpoint (Figure 5-10). We obtain an analytic approximation of the slope $\mathrm{d}\theta^{\mathrm{D}}/\mathrm{d}\sigma_\phi$ as a function of $M$ by performing a least-squares linear fit of the logarithm of the slope to the logarithm of $M$:

$$\log_{10} \frac{\mathrm{d}\theta^{\mathrm{D}}}{\mathrm{d}\sigma_\phi} \approx -0.309 \log_{10} M + 0.051. \tag{5.22}$$

We thereby obtain an expression for the beam width distortion in terms of both the phase noise $\sigma_\phi$ and the array size $M$:

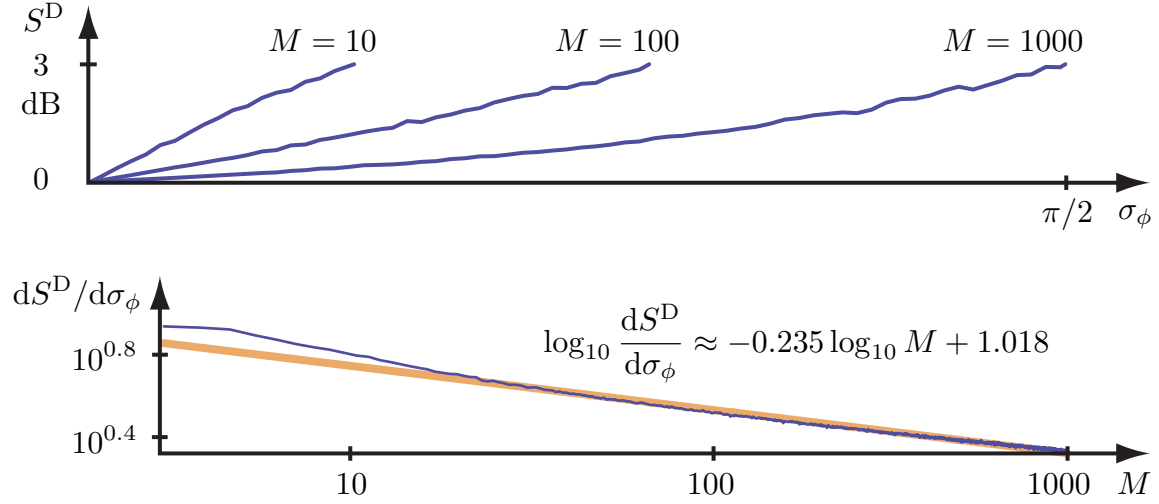$$\theta^{\mathrm{D}} \approx 1.123 \frac{\sigma_\phi}{M^{0.309}}. \tag{5.23}$$
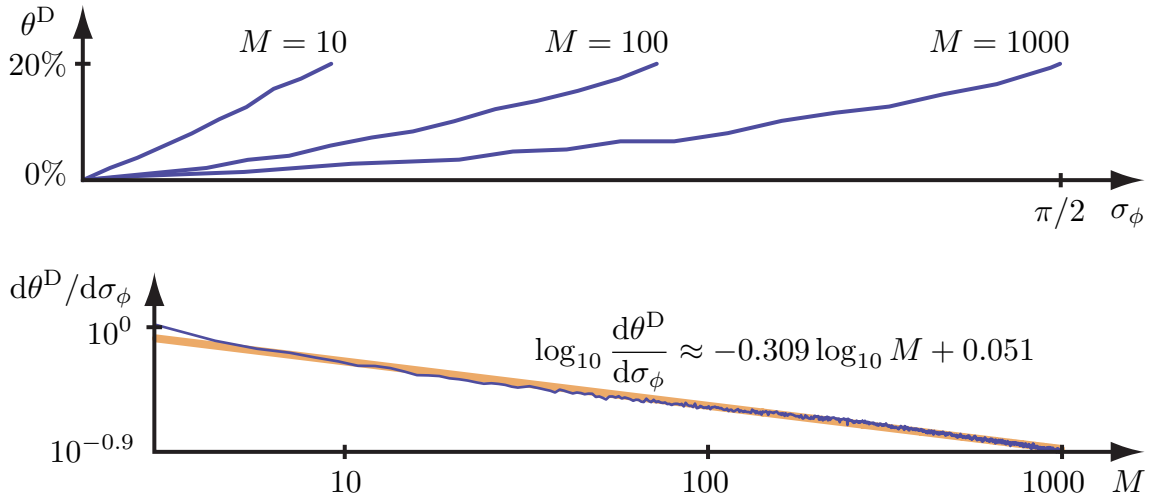


Figure 5-10: The beam width distortion slope decays with array size $M$, so that we may approximate the relationship with a line on a log-log plot.

We now relate $\theta^{\mathrm{D}}$ to the beamforming interval $N$ to understand the impact of time averaging. The standard deviation $\sigma_\theta$ decays as $1/\sqrt{N}$ at each angle $\theta$. Thus the vertical spread between the shifted beam patterns $E[\|B_\theta\|] - \sigma_\theta$ and $E[\|B_\theta\|] + \sigma_\theta$ decays as $1\sqrt{N}$. We assume that the slope of the expected beam pattern is approximately

constant over the region of variation where the endpoints of the noisy beam width are defined, so that $|\theta_{\mathrm{NB}} - \theta_{\mathrm{BW}}|$ decays as $1/\sqrt{N}$, as well. Together with (5.23), we relate beam width distortion to the design parameters by

$$\theta^{\mathrm{D}} \approx 1.123 \frac{\sigma_\phi}{M^{0.309}\sqrt{N}}. \tag{5.24}$$

## 5.4 Summary and future challenges

We have characterized conventional beamformer performance in terms of the number of sensors $M$, the phase noise standard deviation $\sigma_\phi$, and the beamforming interval $N$. As the phase noise increases, performance degrades over four stages, during which the sidelobes rise until they compete with the main lobe, causing the beam to wobble and ultimately collapse. We ideally operate the beamformer in the first stage, below the sidelobe breakpoint, where the expected beam pattern is stable. We then need only ensure that $M$ is large enough so that $\sigma_\phi$ lies below $\sigma_\phi^{\mathrm{SB}}$, and that $N$ is large enough so that time averaging keeps the sidelobes sufficiently attenuated and the beam sufficiently narrow. The tilt standard deviation is related to the product $\sigma_\phi \theta_{\mathrm{BW}}$, which is typically much smaller than the beam width and therefore not of concern.

Our analysis guides design parameter selection and provides an intuitive understanding of the tradeoffs inherent in real coherent imaging applications, even though we have focused on the specific case of a conventional far-zone beamformer for a uniform array. It is straightforward to adapt our Monte Carlo simulations (Appendix C.2) to more accurately model other configurations.

# Chapter 6

# Digital phase tightening for reducing phase noise

Coherent image formation requires precise control over the phase of each signal measured at the sensors. However, the circuitry used to process the signals introduces phase noise, which can degrade image quality by decreasing the resolution and increasing the susceptibility to interference at each pixel, as seen in Chapter 5. Phase noise is of particular concern in millimeter-wave radar, where timing jitter on the order of a picosecond can negatively impact performance. We propose a low-complexity technique, called phase tightening, to reduce the phase noise in each received signal prior to image formation. Phase tightening is a general technique that oversamples the data to reduce noise, and is not limited to coherent imaging.

We distinguish between three different types of phase noise. First, the introduction of a frequency offset causes the phase to drift away from its true value. Compensating for frequency offsets has been extensively studied, and a phase-lock loop of sufficiently high order can eliminate the resulting phase drift [63]. Second, the introduction of a constant phase offset skews the phase from its true value by a fixed amount. Compensating for fixed phase offsets is a relatively simple problem, because phase offsets can be estimated and corrected for during offline calibration, where the environment is under control and real-time performance is not required. Third, the introduction of timing jitter perturbs the phase about its true value. Compensating for timing jitter can be difficult, which motivates the need for phase tightening. We therefore focus exclusively on phase noise due to timing jitter, assuming that other measures have been taken to ensure that the phase noise has zero mean.

We develop phase tightening in discrete time, which although appropriate for use in a digital coherent imaging system, raises additional concerns about managing high digital data rates. Digital millimeter-wave radar imaging systems are challenging to build [64]. Because of the high carrier frequency, the bandwidth of the received signals typically requires the signals to be sampled at a relatively high rate to faithfully represent them. Therefore, the data rate from the sensors to the central processor that forms the image can be high. While phase tightening can produce an output data rate down to the baseband Nyquist rate for the measured signals, this data rate may still be challenging to manage.

We first motivate our chosen phase tightening architecture by reviewing the classical phase estimation theory and comparing alternative architectures for phase noise reduction (Section 6.1). We then qualitatively describe phase tightening behavior, placing emphasis on the inherent bias in discrete phase estimation (Section 6.2). We analyze the performance of our phase tightening algorithm by characterizing loop stability, computing the state dynamics, and correcting for phase estimation bias (Section 6.3). Our analysis suggests a design methodology, which we apply to a millimeter-wave radar design problem. We include functional circuit diagrams for both phase and amplitude estimation (Section 6.4).

## 6.1   Architecture for phase noise reduction

We wish to reduce the phase noise in the receiver of an active imaging system emitting narrowband radiation at a high carrier frequency $\omega_c$. Each sensor measures a signal $y(t)$ that has a slowly-varying amplitude $A(t)$ and phase $\psi(t)$, but that is corrupted by phase noise $\phi(t)$ and additive noise $v(t)$:

$$y(t) = A(t)\sin\left[\omega_c t + \psi(t) + \phi(t)\right] + v(t). \tag{6.1}$$

We assume that $A(t)$ and $\psi(t)$ are deterministic but unknown, that $v(t)$ is additive white Gaussian noise with zero mean, and that $\phi(t)$ is ergodic noise with zero mean. Phase tightening provides a sequence of estimates of $A(t)$ and $\psi(t)$ at each sensor (Figure 6-1). If the discrete estimates $\hat{A}_n$ and $\hat{\psi}_n$ are updated at the baseband Nyquist rate as $A(t)$ and $\psi(t)$ evolve over time, then the underlying envelopes of the received signals can be reconstructed and used to form an image, as in Chapters 2 and 4.

$$A(t)\sin\left[\omega_c t + \psi(t) + \phi(t)\right] + v(t) \longrightarrow \boxed{\begin{array}{c} \text{phase} \\ \text{tightening} \end{array}} \longrightarrow \hat{A}_n,\ \hat{\psi}_n$$
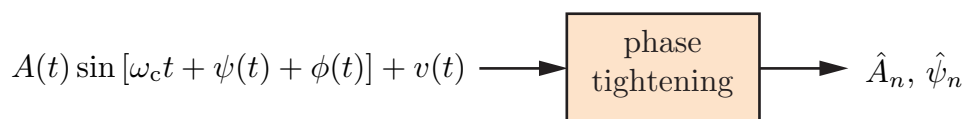
Figure 6-1: Phase tightening estimates the slowly-varying amplitude $A(t)$ and phase $\psi(t)$ of a high-frequency narrowband signal corrupted by both phase noise $\phi(t)$ and additive noise $v(t)$.

Phase tightening thereby results in a digital version of the measured signal, with a lower frequency spectrum and lower phase noise than the original. The phrase *phase tightening* arises because as the data rate decreases, the phase is pulled closer to the true value. The circuit components used to implement phase tightening will themselves introduce phase noise that must be reduced. We incorporate the impact of these internal noise sources into $\psi(t)$ throughout our analysis. We confine most of our analysis to short time scales on which the signal envelopes are approximately constant, and therefore abbreviate $A(t)$ and $\psi(t)$ with the constant amplitude $A$ and true phase $\psi$. We focus on the relationship between phase noise and phase estimation,

and consequently ignore amplitude estimation (by setting $A = 1$) and additive noise (by setting $v(t) = 0$) when they would otherwise unduly complicate the analysis.

Many practical systems exhibit more complicated types of phase noise $\phi(t)$ that can be addressed by alternative techniques. For example, a free-running oscillator may oscillate at a frequency that does not exactly match $\omega_c$. If such an oscillator is mixed with a sensor measurement, it will contribute phase noise with increments $\phi(t) - \phi(t - T)$ of unbounded variance in the time increment $T$. By replacing the free-running oscillator with a phase-lock loop, the variance of the resulting phase increment can be bounded [63]. However, even with a phase-lock loop, each sensor measurement will generally have a different phase offset, so that $\phi(t)$ has nonzero mean. Constant phase offsets can be subtracted out at each sensor by calibrating the imaging system. Specifically, a calibration system can apply a sequence of phase adjustments at each sensor, guided by the fidelity of the resulting beam pattern, until the phase offsets are eliminated when the resulting image is sharp. We assume that phase-lock loops and calibration techniques have been appropriately utilized, so that we can assume that $\phi(t)$ is ergodic noise with zero mean. Our goal is to reduce the variance of $\phi(t)$.

We first present a classical analog solution to phase noise reduction (Section 6.1.1). Implementing a digital version of the classical solution requires analog to digital conversion, which can be done at various stages of processing, with each possibility resulting in a different architecture. One option results in a digital feedback loop that forms the basis for our phase tightening technique (Section 6.1.2). We develop the digital phase tightening algorithm and present a simple state evolution model that we will use in our subsequent analysis (Section 6.1.3).

## 6.1.1  Classical analog solution

We decompose the problem of estimating $\psi$ given $y(t)$ into two classic problems: first, estimating the phase $\zeta(t) = \psi + \phi(t)$ from $y(t)$ in additive white Gaussian noise $v(t)$; and second, estimating the average $\psi$ of the noisy $\zeta(t)$. The phase-lock loop (PLL) is an optimal solution to the phase estimation problem, in both the minimum mean square error and maximum *a posteriori* sense, when the loop is locked and under the appropriate linear approximations [65] (Figure 6-2). The PLL estimates the phase of its input $y(t)$, and then uses the phase estimate $\hat{\zeta}(t)$ to drive a variable phase oscillator to produce a low-noise complement to $y(t)$, given by $\cos[\omega_c t + \hat{\zeta}(t)]$. This complement is modulated with $y(t)$, producing a term oscillating at $2\omega_c$, another at $\omega_c$, and an error term that is a scaled version of $\sin[\zeta(t) - \hat{\zeta}(t)]$. The low-pass filter has a cutoff frequency below $\omega_c$ and therefore eliminates all but the low-frequency error term, upon which it acts as an integrator that aggregates the error. The aggregated error updates the phase estimate, which improves provided $|\zeta(t) - \hat{\zeta}(t)| < \pi$, so that the sign of the error $\zeta(t) - \hat{\zeta}(t)$ is preserved in $\sin[\zeta(t) - \hat{\zeta}(t)]$, guaranteeing that the loop remains locked.

The phase-lock loop operates by comparing the input with a signal generated from a phase estimate, using the result of the comparison to update the phase estimate, and then feeding the new phase estimate back to generate a new signal. Real phase-
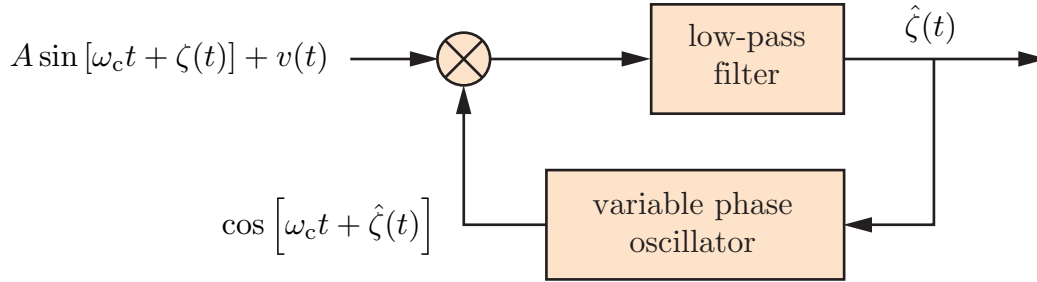
Figure 6-2: The phase-lock loop optimally estimates the phase under the appropriate linear approximations. The PLL produces a phase estimate $\hat{\zeta}(t)$ which is used to construct a low-noise complement to the input. The complement is modulated with the input, and the low-pass filter eliminates all resulting terms except an error term, which it integrates to update $\hat{\zeta}(t)$.

lock loops vary in complexity and implement the overall input-output relationship described above in different ways. For example, practical implementations employ a voltage-controlled oscillator rather than a variable phase oscillator, so that the voltage controls the oscillator frequency instead of the phase [65].

Once we have a phase estimate $\hat{\zeta}(t)$, we need only average it to estimate $\psi$. The averaging can be implemented by appending another filter to the output of the PLL, or even incorporated into the low-pass filter in the loop itself, as long as the resulting loop dynamics are acceptable.

## 6.1.2  Architectural decisions on converting to digital

We seek a digital version of the classical phase-lock loop to use as the basis for phase tightening. There are four clear interfaces in the generalized PLL architecture at which to convert from an analog to a digital representation (Figure 6-3). At one extreme, we can convert the output to obtain the pure analog system from Section 6.1.1, which poses a significant analog circuit design challenge at high frequencies. At the other extreme, we can convert the input to obtain a pure digital system, which could require high bit precision at a high sampling rate to adequately represent the input $y(t)$. Another natural conversion point is after the phase estimation, resulting in an analog PLL followed by digital averaging, which presents similar analog circuit design challenges as the pure analog architecture. We instead make the unconventional choice of converting to digital in the middle of the PLL loop. Although we sample at a high rate to get as much phase information as possible from the input waveform, we do not require high bit precision and can keep the digital logic simple to minimize cost and power consumption. In an imaging system, the circuit components used to implement phase tightening are replicated per antenna array sensor, so that savings in cost and power scale.
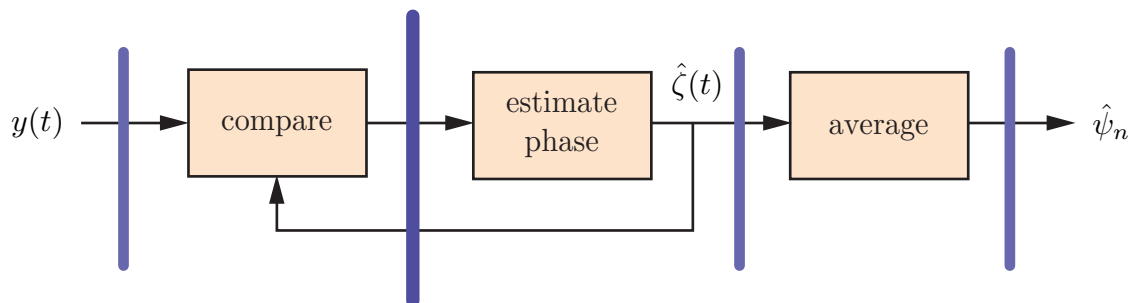
Figure 6-3: Of the four clear interfaces at which to convert the classical phase-lock loop architecture from analog to digital, we convert in the middle of the loop. Although we sample at a high rate, we do not require high bit precision and can keep the digital logic simple.

### 6.1.3 Phase tightening algorithm

We develop the phase tightening algorithm by converting from analog to digital inside the PLL loop and selecting simple digital logic to implement the core PLL functions of comparison, phase estimation, and averaging (Figure 6-4). Converting to digital requires an analog-to-digital converter (ADC). The simplest comparison we can perform is comparing samples of the input $y(t)$ to zero using a 1-bit ADC. When the PLL converges, the comparison should result in equality, so that the PLL locks onto the zero-crossings of the input at the ADC sampling instants. Therefore, the feedback path should adjust the ADC sampling instants to find the zero-crossings. We can equivalently fix the ADC sampling rate at the Nyquist rate, so that the sampling interval is $T = \pi/\omega_c$, and use feedback to delay the input. We accumulate the ADC outputs in a counter with $Q$ states and output $u_n$, which indicates the delay to apply to $y(t)$ for the next sample at index $n + 1$. We need to adjust the delay in different directions for rising and falling zero-crossings, so we toggle the sign of the ADC output by $(-1)^n$. We scale $u_n$ by the quantization step size $\Delta = 2\pi/Q$ to convert to radians and obtain a coarse instantaneous phase estimate $u_n\Delta$, which we average to obtain the improved output phase estimate $\hat{\psi}_n$. We scale $u_n\Delta$ by $1/\omega_c$ to convert to a time delay to apply to $y(t)$. The state update rule is therefore

$$u_{n+1} = u_n + (-1)^n \operatorname{sgn}\left[y\left(nT - u_n\frac{\Delta}{\omega_c}\right)\right]. \tag{6.2}$$

Although the digital logic in the phase tightening algorithm is simple, the analog components are nontrivial. The variable delay, when either applied to the input or the ADC clock, should be consistent while the averaging is being performed. Any drift in delay should correspond to a phase drift significantly smaller than the desired precision of the phase estimation. Both the variable delay block and the ADC clock can introduce timing jitter; as long as the jitter is ergodic noise with zero mean, it can be incorporated into the phase noise $\phi(t)$ and reduced through averaging.
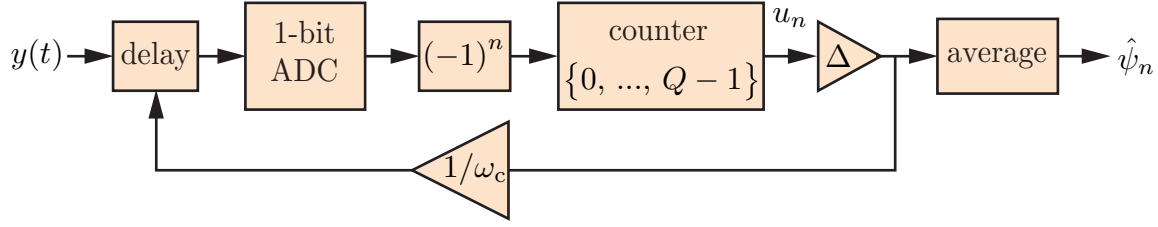
Figure 6-4: The phase tightening algorithm uses a 1-bit ADC to lock onto the zero-crossings of the input $y(t)$ by applying a variable feedback delay. A counter with $Q$ states accumulates the zero-crossing errors to update the delay applied to the input at the next sampling instant. Converting the counter output $u_n$ to radians $u_n \Delta$ and averaging yields the phase estimate $\hat{\psi}_n$.

Digital phase tightening is therefore advantageous when the required cost and power consumption compares favorably with that of an alternative analog architecture that achieves similar overall performance.

Phase tightening operates on the input signal by applying a delay to directly modify the phase, providing us with a simpler model for analysis (Figure 6-5). To derive the simpler model, we substitute the expression in (6.1) for the input $y(t)$ into (6.2), ignore the additive noise $v(t)$, and set $A = 1$ to obtain

$$
\begin{aligned}
u_{n+1} &= u_n + (-1)^n \operatorname{sgn}\left\{ \sin\left[ \psi + \omega_{\mathrm{c}}\left( nT - u_n \frac{\Delta}{\omega_{\mathrm{c}}} \right) + \phi\left( nT - u_n \frac{\Delta}{\omega_{\mathrm{c}}} \right) \right] \right\} \\
&= u_n + \operatorname{sgn}\left\{ \sin\left[ \psi - u_n \Delta + \phi\left( nT - u_n \frac{\Delta}{\omega_{\mathrm{c}}} \right) \right] \right\},
\end{aligned} \tag{6.3}
$$

where we have used the fact that $\omega_{\mathrm{c}} T = \pi$ at the Nyquist sampling rate. The samples

$$
\phi_n = \phi\left( nT - u_n \frac{\Delta}{\omega_{\mathrm{c}}} \right) \tag{6.4}
$$

form a discrete-time ergodic noise process with zero mean. We thereby simplify (6.3) to

$$
u_{n+1} = u_n + \operatorname{sgn}(\psi + \phi_n - u_n \Delta), \tag{6.5}
$$

keeping in mind that all calculations with phases are performed modulo $2\pi$ and all calculations with counter values are performed modulo $Q$. We use the model in (6.5) extensively in our subsequent analysis.

## 6.2   Phase tightening behavior

For the phase tightening algorithm to be effective, the coarse phase estimates $u_n \Delta$ must track the true phase $\psi$ so that enough averaging can reduce the noise variance to an acceptable level. We first require the system to lock onto the zero-crossings of the
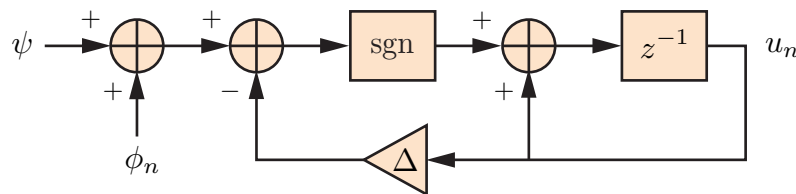
Figure 6-5: The evolution of $u_n$ in the phase tightening algorithm is governed by a simple model that uses the phase $\psi$, instead of the received sinusoid $y(t)$, as input.

input signal $y(t)$, so that successive values of $u_n\Delta$ move towards $\psi$. Then, averaging will converge to the expected value $E[u_n]\Delta$. Ideally, we would like $E[u_n]\Delta = \psi$, so that the bias of the phase estimate is zero. While a constant bias that we can estimate or calibrate is acceptable, a large bias that depends on $\psi$ or drifts on a time scale shorter than the time window required to average out the phase noise is not acceptable.

The state update equation in (6.5) embodies the phase tightening algorithm's behavior. Although phase tightening utilizes feedback in similar manner as other algorithms (Section 6.2.1), there are important subtleties, particularly involving bias. We describe how phase estimation bias is inherent in our discrete phase tightening system (Section 6.2.2), and then describe how to compensate for the bias by operating in the low-bias noise range (Section 6.2.3).

## 6.2.1  Similar feedback systems

Our update equation in (6.5) shares the recursive structure of the Kalman filter [66]. Though unlike the Kalman filter, our update is constrained to be one of a small set of quantized values. When applied to a constant signal, the Kalman filter converges to the true value by making successively finer updates, while our updates remain coarse, and we rely on further averaging downstream for accuracy.

Our update equation applies one bit of feedback at a high rate, similar to the operation of a Delta-Sigma converter [67]. Though unlike a Delta-Sigma converter, in phase tightening the ADC quantizes the error before the counter integrates it, rather than after. Consequently, the feedback loop in phase tightening does not shape the noise present in the input or the noise introduced by the ADC clock. The phase tightening loop processing is deliberately simple; alternative designs may add complexity to increase accuracy, by basing the error computation at each sample on a more accurate phase estimate rather than the coarse $u_n\Delta$.

## 6.2.2  Inherent phase estimation bias

The stochastic process $u_n\Delta$ makes a discrete random walk about the true phase value $\psi$, and, when averaged, provides an estimate of $\psi$. The phase estimate would ideally

be unbiased,

$$E[\hat{\psi}_n] = E[u_n]\Delta = \psi, \tag{6.6}$$

so that $\hat{\psi}_n \to \psi$ provided we average over enough samples. Unfortunately, our discretization of the phase estimation problem imposes a nonzero bias that can render phase tightening ineffective if not properly accounted for. The bias is determined by the noise distribution $p_\phi$, the true phase $\psi$, and the number of quantization steps $Q$. The bias we discuss is restricted to the inherent bias introduced by quantizing a continuous phase estimation problem; other sources of bias in a phase tightening implementation, such as phase noise with nonzero mean, can have a similarly catastrophic effect on phase tightening and must be addressed separately.

The bias is a result of our imperfect and asymmetric discretization of a continuous phenomenon, and is most easily observed when there is no phase noise. From our simplified phase tightening model (Figure 6-5), the update equation for $u_n$ in (6.5) when $\phi_n = 0$ is

$$u_{n+1} = u_n + \text{sgn}(\psi - u_n\Delta). \tag{6.7}$$

Unless $\psi$ is an exact multiple of $\Delta$, the instantaneous phase estimates $u_n\Delta$ will oscillate about $\psi$, and averaging enough samples produces a phase estimate midway between (Figure 6-6). The resulting bias can be as high as $\Delta/2$.

bias error

$$u_n\Delta = q\Delta \qquad\qquad \psi \qquad \hat{\psi}_n \to (q+1/2)\Delta \qquad\qquad u_n\Delta = (q+1)\Delta$$
$$n \text{ even} \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad n \text{ odd}$$
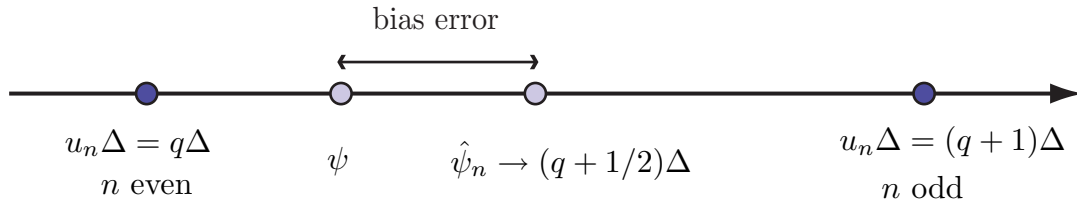
Figure 6-6: In the steady state with no phase noise when $\psi$ is not an exact multiple of $\Delta$, the instantaneous phase estimates oscillate about the true value and result in a bias error as high as $\Delta/2$.

There are two important properties evident in the simple case of no phase noise that also hold in more complicated scenarios. First, the bias is inherently a result of the discrete quantization. As we increase the number of quantization steps to approach a continuous model, $\Delta \to 0$ and the bias disappears. Second, the bias is minimized when a certain minimum level of phase noise is present, which is similar to the dithering effect observed in many-threshold systems [69]. Without noise, the simple update rule for $u_n$ does not account for the fine positioning of $\psi$; we observe the same oscillatory behavior regardless of where $\psi$ lies between the two endpoints. But noise may push us to different estimates, in a manner that takes this fine positioning into account. The bias can be improved by adding noise, and for some noise profile edge cases, eliminated completely.

### 6.2.3 Intuition for the low-bias noise range

Before proceeding with detailed analysis, we describe intuitively how the bias depends on the noise, true phase, and quantization step size. The quantization points are spaced every $\Delta$ in distance. Samples are drawn from the noisy distribution of $\psi + \phi_n$, but we only determine if each sample falls to the left or right side of the current quantization point. If the phase noise is very low, we will always make the same determination with respect to a particular quantization point, so that the best estimate of $\psi$ is midway between the two nearest quantization points (Figure 6-6). As the phase noise increases, samples of $\psi + \phi_n$ will begin to fall on the other side of the nearest quantization point (for example, to the left of the $q^{\text{th}}$ quantization point in Figure 6-6). The frequency with which the sample falls on the opposing side of the nearest quantization point provides information about how close $\psi$ is. Therefore, when the noise is large enough to reach the nearby quantization points, $\psi$ can be accurately estimated and the bias is low (Figure 6-7). However, once the phase noise is large enough to wrap around by $\pi$ radians so that samples fall on both sides of all quantization points, the samples become misleading and the bias increases. When the phase noise is very high, the coarse estimates $u_n\Delta$ spend roughly equal time at each quantization point, with a slight preference for the point $q\Delta$ closest to $\psi$. Averaging a uniform distribution centered at $q\Delta$ results in a phase estimate of $q\Delta$. Thus, for high phase noise, $\hat{\phi}_n$ converges to the quantization point closest to $\phi$. Therefore, the phase estimate has low bias for a significant range of noise levels that depends on both the true phase $\psi$ and the quantization step size $\Delta$, called the low-bias noise range.
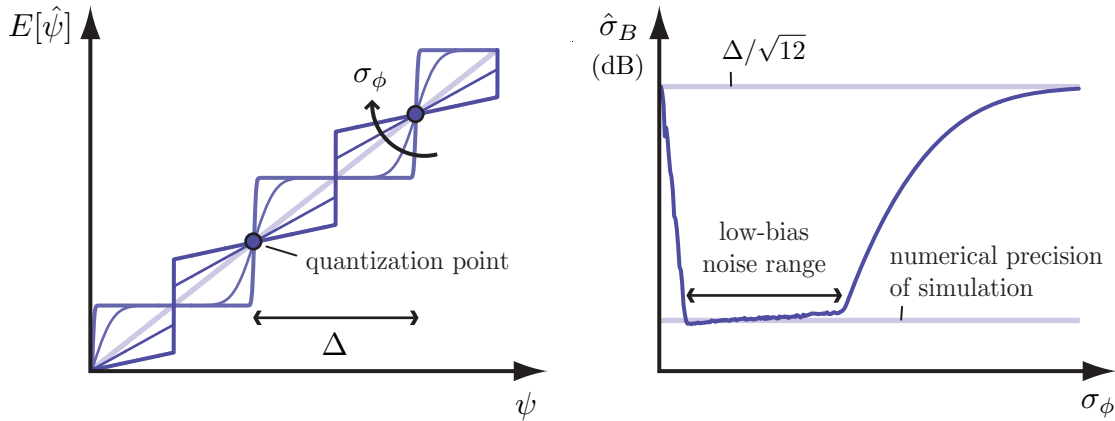


Figure 6-7: The phase tightening phase estimates only exhibit low bias for moderate levels of phase noise relative to the distance from the true phase $\psi$ to the nearest quantization point. For low phase noise $\sigma_\phi$, the asymptotic estimate $E[\hat{\psi}]$ lies midway between the nearest quantization points, while for high phase noise, the asymptotic estimate lies at the nearest quantization point. At either extreme of low or high phase noise, the sample standard deviation $\hat{\sigma}_B$ of the bias as a function of $\psi$ is bound by the quantization step size.

The quantization step size $\Delta$ is the primary design parameter, and our intuition suggests to select $\Delta$ no larger than the expected noise standard deviation $\sigma_\phi$ to achieve

low bias, so that samples fall on the opposing side of at least one quantization point regardless of where $\psi$ is located. However, it may be difficult to implement an accurate delay of such a small time step $\Delta/\omega_{\mathrm{c}}$. Effective phase tightening is still possible in this case, by using a coarse quantization step size $\Delta$ but also adding noise to ensure low bias. Dithering comes with a cost, as any additional phase noise will require more averaging to achieve the same noise target level at the output, requiring either a greater sampling rate or a longer averaging window.

## 6.3  Performance analysis of phase estimation

We analyze phase tightening performance in three stages. First, we assess the stability of the loop and the conditions required to maintain a lock on the zero-crossings (Section 6.3.1). Provided that the phase noise is not too large, the loop is stable. Second, we analyze the dynamics of the state stored in the counter (Section 6.3.2). When $Q$ is even and under modest assumptions on the support of the noise distribution, the evolution of $u_n$ is accurately modeled by an ergodic Markov chain. We provide analytic solutions for the steady-state probability vector of the Markov chain. Third, we compute and correct for the bias inherent in discrete phase estimation (Section 6.3.3). The bias is only low over a moderate range of $\sigma_\phi$. The bias is a complicated function of the true phase $\psi$, but can be simplified for a special noise edge configuration. We show that the bias decays exponentially in $Q$ for the noise edge configuration. Consequently, our analysis provides guidance for selecting the number of quantization steps $Q$ to provide an operating range for $\sigma_\phi$ under which phase tightening achieves the desired reduction in phase noise.

### 6.3.1  Stability of feedback loop

We analyze the state update equation to assess loop stability. We ignore the additive noise $v(t)$ and set $A = 1$ to obtain the state update equation from (6.3),

$$u_{n+1} = u_n + \mathrm{sgn}\left[\sin\left(\psi - u_n\Delta + \phi_n\right)\right]. \tag{6.8}$$

The fixed-point solutions for $u_n$ occur when the argument of the sine is a multiple of $\pi$. In the absence of phase noise, the fixed points are therefore at

$$u_n = \frac{\psi + m\pi}{\Delta}, \tag{6.9}$$

for any integer $m$. The fixed point is stable for even $m$: if $u_n\Delta$ is slightly below $\psi + m\pi$, our update will increase $u_n$ per (6.8), and similarly for $u_n\Delta$ slightly above $\psi + m\pi$. However, the fixed point is unstable for odd $m$, as then the update will push $u_n$ in the opposite directions. Therefore, modulo $2\pi$, $u_n\Delta$ will step steadily towards $\psi$ until reaching it. Assuming that $\psi$ is not an exact multiple of $\Delta$, the $u_n\Delta$ will oscillate about $\psi$, locked onto this target phase. Once locked, the phase noise must exceed $\pi - \Delta$ to cause $u_n$ to step in the wrong direction.

## 6.3.2   State dynamics

The quality of the phase estimate provided by phase tightening depends on the probability distribution of the state $u_n$. We model the evolution of $u_n$ with a cyclic Markov chain, and compute the steady-state distribution of $u_n$. We first prove general properties about the symmetry and support of the transition probabilities of the Markov chain. These properties allow us to both classify the recurrence of the chain and explicitly compute the steady-state distribution in an important special case.
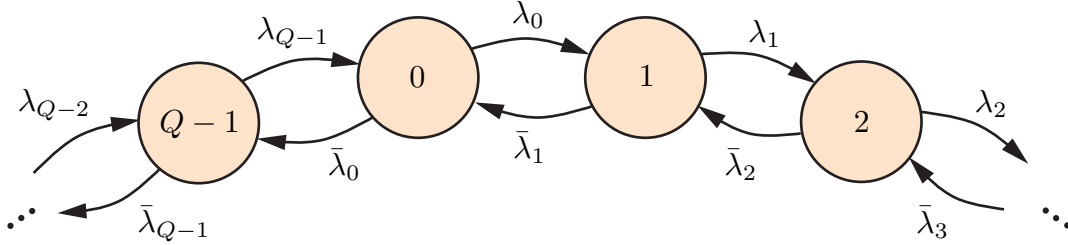


Figure 6-8: A Markov chain of order $Q$ describes the state evolution of $u_n$.

We model the evolution of $u_n$ with a finite state Markov chain (Figure 6-8). We denote the probability of increasing the counter and moving from state $k$ to state $k+1$ by $\lambda_k$, and the probability of decreasing the counter and moving from state $k$ to state $k-1$ by $\bar{\lambda}_k$. From the update equation in (6.5) for our simplified phase tightening model (Figure 6-5), the transition probabilities are

$$
\begin{aligned}
\lambda_k &= \sum_{m \in \mathbb{Z}} \Pr\left(2\pi m < \psi + \phi_n - k\Delta < 2\pi m + \pi\right) \\
&= \sum_{m \in \mathbb{Z}} \Pr\left(2\pi m + k\Delta < \phi_n + \psi < 2\pi m + k\Delta + \pi\right), \\
\bar{\lambda}_k &= 1 - \lambda_k,
\end{aligned}
\tag{6.10}
$$

for $k \in \{0, \ldots, Q-1\}$.

We conveniently express the transition probabilities as samples of a circular convolution of period $2\pi$ between the probability density for $\phi_n + \psi$ and a rectangle function, as

$$
\lambda_k = p_\phi(x - \psi) \circledast \left[u\left(x + \pi\right) - u(x)\right]\big|_{x = k\Delta},
\tag{6.11}
$$

where $p_\phi$ denotes the probability density for $\phi_n$ and $u(\cdot)$ denotes the unit step function. We visualize the transition probabilities for an example where $\phi_n$ is uniformly distributed between $-W$ and $W$, and $Q = 10$, in Figure 6-9.

### Symmetry and support of transition probabilities

The transition probabilities exhibit a symmetry and have concentrated support. We prove two such properties using the convolution interpretation in (6.11). Recall that
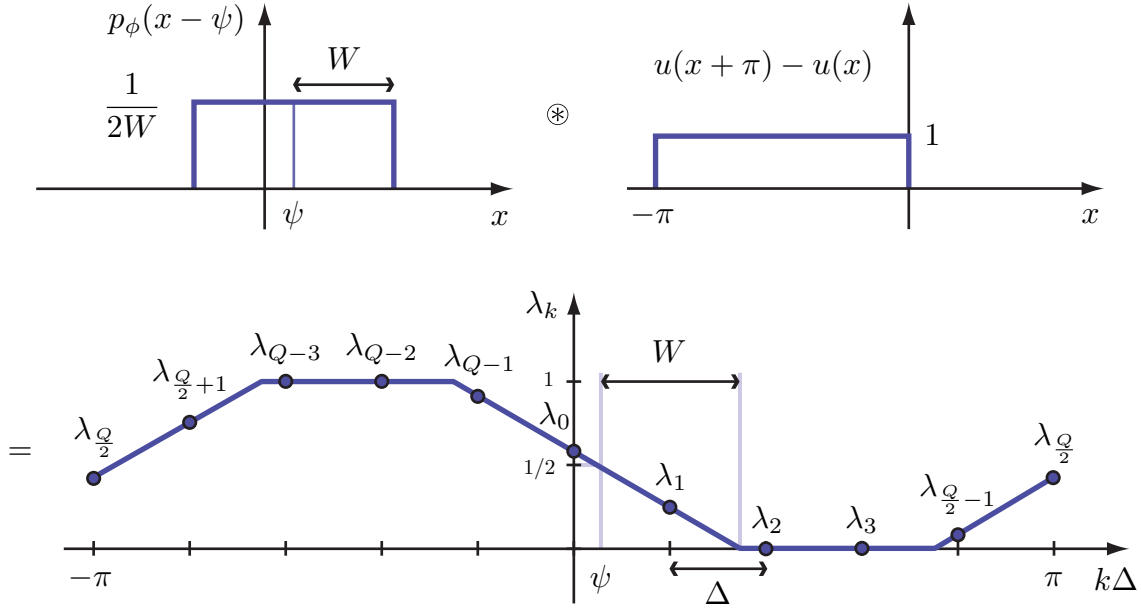
Figure 6-9: The transition probabilities $\lambda_k$ for the Markov chain are samples of the circular convolution of the noise distribution and a rectangle function. In this example, $\phi_n$ is uniformly distributed between $-W$ and $W$, and $Q = 10$.

all index arithmetic and comparisons with elements from the set $\{0, \ldots, Q - 1\}$ are made modulo $Q$, so that, for example, $k + Q/2 \in \{0, \ldots, Q - 1\}$.

**Claim 6.3.1.** *If $Q$ is even, then $\bar{\lambda}_k = \lambda_{k + \frac{Q}{2}}$.*

*Proof.* We compute

$$
\begin{aligned}
\lambda_{k + \frac{Q}{2}} &= p_\phi(x - \psi) \circledast \left[ u\left(x + \pi\right) - u(x) \right]\big|_{x = k\Delta + \pi} \\
&= p_\phi(x - \psi) \circledast \left[ u\left(x + 2\pi\right) - u\left(x + \pi\right) \right]\big|_{x = k\Delta} \\
&= p_\phi(x - \psi) \circledast \left[ u\left(x + 2\pi\right) - u(x) \right]\big|_{x = k\Delta} \\
&\qquad - p_\phi(x - \psi) \circledast \left[ u\left(x + \pi\right) - u(x) \right]\big|_{x = k\Delta} \\
&= 1 - \lambda_k \\
&= \bar{\lambda}_k.
\end{aligned}
$$

$\square$

**Claim 6.3.2.** *The indices where $\lambda_k = 0$ are consecutive. The indices where $\bar{\lambda}_k = 0$ are also consecutive. These two sets of indices do not overlap.*

*Proof.* From (6.11), we have $\lambda_k = 0$ when the support of $p_\phi$ lies outside an interval of length $\pi$. Two different intervals of length $\pi$ that are wrapped into $[-\pi, \pi)$ must either abut each other or overlap. Hence if the corresponding indices are $k_1$ and $k_2$, and $\lambda_{k_1} = \lambda_{k_2} = 0$, then $\lambda_k = 0$ for $k_1 \leq k \leq k_2$. That is, the indices where $\lambda_k = 0$

are consecutive. A similar argument shows that the indices where $\lambda_k = 1$, and hence where $\bar{\lambda}_k = 0$, are also consecutive. Since $\lambda_k + \bar{\lambda}_k = 1$, these sets of indices do not overlap. $\qquad\square$

## Classifying the Markov chain

The properties of the transition probabilities allow us to classify the Markov chain and deduce its recurrence structure. First, we recall some useful terminology for Markov chains [68]. A *class* of states is a maximal set where there is a path between the states in each pair. A state is *recurrent* if there is some return path from any other accessible state. A class is recurrent if all its states are. A state that is not recurrent is called *transient*, and a class is transient if all its states are. Finally, recurrent and transient states cannot mix in a single class; every class is either recurrent or transient.

**Claim 6.3.3.** *The Markov chain contains a single recurrent class. This class is either aperiodic or periodic with period 2, depending on whether it contains an odd or even number of states.*

*Proof.* If all the $\lambda_k$ or all the $\bar{\lambda}_k$ are nonzero, then there is a path from every state to every other, and the chain forms a single recurrent class. Otherwise, Claim 6.3.2 implies that we have a configuration similar to the one shown in Figure 6-10. The states between the consecutive zeros for $\bar{\lambda}_k$ and those for $\lambda_k$ form a single recurrent class, and all other states are transient. Equivalently, we can identify the recurrent class from the transition between 1 and 0 in the convolution diagram. In Figure 6-9, the indices from $Q - 2$ through 2 correspond to the states that form the recurrent class. $\qquad\square$
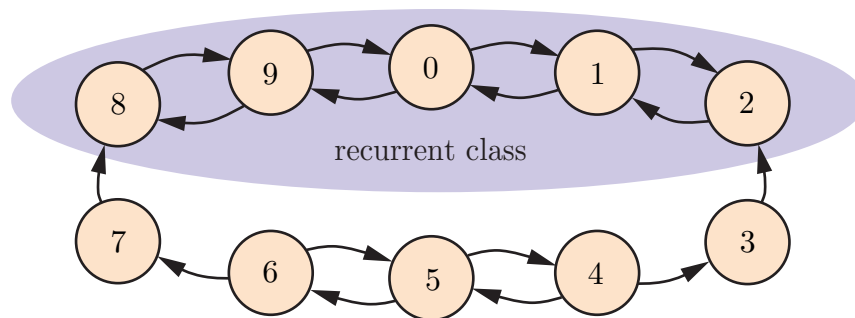


Figure 6-10: The Markov chain contains a single recurrent class. The chain shown here corresponds to the uniform noise example in Figure 6-9.

The noise distribution determines the number of states in the recurrent class, and thereby whether the class is periodic.

**Steady-state distribution formulas**

We now compute the steady-state distribution for $u_n$ for a special case of interest. Since the Markov chain contains a single recurrent class, the Perron-Frobenius theory [68] guarantees the existence of a unique steady-state probability *row* vector solution $\boldsymbol{\pi}$ to

$$\boldsymbol{\pi} = \boldsymbol{\pi}\,\mathbf{P}, \tag{6.12}$$

where $\mathbf{P}$ is the $Q \times Q$ matrix of transition probabilities:

$$\mathbf{P} = \begin{bmatrix} 0 & \lambda_0 & & & & \bar{\lambda}_0 \\ \bar{\lambda}_1 & 0 & \lambda_1 & & & \\ & & \ddots & & & \\ & & \bar{\lambda}_{Q-2} & 0 & \lambda_{Q-2} \\ \lambda_{Q-1} & & & \bar{\lambda}_{Q-1} & 0 \end{bmatrix}. \tag{6.13}$$

We can explicitly solve (6.12) under certain conditions.

**Claim 6.3.4.** *When*

$$\prod_{i=0}^{Q-1} \lambda_i = \prod_{i=0}^{Q-1} \bar{\lambda}_i, \tag{6.14}$$

*then*

$$\pi_k = \begin{cases} \prod_{i=M}^{Q+k-1} \lambda_i \prod_{i=k+1}^{M-1} \bar{\lambda}_i, & k < M \\[2ex] \prod_{i=M}^{k-1} \lambda_i \prod_{i=k+1}^{M+Q-1} \bar{\lambda}_i, & k \geq M \end{cases}, \tag{6.15}$$

*for $k \in \{0, \ldots, Q-1\}$, is a scaled version of the steady-state probability vector solution to (6.12). In (6.15), $M \in \{0, \ldots, Q-1\}$ denotes the index $k$ of the nonzero $\lambda_k$ after one that is zero: the $k$ such that $\lambda_k \neq 0$ and $\lambda_{k-1} = 0$. If no such $k$ exists, we set $M = 0$.*

*Proof.* First we argue that $M$ is well-defined. Due to Claim 6.3.2, the indices for which $\lambda_k = 0$ are consecutive. It is impossible for all the $\lambda_k$ to be zero, as that would contradict the convolution relationship in (6.11). Thus if $\lambda_k = 0$ for some $k$, then there is a unique index $M$ such that $\lambda_M \neq 0$ and $\lambda_{M-1} = 0$.

Similar reasoning shows that the $\boldsymbol{\pi}$ specified by (6.15) is nonzero. For if all the $\lambda_k$ are nonzero, then all the $\pi_k$ are nonzero. Otherwise, let $M'$ denote the index $k$ of the zero $\lambda_k$ after one that is nonzero, $\lambda_{k-1} \neq 0$. $M'$ is uniquely determined, just like $M$. Now, $\pi_{M'}$ is equal to the product of nonzero numbers. Thus $\boldsymbol{\pi}$ is nonzero. All the $\pi_k$ are nonnegative, so $\boldsymbol{\pi}$ can be scaled to be a probability vector.

Finally, we show that (6.15) solves (6.12), which will complete the proof. The matrix equation in (6.12) aggregates the scalar equations

$$\pi_k = \lambda_{k-1}\pi_{k-1} + \bar{\lambda}_{k+1}\pi_{k+1}. \tag{6.16}$$

There are 9 different cases to consider, and we defer the calculations to Appendix B.4.1. $\qquad\square$

The condition in (6.14) encompasses two special cases of interest. First, (6.14) is automatically satisfied when $Q$ is even, by the symmetry in Claim 6.3.1. Second, (6.14) is also satisfied if the support of the noise distribution is not too large, so that some $\lambda_{k_1} = 0$ and another $\lambda_{k_2} = 1$. The first special case where $Q$ is even allows us to simplify our expression for the steady-state distribution.

**Claim 6.3.5.** *If $Q$ is even, then*

$$\pi_k = \prod_{i=k-\frac{Q}{2}+1}^{k-1} \lambda_i, \tag{6.17}$$

*for $k \in \{0, \ldots, Q-1\}$, is a scaled version of the steady-state probability vector solution to (6.12).*

*Proof.* We apply the symmetry result in Claim 6.3.1 to the general solution in (6.15) and notice that all of the $\pi_k$ have the factor

$$\prod_{i=M}^{M+\frac{Q}{2}-1} \lambda_i \tag{6.18}$$

in common. This factor cannot be zero, because we proved that the resulting $\boldsymbol{\pi}$ is nonzero. Dividing (6.15) by (6.18) yields the desired result, (6.17). We present an alternative proof using direct calculation in Appendix B.4.1. □

Based on our results, it is most convenient work with cases where $Q$ is even and where the noise distribution is such that the recurrent class is aperiodic. An aperiodic recurrent class is called ergodic, and guarantees that the state will eventually be distributed according to the steady-state probability vector, regardless of the starting state [68]:

$$\lim_{n \to \infty} \mathbf{P}^n = \frac{\mathbf{e}\boldsymbol{\pi}}{\boldsymbol{\pi}\mathbf{e}}, \tag{6.19}$$

where $\mathbf{e} = [1 \ \cdots \ 1]^{\mathrm{t}}$.

### 6.3.3    Correcting for bias due to discrete phase estimation

We analyze bias in three stages. First, we explicitly compute the phase estimation bias. The periodic nature of phase estimation requires us to unwrap the phase at an appropriate point prior to averaging the steady-state probability distribution. We show that the bias is a periodic and symmetric function of the true phase $\psi$. Second, we define the noise edge configuration for uniform phase noise, where the edge of the distribution of the noisy phase falls on a quantization point. Although the noise edge configuration does not always result in the worst-case bias, it is a good indicator for bias behavior and is simple to analyze because the steady-state probability vector solution generalizes the binomial distribution. Third, we prove that in the noise edge configuration the bias decays to zero exponentially in $Q$.

**Computing bias**

To compute the bias of the phase tightening algorithm, we must address the periodic nature inherent in phase estimation, and unwrap the phase by centering on the most likely index. We define the bias by

$$\begin{aligned} B &= E[\hat{\psi}_n] - \psi \\ &= E[u_n]\Delta - \psi. \end{aligned} \tag{6.20}$$

By arbitrarily associating index zero with a zero delay applied to the phase tightening input, we give the system's reference phase preferential treatment. The resulting impact on phase estimation can be significant. For example, suppose that there is a high level of phase noise so that the steady-state distribution $\boldsymbol{\pi}$ is approximately uniform. If we constrain $u_n$ to $\{0, \ldots, Q-1\}$, then the phase estimate is always

$$\frac{\sum_{k=0}^{Q-1} k\pi_k}{\sum_{k=0}^{Q-1} \pi_k}\Delta = \frac{2\pi}{Q^2}\sum_{k=0}^{Q-1} k = \pi - \frac{\pi}{Q}, \tag{6.21}$$

which can lead to a large bias magnitude over $\pi$.

To improve performance, we instead choose indices for $u_n$ by centering on the index $k$ that maximizes $\pi_k$. Therefore, we compute the bias as

$$B = \frac{\sum_{k=L}^{L+Q-1} k\pi_k}{\sum_{k=0}^{Q-1} \pi_k}\Delta - \psi, \tag{6.22}$$

where the cut point $L$ is chosen to unwrap the periodic phase to be centered on the maximum $\pi_k$ (Figure 6-11). Now, for high phase noise where the steady-state distribution $\boldsymbol{\pi}$ is approximately uniform, the phase estimate will be centered on the quantized $u_k\Delta$ closest to the true phase $\psi$, resulting in a bias no greater than $\Delta/2$. Implementing a sliding window for index averaging requires more work, as an implementation does not have access to the steady-state probabilities $\boldsymbol{\pi}$ used to determine $L$. However, an implementation can adaptively update the window to be centered on the latest phase estimate.

The bias defined in (6.22) is periodic and symmetric with respect to the true phase $\psi$. Specifically, $B(\psi)$ is periodic with period $\Delta$, as $B(\psi + m\Delta) = B(\psi)$ since translating $\psi$ by $\Delta$ shifts the $\pi_k$ by one index. Also, $B(\psi)$ is an odd function, as $B(-\psi) = B(\psi)$ since inverting the sign of $\psi$ exchanges $\pi_k$ with $\pi_{-k}$. Finally, $B(\Delta/2) = B(0) = 0$, since the steady-state distribution is symmetrically centered when $\psi = \Delta/2$ and $\psi = 0$. Consequently, we only need to evaluate the bias for $0 < \psi < \Delta/2$, as the bias for all other values of $\psi$ can be determined by periodicity and symmetry.

**Bias for the noise edge configuration**

While operating so that the phase noise lies in the low-bias noise range is a good first step in phase tightening design, the application requirements may impose strict
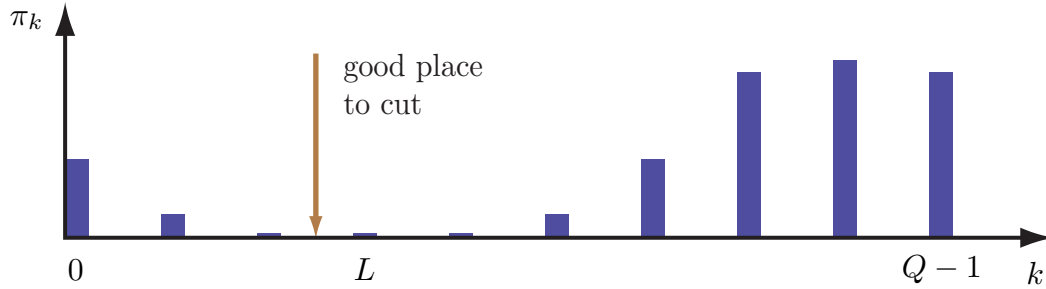
Figure 6-11: When averaging the steady-state probability distribution of $u_n$ to compute the phase estimation bias, we cut and unwrap the periodic phase at a point $L$ so that the averaging window is centered on the index $k$ where $\pi_k$ is maximized.

restrictions on the allowable bias, requiring further analysis. The dependence of bias on the true phase $\psi$ is complicated, and we advise examining all values $0 < \psi < \Delta/2$ to completely characterize performance. However, there is a special case for which the analysis simplifies significantly: assuming uniform noise of width $2W$, where one of the noise edges $\psi \pm W$ is just touching a quantization point. We describe why this noise edge configuration is an important indicator of bias performance, suggesting that we examine the bias for $\psi = W$, where the noise edge $\psi - W$ falls on the quantization point at the origin, before other values of $\psi$. We then show that the steady-state probability vector solution generalizes the binomial distribution.
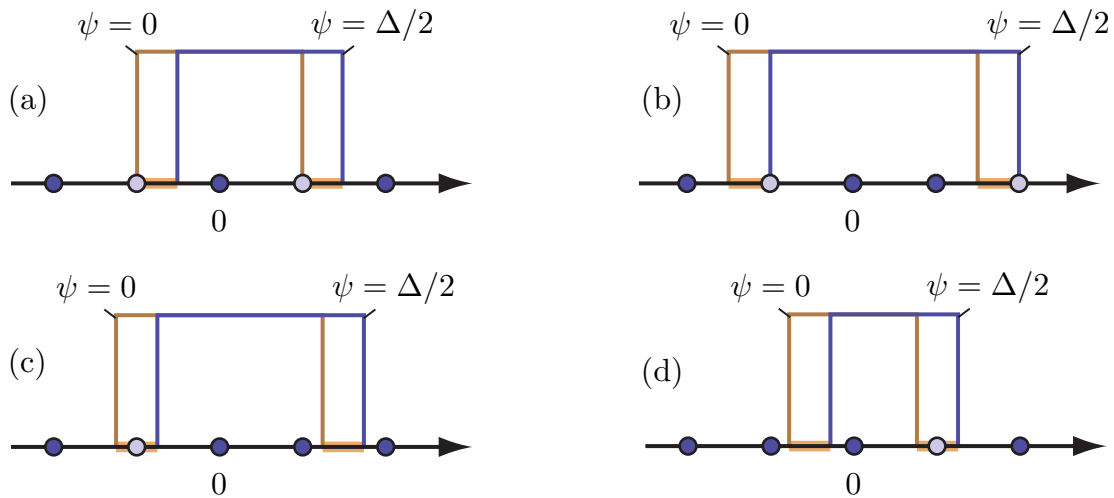


Figure 6-12: There are four scenarios where a noise edge touches a quantization point as $\psi$ varies from 0 to $\Delta/2$. The first two are edge cases where both noise edges touch quantization points when $\psi = 0$ (a) and where both noise edges touch quantization points when $\psi = \Delta/2$ (b). Otherwise, there is a unique value of $\psi$ between 0 and $\Delta/2$ for which either the left edge (c) or the right edge (d) touches a quantization point.

With the exception of two edge cases, there is a unique value of $\psi$ between 0 and

$\Delta/2$ for which a noise edge touches a quantization point (Figure 6-12). We call this the noise edge value of $\psi$, denoted by $\psi^{\mathrm{NE}}$.

**Claim 6.3.6.** *For fixed $\Delta$ and uniform noise width $2W$, exactly one the following cases holds.*

- *$W$ is a multiple of $\Delta$ so that when $\psi = 0$ both noise edges touch quantization points, but when $0 < \psi \le \Delta/2$ no noise edges touch quantization points.*

- *$W$ is an odd multiple of $\Delta/2$ so that when $\psi = \Delta/2$ both noise edges touch quantization points, but when $0 \le \psi < \Delta/2$ no noise edges touch quantization points.*

- *There is a unique $\psi$ between 0 and $\Delta/2$ so that exactly one noise edge touches a quantization point.*

*Proof.* The first two edge cases are straightforward to verify directly. Otherwise, we may assume that no noise edges touch quantization points at $\psi = 0$ or $\psi = \Delta/2$. Over the range $0 < \psi < \Delta/2$, each noise edge can only touch one quantization point, as the quantization points are spaced every $\Delta$. First, assume towards a contradiction that both edges touch quantization points, so that

$$W < n_1 \Delta < W + \Delta/2 \quad \text{and} \quad -W < n_2 \Delta < -W + \Delta/2, \qquad (6.23)$$

for integers $n_1$ and $n_2$. But adding the equations in (6.23) together yields $0 < n_1 + n_2 < 1$, a contradiction. Next, assume towards a contradiction that neither edge touches a quantization point, so that there is no quantization point on the interval $[W, W + \Delta/2]$, and none on $[-W, -W + \Delta/2]$. By symmetry there is no quantization point on $[W - \Delta/2, W + \Delta/2]$, which being an interval of length $\Delta$, yields another contradiction. Thus, exactly one noise edge touches a quantization point, for some value of $\psi$ between 0 and $\Delta/2$. $\qquad\square$

Intuitively, the bias at a noise edge can be large. On the one hand, moving $\psi$ from $\psi^{\mathrm{NE}}$ closer to the source at the noise edge will provide the source with information to estimate $\psi$. On the other hand, moving $\psi$ from $\psi^{\mathrm{NE}}$ further away from the source at the noise edge sometimes allows the other sources to better estimate $\psi$. Unfortunately, the precise behavior of the phase estimate is complicated (Figure 6-13). Although the noise edge introduces a kink in $B(\psi)$ at $\psi^{\mathrm{NE}}$ that may maximize the bias, there are some parameter settings for which the bias is even larger elsewhere (for example, when $W = 0.145$ and $\Delta = 0.1$ in Figure 6-13). Even though $B(\psi^{\mathrm{NE}})$ does not provide a bound for the bias, we will consider the noise edge configuration further because of its analytic tractability.

We now compute the steady-state probability distribution for the noise edge configuration. We set $\psi = W$, so that the noise edge $\psi - W$ falls on the quantization point at the origin. We illustrate a portion of the resulting convolution diagram in Figure 6-14. We introduce three new variables for convenience. First, the height of the uniform noise distribution is $w = 1/2W$, which is also the negative of the slope
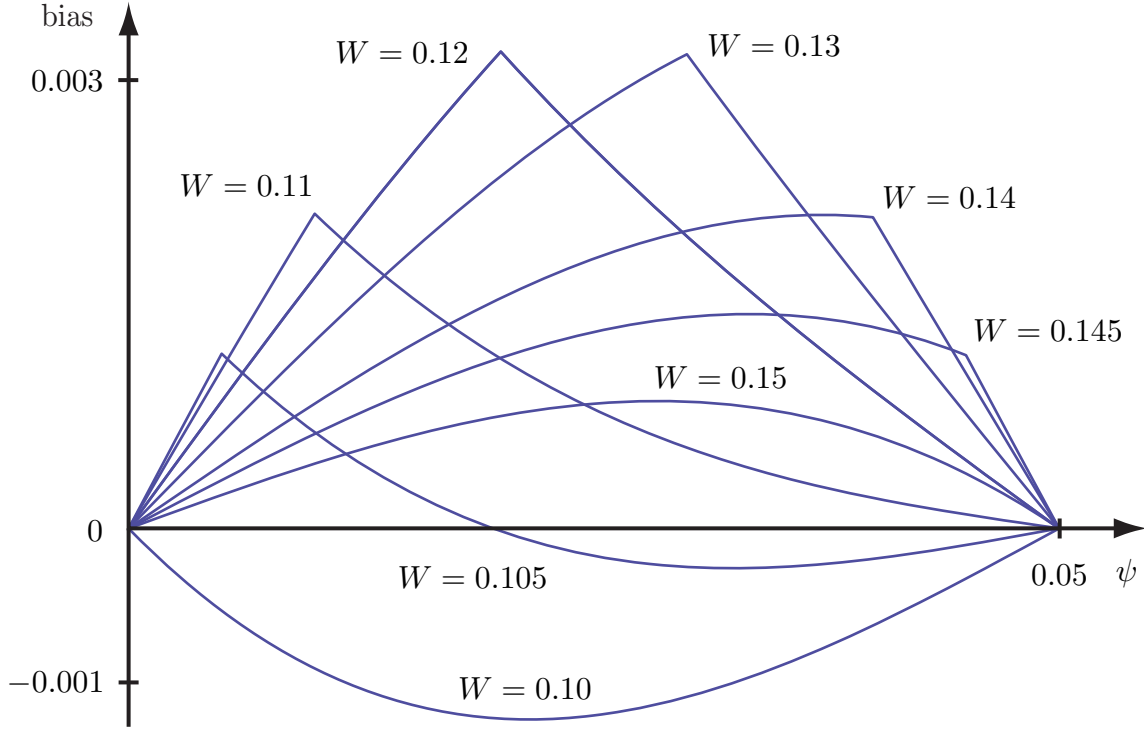
Figure 6-13: The bias exhibits a kink at $\psi^{\mathrm{NE}}$, the noise edge value of $\psi$, that in many but not all cases indicates the maximum bias. In these examples, the quantization step size is $\Delta = 0.1$ and the phase noise is uniformly distributed with width $2W$. The setting $W = 0.10$ results in the edge case in Figure 6-12a, so there is no kink. When $W = 0.145$, the maximum bias is not at $\psi^{\mathrm{NE}}$.

of the curve in the convolution diagram. Second, $P = \lfloor 2W/\Delta \rfloor$ is the number of quantization points under the noise distribution. Third, $\epsilon = 2W/\Delta - P$ indicates the remainder when dividing the noise width by $\Delta$. Note that $0 \le \epsilon < 1$.

**Claim 6.3.7.** *For the noise edge configuration where the phase noise $\phi_n$ is uniformly distributed with width $2W$ and $\psi = W$, the steady-state probability distribution is a scaled version of*

$$\pi_k^P = \begin{cases} (P+\epsilon)_k \, (P)_{P-k} & 0 \le k \le P \\ (P+\epsilon-1)_P & k = P+1 \\ 0 & otherwise \end{cases} , \tag{6.24}$$

*where $(x)_n = x(x-1)\cdots(x-n+1)$ denotes the falling factorial function.*

*Proof.* By Claim 6.3.5 and the symmetry of the transition probabilities (Claim 6.3.1), the steady-state probability distribution is given by

$$\pi_k = \prod_{i=k-\frac{Q}{2}+1}^{k-1} \lambda_i$$
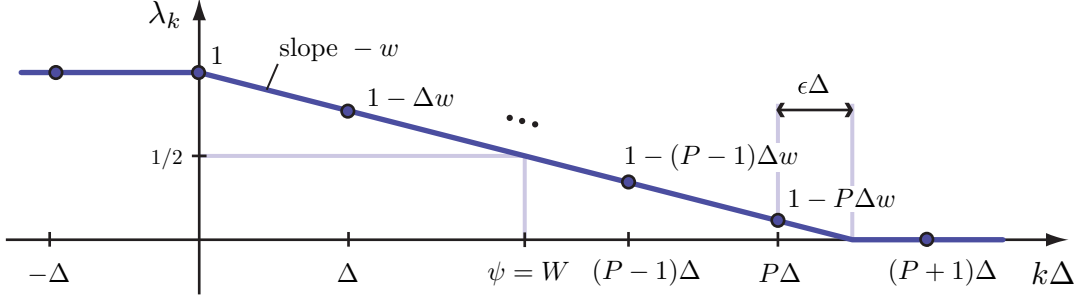
Figure 6-14: A portion of the convolution diagram for uniform noise of width $2W$ where $\psi = W$, so that the noise edge $\psi - W$ falls on the quantization point at the origin.

$$= \prod_{j=k-\frac{Q}{2}+1}^{-1} \lambda_j \prod_{i=0}^{k-1} \lambda_i$$

$$= \prod_{i=0}^{k-1} \lambda_i \prod_{j=k+1}^{\frac{Q}{2}-1} \bar{\lambda}_j. \tag{6.25}$$

Evidently $\pi_k = 0$ unless $0 \leq k \leq P + 1$.

From the convolution diagram, we compute

$$\lambda_k = 1 - k\Delta w, \tag{6.26}$$

for $0 \leq k \leq P$. We may limit the upper product index $Q/2 - 1$ in (6.25) to $P$, since the remaining values of $\bar{\lambda}_j$ are 1. We compute

$$
\begin{aligned}
\pi_k &= \prod_{i=0}^{k-1} \lambda_i \prod_{i=k+1}^{P} \bar{\lambda}_i \\
&= \prod_{i=0}^{k-1}(1 - i\Delta w) \prod_{i=k+1}^{P} i\Delta w \\
&= (\Delta w)^P \prod_{i=0}^{k-1}\left(\frac{1}{\Delta w} - i\right) \prod_{i=k+1}^{P} i \\
&= (\Delta w)^P \left(\frac{1}{\Delta w}\right)_k (P)_{P-k} \\
&= (\Delta w)^P (P + \epsilon)_k (P)_{P-k}, \tag{6.27}
\end{aligned}
$$

for $0 \leq k \leq P$. Finally, for $k = P + 1$, we have

$$\pi_{P+1} = (\Delta w)^P (P + \epsilon - 1)_P. \tag{6.28}$$

We divide out $(\Delta w)^P$ to complete the proof. $\qquad\square$

We will ignore $\pi_{P+1}$, which becomes exponentially negligible compared with most

of the other entries for large $Q$. When $\epsilon = 0$, the steady-state probability vector is a (scaled) binomial distribution with $P$ trials and success probability $1/2$. The steady-state probability distribution for the noise edge configuration thereby generalizes the binomial distribution.

**Reducing bias by increasing quantization steps**

Once we ensure that the phase noise standard deviation is in the low-bias noise range, we may still need to decrease the quantization step size $\Delta$ to ensure that the bias is small enough to meet application requirements. We prove that in the noise edge configuration the bias decays exponentially in $Q$. Our proof consists of three steps. First, we identify an appropriate decreasing sequence of step sizes $\{\Delta_n\}$ corresponding to linearly increasing $Q$. Second, we express the bias in terms of a generalized binomial sum $S_P$. And third, we use a recurrence relation for $S_P$ to show that the bias decays exponentially in $\Delta_n$, as $n \to \infty$. Although it is theoretically possible to eliminate bias by increasing $Q$, it may be impractical to implement a large number of quantization steps.

We begin our proof by identifying an appropriate decreasing sequence of step sizes $\{\Delta_n\}$. Although we want to know how bias decays with the number of quantization points $Q$, the steady-state probability distribution only depends on those quantization points under the noise distribution, through $P$. Clearly $Q$ and $P$ grow at the same rate, because the noise width is fixed at $2W$. Therefore, instead of taking the limit as $Q \to \infty$, we will equivalently decrease $\Delta$ so that at each step $P$ increments by 1 and $\epsilon$ is held constant.

**Claim 6.3.8.** *There is a decreasing sequence $\{\Delta_n\}$ such that $P_{n+1} = P_n + 1$ and $\epsilon_{n+1} = \epsilon$.*

*Proof.* We define

$$\Delta_{n+1} = \frac{\Delta_n}{\Delta_n w + 1}. \tag{6.29}$$

Then

$$P_{n+1} = \left\lfloor \frac{2W}{\Delta_{n+1}} \right\rfloor = \left\lfloor \frac{2W}{\Delta_n} + 1 \right\rfloor = P_n + 1. \tag{6.30}$$

Furthermore,

$$\epsilon_{n+1} = \frac{1}{\Delta_{n+1} w} - P_{n+1} = \frac{1}{\Delta_n w} - P_n = \epsilon_n. \tag{6.31}$$

$\square$

We now express the bias in terms of a generalized binomial sum $S_P$.

**Claim 6.3.9.** *The bias for the noise edge configuration is*

$$B_P = W \left( \frac{2PS_{P-1}}{S_P} - 1 \right), \tag{6.32}$$

*where*

$$S_P = \sum_{k=0}^{P} \pi_k^P$$

$$= \sum_{k=0}^{P} (P + \epsilon)_k \, (P)_{P-k}. \tag{6.33}$$

*Proof.* First, we apply Claim 6.3.7 to compute

$$\sum_{i=0}^{P} i\pi_i^P = \sum_{i=1}^{P} i \, (P + \epsilon)_i \, (P)_{P-i}$$

$$= (P + \epsilon)P \sum_{i=1}^{P} (P - 1 + \epsilon)_{i-1} \, (P - 1)_{P-i}$$

$$= (P + \epsilon)P \sum_{i=0}^{P-1} (P - 1 + \epsilon)_i \, (P - 1)_{P-1-i}$$

$$= (P + \epsilon)P S_{P-1}. \tag{6.34}$$

Because the support of the steady-state distribution begins at index 0, we can set the cut point $L = 0$. The bias in (6.22) is then

$$B_P = \frac{\sum_{i=0}^{P} i\pi_i^P}{\sum_{i=0}^{P} \pi_i^P} \Delta - \psi$$

$$= (P + \epsilon)P \frac{S_{P-1}}{S_P} \Delta - W$$

$$= W \left( \frac{2P S_{P-1}}{S_P} - 1 \right). \tag{6.35}$$

$\square$

Finally, we use a recurrence relation for $S_P$ to show that the bias decays exponentially in $Q$.

**Claim 6.3.10.** *The bias for the noise edge configuration decays to zero exponentially in $Q$.*

*Proof.* We equivalently show that $|B_P| \to 0$ exponentially in $P$. $S_P$ satisfies the recurrence relation (Appendix B.4.2)

$$S_P = 2P S_{P-1} + (P - 1 + \epsilon)_P. \tag{6.36}$$

Expanding, we see that $S_P \geq 2^P P!$. We substitute (6.36) into (6.32) to obtain

$$|B_P| = W \left| \frac{-(P - 1 + \epsilon)_P}{S_P} \right|$$

$$\leq W \frac{(P - 1 + \epsilon)_P}{2^P P!}$$

$$< \ \frac{W}{2^P}, \tag{6.37}$$

which decays exponentially in $P$. By Claim 6.3.8, $|B_P|$ also decays exponentially in $Q$. □

Although there is no closed-form expression for $S_P$, we may approximate $S_P \to 2^{P+\epsilon} P!$ for large $P$ (Appendix B.4.2).

## 6.4  Millimeter-wave radar design example

We have described the functional components of a phase tightening system and have presented performance analysis that suggests a design methodology. We now describe how to implement phase tightening for a millimeter-wave radar imaging system. First, we present a circuit diagram and describe the internal sources of phase noise (Section 6.4.1). Next, we apply our design methodology to select appropriate parameters (Section 6.4.2). Finally, we augment the circuit diagram to incorporate amplitude estimation and realize the full phase tightening implementation (Section 6.4.3).

### 6.4.1  Circuit implementation for phase estimation

The application requirements for millimeter-wave radar imaging typically specify the carrier frequency, bandwidth, and target phase noise. The carrier frequency is approximately 100 GHz, which is too high to sample directly. We therefore modulate the input down to an intermediate frequency $\omega_I$ that is high enough to support a modest bandwidth yet still make a narrowband assumption. The implementation is therefore divided into three regions: analog circuitry operating at the carrier frequency, analog circuitry operating at the intermediate frequency, and digital circuitry (Figure 6-15).

Instead of delaying the input signal directly, we drive the sampling circuitry of the ADC with a delay-locked loop (DLL). We use a common clock source for the modulation and sampling. A phase-lock loop (PLL) multiplies the clock frequency to the appropriate value for demodulation. The DLL performs a similar function for the ADC, additionally using input from the counter to apply a delay. We use a cyclic counter to store the phase tightening state, and a lookup table to translate the state to a control input for the DLL that indicates the appropriate delay to apply. We average the counter output and downsample by a factor of $L$ to obtain a phase estimate at a low data rate and with low phase noise.

Timing jitter is added by the PLL, modulator, DLL, and clock. The jitter contributes additional phase noise at the time scale in the region that the jitter is introduced. For example, 1 ps of jitter at the PLL output will contribute as much phase noise as $\omega_c/\omega_I$ ps of jitter at the ADC clock. For this reason, the aggregate phase noise is typically dominated by the components in the carrier region. We add all internal phase noise contributions to determine $\sigma_\phi$ in our phase tightening model.

The phase noise introduced by different oscillators and phase-lock loops has been extensively studied and characterized [63]. While our phase tightening model assumes
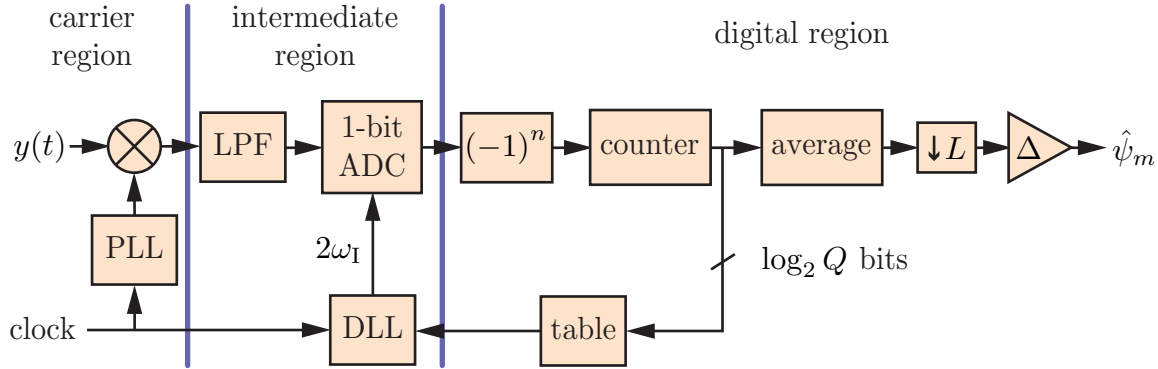
Figure 6-15: Our phase tightening implementation comprises (i) analog circuitry that operates at the carrier frequency to modulate the input, (ii) analog circuitry that operates at an intermediate frequency to apply a controlled delay prior to computing each 1-bit sample, and (iii) digital circuitry that accumulates and averages the samples to produce a phase estimate.

a fixed phase noise variance $\sigma_\phi^2$, the phase noise variance for free-running oscillators increases linearly with time, and the phase noise variance for PLLs of different orders evolves in more complicated ways. We must ensure that the phase noise stays within the low-bias noise range over the majority of the averaging window.

## 6.4.2   Methodology for parameter selection

We now select parameters for our circuit implementation based on hypothetical requirements. Specifically, suppose the carrier frequency is 100 GHz, the bandwidth is 1 kHz, and we desire phase noise with standard deviation less than $2\pi/10^4$, corresponding to 1 fs at 100 GHz. Also suppose that the total phase noise at the input has a maximum standard deviation of $\sigma_\phi = 2\pi/10^1$, corresponding to 1 ps at 100 GHz, after adding the contributions from the various internal analog components. We select the system parameters $Q$, $\omega_I$, and $L$ using the following methodology.

1. Decide how much averaging is required to tighten the phase noise to the desired target, thereby determining the downsampling factor $L$.

2. Select $\omega_I$ high enough to accommodate the bandwidth after downsampling, but low enough so that sampling at the Nyquist rate $2\omega_I$ is practical. A large $\omega_I$ may increase the phase noise contribution of the analog components in the intermediate region, in which case we may need to increase $L$.

3. Select an appropriate distribution $p_\phi$ to model the phase noise.

4. Using a noise edge configuration, estimate the minimum number of quantization steps $Q$ required to achieve a small bias relative to the target phase noise. We may prefer $Q$ to be a power of 2, so that no bits are wasted in a digital representation.

5. Simulate the bias for various $\psi$, to confirm the choice of $Q$.

6. Simulate the bias at the noise edge for various $\sigma_\phi$ to determine the low-bias noise range to operate over. If the range is not acceptable, increase $Q$ or dither as necessary.
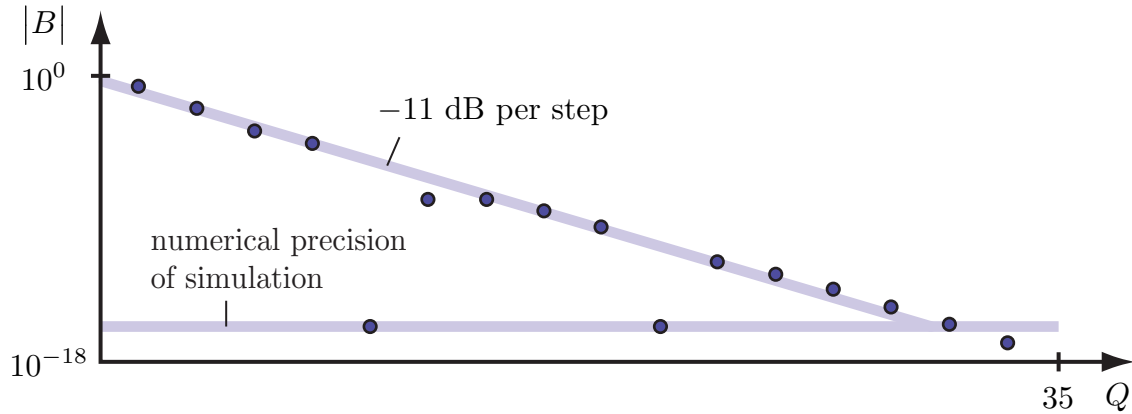


Figure 6-16: The bias decays exponentially in $Q$, at approximately $-11$ dB per quantization step.

We wish to reduce the standard deviation of the phase noise by 3 orders of magnitude, from $2\pi/10^1$ to $2\pi/10^4$. Averaging $n$ independent identically distributed zero-mean random variables reduces the standard deviation by $1/\sqrt{n}$, requiring approximately $10^6$ samples per phase estimate. Note this is an approximation because the samples form a Markov process and are not independent. Nonetheless, we set $L = 10^6$.

The lowest output data rate we can tolerate is 2 kHz, which is twice the bandwidth of our input signal. Therefore, we set $\omega_I = 2\pi \times 1$ GHz. We have $\omega_c/\omega_I = 100$, so the phase noise is dominated by contributions from the analog components in the carrier region.

We assume a Gaussian noise model for the phase noise [63]. However, our analysis of the noise edge configuration was based on a uniform distribution. We set $\psi = \sigma_\phi$ for our simulations of the noise edge configuration, in order to compare with our results above (Appendix C.2.4).

The bias decreases exponentially at the noise edge for Gaussian noise, just as it did for uniform noise (Figure 6-16). The bias decreases by 11 dB per quantization step. We judge the bias for $Q = 8$ to be too close to our phase noise target of $2\pi/10^4$, so we select $Q = 16$, the next power of 2.

We simulate the bias for various $\psi$ to determine that the bias is bound below $5 \times 10^{-9}$, which is well below our phase noise target, confirming our selection of $Q$ (Figure 6-17). Finally, we simulate the bias in the noise edge configuration for a range of $\sigma_\phi$, and determine that our operating range is for phase noise between $0.12\pi$ and $0.28\pi$, which is acceptable.
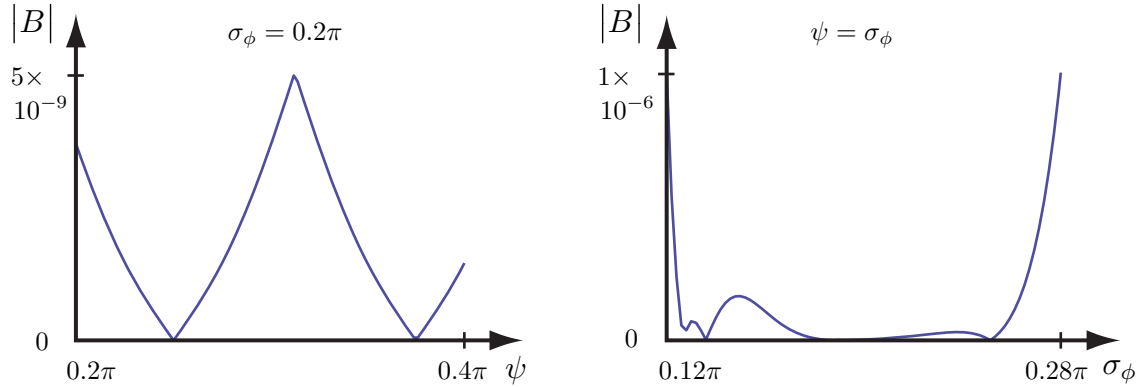
Figure 6-17: For $Q = 16$, the bias is acceptable for all values of the true phase $\psi$ and over a wide range of values for the standard deviation $\sigma_\phi$ of Gaussian phase noise.

## 6.4.3  Incorporating amplitude estimation

We have shown that by using simple digital logic with feedback, we can extract the information necessary to estimate the phase. By averaging and downsampling the phase estimates, we can reduce the effective phase noise and compress the data rate. To form a coherent image using a conventional beamformer or another quasi light field, however, we also need to estimate the received signal amplitude. We present three different ways of implementing amplitude estimation.

The simplest approach is to treat amplitude estimation as a separate problem and implement a separate system to provide amplitude estimates that are later synchronized with the phase estimates at a low data rate. There is some merit to this approach, as the phase estimation implementation we have described is somewhat specialized. However, if we are willing to incorporate an ADC with more than 1 bit of resolution, we can leverage the phase estimation circuitry in two different ways: by interleaving samples and timesharing.

**Interleaving samples for phase and amplitude estimation**

One way to incorporate amplitude estimation is to double the sampling rate to twice Nyquist. Then, once the phase tightening loop locks onto the zero-crossings of the input waveform, the two additional samples per period will align with the extrema of the sinusoid. We can obtain reliable amplitude estimates by computing the maximum absolute sample values at the extrema, taking care to account for outliers due to noise. The amplitude estimates can be downsampled and paired with the corresponding phase information (Figure 6-18). The ADC bit resolution $d$ should be selected based on application requirements. For example, in an imaging system, $d$ must be comparable to the desired bit depth of the resulting image.
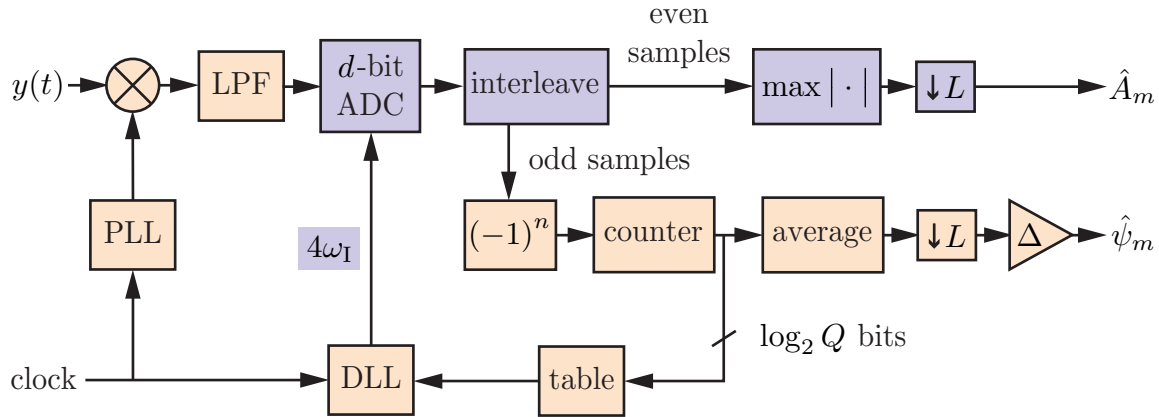
Figure 6-18: One way to incorporate amplitude estimation into phase tightening is to double the sampling rate and interleave the samples, so that the phase estimation logic processes the zero-crossing estimates, and the amplitude estimation logic processes the extrema.

### Timesharing samples for phase and amplitude estimation

One downside of interleaving is that fully half of the samples are devoted to amplitude estimation, which may benefit from the oversampled data substantially less than phase estimation. Timesharing allows us to more finely allocate samples between amplitude and phase estimation. By timesharing, we switch between two different modes of operation: the usual phase tightening and a programmable sweep mode, where we apply a rapidly varying controlled delay over several cycles to scan for the peaks of the input waveform (Figure 6-19). The ADC bit resolution $d$ should be selected based on application requirements, just as with interleaving. Instead of downsampling by $L$, we downsample both amplitude and phase estimates by a larger factor $L'$, to account for the additional samples used for amplitude estimation that do not contribute to reducing the phase noise.

## 6.5   Summary and future challenges

We have presented an architecture for phase tightening that uses simple digital logic to compute amplitude and phase estimates of a narrowband input signal, which we can digitally reconstruct at a lower carrier frequency and with less phase noise. Our phase tightening algorithm is one way to approximate the optimal phase estimator, embodied by the phase-lock loop, in the digital domain. We provide extensive performance analysis to characterize the stability, state dynamics, and bias of phase tightening. Provided that the phase noise is not too large, the loop is stable. When the number of quantization steps $Q$ is even and under modest assumptions on the support of the noise distribution, the evolution of the phase tightening state is accurately modeled by an ergodic Markov chain that has analytic expressions for its steady-state probability distribution. The bias is only low over a moderate range
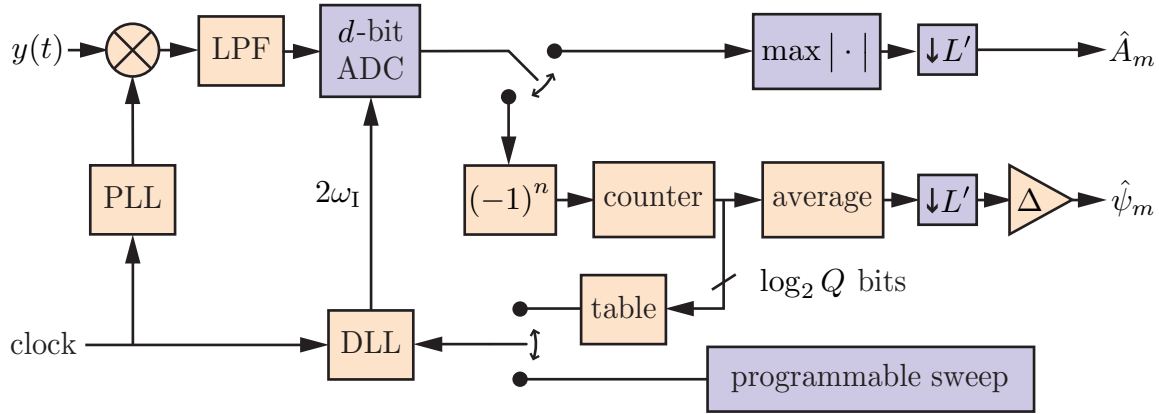
Figure 6-19: Timesharing is an alternative way to incorporate amplitude estimation into phase tightening. Periods of phase estimation are interspersed with amplitude estimation, where the DLL is programmed to sweep over several cycles of the input waveform to scan for the peaks.

of values of the phase noise standard deviation $\sigma_\phi$, and decays exponentially in $Q$ for the noise edge configuration. Consequently, our analysis provides guidance for selecting the number of quantization steps $Q$ to provide an operating range for $\sigma_\phi$ under which phase tightening achieves the desired reduction in phase noise. We have further synthesized our results to support a design methodology that we applied to a millimeter-wave radar imaging system. Phase tightening thereby enables the production of coherent images in scenarios where the phase noise of the received signals is otherwise too large.

While phase tightening provides an architecture for phase noise reduction, there are further implementation challenges in building a low phase noise millimeter-wave imaging system that scales well in cost and power consumption. First, any phase drift introduced by the analog circuitry must be controlled over the phase tightening averaging interval. Second, the imaging system must be calibrated to correct for any introduced constant phase offsets. Finally, the data rate from the individual sensors to the central processor that forms the image must be managed.

# Chapter 7

# Active imaging and opportunities in MIMO ultrasound

We have effectively been studying passive imaging systems, that treat the scene illumination as fixed and form a representation of the environment without modifying it. In contrast, active imaging systems create and propagate waves to illuminate the scene, and then form an image from the reflected and scattered response. Strictly speaking, the millimeter-wave radar imaging system in Chapter 6 is an active system, as the radar emits a narrowband signal whose reflections stand out over the relatively low levels of ambient black-body radiation. But this is about as sophisticated a use of active imaging technology as taking a picture with a flash on a conventional camera; we simply provide the illumination because a natural source of sufficient intensity is not already present. It is possible to capture more information from a scene by coordinating applied illumination with signal measurement and image formation, for both incoherent and coherent imaging.

The processing implicitly performed in active imaging is similar to that done explicitly in passive imaging. Signals emitted from multiple transducers propagate through a medium and combine at a target, according to the laws of physics. When the medium is free space, we model the propagation of each signal by a time delay and an attenuation factor, and the resulting signals superimpose linearly at the target. The incident wave at the target is thereby obtained from the transmitted signals by a beamforming operation, mathematically identical to the receiver processing we originally considered in Chapter 2. The connection between the transmission physics and the receiver processing is not a coincidence; we derived the beamformer from the same physical principles and approximations. Although we have demonstrated that quasi light fields offer more tradeoffs than conventional beamforming at the receiver, there may not be a physical interpretation for the generally complex-valued quasi light fields. Therefore, quasi light fields cannot be applied to signal transmission in the same way that they can be applied to signal reception.

A multiple-input multiple-output (MIMO) architecture [15] enables us to simulate the effects of different physical signal transmission strategies through processing at the receiver. The key idea is to design the transmitted signals so that their responses can be separated, providing us with a set of measurements for each transmitter act-

ing in isolation from the others. For example, in a MIMO radar application, each narrowband signal could be transmitted at a different carrier frequency. The receiver could apply a filter bank to undo the physical superposition at the target, allowing us to treat the target's response to each transmitter independently. Since each response is a convolution of the corresponding transmitted signal, and since convolution is a commutative operation, the receiver can filter each individual response to obtain the same result as if the transmitter had previously applied the same filter to the transmitted signal. Therefore, in a MIMO system, the benefits of focusing on both signal transmission and reception can be realized through processing performed exclusively at the receiver. Moreover, by storing the received waveforms, the receiver can simultaneously focus at different points using the same set of measurements. Although we will focus on coherent imaging, these ideas are also applicable to incoherent imaging: we may augment the plenoptic function to represent all potential incoherent pictures for all possible locations of a single illuminating point source, thereby enabling the relighting of scenes [70].

The ability of MIMO imaging systems to simultaneously focus on both transmission and reception is of importance to medical ultrasound, because of the slow speed of sound. In a traditional pulsed medical ultrasound imaging system, an array of transducers emits a series of pulses each focused at a different point in the tissue with a conventional beamformer. The transducer array receives the echo from each pulse before transmitting the next pulse. Once the ultrasound system has processed enough pulses to cover the region of interest, it forms an image. The resulting frame rate is thereby limited by the pulse propagation time and the size of the region of interest relative to the transmit beamformer resolution. Consequently, ultrasound systems used in applications such as blood velocity imaging that demand both high resolution and high frame rate forgo beamforming on transmit in order to generate an image from a single pulse [71]. MIMO processing can theoretically enable ultrasound systems to enjoy the benefits of beamforming on transmit while still generating an image from each pulse, to achieve both a high resolution and a high frame rate.

Medical ultrasound imaging is more challenging than the other applications we have considered, because the medium is filled with tissue to be imaged, rather than a handful of point targets in free space. Consequently, to accommodate ultrasound, we generalize the theory in two stages. First, we generalize from passive to active imaging, updating our formulation in Chapter 2 accordingly (Section 7.1). Second, we generalize from point targets to volume imaging, by framing ultrasound imaging as a classic system identification problem with novel physical constraints (Section 7.2). At such a high level of generality, specific imaging applications are no longer tightly coupled to each other, so that our quasi light field results do not immediately carry over to the specific challenges in ultrasound. Nonetheless, our general system identification perspective allows us to identify a more pressing area of improvement for ultrasound. After reviewing existing applications of transmit signal design to ultrasound (Section 7.3), we formulate a discrete estimation problem to compare the existing techniques with time-varying filters (Section 7.4).

## 7.1 MIMO processing and active imaging

In a multiple-input multiple-output (MIMO) architecture, multiple transmitters input signals into a physical process, and multiple receivers record the output. In radar, as in ultrasound, the objective is to infer something about the physical process from the inputs and outputs. In wireless communications, the objective is instead to recover the input from the output, in spite of the physical process. All three fields are concerned with how to best encode the inputs, estimate parameters from the outputs, and exploit the spatial diversity of the transmitters and receivers. Research in MIMO techniques has flourished in wireless communications [72] and radar [15]. We extend our model for source localization in Chapter 2 to describe the benefits of MIMO processing. We focus on the familiar scenario for radar before introducing the unique complications of ultrasound in Section 7.2.

We generalize our coherent wave processing model to accommodate the MIMO architecture. The $M$ sensors are now transducers additionally capable of transmitting signals $z_i(t)$. The $P$ sources are no longer sources, but rather targets at positions $\mathbf{r}_j^{\mathrm{P}}$ with reflectivities $\alpha_j$. The signal emanating from each target is a superposition of the signals $z_i(t)$ transmitted by each transducer:

$$x_j(t) = \alpha_j \mathbf{a}^{\mathrm{t}}(\mathbf{r}_j^{\mathrm{P}})\mathbf{z}(t), \tag{7.1}$$

where $\mathbf{z}(t) = [z_1(t) \cdots z_M(t)]^{\mathrm{t}}$. The array steering vector $\mathbf{a}(\mathbf{r}_j^{\mathrm{P}})$ in (7.1) is identical to the one in Chapter 2 by reciprocity: wave propagation from a source to a destination is identical to the reverse propagation from the destination to the source. Consequently, the sensor output due to a single target is

$$\mathbf{y}(t) = \alpha_j \mathbf{a}(\mathbf{r}_j^{\mathrm{P}})\mathbf{a}^{\mathrm{t}}(\mathbf{r}_j^{\mathrm{P}})\mathbf{z}(t) + \mathbf{v}(t). \tag{7.2}$$

Adding contributions due to all $P$ targets results in

$$
\begin{aligned}
\mathbf{y}(t) &= \left(\sum_{j=1}^{P} \alpha_j \mathbf{a}(\mathbf{r}_j^{\mathrm{P}})\mathbf{a}^{\mathrm{t}}(\mathbf{r}_j^{\mathrm{P}})\right)\mathbf{z}(t) + \mathbf{v}(t) \\
&= \mathbf{A}'\mathbf{z}(t) + \mathbf{v}(t).
\end{aligned} \tag{7.3}
$$

The MIMO model in (7.3) is similar to our previous passive imaging model in (2.7), with two important differences. First, the MIMO model trades the $P$ unknown source signals $\mathbf{x}(t)$ in the passive model for $M$ known transmit signals $\mathbf{z}(t)$ and $P$ new unknown target reflectivities $\alpha_j$. Second, the MIMO transducer array geometry involves the outer product of the steering vectors, $\mathbf{a}(\mathbf{r}_j^{\mathrm{P}})\mathbf{a}^{\mathrm{t}}(\mathbf{r}_j^{\mathrm{P}})$. In the far zone, the $M^2$ entries of this outer product correspond to having virtual sensors at all positions $\mathbf{r}_{i_1}^{\mathrm{M}} + \mathbf{r}_{i_2}^{\mathrm{M}}$, where $1 \leq i_1, i_2 \leq M$. When the transmit signals are orthogonal, MIMO thereby enables coherent image processing using a virtual sensor array that is larger than the physical array [73]. Among other things, the virtual array allows us to effectively focus on both transmission and reception by processing the received signal in different ways [74]. We gain the benefit of sweeping a focused beam across the

entire scene without having to wait for a wave to physically propagate to and from each area of interest sequentially. Intuitively, if the transmit signals are orthogonal, we can cleanly separate the system response that the physics of the scene imposes on each signal at the receiver. Furthermore, if the system is linear, we can recombine the separated responses to precisely simulate what would have happened had we instead coherently combined the transmitted signals at the system's input.

Quasi light fields are compatible with MIMO processing, allowing us to effectively focus both on signal transmission and reception, entirely through processing at the receiver. We illustrate the application of quasi light fields in a simple, symmetric scenario (Figure 7-1). We use a 5-element array to focus at an empty target with reflectivity $\alpha_1 = 0$, in the presence of an interfering target with reflectivity $\alpha_2 = 1$. We compare imaging performance using the spectrogram quasi light field in (4.26) with using the Wigner quasi light field in (4.27). The Wigner samples the scalar field every distance $d/2$, which is twice as fine as the spectrogram spacing $d$. We fix the width of the array in order to make a fair comparison. Therefore, the Wigner will make use of all five sensor measurements $y_{-1}$, $y_{-1/2}$, $y_0$, $y_{1/2}$, and $y_1$, while the spectrogram will only use $y_{-1}$, $y_0$, and $y_1$. Although a practical implementation might use a larger transducer array and apply weights to the sensor outputs to shape the response, our simple model suffices to illustrate the effect of quasi light field choice.
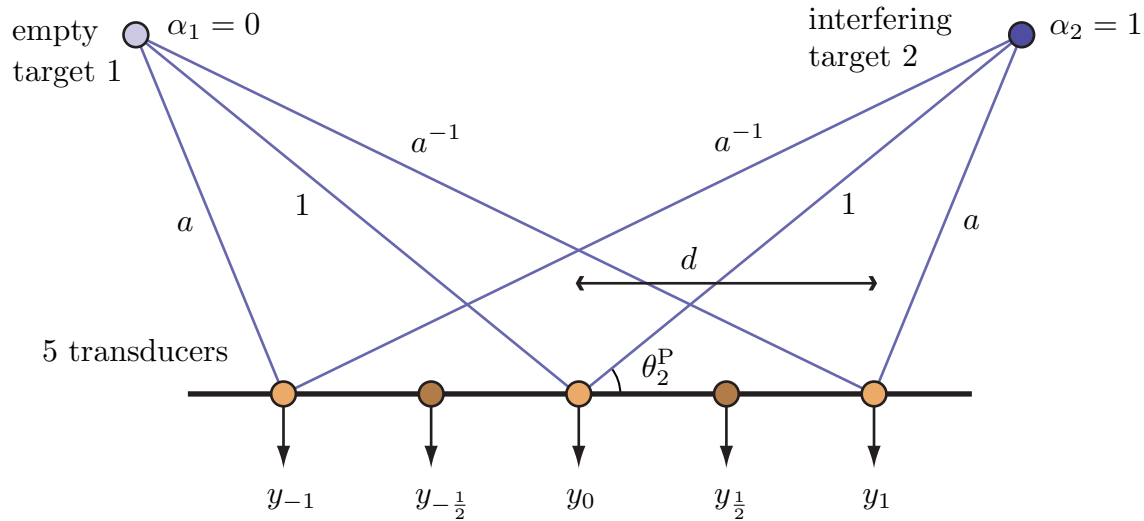


Figure 7-1: Quasi light fields can be applied at the receiver to realize the benefits of focusing on both signal transmission and reception. We compare different signal processing strategies using a simple 5-element array and two targets in a symmetric configuration. We judge the effectiveness of each strategy by comparing the extent to which a target with reflectivity $\alpha_2 = 1$ interferes in the pixel computed when the array is focused at an empty target with reflectivity $\alpha_1 = 0$.

We make several simplifying geometric assumptions. Both the empty target and interfering target are in the far zone, and are located symmetrically opposite each other with respect to the plane perpendicular to the array at its center. Therefore,

the array steering vector at the two target positions is, apart from a constant phase factor, a function of a single scalar parameter $a$. Specifically,

$$\mathbf{a}\left(\mathbf{r}_1^P\right) = \begin{bmatrix} a & a^{1/2} & 1 & a^{-1/2} & a^{-1} \end{bmatrix}^t \tag{7.4}$$

$$\mathbf{a}\left(\mathbf{r}_2^P\right) = \begin{bmatrix} a^{-1} & a^{-1/2} & 1 & a^{1/2} & a \end{bmatrix}^t, \tag{7.5}$$

where

$$a = \exp\left(-ikd\cos\theta_2^P\right) \tag{7.6}$$

indicates the relative phase offset between transducer 1 and transducer 0 with respect to the interfering target 2, as in (2.6). With these geometric assumptions and ignoring additive noise and the time dependence in our model in (7.3), the sensor outputs are a single-parameter function of the inputs:

$$\mathbf{y} = \mathbf{a}\left(\mathbf{r}_2^P\right)\mathbf{a}^t\left(\mathbf{r}_2^P\right)\mathbf{z}. \tag{7.7}$$

Conventional beamforming corresponds to computing the spectrogram quasi light field with a rectangular window. The discrete spectrogram quasi light field in (4.26) for pixels centered at each target position is

$$\ell_1^S = \left|\begin{bmatrix} a^{-1} & 0 & 1 & 0 & a \end{bmatrix}\mathbf{y}\right|^2, \tag{7.8}$$

$$\ell_2^S = \left|\begin{bmatrix} a & 0 & 1 & 0 & a^{-1} \end{bmatrix}\mathbf{y}\right|^2. \tag{7.9}$$

If the supplied input $\mathbf{z}$ is the same when computing each pixel, the intensity ratio for applying the spectrogram to signal reception is

$$\Gamma^{SR} = \frac{\ell_1^S}{\ell_2^S} = \frac{1}{9}\left|a^{-2} + 1 + a^2\right|^2$$

$$= \frac{a^{-4} + 2a^{-2} + 3 + 2a^2 + a^4}{9}. \tag{7.10}$$

Equation (7.10) corresponds to conventional beamforming at the receiver. If instead the inputs are delayed to focus on each target when the corresponding pixel is computed, so that

$$\mathbf{z} = \begin{bmatrix} a^{-1} & 0 & 1 & 0 & a \end{bmatrix}^t, \tag{7.11}$$

and

$$\mathbf{y} = \mathbf{a}\left(\mathbf{r}_2^P\right)\mathbf{a}^t\left(\mathbf{r}_2^P\right)\begin{bmatrix} a^{-1} & 0 & 1 & 0 & a \end{bmatrix}^t \tag{7.12}$$

for target 1, and similarly for target 2, then the intensity ratio is quadratically compressed:

$$\Gamma^{SS} = \frac{\ell_1^S}{\ell_2^S} = \left(\frac{a^{-4} + 2a^{-2} + 3 + 2a^2 + a^4}{9}\right)^2. \tag{7.13}$$

Equation (7.13) corresponds to applying the spectrogram quasi light field to both signal transmission and reception, and is traditionally achieved by conventional beam-

forming at both the transmitter and receiver. However, the same result can be obtained through MIMO processing exclusively at the receiver, by separating each transmitter response and recombining:

$$\mathbf{y} = a^{-1}\mathbf{a}\left(\mathbf{r}_2^{\mathrm{P}}\right)\mathbf{a}^{\mathrm{t}}\left(\mathbf{r}_2^{\mathrm{P}}\right)\mathbf{e}_1 + \mathbf{a}\left(\mathbf{r}_2^{\mathrm{P}}\right)\mathbf{a}^{\mathrm{t}}\left(\mathbf{r}_2^{\mathrm{P}}\right)\mathbf{e}_3 + a\mathbf{a}\left(\mathbf{r}_2^{\mathrm{P}}\right)\mathbf{a}^{\mathrm{t}}\left(\mathbf{r}_2^{\mathrm{P}}\right)\mathbf{e}_5. \tag{7.14}$$

Although the output expressions in (7.12) and (7.14) are mathematically identical, (7.12) is physically produced by wave propagation, while (7.14) is computed at the receiver.

Other quasi light fields can achieve better target isolation than the spectrogram, and are compatible with the MIMO processing architecture that enables simultaneous focusing on both signal transmission and reception. The discrete Wigner quasi light field in (4.27) for pixels centered at each target position is

$$\ell_1^{\mathrm{W}} = \begin{bmatrix} a^{-2}y_1^* & a^{-1}y_{1/2}^* & y_0^* & ay_{-1/2}^* & a^2y_{-1}^* \end{bmatrix}\mathbf{y} \tag{7.15}$$

$$\ell_2^{\mathrm{W}} = \begin{bmatrix} a^2y_1^* & ay_{1/2}^* & y_0^* & a^{-1}y_{-1/2}^* & a^{-2}y_{-1}^* \end{bmatrix}\mathbf{y}. \tag{7.16}$$

If the supplied input $\mathbf{z}$ is the same when computing each pixel, the intensity ratio for applying the Wigner to signal reception is

$$\Gamma^{\mathrm{WR}} = \frac{\ell_1^{\mathrm{W}}}{\ell_2^{\mathrm{W}}} = \frac{a^{-4} + a^{-2} + 1 + a^2 + a^4}{5}. \tag{7.17}$$

Although the spectrogram and Wigner intensity ratios in (7.10) and (7.17) are averages of the same conjugate pairs of complex numbers, the spectrogram average is weighted towards larger values, so that the Wigner can better suppress the interfering target. Furthermore, quasi light fields are compatible with MIMO processing, in that we can substitute (7.14) for $\mathbf{y}$ into (7.15), and similarly for (7.16), to obtain a hybrid system that applies the spectrogram to signal transmission and the Wigner to signal reception, achieving an even lower intensity ratio

$$\Gamma^{\mathrm{SW}} = \frac{\ell_1^{\mathrm{W}}}{\ell_2^{\mathrm{W}}} = \left(\frac{a^{-4} + a^{-2} + 1 + a^2 + a^4}{5}\right)\left(\frac{a^{-4} + 2a^{-2} + 3 + 2a^2 + a^4}{9}\right). \tag{7.18}$$

It is not as straightforward to apply other quasi light fields to signal transmission. There is a fundamental asymmetry in MIMO processing; non-physical quasi light fields cannot be directly applied to signal reception as they can to signal transmission.

Because of the extra degrees of freedom in designing signals at each transmitter, signal design for MIMO radar is more complicated than for conventional radar. Conventional radar tools have been extended to accommodate MIMO radar, but some are otherwise limited in their applicability. For example, MIMO extensions to the classic ambiguity function visually weight the tradeoffs between range estimation, resolution, and velocity estimation [75]. But despite their generality in accommodating multiple transmit signals, MIMO ambiguity functions are based on the assumption of a time-invariant matched filter receiver architecture, just like the classic ambiguity

function.

## 7.2 Ultrasound imaging as system identification

Medical ultrasound imaging exhibits unique characteristics and challenges that prevent the direct application of MIMO solutions from radar. The coherent waves in medical ultrasound do not propagate in free space, but rather traverse tissue that absorbs and scatters the waves at different depths [17]. We may model the measured scattered signal at a transducer element due to a transmitted signal from another transducer element as a linear time-invariant filtered version of the transmitted signal [76]. We superimpose the signals and include additive noise to obtain the analog model for the waveform measured at the $i^{\text{th}}$ transducer element

$$y_i(t) = \sum_{j=1}^{M} h_{ij}(t) * z_j(t) + v_i(t). \tag{7.19}$$

The different system impulse responses $h_{ij}(t)$ are defined by the inhomogeneities in the tissue as well as the electromechanical response and spatial geometry of the transducers. Unlike the MIMO radar formulation in Section 7.1 that models a finite number of unknown target reflectivities, the MIMO ultrasound formulation models a finite number of unknown impulse responses.

Conventional ultrasound performs single-input single-output system identification. Specifically, conventional ultrasound systems employ beamformers to focus along the same direction on both transmit and receive. The transmitter selects the $z_j(t)$ to be weighted versions of the same pulse signal $s(t)$, and the receiver coherently combines the $y_i(t)$ to form a single received signal $r(t)$. The receiver then convolves $r(t)$ with a matched filter with impulse response $s(-t)$ and displays the output. The transmit signal $s(t)$ is designed to have a narrow impulse-like autocorrelation function, so that the matched filter estimates the impulse response of the tissue along the beam of focus, which is a superposition of the $h_{ij}(t)$ applied with the beamformer weights. Each time sample of the matched filter output corresponds to the tissue at the appropriate depth along the beam, according to the propagation time of the focused pulse. The matched filter output thereby defines a scan line that begins at the transducer-tissue interface and extends deeper into the tissue along the steering angle. To keep the entire scan line in focus, ultrasound systems employ a time-varying beamformer at the receiver, to slide the point of focus along the line at the same rate that scattered waves from tissue at those points arrive back at the transducer. Ultrasound systems assemble scan lines at different angles to produce a B-mode (brightness-mode) image, which medical practitioners use to visualize a cross section of the tissue.

MIMO ultrasound performs multiple-input multiple-output system identification. Specifically, we choose the inputs $\mathbf{z}(t)$ and estimate each $h_{ij}(t)$ from the measured outputs $\mathbf{y}(t)$. We can weight and combine the estimated $h_{ij}(t)$ to produce any of the scan lines that a conventional ultrasound system estimates. General-purpose system identification techniques [77] are not well-suited to MIMO ultrasound because

of the unique constraints placed on the allowable system inputs.  The Food and
Drug Administration (FDA) imposes limits on the allowable peak spatial pressure
and power in human tissue, and both physical and application requirements further
impose limits on transmit signal bandwidth and time duration. The FDA limits the
mechanical index (MI), which is the peak rarefactional pressure divided by the square
root of the center frequency.  A high MI can result in microbubble cavitation [78].
The FDA also limits the average pulse power through the spatial peak pulse average
intensity $I_{\mathrm{sppa}}$ and the maximum intensity $I_{\mathrm{m}}$, as well as the time-averaged power
through the spatial peak temporal average intensity $I_{\mathrm{spta}}$ [2]. Ultrasound transducers
have a bandpass response and thereby limit the transmit signal bandwidth. Tissue
attenuation due to absorption and scattering is frequency-dependent, so that higher
frequencies decay more rapidly with depth [17]. Finally, the application's frame rate
limits the maximum time duration of the transmit signals.

The input signal constraints in ultrasound result in a rich set of engineering trade-
offs. To maximize the signal-to-noise ratio of the resulting image, we deliver as much
power to the tissue as the constraints allow [79]. All FDA pressure and power con-
straints are applied at the spatial peak, so the power should be equally distributed
throughout the tissue, rather than focused at a particular point on transmit.  The
constraints on pulse design interfere with each other.  The MI constraint is applied
to the temporal peak, suggesting that the power should be equally distributed over
the duration of the pulse.  Additionally, the pulse should use the available bandwidth
of the transducer and be short enough in duration to achieve the desired frame rate
for the application, and should otherwise be shaped to satisfy the FDA average pulse
power constraints on $I_{\mathrm{sppa}}$ and $I_{\mathrm{m}}$.  While it is possible to numerically optimize the
pulse to maximize power transport given these constraints, tissue attenuation fur-
ther complicates the tradeoffs. Specifically, power delivery decays with tissue depth,
with higher frequencies decaying more rapidly.  A pulse should therefore contain more
power in lower frequency components to achieve higher depth penetration.  Finally,
the FDA time-averaged power constraint on $I_{\mathrm{spta}}$ limits the frame rate through the
pulse repetition rate.

## 7.3  Existing ultrasound transmit signal design

Existing approaches to MIMO ultrasound do not systematically evaluate the entire
tradeoff space, but rather sequentially optimize over local constraints until the design
possibilities are exhausted. We describe three stages of conceptual development of
MIMO ultrasound ideas. First, a strict matched filter architecture coupled with a
desire for excellent axial resolution leads to linear frequency modulated, or chirp,
transmit waveforms. Second, slightly relaxing the matched filter architecture enables
perfect axial resolution using complementary Golay encoded transmit waveforms.
Third, multiple waveforms can be transmitted so that their responses are separated
at the receiver provided that the cross-correlation between the waveforms is low.

The matched filter receiver architecture is widely justified for use in ultrasound
because it is the linear time-invariant filter that maximizes the signal-to-noise ratio

at the output in the presence of white noise [79]. The matched filter presupposes a linear time-invariant filter architecture [80]. With a matched filter, the width of the autocorrelation function of the transmitted signal, $s(t) * s(-t)$, determines the axial resolution of the imaging system. Perfect axial resolution, corresponding to $s(t) * s(-t) = \delta(t)$, is impossible with a matched filter architecture, since (i) the FDA limits the peak pressure through the MI and (ii) the transducer imposes a finite bandwidth on the transmitted pulse. Axial resolution is instead maximized by using all the available bandwidth and spreading the transmitted power out over a longer time interval, which is a technique called pulse compression. Researchers have determined that under these design criteria, an appropriately weighted chirp waveform performs better than many alternatives [81].

Complementary codes, such as Golay codes, enable perfect axial resolution with a different receiver architecture. Instead of transmitting a single pulse with an impulse-like autocorrelation function, the transmitter sequentially transmits two different pulses whose autocorrelation functions sum to an impulse. For example, the sequence $(1, 1)$ has autocorrelation $(1, 2, 1)$ and the sequence $(-1, 1)$ has autocorrelation $(-1, 2, -1)$. When added together, the autocorrelation functions sum to the impulse $(0, 4, 0)$. The transmitter sends the first pulse $(1, 1)$ followed by the second pulse $(1, -1)$. The receiver applies the corresponding matched filters to each of the two responses and then sums their outputs [82]. The pulses must be separated in time so that the responses do not overlap, but cannot be too far apart lest nonstationarities in the tissue distort the output. Unlike the conventional perspective just described, we instead view both transmitted waveforms $(1, 1)$ and $(1, -1)$ as different portions of the same single transmitted pulse. The receiver applies a linear time-varying filter, rather than a matched filter, to this composite pulse, by changing the filter impulse response from $(1, 1)$ to $(-1, 1)$ during the processing. Our perspective highlights the facts that (i) a time-varying filter, unlike a time-invariant filter, can achieve perfect axial resolution and (ii) the separation between $(1, 1)$ and $(1, -1)$ is a deliberate design decision not to transmit more power in between.

Both the matched filter chirp architecture and the complementary Golay code architecture are formulated using a subset of the constraints and optimization criteria for the entire ultrasound system design problem. It is consequently challenging to compare these two architectures with respect to the remaining criteria, such as depth penetration. For example, it is straightforward to distribute the energy of a chirp waveform evenly over the bandwidth of the transducer by sweeping over the corresponding frequency range at a constant rate. In contrast, a Golay code is a binary representation that must be modulated to produce an analog waveform to transmit, allowing considerable design freedom to shape the pulse and define the frequency content to compare with a chirp. A naïve modulation scheme instead results in a concentrated frequency spectrum with limited depth penetration that is inferior to that of the chirp [79]. Furthermore, the matched filter should ideally be matched to the attenuated signal interrogating the tissue of interest, rather than the originally transmitted waveform. To first order, the attenuated signal resembles a Doppler-shifted version of the transmitted signal [2]. Filtering with a chirp is naturally robust to Doppler shifts, while additional pulses must be added to make the complementary

Golay receiver similarly robust [83]. Therefore, a broader perspective of the entire ultrasound tradeoff space is necessary to make fair comparisons between different architectures.

Existing MIMO ultrasound implementations simultaneously transmit signals in different ways to achieve different tradeoffs. The simplest approach is synthetic transmit aperture (STA) imaging, where the transducer elements transmit waveforms sequentially over time [84]. STA imaging thus achieves signal orthogonality by ensuring that the received signals do not overlap in time. In comparison with conventional ultrasound systems that emit one pulse per scan line and beamform on both transmit and receive, STA systems emit one pulse per transducer element. Thus although an STA system may achieve a higher frame rate than conventional ultrasound, an STA system exhibits a lower signal-to-noise ratio. Some ultrasound systems forgo beamforming on transmit entirely, along with the corresponding improvement in resolution, to obtain both a higher frame rate and signal-to-noise ratio than an STA system [71]. Other ultrasound systems transmit signals from several transducer elements at the same time and apply matched filters at the receiver, achieving both higher frame rates and higher signal-to-noise ratios than STA imaging, but suffering from interference due to poor signal cross-correlation [84, 85].

## 7.4   Discrete estimation with time-varying filters

We recognize that both the matched filter chirp architecture and the complementary Golay code architecture are two specific instances of a more general time-varying filter architecture. We describe a framework that incorporates both approaches and suggests new alternatives. We do not systematically address all aspects of the entire ultrasound tradeoff space, but rather make heuristic decisions to arrive at a simple framework that describes both architectures. First we describe how to modulate digital codewords to produce analog waveforms to transmit, to compare the frequency content of chirps and Golay codes in the same signal space (Section 7.4.1). Then we formulate a discrete estimation problem using samples of the measured signals (Section 7.4.2). We show how the matched filter and complementary Golay code architectures solve the discrete estimation problem, and consider the performance of alternative time-varying filters (Section 7.4.3).

### 7.4.1   Modulating digital codewords

We modulate digital codewords to produce analog waveforms for ultrasound transmission in order to compare chirps and Golay codes in the same signal space. Golay codes are sequences of binary digits that we represent by $-1$ and $1$. Between codewords in a complementary pair, there is no transmission, which we represent by a sequence of 0s. Therefore, we construct a class of analog transmit waveforms by modulating a signal $p(t)$ with a sequence $\mathbf{z}_j = [z_j^0 \cdots z_j^{L-1}]^{\mathrm{t}}$ of length $L$ by scaling and replicating $p(t)$ every time $T$. Each $z_j^l$ is either $-1$, $0$, or $1$. The signal transmitted by the $j^{\mathrm{th}}$

transducer element is thereby

$$z_j(t) = \sum_{l=0}^{L-1} z_j^l p(t - lT).$$  (7.20)

The frequency spectrum of the transmitted waveforms is

$$\int z_j(t)e^{-i2\pi ft}\mathrm{d}t = P(f)\sum_{l=0}^{L-1} z_j^l \exp(-i2\pi flT)$$
$$= P(f)Z_j(f),$$  (7.21)

where $P(f)$ is the Fourier transform of $p(t)$, and $Z_j(f)$ is the discrete-time Fourier transform of $\mathbf{z}_j$.

Designing $p(t)$ is a nontrivial problem. The FDA constraint on MI limits the extent to which shifted copies of $p(t)$ may overlap. Energy outside the transducer bandwidth will be wasted, and the frequency distribution of the transmitted energy determines the penetration depth. Both $p(t)$ and the digital codewords $\mathbf{z}_j$ determine the spectral content of the transmitted signals. We can therefore first design the codewords and subsequently design $p(t)$ to achieve the desired tradeoffs between energy efficiency and depth penetration.

### 7.4.2   Imaging as discrete estimation

We sample the measured signals in (7.19) to formulate ultrasound imaging as a discrete estimation problem. Recall that MIMO imaging corresponds to estimating the $h_{ij}(t)$ given $\mathbf{y}(t)$. Since each additional $y_i(t)$ adds $M$ new systems $h_{ij}(t)$ driven by the same inputs $\mathbf{z}(t)$ to estimate, we can focus on the estimation problem for fixed $i$ without loss of generality. We consequently omit the index $i$ from (7.19) and process a single transducer element measurement $y(t)$.

We discretize the model by sampling $y(t)$. We first substitute the modulated signal in (7.20) into the analog model in (7.19) to obtain

$$y(t) = \left\{ \sum_{j=1}^{M} \sum_{l=0}^{L-1} z_j^l \int h_j(t - \tau - lT)p(\tau)\,\mathrm{d}\tau \right\} + v(t).$$  (7.22)

We sample the measured signal assuming that it decays to zero over $Q + L$ samples, where $Q$ is the effective number of filter taps, so that

$$y^k = y(kT) = \sum_{j=1}^{M} \sum_{l=0}^{L-1} z_j^l h_j^{k-l} + v(kT)$$
$$= \sum_{j=1}^{M} (\mathbf{z}_j * \mathbf{h}_j)^k + v^k,$$  (7.23)

where $\mathbf{h}_j = [h_j^0 \cdots h_j^{Q-1}]^{\mathrm{t}}$, $v^k = v(kT)$, $(\mathbf{z}_j * \mathbf{h}_j)^k$ is the $k^{\mathrm{th}}$ sample of $\mathbf{z}_j * \mathbf{h}_j$, and

$$h_j^k = \int h_j(kT - \tau)p(\tau)\,\mathrm{d}\tau. \tag{7.24}$$

In matrix form, our discrete model becomes

$$\mathbf{y} = \mathbf{Zh} + \mathbf{v}, \tag{7.25}$$

where

$$
\begin{aligned}
\mathbf{y} &= [y^0 \cdots y^{Q+L-1}]^{\mathrm{t}}, \\
\mathbf{Z} &= [\mathbf{Z}_1 \cdots \mathbf{Z}_M], \\
\mathbf{h} &= [\mathbf{h}_1^{\mathrm{t}} \cdots \mathbf{h}_M^{\mathrm{t}}]^{\mathrm{t}}, \\
\mathbf{v} &= [v^0 \cdots v^{Q+L-1}]^{\mathrm{t}},
\end{aligned}
\tag{7.26}
$$

and

$$
\mathbf{Z}_j = \begin{bmatrix}
z_j^0 & & & \\
\vdots & \ddots & & \\
z_j^{L-1} & & z_j^0 & \\
& \ddots & & \vdots \\
& & & z_j^{L-1}
\end{bmatrix}. \tag{7.27}
$$

The discrete version of the MIMO ultrasound imaging system identification problem is to choose the inputs $\mathbf{Z}$ and estimate $\mathbf{h}$ from the measured outputs $\mathbf{y}$ in (7.25).

### 7.4.3   Comparing receiver architectures

We now compare the matched filter architecture with the complementary Golay code architecture using our discrete system identification formulation. We do not consider the entire ultrasound tradeoff space, but rather focus on the specific tradeoff between axial resolution and signal-to-noise ratio. We consider a simple example with $M = 1$ transducer element, $Q = 2$ filter taps, a length $L = 6$ transmit sequence, and unit noise variance $\sigma_v^2 = 1$. We treat the unknown system $\mathbf{h}$ as a deterministic pair of parameters to estimate. Although a more sophisticated Bayesian estimation framework could incorporate prior information known about $\mathbf{h}$, our nonrandom parameter estimation framework captures the tradeoff between axial resolution and signal-to-noise ratio.

We derive the minimum variance unbiased estimator for $\mathbf{h}$ to serve as a baseline for comparison with the matched filter and Golay code architectures. Estimation from a linear transformation of unknown parameters corrupted by Gaussian noise, as in (7.25), is a classic problem for which the maximum likelihood estimator achieves the Cramér-Rao bound [80]. The maximum likelihood estimator is therefore the minimum variance unbiased estimator. We have already derived this estimator in Chapter 2, in the context of maximum likelihood source location estimation. From

(2.11), the minimum variance unbiased estimator of $\mathbf{h}$ is

$$\hat{\mathbf{h}}_{\text{MVU}} = \left( \mathbf{Z}^{\text{t}} \mathbf{Z} \right)^{-1} \mathbf{Z}^{\text{t}} \mathbf{y} \tag{7.28}$$

and has error covariance matrix

$$\mathbf{\Lambda}_{\text{MVU}} = \left( \mathbf{Z}^{\text{t}} \mathbf{Z} \right)^{-1}. \tag{7.29}$$

The variance of each filter tap estimate is given by the diagonal entries of $\mathbf{\Lambda}_{\text{MVU}}$. By computing $\mathbf{\Lambda}_{\text{MVU}}$ for all sequences for $\mathbf{Z}$ of length $L = 6$, we determine that the minimum variance for each filter tap estimate is $6/35$. One of the codewords that achieves the minimum variance is the truncated Barker codeword $(1, 1, 1, -1, -1, 1)$ corresponding to

$$\mathbf{Z}_{\text{B}} = \begin{bmatrix} 1 & 1 & 1 & -1 & -1 & 1 & 0 \\ 0 & 1 & 1 & 1 & -1 & -1 & 1 \end{bmatrix}^{\text{t}}. \tag{7.30}$$

In comparison with the minimum variance unbiased estimator for $\mathbf{h}$ in (7.28), the matched filter processes the measured signals with $\mathbf{Z}^{\text{t}}$:

$$\hat{\mathbf{h}}_{\text{MF}} = \frac{1}{L} \mathbf{Z}^{\text{t}} \mathbf{y}. \tag{7.31}$$

With the truncated Barker codeword as the system input, the matched filter output is

$$\hat{\mathbf{h}}_{\text{MF}} = \begin{bmatrix} 1 & 1/6 \\ 1/6 & 1 \end{bmatrix} \mathbf{h} + \frac{1}{6} \mathbf{Z}_{\text{B}}^{\text{t}} \mathbf{v}. \tag{7.32}$$

The matched filter estimate is biased, so that $E[\hat{h}_{\text{MF}}^0] = h^0 + h^1/6$ and $E[\hat{h}_{\text{MF}}^1] = h^1 + h^0/6$. The variance of each filter tap estimate is $1/6$, which is slightly lower than that for the minimum variance unbiased estimator:

$$\hat{\mathbf{h}}_{\text{MVU}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{h} + \left( \mathbf{Z}_{\text{B}}^{\text{t}} \mathbf{Z}_{\text{B}} \right)^{-1} \mathbf{Z}_{\text{B}}^{\text{t}} \mathbf{v}, \tag{7.33}$$

where

$$\left( \mathbf{Z}_{\text{B}}^{\text{t}} \mathbf{Z}_{\text{B}} \right)^{-1} \mathbf{Z}_{\text{B}}^{\text{t}} = \frac{1}{35} \begin{bmatrix} 6 & 5 & 5 & -7 & -5 & 7 & -1 \\ -1 & 5 & 5 & 7 & -5 & -7 & 6 \end{bmatrix}. \tag{7.34}$$

The bias of each estimator corresponds to the axial resolution of the ultrasound imaging system, and the variance of each estimator corresponds to the signal-to-noise ratio of the ultrasound imaging system. An unbiased estimator corresponds to perfect axial resolution, and low estimator variance corresponds to high signal-to-noise ratio. By using the time-varying filter in (7.34), the minimum variance unbiased estimator sacrifices little signal-to-noise ratio (a variance of $6/35$ versus $6/36$) to achieve perfect axial resolution (zero bias), compared with the matched filter.

The Golay code architecture offers a different comparison. We select the pair of

codewords $(1, 1)$ and $(1, -1)$, separated by two 0s

$$\mathbf{Z}_{\mathrm{G}} = \left[ \begin{array}{ccccccc} 1 & 1 & 0 & 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & -1 \end{array} \right]^{\mathrm{t}}. \tag{7.35}$$

The separation is such that the Golay code receiver described in Section 7.3 can be implemented by processing the measured signals with $\mathbf{Z}_{\mathrm{G}}^{\mathrm{t}}$:

$$\hat{\mathbf{h}}_{\mathrm{GC}} = \left[ \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right] \mathbf{h} + \frac{1}{4} \mathbf{Z}_{\mathrm{G}}^{\mathrm{t}} \mathbf{v}. \tag{7.36}$$

The Golay code estimator is unbiased, but the filter tap estimates have a high variance of $1/4$. Although the Golay code architecture achieves perfect axial resolution, so does the minimum variance unbiased estimator, which additionally achieves a significantly higher signal-to-noise ratio. Therefore, the benefits of the Golay code architecture can be improved upon by inserting more power in between the codeword pair, at the cost of implementing a time-varying filter at the receiver.

## 7.5   Summary and future challenges

We have generalized our imaging framework to accommodate active illumination, allowing us to jointly design focusing strategies on both signal transmission and reception. While traditional signal transmission focusing relies on the physical combination of waves incident at the target, a MIMO architecture isolates the contribution from each transmitter and recombines the pieces at the receiver, by appropriately coding the transmitted signals. A MIMO architecture thereby enables us to apply complex-valued quasi light fields to focus on signal reception in conjunction with a physical quasi light field to focus on signal transmission.

We have further generalized our imaging framework to accommodate medical ultrasound by formulating imaging as a system identification problem. Medical ultrasound involves the characterization of a volume of tissue mass, which is more naturally modeled as an unknown system through which acoustic waves propagate, rather than as a collection of point targets. At this high level of generality, ultrasound is not tightly coupled to other imaging applications such as millimeter-wave radar, and our quasi light field results do not immediately carry over to the challenges of ultrasound. However, the system identification perspective has allowed us to identify a more pressing area of improvement.

The physical constraints in medical ultrasound define a rich tradeoff space that research has not yet fully explored. Many existing ultrasound systems presume a linear time-invariant filter architecture, and consequently employ matched filters to maximize the output signal-to-noise ratio. Other ultrasound systems employ complementary Golay codes and a corresponding receiver architecture to achieve perfect axial resolution. In order to compare these two architectures, we have described how to modulate digital codewords to form analog transmit signals, formulated a discrete version of the system identification problem, and proposed a time-varying filter re-

ceiver architecture. We have shown how a time-varying filter can achieve the perfect axial resolution of a Golay code with a signal-to-noise ratio that is only slightly worse than the matched filter.

Our time-varying filter results do not exploit the full power of MIMO processing, but rather make a first step in this direction. Future work is required to navigate a complex tradeoff space and provide a methodology to explore new receiver architectures, design codewords, and craft modulation schemes to achieve specific tradeoffs in a systematic, comprehensive manner.

# Chapter 8

# Conclusions and perspectives

We have presented a unified theory of image formation that highlights opportunities at the intersection of coherent and incoherent wave processing. On the one hand, traditional coherent images result from visualizing the objective function for the maximum likelihood estimator of the locations of multiple sources. On the other hand, traditional incoherent images result from a physical model of the operation of the human visual system. Both coherent and incoherent imaging systems process directional illumination information, by directly measuring phase and capturing the traditional light field, respectively. Quasi light fields relate the different image formation perspectives and enable the formation of coherent images by integrating bundles of rays, so that the coherent image formation process converges to the incoherent image formation process in the limit of small wavelength and globally incoherent light. A radar system searching for well-separated airborne targets thereby performs the same function as a human eye that forms an image, under the limiting process applied to quasi light fields.

As technology improves, coherent millimeter-wave imaging systems are uniquely positioned to benefit from a deeper understanding of the relationship between coherent and incoherent image formation. This potential is due to the ability to fabricate both the circuitry for millimeter-wave radar and low-power digital logic for millimeter-wave ultrasound at low cost. Opportunities for millimeter-wave imaging include using appropriate quasi light fields to surpass the Fourier uncertainty limitation in the conventional beamformer to better jointly localize energy in position and direction. More generally, quasi light fields enable the application of traditional light field processing techniques to coherent imaging problems. For example, landmine detection from coherent radiation [86] may directly benefit from light field pattern recognition techniques [87].

We have pursued opportunities at the intersection of coherent and incoherent wave processing by addressing the practical challenges faced when implementing high-performance millimeter-wave systems. Phase noise can degrade image quality, especially when using arrays with a small number of sensors. We have characterized the effect of the number of sensors, beamforming interval, and phase noise on the beam pattern tilt, sidelobe suppression, and main lobe width. We have provided recommendations on how to provision an imaging system to operate in a stable region where

sufficient time averaging maintains the integrity of the beam pattern in the presence of phase noise. We have additionally introduced a digital phase tightening algorithm to reduce excessive phase noise at each sensor. Phase tightening oversamples the received signals and applies simple digital logic in a feedback loop to average the phase noise down to a lower value. We have synthesized our results to support a design methodology for a millimeter-wave radar system.

We have identified future research opportunities in MIMO ultrasound, which we have formulated as a system identification problem with unique constraints and trade-offs. We have characterized the traditional matched filter and complementary Golay code architectures as estimators that achieve different tradeoffs between bias and variance. We have shown how a time-varying filter can achieve the perfect axial resolution of a Golay code with a signal-to-noise ratio that is only slightly worse than the matched filter. Additional research opportunities lie in other implementation challenges for millimeter-wave technology, such as sensor array calibration. For example, recent results in distributed beamforming show that feedback can be employed to guide random phase changes to convergence [88]. These results may be extended to our scenario, where the high-frequency content of a formed image may guide random phase changes applied to each sensor to ultimately converge to the correct calibration. Finally, quasi light fields allow us to translate future techniques between coherent and incoherent imaging. More research and experience will provide more guidance on the selection of quasi light fields for future applications.

Our investigation supports five architectural principles relevant to implementing imaging systems. First, push all image formation logic from the transmitter to the receiver, as MIMO processing provides complete array steering flexibility that scales with processing power. Second, use digital logic to compensate for the imperfections of high-frequency analog circuitry, as digital phase tightening can reduce the phase noise introduced by a PLL operating at 100 GHz. Third, achieve high precision using a large number of low-precision samples, as both a large number of sensors and a high sampling rate reduce the impact of phase noise through averaging. Fourth, process the data where it resides to minimize inter-component communication, as compressing sensor measurements to the minimum required samples puts less strain on the communication links to the central processor that forms the image. Fifth, find common points of comparison to relate technologies, as the light field allows us to compare the operation and performance of coherent and incoherent imaging systems.

# Appendix A

# Optics toolbox

For completeness, we describe the background we presume and the tools we use from optics. We first describe the complex analytic representation of signals (Section A.1). Then we describe the theoretical foundation of geometric optics (Section A.2). We outline scalar diffraction theory and key results from Fourier optics (Section A.3). We briefly introduce the tools used in optical coherence theory (Section A.4). Finally, we describe the operators that Agarwal et al. used to generate extended light fields (Section A.5) [49].

## A.1  Complex analytic representation

Although the physical phenomena we are interested in are naturally modeled with real-valued signals, complex-valued signals are more convenient to work with. We obtain a complex analytic representation of a real signal by eliminating redundancy in the frequency domain. Specifically, suppose that $U^{\mathrm{r}}(\mathbf{r}, t)$ is a real signal that represents the physical phenomenon of interest, such as a scalar wave. We can think of $U^{\mathrm{r}}(\mathbf{r}, t)$ as a Cartesian component of the electric field, or the air pressure, at position $\mathbf{r}$ and time $t$. The complex analytic representation is defined as

$$U(\mathbf{r}, t) = 2 \int_0^\infty \mathscr{U}^{\mathrm{r}}(\mathbf{r}, \nu) \exp(-i 2\pi \nu t) \, \mathrm{d}\nu, \tag{A.1}$$

where $\mathscr{U}^{\mathrm{r}}(\mathbf{r}, \nu)$ is the temporal Fourier transform of $U^{\mathrm{r}}(\mathbf{r}, t)$. We define the temporal Fourier transform by

$$\mathscr{U}(\mathbf{r}, \nu) = \int_{-\infty}^\infty U(\mathbf{r}, t) \exp(i 2\pi \nu t) \, \mathrm{d}t, \tag{A.2}$$

so that the inverse transform is

$$U(\mathbf{r}, t) = \int_{-\infty}^\infty \mathscr{U}(\mathbf{r}, \nu) \exp(-i 2\pi \nu t) \, \mathrm{d}\nu. \tag{A.3}$$

Therefore, $\mathscr{U}$ is simply twice the positive frequency portion of $\mathscr{U}^{\mathrm{r}}$, and $U^{\mathrm{r}}(\mathbf{r}, t)$ is the real part of $U(\mathbf{r}, t)$. Alternatively, the complex analytic representation can be defined

purely in the time domain [89] by

$$U(\mathbf{r}, t) = \left[\delta(t) - \frac{i}{\pi t}\right] * U^{\mathrm{r}}(\mathbf{r}, t). \tag{A.4}$$

The time domain definition extends from deterministic waveforms to stochastic processes. The stochastic processes we are interested in are stationary and do not generally have temporal Fourier transforms, as the tails do not die out.

We now present two examples. First, a monochromatic or harmonic signal at a fixed frequency $\nu_0$ is given by

$$U^{\mathrm{r}}(\mathbf{r}, t) = A(\mathbf{r}) \cos\left[2\pi\nu_0 t + \psi(\mathbf{r})\right], \tag{A.5}$$

so that

$$U(\mathbf{r}, t) = U_0(\mathbf{r}) \exp(-i2\pi\nu_0 t), \tag{A.6}$$

where

$$U_0(\mathbf{r}) = A(\mathbf{r}) \exp\left[-i\psi(\mathbf{r})\right]. \tag{A.7}$$

Although monochromatic signals are ideals that do not exists in practice, a broad class of signals can be decomposed into monochromatic components.

Second, a quasi-monochromatic signal is narrowband so that its amplitude and phase vary, but with a bandwidth significantly smaller than the carrier frequency $\nu_0$:

$$U^{\mathrm{r}}(\mathbf{r}, t) = A(\mathbf{r}, t) \cos\left[2\pi\nu_0 t + \psi(\mathbf{r}, t)\right]. \tag{A.8}$$

The complex analytic representation is

$$U(\mathbf{r}, t) = U_0(\mathbf{r}, t) \exp(-i2\pi\nu_0 t), \tag{A.9}$$

where

$$U_0(\mathbf{r}, t) = A(\mathbf{r}, t) \exp\left[-i\psi(\mathbf{r}, t)\right]. \tag{A.10}$$

## A.2   Geometric optics

Geometric optics enables us to model optical behavior using rays and tools from geometry, and is valid in the limit as the wavelength tends to zero. In order to highlight the assumptions made, we present the rigorous derivation of geometric optics from Born and Wolf [18]. We assume that the light is monochromatic, oscillating at an angular frequency $\omega = 2\pi\nu$. The complex analytic representation of the electric and magnetic fields are

$$\begin{align}
\mathbf{E}(\mathbf{r}, t) &= \mathbf{E}_0(\mathbf{r}) \exp(-i\omega t) \tag{A.11} \\
\mathbf{H}(\mathbf{r}, t) &= \mathbf{H}_0(\mathbf{r}) \exp(-i\omega t). \tag{A.12}
\end{align}$$

We assume that the medium is linear, isotropic, dispersion-free, and source-free. In this medium, Maxwell's equations are, in SI units, [90]

$$\nabla \times \mathbf{E} + \mu \frac{\partial}{\partial t}\mathbf{H} = 0 \tag{A.13}$$

$$\nabla \times \mathbf{H} - \epsilon \frac{\partial}{\partial t}\mathbf{E} = 0 \tag{A.14}$$

$$\nabla \cdot \epsilon \mathbf{E} = 0 \tag{A.15}$$

$$\nabla \cdot \mu \mathbf{H} = 0, \tag{A.16}$$

where $\epsilon$ and $\mu$ are the dielectric constant and permeability of the medium. We substitute (A.11) and (A.12) into (A.13)–(A.16) to obtain

$$\nabla \times \mathbf{E}_0 - i\omega\mu\mathbf{H}_0 = 0 \tag{A.17}$$

$$\nabla \times \mathbf{H}_0 + i\omega\epsilon\mathbf{E}_0 = 0 \tag{A.18}$$

$$\nabla \cdot \epsilon\mathbf{E}_0 = 0 \tag{A.19}$$

$$\nabla \cdot \mu\mathbf{H}_0 = 0. \tag{A.20}$$

We assume that $\mathbf{E}_0$ and $\mathbf{H}_0$ can be written as

$$\mathbf{E}_0 = \mathbf{e}(\mathbf{r})\exp\left[i\omega S(\mathbf{r})\right] \tag{A.21}$$

$$\mathbf{H}_0 = \mathbf{h}(\mathbf{r})\exp\left[i\omega S(\mathbf{r})\right], \tag{A.22}$$

where $S(\mathbf{r})$ is a scalar function of position known as the optical path. Using the product rules for the curl and divergence operations, we compute

$$\nabla \times \mathbf{H}_0 = (\nabla \times \mathbf{h} + i\omega\nabla S \times \mathbf{h})\exp(i\omega S) \tag{A.23}$$

$$\nabla \times \mathbf{E}_0 = (\nabla \times \mathbf{e} + i\omega\nabla S \times \mathbf{e})\exp(i\omega S) \tag{A.24}$$

$$\nabla \cdot \mu\mathbf{H}_0 = (\mu\nabla \cdot \mathbf{h} + \mathbf{h} \cdot \nabla\mu + i\omega\mu\mathbf{h} \cdot \nabla S)\exp(i\omega S) \tag{A.25}$$

$$\nabla \cdot \epsilon\mathbf{E}_0 = (\epsilon\nabla \cdot \mathbf{e} + \mathbf{e} \cdot \nabla\epsilon + i\omega\epsilon\mathbf{e} \cdot \nabla S)\exp(i\omega S). \tag{A.26}$$

Using the relations (A.21)–(A.22) and (A.23)–(A.26) in equations (A.17)–(A.20), we obtain

$$\nabla S \times \mathbf{e} - \mu\mathbf{h} = \frac{i}{\omega}\nabla \times \mathbf{e} \tag{A.27}$$

$$\nabla S \times \mathbf{h} + \epsilon\mathbf{e} = \frac{i}{\omega}\nabla \times \mathbf{h} \tag{A.28}$$

$$\mathbf{e} \cdot \nabla S = \frac{i}{\omega}(\nabla \cdot \mathbf{e} + \mathbf{e} \cdot \nabla\log\epsilon) \tag{A.29}$$

$$\mathbf{h} \cdot \nabla S = \frac{i}{\omega}(\nabla \cdot \mathbf{h} + \mathbf{h} \cdot \nabla\log\mu). \tag{A.30}$$

In the geometric optics approximation, $\omega \to \infty$ so that the righthand sides of (A.27)–(A.30) vanish. Roughly speaking, the geometric optics approximation is valid when the changes in $\mathbf{e}$ and $\mathbf{h}$ are proportionally small over domains for $\mathbf{r}$ with linear dimension on the same order as the wavelength $\lambda$. That is, the approximation breaks down at points $\mathbf{r}$ where the fields vary substantially over a neighborhood of $\mathbf{r}$ of radius $\lambda$. Geometric optics therefore fails to accurately describe optical behavior at shadow edges, for example.

Now having made the geometric optics approximation, we simplify the equations to clarify the dynamics. Equations (A.27)–(A.30) become

$$\nabla S \times \mathbf{e} - \mu\mathbf{h} = 0 \tag{A.31}$$

$$\nabla S \times \mathbf{h} + \epsilon\mathbf{e} = 0 \tag{A.32}$$

$$\mathbf{e} \cdot \nabla S = 0 \tag{A.33}$$

$$\mathbf{h} \cdot \nabla S = 0. \tag{A.34}$$

We drop equations (A.33)–(A.34), since they follow from (A.31)–(A.32) by taking the dot product with $\nabla S$. Solving for $\mathbf{h}$ in (A.31), substituting into (A.32), and using a triple product identity yields

$$\nabla S \left( \nabla S \cdot \mathbf{e} \right) - \mathbf{e} \left( \nabla S \cdot \nabla S \right) + \epsilon\mu\mathbf{e} = 0. \tag{A.35}$$

The first term of (A.35) vanishes by (A.33), so that we conclude

$$\nabla^2 S = n^2, \tag{A.36}$$

where $n = \sqrt{\epsilon\mu}$ is the index of refraction. Equation (A.36) is the eikonal equation that determines how the optical path evolves as the index of refraction changes.

The rays in geometric optics point in the direction of the time-averaged Poynting vector, which we now compute. We integrate over a long time interval $T \gg 1/\omega$:

$$\langle \mathbf{S} \rangle = \frac{1}{2T} \int_{-T}^{T} \mathbf{S} \, dt \tag{A.37}$$

$$= \frac{1}{2T} \int_{-T}^{T} \mathrm{Re}\{\mathbf{E}\} \times \mathrm{Re}\{\mathbf{H}\} \, dt \tag{A.38}$$

$$= \frac{1}{2} \mathrm{Re}\{\mathbf{E}_0 \times \mathbf{H}_0^*\} \tag{A.39}$$

$$= \frac{1}{2} \mathrm{Re}\{\mathbf{e} \times \mathbf{h}^*\} \tag{A.40}$$

$$= \frac{1}{2\mu} \mathrm{Re}\left\{ (\mathbf{e} \cdot \mathbf{e}^*)\nabla S - (\mathbf{e} \cdot \nabla S)\mathbf{e}^* \right\} \tag{A.41}$$

$$= \frac{|\mathbf{e}|^2}{2\mu} \nabla S. \tag{A.42}$$

In (A.39) we have neglected the oscillating terms, assuming that they integrate to zero over the long time period. In (A.41), we have calculated $\mathbf{e} \times \mathbf{h}^*$ from (A.31) and

made use of a triple product formula. Finally, in (A.42) we have used (A.33). The rays in geometric optics therefore point along the gradient of the optical path, so that they are orthogonal to the level sets of the optical path, which are the wavefronts.

## A.3  Scalar diffraction theory

Diffraction is the physical basis for wave propagation and is used throughout the thesis. We derive the Rayleigh diffraction formula from Maxwell's equations. First, we approximate the electric field vector by a scalar field that satisfies the scalar wave equation (Section A.3.1). Then, we formulate and solve a boundary value problem to derive the Rayleigh diffraction formula of the first kind (Section A.3.2). Finally, we describe the angular decomposition of light into plane waves that is used extensively in Fourier optics (Section A.3.3). Although Maxwell's equations only apply to electromagnetic radiation, our results apply equally well to acoustic waves, which satisfy a similar scalar wave equation [19].

### A.3.1  A scalar approximation

We make a scalar approximation to the electric field vector $\mathbf{E}(\mathbf{r}, t)$. In a homogeneous medium free of currents and charges, $\mathbf{E}$ satisfies the wave equation [18]

$$\nabla^2 \mathbf{E} - \epsilon\mu \frac{\partial^2}{\partial t^2} \mathbf{E} = 0. \tag{A.43}$$

Thus, each Cartesian component $U$ satisfies the homogeneous scalar wave equation

$$\nabla^2 U - \epsilon\mu \frac{\partial^2}{\partial t^2} U = 0. \tag{A.44}$$

We can consequently describe the electromagnetic field by applying scalar wave theory to each component.

It is more convenient to work with a single scalar quantity instead of three. Incoherent light is primarily characterized by its intensity, which can be approximated by a single scalar wave function [18]. The scalar diffraction approximation is not as well-justified in coherent wave applications, although we nonetheless assume that the electric field is linearly polarized. A more precise treatment would leverage stochastic models for coherence and polarization [91].

We assume that $U$ is monochromatic, so that

$$U(\mathbf{r}, t) = U_0(\mathbf{r}) \exp(-i2\pi\nu t) \tag{A.45}$$

for some frequency $\nu$. By (A.44), $U_0$ satisfies the homogeneous Helmholtz equation

$$\nabla^2 U_0(\mathbf{r}) + k^2 U_0(\mathbf{r}) = 0, \tag{A.46}$$

where $k = 2\pi\nu\sqrt{\mu\epsilon}$.

## A.3.2   The Rayleigh diffraction formula

We obtain the Rayleigh diffraction formula by solving for the scalar field within the half-space $z \geq 0$ in terms of the boundary conditions on the plane $z = 0$ [92]. We assume that the field on any cross section of a rectangular slab occupying the region $0 \leq z \leq Z$ has a two-dimensional spatial Fourier transform denoted by $\mathscr{U}_0(u, v; z)$, so that

$$U_0(x, y, z) = \iint \mathscr{U}_0(u, v; z) \exp\left[i(ux + vy)\right] \mathrm{d}u \, \mathrm{d}v. \tag{A.47}$$

By substituting (A.47) into (A.46), we obtain a differential equation for $\mathscr{U}_0(u, v; z)$, whose general solution yields four different types of plane waves. Two types are homogeneous plane waves of low spatial frequencies $u^2 + v^2 \leq k^2$; one propagates from the plane $z = 0$ towards the plane $z = Z$, and the other propagates in the opposite direction. The other two types are evanescent waves of high spatial frequencies $u^2 + v^2 > k^2$, again, propagating in opposite directions. The evanescent waves decay in amplitude exponentially with distance, so we neglect them. As $Z \to \infty$ so that the slab models the half-space $z \geq 0$, there is no longer a physical basis for a backward propagating wave. Therefore, we are left with a single type of plane wave solution, so that

$$U_0(x, y, z) = \iint A(u, v) \exp\left[i(ux + vy + \sqrt{k^2 - u^2 - v^2}z)\right] \mathrm{d}u \, \mathrm{d}v \tag{A.48}$$

is the superposition of plane waves traveling along the direction $(u, v, \sqrt{k^2 - u^2 - v^2})$. We recognize $A(u, v)$ in (A.48) as the two-dimensional spatial Fourier transform of $U_0(x, y, 0)$. We substitute the Fourier transform formula for $A(u, v)$, set $u = kp$, $v = kq$, and interchange the order of integration to obtain

$$U_0(x, y, z) = \iint U_0(x', y', 0)$$
$$\times \left(\frac{k}{2\pi}\right)^2 \iint \exp\left\{ik\left[p(x - x') + q(y - y') + \sqrt{1 - p^2 - q^2}z\right]\right\} \mathrm{d}p \, \mathrm{d}q \, \mathrm{d}x' \, \mathrm{d}y'. \tag{A.49}$$

The inner integral in (A.49) is the derivative of the Weyl representation of a diverging spherical wave:

$$\frac{\partial}{\partial z}\left[\frac{\exp(ik|\mathbf{r} - \mathbf{r}'|)}{|\mathbf{r} - \mathbf{r}'|}\right] =$$
$$-\frac{k^2}{2\pi} \iint \exp\left\{ik\left[p(x - x') + q(y - y') + \sqrt{1 - p^2 - q^2}z\right]\right\} \mathrm{d}p \, \mathrm{d}q, \tag{A.50}$$

where $\mathbf{r} = (x, y, z)$ and $\mathbf{r}' = (x', y', 0)$. We obtain the Rayleigh diffraction formula of the first kind by substituting (A.50) into (A.49):

$$U_0(\mathbf{r}) = -\frac{1}{2\pi} \int U_0^0(\mathbf{r}') \frac{\partial}{\partial z}\left[\frac{\exp(ik|\mathbf{r} - \mathbf{r}'|)}{|\mathbf{r} - \mathbf{r}'|}\right] \mathrm{d}^2r', \tag{A.51}$$

where the superscript 0 is a reminder to evaluate $U_0(\mathbf{r}')$ on the plane $z = 0$. Formula (A.51) completely specifies the scalar field in a region of free space, using just measurements made on a plane.

### A.3.3  Angular decomposition into plane waves

A two-dimensional spatial Fourier transform decomposes radiation into plane waves. To express the decomposition in compact notation, we start with (A.48) and substitute $u = kp$, $v = kq$, $a(p, q) = k^2 A(kp, kq)$, and $\mathbf{s} = (p, q, \sqrt{1 - p^2 - q^2})$ to obtain

$$U_0(\mathbf{r}) = \int a(\mathbf{s}) \exp(ik\mathbf{s} \cdot \mathbf{r}) \, \mathrm{d}^2 s, \tag{A.52}$$

where

$$a(\mathbf{s}) = \left( \frac{k}{2\pi} \right)^2 \int U_0^0(\mathbf{r}') \exp(-ik\mathbf{s} \cdot \mathbf{r}') \, \mathrm{d}^2 r'. \tag{A.53}$$

We may equivalently integrate over any plane parallel to the $xy$-plane:

$$a(\mathbf{s}) = \left( \frac{k}{2\pi} \right)^2 \int U_0(\mathbf{r}) \exp(-ik\mathbf{s} \cdot \mathbf{r}) \, \mathrm{d}^2 r. \tag{A.54}$$

We interpret the Rayleigh diffraction formula in (A.51) as a recipe for computing the field by summing the radiation patterns from secondary spherical waves originating at points across the plane $z = 0$, which is a construction known as the Huygens-Fresnel principle. The angular decomposition in (A.52) provides an alternative interpretation, so that the field is instead computed by summing contributions from plane waves heading in different directions into the positive half-space $z > 0$.

In the far zone, only a single plane wave component survives. To demonstrate this, we make a far-zone approximation to the Rayleigh diffraction formula. We denote the unit vector pointing in the direction of $\mathbf{r}$ by $\hat{\mathbf{r}} = \mathbf{r}/|\mathbf{r}|$. When $|\mathbf{r}|$ is sufficiently large, we may approximate

$$\frac{\exp(ik|\mathbf{r} - \mathbf{r}'|)}{|\mathbf{r} - \mathbf{r}'|} \approx \frac{\exp(ik|\mathbf{r}|)}{|\mathbf{r}|} \exp(-ik\hat{\mathbf{r}} \cdot \mathbf{r}') \tag{A.55}$$

so that

$$\frac{\partial}{\partial z} \left[ \frac{\exp(ik|\mathbf{r} - \mathbf{r}'|)}{|\mathbf{r} - \mathbf{r}'|} \right] \approx ik \frac{z}{|\mathbf{r}|} \frac{\exp(ik|\mathbf{r}|)}{|\mathbf{r}|} \exp(-ik\hat{\mathbf{r}} \cdot \mathbf{r}'). \tag{A.56}$$

We denote the angle that $\mathbf{r}$ makes with the $z$-axis by $\theta$. Using the approximation in (A.56) in the Rayleigh diffraction formula in (A.51), we obtain the field in the far zone:

$$U_0^\infty(\mathbf{r}) = -\frac{ik}{2\pi} \cos\theta \frac{\exp(ik|\mathbf{r}|)}{|\mathbf{r}|} \int U_0^0(\mathbf{r}') \exp(-ik\hat{\mathbf{r}} \cdot \mathbf{r}') \, \mathrm{d}^2 r'. \tag{A.57}$$

The integral in (A.57) is a plane wave component in direction $\hat{\mathbf{r}}$, so that

$$U_0^\infty(\mathbf{r}) = -\frac{2\pi i}{k} \cos\theta \frac{\exp(ik|\mathbf{r}|)}{|\mathbf{r}|} a(\hat{\mathbf{r}}). \tag{A.58}$$

## A.4  Coherence theory

Optical coherence theory uses stochastic processes to predict interference effects. We describe how the second-order statistics of the scalar field model phase correlation (Section A.4.1) and how to handle stochastic processes that lack well-defined temporal Fourier transforms (Section A.4.2).

### A.4.1  Second-order statistics

Let the scalar field $U(\mathbf{r}, t)$ be a zero-mean, complex, stationary, ergodic, stochastic process. We define the mutual coherence function [89] by

$$\Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau) = \langle U(\mathbf{r}_1, t+\tau) U^*(\mathbf{r}_2, t) \rangle, \tag{A.59}$$

where the operation $\langle \cdot \rangle$ denotes taking an ensemble average. $\Gamma$ tells us how well-correlated the field is at different points in space and at different instants in time. We call $\Gamma(\mathbf{r}, \mathbf{r}, 0)$ the average intensity at $\mathbf{r}$. Several quantities are derived from the mutual coherence function. The complex degree of coherence is a normalized version of $\Gamma$:

$$\gamma(\mathbf{r}_1, \mathbf{r}_2, \tau) = \frac{\Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)}{\sqrt{\Gamma(\mathbf{r}_1, \mathbf{r}_1, 0)}\sqrt{\Gamma(\mathbf{r}_2, \mathbf{r}_2, 0)}}. \tag{A.60}$$

As with a generic correlation coefficient, the absolute value of the complex degree of coherence is bounded

$$0 \le |\gamma(\mathbf{r}_1, \mathbf{r}_2, \tau)| \le 1, \tag{A.61}$$

where a value of 0 implies that the indicated vibrations are completely incoherent, and a 1 implies they are completely coherent.

The temporal Fourier transform of $\Gamma$ is called the cross-spectral density function

$$W(\mathbf{r}_1, \mathbf{r}_2, \nu) = \int \Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau) \exp(i2\pi\nu\tau)\, \mathrm{d}\tau, \tag{A.62}$$

which, when normalized, is called the spectral degree of coherence:

$$\mu(\mathbf{r}_1, \mathbf{r}_2, \nu) = \frac{W(\mathbf{r}_1, \mathbf{r}_2, \nu)}{\sqrt{W(\mathbf{r}_1, \mathbf{r}_1, \nu)}\sqrt{W(\mathbf{r}_2, \mathbf{r}_2, \nu)}}. \tag{A.63}$$

The absolute value of the spectral degree of coherence is bounded

$$0 \le |\mu(\mathbf{r}_1, \mathbf{r}_2, \nu)| \le 1, \tag{A.64}$$

with 0 indicating complete incoherence and 1 complete coherence. Also note that

$W(\mathbf{r}, \mathbf{r}, \nu) = S(\mathbf{r}, \nu)$, the standard power spectral density of the time series $U(\mathbf{r}, t)$ at a particular position $\mathbf{r}$.

### A.4.2  Coherent-mode representation

The coherent-mode representation decomposes the stochastic process $U(\mathbf{r}, t)$ by temporal frequency [92]. We fix the frequency $\nu$. By Mercer's theorem, $W(\mathbf{r}_1, \mathbf{r}_2, \nu)$ is a convergent sum of eigenfunctions and eigenvalues of the associated integral equation. We construct a new stochastic process $U_\nu(\mathbf{r})$ from these eigenfunctions, by introducing random variables with appropriate distributions, so that the spatial cross-correlation function of $U_\nu$ is precisely the cross-spectral density:

$$W(\mathbf{r}_1, \mathbf{r}_2, \nu) = \langle U_\nu(\mathbf{r}_1) U_\nu^*(\mathbf{r}_2) \rangle. \tag{A.65}$$

Equation (A.65) expresses the Fourier transform of a correlation function as a correlation function itself, of an entirely different stochastic process. In effect, we have performed a Karhunen-Loève expansion, in reverse. Instead of decomposing a stochastic process using its covariance function to determine the appropriate eigenvectors, we have synthesized a stochastic process with a covariance function that has been predetermined. To employ the coherent-mode representation, we simply work with the stochastic process $U_\nu(\mathbf{r}) \exp(-i2\pi\nu t)$ instead of $U(\mathbf{r}, t)$ to understand the behavior of radiation at the fixed frequency $\nu$.

## A.5  Operators and Dirac notation

We describe the three main Hilbert space operators introduced by Agarwal et al. [49]: the position operator $\hat{\mathbf{r}}_\perp$, the direction operator $\hat{\mathbf{s}}_\perp$, and the coherence operator $\hat{L}$. The position and direction operators are defined component-wise, so that

$$\begin{aligned} \hat{\mathbf{r}}_\perp &= (\hat{x}, \hat{y}) \\ \hat{\mathbf{s}}_\perp &= (\hat{s}_x, \hat{s}_y), \end{aligned} \tag{A.66}$$

where the components operate on a function $f$ by

$$\begin{aligned} \hat{x} f &= x f \\ \hat{s}_x f &= -\frac{i\lambda}{2\pi} \frac{\partial f}{\partial x}, \end{aligned} \tag{A.67}$$

and similarly for $\hat{y}$ and $\hat{s}_y$. The commutator of two operators is defined as [93]

$$[\hat{x}, \hat{s}_x] = \hat{x}\hat{s}_x - \hat{s}_x\hat{x}. \tag{A.68}$$

We compute

$$[\hat{x}, \hat{s}_x] f = (\hat{x}\hat{s}_x - \hat{s}_x\hat{x}) f \tag{A.69}$$

$$= -\frac{i\lambda}{2\pi}\left[x\frac{\partial f}{\partial x} - \frac{\partial}{\partial x}(xf)\right] \tag{A.70}$$

$$= \frac{i\lambda}{2\pi}f, \tag{A.71}$$

from which the commutation relations in (4.11) follow.

The coherence operator $\hat{L}$ is defined in (4.12) by its matrix elements using the Dirac notation [93]

$$\left\langle \mathbf{r}_\perp^{\mathrm{R}} \middle| \hat{L} \middle| \mathbf{r}_\perp^{\mathrm{C}} \right\rangle = U\left(\mathbf{r}^{\mathrm{R}}\right) U^*\left(\mathbf{r}^{\mathrm{C}}\right), \tag{A.72}$$

where $|\mathbf{r}_\perp^{\mathrm{R}}\rangle$ is the eigenfunction of $\hat{\mathbf{r}}_\perp$ with eigenvalue $\mathbf{r}_\perp^{\mathrm{R}}$, and similarly for $|\mathbf{r}_\perp^{\mathrm{C}}\rangle$. In position coordinates, $|\mathbf{r}_\perp^{\mathrm{R}}\rangle$ is given by $\delta(\mathbf{r}_\perp - \mathbf{r}_\perp^{\mathrm{R}})$. Therefore, the lefthand side of equation (A.72) evaluates to

$$\int \delta\left(\mathbf{r}_\perp - \mathbf{r}_\perp^{\mathrm{R}}\right) \hat{L}\, \delta\left(\mathbf{r}_\perp - \mathbf{r}_\perp^{\mathrm{C}}\right)\, \mathrm{d}^2 r, \tag{A.73}$$

which specifies how $\hat{L}$ operates on functions by mapping the Dirac delta basis function at $\mathbf{r}_\perp^{\mathrm{C}}$ to the Dirac delta basis function at $\mathbf{r}_\perp^{\mathrm{R}}$. When applied to all $\mathbf{r}_\perp^{\mathrm{R}}$ and $\mathbf{r}_\perp^{\mathrm{C}}$, equation (A.72) thereby uniquely determines the operator $\hat{L}$ from the scalar field $U$ in the source plane.

# Appendix B

# Additional proofs and results

We include several proofs and results that are highly relevant to the thesis, but whose details are tangential to the main discussion. First, we relate the main lobe width of the beam pattern to the array length (Section B.1). Then, we describe how the Rihaczek quasi light field behaves like the traditional light field for globally incoherent radiation in the limit of zero wavelength (Section B.2). We show how alternative coherent cameras generalize beamforming and how comparing cameras is equivalent to applying tomographic constraints on the light field (Section B.3). Finally, we present proofs and results on the steady-state distribution and bias inherent in phase tightening (Section B.4).

## B.1    Beam width approximation

We prove a stronger version of the half-power beam width approximation in (2.36).

**Claim B.1.1.** *In the absence of phase noise, the angular half-power beam width for a uniform linear array is*

$$\theta_{BW} \approx 0.8859 \, \lambda / Md, \tag{B.1}$$

*for large values of $M$.*

*Proof.* We determine the points at which the conventional beam pattern $B_\theta$ in (2.32) reaches the half-power mark where $|B_\theta|^2 = 1/2$. We assume the array is steered perpendicular to its axis at $\theta_T = \pi/2$ and that $\theta \approx \pi/2$, so that $\tilde{\theta} = \theta - \pi/2$ is small and

$$B_\theta \approx \frac{1}{M} \frac{\sin\left(\frac{M\pi d}{\lambda}\tilde{\theta}\right)}{\sin\left(\frac{\pi d}{\lambda}\tilde{\theta}\right)}. \tag{B.2}$$

We define

$$\gamma = \frac{M\pi d}{\lambda}\tilde{\theta}, \tag{B.3}$$

and note that for large $M$, the value of $\tilde{\theta}$ that marks the end of the main lobe is small

149

so that we may make a linear approximation for the sine in the denominator of (B.2):

$$B_\theta \approx \frac{\sin\gamma}{\gamma}. \tag{B.4}$$

Solving $|B_\theta|^2 = 1/2$ yields $\gamma = 1.3916$. Thus, $\tilde\theta \approx 1.3916\lambda/M\pi d$, and the half-power beam width is twice this quantity.  □

## B.2  Traditional light field properties

The traditional light field satisfies a set of properties that quasi light fields attain for globally incoherent radiation in the limit of zero wavelength. We make this limiting process precise for the Rihaczek conjugate light field, following Mandel and Wolf [92].

We model globally incoherent radiation with a quasi-homogeneous source, which is a special type of Schell-model source. A Schell-model source has a spatially stationary spectral degree of coherence

$$\mu(\mathbf{r}_1, \mathbf{r}_2, \nu) = g(\mathbf{r}_1 - \mathbf{r}_2, \nu), \tag{B.5}$$

for some function $g$. A quasi-homogeneous source additionally (i) has a power spectral density that varies much more slowly with position than $g(\mathbf{r}, \nu)$ and (ii) is physically larger than both the wavelength and the spatial width of $g(\mathbf{r}, \nu)$. Thus quasi-homogeneous sources are said to be globally incoherent because the radiation is only spatial correlated across distances that are small compared with the size of the source. For example, natural sunlight can be modeling with a quasi-homogeneous source.

We now compute the conjugate Rihaczek light field for a quasi-homogeneous source. The partially coherent version of the conjugate Rihaczek quasi light field in (4.28) is

$$L^{\mathrm{R}}(\mathbf{r}, \mathbf{s}) = \left(\frac{k}{2\pi}\right)^2 s_z \exp(ik\mathbf{s}\cdot\mathbf{r}) \int W(\mathbf{r}', \mathbf{r}, \nu) \exp(-ik\mathbf{s}\cdot\mathbf{r}')\, \mathrm{d}^2 r'. \tag{B.6}$$

It follows from the definition of the spectral degree of coherence that the cross-spectral density in the source plane of a quasi-homogeneous source is

$$W^0(\mathbf{r}_1, \mathbf{r}_2, \nu) \approx S^0(\mathbf{r}_2, \nu)g^0(\mathbf{r}_1 - \mathbf{r}_2, \nu). \tag{B.7}$$

Upon substituting (B.7) into (B.6), the light field in the source plane factors into terms depending on position and direction:

$$L^0(\mathbf{r}, \mathbf{s}) = k^2 s_z S^0(\mathbf{r}, \nu)\mathscr{G}^0(k\mathbf{s}, \nu), \tag{B.8}$$

where $\mathscr{G}^0$ is the spatial Fourier transform of $g^0$. All factors in (B.8) are nonnegative [92]. Furthermore, since the power spectral density is zero outside of the source, the light field is zero there too. Therefore the Rihaczek light field for a quasi-homogeneous source has the properties of the traditional light field in the source plane.

We now compute the Rihaczek light field along rays pointing from the source plane into the half-space $z > 0$. First, we use the coherent mode representation to express the Rihaczek light field in (B.6) in terms of a stochastic plane wave $a_\nu$:

$$L^{\mathrm{R}}(\mathbf{r}, \mathbf{s}) = s_z \exp(ik\mathbf{s} \cdot \mathbf{r})\langle U_\nu^*(\mathbf{r})a_\nu(\mathbf{s})\rangle. \tag{B.9}$$

We express the scalar field $U_\nu$ at $\mathbf{r}$ in terms of the scalar field on the source plane using the Rayleigh diffraction formula, in order to express $L^{\mathrm{R}}$ at $\mathbf{r}$ in terms of the light field on the source plane:

$$L^{\mathrm{R}}(\mathbf{r}, \mathbf{s}) = \exp(ik\mathbf{s} \cdot \mathbf{r}) \int L^0(\mathbf{r}', \mathbf{s}) \left\{-\frac{1}{2\pi}\frac{\partial}{\partial z}\left[\frac{\exp(-ik|\mathbf{r} - \mathbf{r}'|)}{|\mathbf{r} - \mathbf{r}'|}\right]\right\} \exp(-ik\mathbf{s} \cdot \mathbf{r}') \, \mathrm{d}^2r'. \tag{B.10}$$

We substitute the light field on the source plane of a quasi-homogeneous source in (B.8) into (B.10):

$$L(\mathbf{r}, \mathbf{s}) = k^2 s_z \mathscr{G}^0(k\mathbf{s}, \nu) \exp(ik\mathbf{s} \cdot \mathbf{r})$$
$$\times \int \left\{-\frac{1}{2\pi}\frac{\partial}{\partial z}\left[\frac{\exp(-ik|\mathbf{r} - \mathbf{r}'|)}{|\mathbf{r} - \mathbf{r}'|}\right]\right\} S^0(\mathbf{r}', \nu) \exp(-ik\mathbf{s} \cdot \mathbf{r}') \, \mathrm{d}^2r'. \tag{B.11}$$

We assume that $\mathbf{r}$ is several wavelengths away from the source plane and approximate

$$\begin{aligned}
\frac{\partial}{\partial z}\left[\frac{\exp(-ik|\mathbf{r} - \mathbf{r}'|)}{|\mathbf{r} - \mathbf{r}'|}\right] &= \left(-ik - \frac{1}{|\mathbf{r} - \mathbf{r}'|}\right)\frac{\exp(-ik|\mathbf{r} - \mathbf{r}'|)}{|\mathbf{r} - \mathbf{r}'|}\frac{z}{|\mathbf{r} - \mathbf{r}'|} \\
&\approx -ik\frac{\exp(-ik|\mathbf{r} - \mathbf{r}'|)}{|\mathbf{r} - \mathbf{r}'|}\frac{z}{|\mathbf{r} - \mathbf{r}'|}
\end{aligned} \tag{B.12}$$

so that

$$L(\mathbf{r}, \mathbf{s}) \approx k^2 s_z \mathscr{G}^0(k\mathbf{s}, \nu) \exp(ik\mathbf{s} \cdot \mathbf{r}) \int \frac{ikz}{2\pi} \frac{S^0(\mathbf{r}', \nu)}{|\mathbf{r} - \mathbf{r}'|^2} \exp\left[-ik(|\mathbf{r} - \mathbf{r}'| + \mathbf{s} \cdot \mathbf{r}')\right] \mathrm{d}^2r'. \tag{B.13}$$

We evaluate the integral in (B.13) for large $k$ using the principle of stationary phase [92]. Intuitively, the exponential oscillates rapidly, so that there is much cancellation upon integration. The points that have a lasting impact on the integral are those where the first-order partial derivatives of $|\mathbf{r} - \mathbf{r}'| + \mathbf{s} \cdot \mathbf{r}'$ are zero. There is one such point, at

$$\mathbf{r}_0 = \left(x - \frac{s_x}{s_z}z, y - \frac{s_y}{s_z}z, 0\right). \tag{B.14}$$

Geometrically, $\mathbf{r}_0$ is the point at which a line passing through $\mathbf{r}$ along $\mathbf{s}$ intersects the source plane $z = 0$. By the principle of stationary phase, the integral in (B.13) sifts out the value of $S^0(\mathbf{r}', \nu)/|\mathbf{r} - \mathbf{r}'|^2$ at $\mathbf{r}' = \mathbf{r}_0$, but with an additional factor based on second-order partial derivatives. The integral in (B.13) ultimately evaluates to $S^0(\mathbf{r}_0, \nu) \exp(-ik\mathbf{s} \cdot \mathbf{r})$. We conclude that the Rihaczek light field for a quasi-

homogeneous source in the limit as $k \to \infty$ is

$$
\begin{aligned}
L^{\mathrm{R}}(\mathbf{r},\mathbf{s}) &= k^2 s_z S^0(\mathbf{r}_0,\nu)\mathscr{G}^0(k\mathbf{s},\nu) \\
&= L^0(\mathbf{r}_0,\mathbf{s}),
\end{aligned}
\tag{B.15}
$$

which is constant along rays.

## B.3   Coherent cameras

We explore two alternative ways to interpret light field design for coherent imaging applications. First, coherent cameras based on generalized Wigner-Ville quasi light fields aggregate and isolate power contributions in different ways (Section B.3.1). Second, instead of comparing how two different cameras capture the light field, we can force a camera to match the pixel values of a specific reference camera, which acts as a constraint on light field design (Section B.3.2).

### B.3.1   Generalized Wigner-Ville beamformers

Coherent cameras based on the Rihaczek quasi light field form image pixels by computing the Hermitian product of two different beamformers that divide the tasks of aggregating and isolating power contributions (Section 4.4.3). We now explore how other quasi light fields lead to different divisions of labor between aggregating and isolating power. Specifically, we consider the generalized Wigner-Ville quasi light fields [34] obtained by selecting

$$
K(\mathbf{a},\mathbf{b}) = \delta\left[\left(\frac{1}{2}+\gamma\right)\mathbf{a} + \left(\frac{1}{2}-\gamma\right)\mathbf{b}\right]
\tag{B.16}
$$

in (4.24), resulting in the light fields

$$
\begin{aligned}
L^{\mathrm{GWV}}(\mathbf{r},\mathbf{s}) = &\left(\frac{k}{2\pi}\right)^2 s_z \\
&\times \int U^*\left[\mathbf{r} - \left(\gamma+\frac{1}{2}\right)\mathbf{r}'\right] U\left[\mathbf{r} - \left(\gamma-\frac{1}{2}\right)\mathbf{r}'\right]\exp(-ik\mathbf{s}_\perp \cdot \mathbf{r}'_\perp)\,\mathrm{d}^2 r',
\end{aligned}
\tag{B.17}
$$

where $\gamma$ is a real parameter restricted to $|\gamma| \le 1/2$. Specific values of $\gamma$ correspond to special quasi light fields: $\gamma = -1/2$ for the conjugate Rihaczek distribution, $\gamma = 1/2$ for the regular Rihaczek distribution, and $\gamma = 0$ for the Wigner distribution.

To form an image, we integrate $\mathbf{r}$ over the aperture to aggregate power, and $\mathbf{r}'$ over the entire sensor array to isolate power. The integral incorporates samples of $U(\mathbf{r}^{\mathrm{R}})$ and $U^*(\mathbf{r}^{\mathrm{C}})$ according to the change of variables

$$
\begin{aligned}
\mathbf{r}^{\mathrm{R}} &= \mathbf{r} - (\gamma - 1/2)\mathbf{r}', \\
\mathbf{r}^{\mathrm{C}} &= \mathbf{r} - (\gamma + 1/2)\mathbf{r}'.
\end{aligned}
\tag{B.18}
$$

For simplicity, we assume the two-dimensional case and integrate over an aperture of width $A$ and a sensor array of length $D$. The resulting region of integration is a rectangular strip rotated according to $\gamma$ and bound by the available sensors (Figure B-1).



(a)  conjugate Rihaczek   $\gamma = -1/2$                   (b)  Wigner   $\gamma = 0$
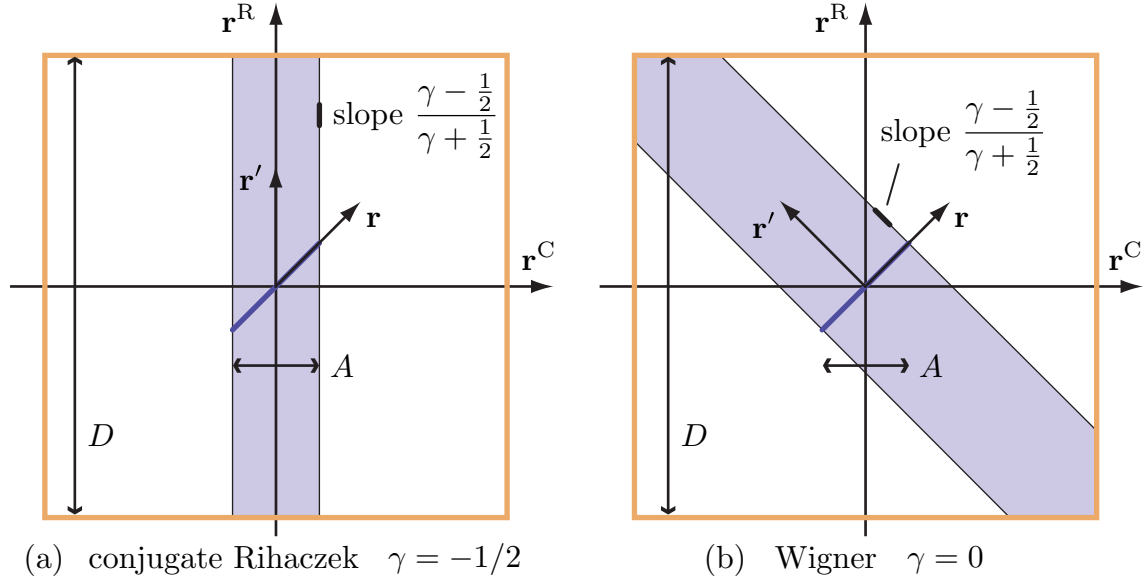
Figure B-1: Coherent cameras based on generalized Wigner-Ville quasi light fields aggregate and isolate power contributions in different ways, depending on the parameter $\gamma$. For the conjugate Rihaczek light field (a), when $\gamma = -1/2$, power isolation along $\mathbf{r}'$ uses the field measurements at all $\mathbf{r}^{\mathrm{R}}$ with a fixed conjugate field measurement at $\mathbf{r}^{\mathrm{C}}$. For the Wigner light field (b), when $\gamma = 0$, power isolation along $\mathbf{r}'$ uses the field and conjugate field measurements symmetrically.

The conjugate field and field measurements contribute to the power aggregation and selection tasks in different ways for each choice of $\gamma$. When $\gamma = -1/2$ and the $\mathbf{r}'$-axis is vertical, the conjugate field measurements are only used for power aggregation. As $\gamma$ increases and the $\mathbf{r}'$-axis tilts, more conjugate field measurements are used for power selection, as the projection of the region of integration onto the $\mathbf{r}^{\mathrm{C}}$-axis increases in length. The conjugate field and field measurement equally contribute to the power aggregation and selection tasks for the symmetric Wigner camera at $\gamma = 0$. As $\gamma$ increases further from 0 to $1/2$, we witness similar behavior, with the roles of the conjugate field and field measurements exchanged.

## B.3.2   Camera comparison and tomography

The light field is a common currency for camera comparison, allowing us to compare the operation of different cameras by comparing the light fields that they capture. We can alternatively view camera comparison as a constraint on quasi light field design: we require the pixel values produced by a coherent camera to match a known reference

camera, and translate this requirement to constraints on the light field. Specifically, pixel value requirements translate to tomographic constraints on the quasi light field.

Inferring a light field from pixel values is mathematically equivalent to the problem in the field of tomography of reconstructing a function from its projections. While in a traditional CAT (computerized axial tomography) scan, the reconstructed function represents the linear attenuation of X-rays at each point in the body [94], in our context, the reconstructed function is the light field. In the context of computational photography, Ng showed that images formed by conventional cameras are projections of the light field [95]. Earlier, in the context of quantum mechanics, Bertrand and Bertrand derived the Wigner-Ville distribution from its projections [96].

We express the pixel values of the ideal single-lens camera in two dimensions as line integrals by applying a change of variables $m = \tan \psi$ to (4.51):

$$P(y_0, R) = \frac{V}{R} \iint \left(m^2 + 1\right)^{-5/2} \mathrm{rect}(y/A) L(y, m) \delta(y - Rm - y_0) \, \mathrm{d}m \, \mathrm{d}y, \quad \text{(B.19)}$$

where rect denotes the unit rectangle function. The integral on the righthand side of (B.19) is the Radon transform of $(m^2 + 1)^{-5/2} \mathrm{rect}(y/A) L(y, m)$ with slope $R$ and offset $y_0$ [94]. Therefore, if the pixel values $P(y_0, R)$ are specified for all $y_0$ and $R$, then (B.19) uniquely determines the light field $L(y, m)$ for $|y| < A/2$. In comparison, the tomographic constraint does not uniquely determine the quasi light field in three dimensions.

Determining light fields by tomographic projection need not result in quasi light fields with desirable properties. To illustrate the potential difficulties, we apply the inverse Radon transform to explicitly solve for $L(y, m)$. First, we multiply both sides of (B.19) by $\exp(-iuy_0)$, integrate over $y_0$, and define $v = Ru$:

$$\int P(y_0, v/u) \exp(-iuy_0) \, \mathrm{d}y_0 =$$
$$\frac{uV}{v} \iint \left(m^2 + 1\right)^{-5/2} \mathrm{rect}(y/A) L(y, m) \exp\left[i(vm - uy)\right] \mathrm{d}m \, \mathrm{d}y. \quad \text{(B.20)}$$

The righthand side of (B.20) is a two-dimensional Fourier transform, which we invert to obtain

$$L(y, m) = \frac{1}{4\pi^2 V} \left(m^2 + 1\right)^{5/2} \iiint \frac{v}{u} P(y_0, v/u) \exp\left\{i\left[u(y - y_0) - vm\right]\right\} \mathrm{d}y_0 \, \mathrm{d}u \, \mathrm{d}v, \quad \text{(B.21)}$$

for $|y| < A/2$.

If the tomographic constraints are too restrictive, we can relax them and just apply pixel value constraints in the far zone, where $R \gg A$. For values of $y$ where $|y| \le A/2$, $m_0 = -y_0/R$ approximately solves $y - Rm_0 - y_0 = 0$. We perform the integration over $m$ in (B.19) to obtain

$$P = \frac{V}{R} \left(m_0^2 + 1\right)^{-5/2} \int_{-A/2}^{A/2} L(y, m_0) \, \mathrm{d}y. \quad \text{(B.22)}$$

For $A \to \infty$, (B.22) specifies the marginal distribution of $L(y, m)$ with respect to $m$. There are many functions $L$ that satisfy (B.22), allowing more flexibility for light field design.

## B.4   Phase tightening analysis

We elaborate on the steady-state distribution (Section B.4.1) and bias calculation for the noise edge configuration (Section B.4.2) in phase tightening.

### B.4.1   Steady state calculations

We provide details for the proofs of two claims that specify the steady-state distribution of the Markov chain in Chapter 6.

**General steady state proof**

We describe the detailed calculations omitted from the proof of Claim 6.3.4. We show that

$$\pi_k = \lambda_{k-1}\pi_{k-1} + \bar{\lambda}_{k+1}\pi_{k+1}, \tag{B.23}$$

for $0 \leq k \leq Q$. There are 9 cases to consider.

**Case 1:** $0 < k < M - 1$. We have

$$
\begin{aligned}
\lambda_{k-1}\pi_{k-1} + \bar{\lambda}_{k+1}\pi_{k+1} &= \lambda_{k-1} \prod_{i=M}^{Q+k-2} \lambda_i \prod_{i=k}^{M-1} \bar{\lambda}_i + \bar{\lambda}_{k+1} \prod_{i=M}^{Q+k} \lambda_i \prod_{i=k+2}^{M-1} \bar{\lambda}_i \\
&= \prod_{i=M}^{Q+k-1} \lambda_i \prod_{i=k}^{M-1} \bar{\lambda}_i + \prod_{i=M}^{Q+k} \lambda_i \prod_{i=k+1}^{M-1} \bar{\lambda}_i \\
&= \bar{\lambda}_k \prod_{i=M}^{Q+k-1} \lambda_i \prod_{i=k+1}^{M-1} \bar{\lambda}_i + \lambda_k \prod_{i=M}^{Q+k-1} \lambda_i \prod_{i=k+1}^{M-1} \bar{\lambda}_i \\
&= \prod_{i=M}^{Q+k-1} \lambda_i \prod_{i=k+1}^{M-1} \bar{\lambda}_i \\
&= \pi_k.
\end{aligned}
$$

The next 3 cases proceed in almost exactly the same manner, so we omit the details.

**Case 2:** $M < k < Q - 1$.

**Case 3:** $k = 0$, $1 < M$.

**Case 4:** $k = Q - 1$, $0 < M < Q - 1$.

Up until this point, we have not used the condition (6.14) in the hypothesis of the claim. This condition is vital in the remaining edge cases.

**Case 5:** $k = M - 1$, $1 < M$**.** We have

$$
\begin{aligned}
\lambda_{k-1}\pi_{k-1} + \bar{\lambda}_{k+1}\pi_{k+1} &= \lambda_{M-2}\prod_{i=M}^{Q+M-3}\lambda_i \prod_{i=M-1}^{M-1}\bar{\lambda}_i + \bar{\lambda}_M\prod_{i=M+1}^{M+Q-1}\bar{\lambda}_i \\
&= \bar{\lambda}_{M-1}\prod_{i=M}^{Q+M-2}\lambda_i + \prod_{i=0}^{Q-1}\bar{\lambda}_i \\
&= \bar{\lambda}_{M-1}\prod_{i=M}^{Q+M-2}\lambda_i + \prod_{i=0}^{Q-1}\lambda_i \\
&= \bar{\lambda}_{M-1}\prod_{i=M}^{Q+M-2}\lambda_i + \lambda_{M-1}\prod_{i=M}^{Q+M-2}\lambda_i \\
&= \prod_{i=M}^{Q+M-2}\lambda_i \\
&= \pi_k.
\end{aligned}
$$

The last 4 cases proceed similarly. We omit the details.

**Case 6:** $k = M$, $0 < M < Q - 1$**.**

**Case 7:** $k = 0$, $M = 1$**.**

**Case 8:** $k = Q - 1$, $M = 0$**.**

**Case 9:** $k = Q - 1$, $M = Q - 1$**.**

**Simple steady state proof**

We present an alternative proof of Claim 6.3.5 using direct calculation. We treat (6.17) as a candidate solution and show that it satisfies (6.12) to conclude that it is a scaled steady-state probability vector solution. We compute

$$
\begin{aligned}
\lambda_{k-1}\pi_{k-1} + \bar{\lambda}_{k+1}\pi_{k+1} &= \lambda_{k-1}\prod_{i=k-\frac{Q}{2}}^{k-2}\lambda_i + \lambda_{k-\frac{Q}{2}+1}\prod_{i=k-\frac{Q}{2}+2}^{k}\lambda_i \\
&= \prod_{i=k-\frac{Q}{2}}^{k-1}\lambda_i + \prod_{i=k-\frac{Q}{2}+1}^{k}\lambda_i \\
&= \lambda_{k-\frac{Q}{2}}\prod_{i=k-\frac{Q}{2}+1}^{k-1}\lambda_i + \lambda_k\prod_{i=k-\frac{Q}{2}+1}^{k-1}\lambda_i \\
&= \bar{\lambda}_k\prod_{i=k-\frac{Q}{2}+1}^{k-1}\lambda_i + \lambda_k\prod_{i=k-\frac{Q}{2}+1}^{k-1}\lambda_i \\
&= \prod_{i=k-\frac{Q}{2}+1}^{k-1}
\end{aligned}
$$

$$= \pi_k.$$

Note that $\boldsymbol{\pi}$ consists of nonnegative entries. We assume towards a contradiction that $\boldsymbol{\pi} = 0$. Since the indices where $\lambda_i = 0$ are consecutive (Claim 6.3.2) and each $\pi_k$ in (6.17) is the product of $Q/2 - 1$ consecutive factors $\lambda_i$, at least $Q - (Q/2 - 1) + 1 = Q/2 + 2$ of the $\lambda_i$ must be zero. But that is more than half of them, which is impossible since the convolution relationship in (6.11) would force the continuous density $p_v$ to be identically zero. Thus $\boldsymbol{\pi}$ is a scaled steady-state probability vector solution.

## B.4.2   Analysis of $S_P$ for bias

In digital phase tightening, the summation

$$
\begin{aligned}
S_P &= \sum_{k=0}^{P} \pi_k^P \\
&= \sum_{k=0}^{P} (P + \epsilon)_k \, (P)_{P-k}
\end{aligned}
\tag{B.24}
$$

in (6.33) appears in our analysis of bias for the noise edge configuration. We now further analyze $S_P$. We first prove that there is no closed-form expression for $S_P$. Then, we establish two recurrence relations. Finally, we derive an asymptotic approximation to $S_P$ as $P \to \infty$. We base our analysis on the observation that $S_P$ resembles a hypergeometric series, which is a series where the ratio of two consecutive terms is a rational function of the summation index. The theory of hypergeometric functions is well-developed [97].

### No closed-form expression

We define a closed-form expression as a linear combination of a finite number of hypergeometric terms [97]. We denote each term in $S_P$ by

$$t_k = (P - 1 + \epsilon)_k (P - 1)_{P-1-k}, \tag{B.25}$$

so that

$$\sum_{k=0}^{P-1} t_k = S_{P-1}. \tag{B.26}$$

We seek a closed-form expression for when $0 < \epsilon < 1$. We define the consecutive term ratio by

$$r(k) = \frac{t_{k+1}}{t_k} = \frac{P - 1 - k + \epsilon}{k + 1}. \tag{B.27}$$

$S_P$ is not an indefinite hypergeometric series, since $r(k)$ depends on $P$ and is therefore not purely a rational function of $k$. To apply the hypergeometric theory, we treat $P$ as a constant $P_0$. Specifically, we decouple the $P$ in (B.25) from the $P$ in

(B.26). The terms become

$$t_k = (P_0 - 1 + \epsilon)_k (P_0 - 1)_{P_0 - 1 - k}, \tag{B.28}$$

and the consecutive term ratio becomes

$$r(k) = \frac{t_{k+1}}{t_k} = \frac{P_0 - 1 - k + \epsilon}{k + 1}. \tag{B.29}$$

The summation $\sum_{k=0}^{P-1} t_k$ is hypergeometric, so we may apply Gosper's algorithm to determine if there is a closed-form expression [97].

To apply Gosper's algorithm, we must write $r(k)$ in the form

$$r(k) = \frac{a(k)}{b(k)} \frac{c(k + 1)}{c(k)}, \tag{B.30}$$

where $a(k)$, $b(k)$, and $c(k)$ are polynomials such that $\gcd\left[a(k), b(k + h)\right] = 1$ for all nonnegative integers $h$. This is done by setting $a(k) = P_0 - 1 - k + \epsilon$, $b(k) = k + 1$, and $c(k) = 1$. Next, we ask if there is a polynomial solution $x(k)$ to

$$a(k)x(k + 1) - b(k - 1)x(k) = c(k), \tag{B.31}$$

which is

$$(P_0 - 1 - k + \epsilon)x(k + 1) - kx(k) = 1. \tag{B.32}$$

Say $x(k)$ has leading term $\alpha k^n$. From the coefficient of highest degree on the lefthand side of (B.32), we conclude that $\alpha = 0$. Thus $x(k) = 0$, which does not solve (B.32). Thus, Gosper's algorithm implies there is no closed-form expression for $\sum_{k=0}^{P-1} t_k$.

**Recurrence relations**

First, we prove the recurrence relation for $S_P$ in (6.36).

**Claim B.4.1.**
$$S_P = 2PS_{P-1} + (P - 1 + \epsilon)_P. \tag{B.33}$$

*Proof.*

$$
\begin{aligned}
S_P &= \sum_{i=0}^{P} (P + \epsilon)_i \, (P)_{P-i} \\
&= P! + (P + \epsilon)_P + P \sum_{i=1}^{P-1} (P - 1 + \epsilon)_{i-1} \, (P - 1)_{P-1-i}(P + \epsilon - i + i) \\
&= P! + (P + \epsilon)_P + P \left[ \sum_{i=1}^{P-1} (P - 1 + \epsilon)_i \, (P - 1)_{P-1-i} \right. \\
&\qquad\qquad \left. + \sum_{i=1}^{P-1} (P - 1 + \epsilon)_{i-1} \, (P - 1)_{P-i} \right]
\end{aligned}
$$

$$
\begin{aligned}
&= P\left[(P-1)! + \frac{P+\epsilon}{P}(P-1+\epsilon)_{P-1}\right.\\
&\qquad\qquad + \sum_{i=1}^{P-1}(P-1+\epsilon)_i\,(P-1)_{P-1-i}\\
&\qquad\qquad \left. + \sum_{i=0}^{P-2}(P-1+\epsilon)_i\,(P-1)_{P-1-i}\right]\\
&= P\left[2\sum_{i=0}^{P-1}(P-1+\epsilon)_i\,(P-1)_{P-1-i} + \frac{\epsilon}{P}(P-1+\epsilon)_{P-1}\right]\\
&= 2P\sum_{i=0}^{P-1}(P-1+\epsilon)_i\,(P-1)_{P-1-i} + \epsilon(P-1+\epsilon)_{P-1}\\
&= 2PS_{P-1} + (P-1+\epsilon)_P.
\end{aligned}
$$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Second, we develop a recurrence relation for the terms of $S_P$ using Zeilberger's algorithm [97]. We denote

$$
F(P,k) = (P+\epsilon)_k (P)_{P-k}, \tag{B.34}
$$

so that $S_P = \sum_{k=0}^{P} F(P,k)$. Zeilberger's algorithm results in the recurrence

$$
-2(P+1)F(P,k) + F(P+1,k) = G(P,k+1) - G(P,k), \tag{B.35}
$$

where

$$
\begin{aligned}
G(P,k) &= -(P+1)_{P-k+2}(P+\epsilon)_{k-1}\\
&= -(P+1)F(P,k-1).
\end{aligned} \tag{B.36}
$$

This recurrence is equivalent to

$$
F(P+1,k) = (P+1)\left[F(P,k) + F(P,k-1)\right]. \tag{B.37}
$$

Zeilberger would then have us sum (B.37) over all $k$ to recover the series we are interested in. While we recover $S_P$ in the case of a binomial series where $\epsilon = 0$, we do not recover $S_P$ in general.

## Asymptotic approximation

If we fix $P$ in each term of $S_P$ and take an infinite sum, we obtain a familiar hypergeometric series

$$
\sum_{k=0}^{\infty}(P+\epsilon)_k\,(P)_{P-k} = {}_1F_0\left[\begin{array}{c}-(P+\epsilon)\\-\end{array}; -1\right]P!. \tag{B.38}
$$

There is a well-known identity

$$
{}_1F_0\left[\begin{array}{c} a \\ - \end{array}; z\right] = \frac{1}{(1-z)^a},
\tag{B.39}
$$

which implies that

$$
\sum_{k=0}^{\infty} (P+\epsilon)_k \, (P)_{P-k} = 2^{P+\epsilon} P!.
\tag{B.40}
$$

It is therefore reasonable to suspect that asymptotically

$$
S_P \to 2^{P+\epsilon} P!.
\tag{B.41}
$$

We now prove (B.41). First, we use the recurrence relation in (B.4.1) to completely expand $S_P$ down to $S_0 = 1$:

$$
\begin{aligned}
S_P &= 2^P P! + \sum_{i=0}^{P-1} 2^{P-1-i}(i+\epsilon)_{i+1}(P)_{P-1-i} \\
&= 2^P P! + 2^P P! \sum_{i=0}^{P-1} \frac{(i+\epsilon)_{i+1}}{2^{i+1}(i+1)!} \\
&= 2^P P! + 2^P P! \sum_{i=1}^{P} \frac{(i-1+\epsilon)_i}{2^i i!} \\
&= 2^P P! + 2^P P! \sum_{i=1}^{P} \frac{1}{i2^i}\left(1+\frac{\epsilon}{i-1}\right)\left(1+\frac{\epsilon}{i-2}\right)\cdots(1+\epsilon)\,\epsilon.
\end{aligned}
\tag{B.42}
$$

We have eliminated $P$ from the summand in (B.42). We now expand the summation in terms of powers of $\epsilon$. As $P \to \infty$, the coefficient $C_k$ of $\epsilon^k$ in the summation in (B.42) tends to

$$
\lim_{P\to\infty} C_k = \sum_{i=1}^{\infty} \frac{1}{i2^i} \sum_{l_1=1}^{i-1} \sum_{l_2=l_1+1}^{i-1} \cdots \sum_{l_{k-1}=l_{k-2}+1}^{i-1} \frac{1}{l_1 \cdots l_{k-1}}.
\tag{B.43}
$$

We compute the summation in (B.43) using generating functions.

**Claim B.4.2.** *For $k \geq 1$ define*

$$
f_k(x) = \sum_{i=1}^{\infty} \frac{x^i}{i2^i} \sum_{l_1=1}^{i-1} \sum_{l_2=l_1+1}^{i-1} \cdots \sum_{l_{k-1}=l_{k-2}+1}^{i-1} \frac{1}{l_1 \cdots l_{k-1}}.
\tag{B.44}
$$

*Then*

$$
f_k(x) = \frac{\log^k \frac{2}{2-x}}{k!}.
\tag{B.45}
$$

*Proof.* We proceed by induction. For the base case,

$$
\begin{aligned}
f_1(x) &= \sum_{i=1}^{\infty} \frac{x^i}{i2^i} = \int_0^x \sum_{i=1}^{\infty} \frac{t^{i-1}}{2^i} \mathrm{d}t \\
&= \int_0^x \frac{1}{2-t} \mathrm{d}t = -\log(2-t)\big|_0^x \\
&= \log \frac{2}{2-x}.
\end{aligned}
$$

Next,

$$
\begin{aligned}
f_{k+1}(x) &= \sum_{i=1}^{\infty} \frac{x^i}{i2^i} \sum_{l_1=1}^{i-1} \sum_{l_2=l_1+1}^{i-1} \cdots \sum_{l_k=l_{k-1}+1}^{i-1} \frac{1}{l_1 \cdots l_k} \\
&= \sum_{l_1=1}^{\infty} \sum_{l_2=l_1+1}^{\infty} \cdots \sum_{l_k=l_{k-1}+1}^{\infty} \frac{1}{l_1 \cdots l_k} \sum_{i=l_k+1}^{\infty} \frac{x^i}{i2^i} \\
&= \sum_{l_1=1}^{\infty} \sum_{l_2=l_1+1}^{\infty} \cdots \sum_{l_k=l_{k-1}+1}^{\infty} \frac{1}{l_1 \cdots l_k} \int_0^x \sum_{i=l_k+1}^{\infty} \frac{t^{i-1}}{2^i} \mathrm{d}t \\
&= \sum_{l_1=1}^{\infty} \sum_{l_2=l_1+1}^{\infty} \cdots \sum_{l_k=l_{k-1}+1}^{\infty} \frac{1}{l_1 \cdots l_k} \int_0^x \frac{\left(\frac{t}{2}\right)^{l_k}}{2-t} \mathrm{d}t \\
&= \int_0^x \frac{1}{2-t} \sum_{l_1=1}^{\infty} \sum_{l_2=l_1+1}^{\infty} \cdots \sum_{l_k=l_{k-1}+1}^{\infty} \frac{1}{l_1 \cdots l_k} \left(\frac{t}{2}\right)^{l_k} \mathrm{d}t \\
&= \int_0^x \frac{1}{2-t} \sum_{l_k=1}^{\infty} \frac{t^{l_k}}{l_k 2^{l_k}} \sum_{l_1=1}^{l_k-1} \sum_{l_2=l_1+1}^{l_k-1} \cdots \sum_{l_{k-1}=l_{k-2}+1}^{l_k-1} \frac{1}{l_1 \cdots l_{k-1}} \mathrm{d}t \\
&= \int_0^x \frac{1}{2-t} f_k(t) \mathrm{d}t = \int_0^x \frac{1}{2-t} \frac{\log^k \frac{2}{2-t}}{k!} \mathrm{d}t \\
&= \frac{\log^{k+1} \frac{2}{2-x}}{(k+1)!}.
\end{aligned}
$$

$\square$

We evaluate $f_k(1)$ to conclude that

$$
C_k = \frac{\log^k 2}{k!}. \tag{B.46}
$$

Returning to (B.42), we compute

$$
\begin{aligned}
\sum_{i=1}^{\infty} \frac{1}{i2^i} \left(1 + \frac{\epsilon}{i-1}\right) \left(1 + \frac{\epsilon}{i-2}\right) \cdots (1+\epsilon) \epsilon &= \sum_{i=1}^{\infty} C_i \epsilon^i \\
&= \sum_{i=1}^{\infty} \frac{\log^i 2}{i!} \epsilon^i \\
&= 2^\epsilon - 1,
\end{aligned}
$$

so that

$$S_P \rightarrow 2^{P+\epsilon} P!. \tag{B.47}$$

# Appendix C

# Simulation toolbox

We have used two types of simulations in the thesis. First, we have simulated wave diffraction to physically model quasi light field performance for millimeter-wave imaging in Chapter 4. Second, we have used Monte Carlo techniques to estimate expected values of beam pattern metrics and phase tightening bias, in Chapters 5 and 6. We describe our wave simulation technique (Section C.1) and include Matlab scripts for selected Monte Carlo simulations (Section C.2).

## C.1  Wave propagation in free space

While the search for general-purpose numerical solutions to the wave equation is an active area of research [98], specialized approximations are often leveraged in commercial software [56]. For our purposes, evaluating the Rayleigh diffraction formula is too computationally complex. While the Fresnel diffraction approximation allows us to simplify the formula, the errors introduced are too severe for the near-zone behavior we are interested in. We therefore instead implement the plane-wave propagation method.

The plane-wave propagation method makes use of the fact that it is simple to compute the propagation of a plane wave across free space, as the appropriate phase adjustment need only be applied. Therefore, the propagation of an arbitrary scalar field can be determined by using a spatial Fourier transform to decompose the field into component plane waves, advancing each plane wave by the appropriate phase, and then using an inverse Fourier transform to obtain the resulting scalar field. We first describe the method (Section C.1.1) and then discuss implementation issues (Section C.1.2).

### C.1.1  The plane-wave propagation method

The plane-wave propagation method determines the scalar field $U_2$ on the plane $z = z_2$ from the scalar field $U_1$ on the plane $z = z_1 < z_2$. The angular decomposition on the plane $z = z_2$ is

$$U_2(\mathbf{r}_2) = \int a(\mathbf{s}) \exp(ik\mathbf{s} \cdot \mathbf{r}_2) \, \mathrm{d}^2 s, \tag{C.1}$$

where we express the plane wave component $a(\mathbf{s})$ in terms of the field on the plane $z = z_1$:

$$a(\mathbf{s}) = \left(\frac{k}{2\pi}\right)^2 \int U_1(\mathbf{r}_1) \exp(-ik\mathbf{s} \cdot \mathbf{r}_1) \, \mathrm{d}^2 r_1. \tag{C.2}$$

We choose the coordinates $\mathbf{r}_i = (x_i, y_i, z_i)$ and $\mathbf{s} = (p, q, \sqrt{1 - p^2 - q^2})$, and combine (C.1) with (C.2) to obtain

$$U_2(\mathbf{r}_2) = \iint \left(\frac{k}{2\pi}\right)^2 \exp\left[ik(px_2 + qy_2)\right] \exp\left[ik(z_2 - z_1)\sqrt{1 - p^2 - q^2}\right]$$
$$\times \iint U_1(x_1, y_1) \exp\left[-ik(px_1 + qy_1)\right] \, \mathrm{d}x_1 \, \mathrm{d}y_1 \, \mathrm{d}p \, \mathrm{d}q. \tag{C.3}$$

We substitute $u = kp$ and $v = kq$ into (C.3), so that

$$U_2(\mathbf{r}_2) = \iint \mathscr{U}_1(u, v; z_1) \exp\left[i(z_2 - z_1)\sqrt{k^2 - u^2 - v^2}\right] \exp\left[i(ux_2 + vy_2)\right] \, \mathrm{d}u \, \mathrm{d}v, \tag{C.4}$$

where $\mathscr{U}_1(u, v; z_1)$ is the two-dimensional spatial Fourier transform of $U_1$. Formula (C.4) quantifies our earlier intuition: to compute the field on $z = z_2$, we compute the Fourier transform of the field on $z = z_1$, apply a phase factor, then compute the inverse transform. The primary shortcoming of the plane wave propagation method is evident in (C.4). Specifically, when using a discrete Fourier transform (DFT) to compute (C.4), the samples of $u$ and $v$ must be dense enough to capture the rapid oscillations of the phase factor, which increases with distance $z_2 - z_1$. Therefore, the plane-wave propagation method is ideally suited for making near-zone propagation calculations.

We now specialize the plane-wave propagation method to two dimensions by assuming that the scalar field $U_1(\mathbf{r}_1)$ is independent of $x_1$. We compute

$$\mathscr{U}_1(u, v; z_1) = \delta(u)\mathscr{U}_1(v; z_1), \tag{C.5}$$

where $\mathscr{U}_1(v; z_1)$ is the one-dimensional spatial Fourier transform of $U_1$ with respect to $y_1$. We integrate over $u$ in (C.4) to obtain

$$U_2(\mathbf{r}_2) = \int \mathscr{U}_1(v; z_1) \exp\left[i(z_2 - z_1)\sqrt{k^2 - v^2}\right] \exp(ivy_2) \, \mathrm{d}v. \tag{C.6}$$

The plane-wave propagation method thereby reduces to a one-dimensional procedure of computing the (one-dimensional) Fourier transform, applying a phase delay, and then computing the inverse transform.

## C.1.2   Discrete Fourier transform implementation

We implement the plane-wave propagation method with the discrete Fourier transform. The spatial sample spacing $d_0$ for $U_1$ should be fine enough to capture the high-frequency details in the scalar field and to avoid aliasing. We choose $d_0 = d$,

the sensor array spacing for light field measurement, for two reasons. First, since $d_0$ divides $d$, we do not need to interpolate the simulated scalar field to make sensor measurements. Second, the discrete-time Fourier transform of $U_1$ consists of scaled copies of $\mathscr{U}_1(v; z_1)$ shifted every $v = 2\pi/d_0$ [25]. Since $|v| = k|q| \leq k|\mathbf{s}| = k$, aliasing is eliminated when $2\pi/d_0 \geq 2k$, which is equivalent to $d_0 \leq \lambda/2 = d$.

The DFT size $N$ must be large enough to include the support of the scalar field. Specifically, the aliasing introduced by sampling the phase factor in (C.6) must be negligible, which is equivalent to saying that the scalar field should fall off to zero at the boundary of the sampling region on both the source plane $z = z_1$ and the destination plane $z = z_2$. Any residual field falling outside of the sampling region will be aliased, as multiplying the $N$-point DFTs of two sequences is equivalent to circularly convolving their space domain representations, wrapping every $N$ samples. We choose $N$ large enough to keep the aliasing error below a desirable level.

## C.2   Monte Carlo code

We use Monte Carlo techniques to compute the expected beam pattern to analyze the main lobe tilt (Section C.2.1), sidelobe suppression (Section C.2.2), and beam width (Section C.2.3) in the presence of phase noise. We also use Monte Carlo techniques to compute the expected bias of the phase tightening phase estimate (Section C.2.4).

### C.2.1   Main lobe tilt

One challenge in simulating the main lobe tilt is that the combination of a large array (large $M$) and modest phase noise (small $\sigma_\phi$) requires high beam pattern resolution in order to estimate the distribution of $\tau$. We therefore concentrate on a narrow band about the center of the main lobe. The Matlab function `computeTiltStats` is parameterized for this purpose, and computes both the $p$-value of the Shapiro-Wilk test as well as an estimate of $\sigma_\tau$.

```
function [oTauPval, oTauStd] = computeTiltStats(aN, aPhaseStd, aNumSamps, aAngRes, aAngWid)
%
%   [oTauPval, oTauStd] = computeTiltStats(aN, aPhaseStd, aNumSamps, aAngRes, aAngWid)
%
%   Computes statistics on the tilt of a beam pattern when beamforming
%   in the presence of phase noise. This is done via Monte Carlo
%   simulation.
%
%   aN        : Number of sensors in the array.
%   aPhaseStd : Standard deviation of the phase noise at each sensor.       10
%   aNumSamps : Number of samples to use for Monte Carlo simulation.
%   aAngRes   : Angular resolution for the beam pattern.
%   aAngWid   : Width (in radians) for the beam pattern, centered about pi/2.
%
%   oTauPval  : P-value for Shapiro-Wilk (2-sided) test for normality for tilt.
%   oTauStd   : Estimated standard deviation of tilt.

    ang = pi/2 - aAngWid/2 : aAngWid/aAngRes : pi/2 + aAngWid/2 - aAngWid/aAngRes;

    %   delays(i,j) is the delay to apply to sensor i to steer along angle ang(j)    20
    delays = (0:(aN-1))' * pi * cos(ang);

    empDist = zeros(1, aNumSamps);
    oTauStd = 0;

    for idx = 1:aNumSamps
        noisyDelays = delays + aPhaseStd * randn(size(delays));
        beamPattern = abs( sum(exp(j * noisyDelays)) );
```

```
    [a, maxIdx] = max(beamPattern);                                        30

    empDist(idx) = ang(maxIdx);
    oTauStd = oTauStd + (ang(maxIdx) − pi/2)^2;
  end;

  %   swtest.m is available from www.mathworks.com, on the Matlab Central
  %   File Exchange. It was published by Ahmed Ben Saida on 2/15/2007.

  [a, oTauPval] = swtest(empDist, 0.01, 0);
  oTauStd = sqrt(oTauStd / aNumSamps);                                     40

end
```

Since the tilt breakpoint is monotonically increasing in array size $M$, we use the $\sigma_\phi^{\mathrm{TB}}$ for $M$ as an initial starting point for estimating the $\sigma_\phi^{\mathrm{TB}}$ for $M+1$. The Matlab script computeTiltBreakPoints.m incorporates this optimization.

```
% computeTiltBreakPoints.m
%
% Compute tilt breakpoints by determining the noise level at which
% the empirical distribution of the tilt no longer appears Gaussian.

sigma = 0.4;
sigStep = pi/1000;

N = 1;
maxN = 1000;                                                               10

TiltBreakPoints = zeros(maxN, 1);

while (N <= maxN)
    pval = 1;
    strikes = 0;

    while (pval >= 0.01 || strikes < 3)
        pval = computeTiltStats(N, sigma, 200, 500, 2*pi/N);              20

        if (pval < 0.01)
            strikes = strikes + 1;
        else
            strikes = 0;
        end;

        sigma = sigma + sigStep;
    end;

    % Take a consevative step back.                                       30
    sigma = sigma − 2 * 4*sigStep;

    TiltBreakPoints(N) = sigma;
    N = N + 1;

end;
```

The tilt breakpoint estimates enable us to select appropriate values for $\sigma_\phi$ to estimate the slope $\mathrm{d}\sigma_\tau/\mathrm{d}\sigma_\phi$, as done in the Matlab script computeTiltSlopes.m.

```
% computeTiltSlopes.m
%
% Compute slope of tilt standard deviation with respect to phase noise
% standard deviation.

sigLen = 30;
N = 1:1000;
TiltStd = zeros(sigLen, length(N));
TiltSlopes = zeros(length(N), 1);
                                                                          10
for nidx = 1:length(N)
    sigStart = pi/50;
    sigEnd = 0.9 * TiltBreakPoints(N(nidx));
    sigma = sigStart : (sigEnd−sigStart)/sigLen : sigEnd − (sigEnd−sigStart)/sigLen;

    for sidx = 1:sigLen
        [pval, tauStd] = computeTiltStats(N(nidx), sigma(sidx), 100, 500, pi/N(nidx));
        TiltStd(sidx, nidx) = tauStd;
    end;
                                                                          20
    coefs = polyfit(sigma, TiltStd(:, nidx)', 1);
    TiltSlopes(nidx) = coefs(1);

end;
```

## C.2.2    Sidelobe suppression

To analyze the sidelobes of the beam pattern, we first determine the main lobe and sidelobe peak locations in the absence of noise. Then we estimate the mean and standard deviation of the beam pattern at the peak locations. We use the same technique to compute expected beam patterns and more complicated variance estimates. The Matlab script `computeSidelobeStats.m` estimates the key sidelobe suppression statistics: $\mu_S$, $\mu_M$, $\sigma_S$, $\sigma_M$, and $S^{\mathrm{SL}}$.

```matlab
% computeSidelobeStats.m
%
% Compute the expected value and standard deviation of the beam
% pattern at the main lobe and sidelobe peaks.

sigma = 0 : 2.4/100 : 2.4 - 2.4/100;
maxN = 1000;
numSamps = 1000;

mu_S = zeros(length(sigma), maxN);
mu_M = zeros(length(sigma), maxN);
sig_S = zeros(length(sigma), maxN);
sig_M = zeros(length(sigma), maxN);
S_SL = zeros(length(sigma), maxN);

for N = 2:maxN
    angWidth = 3*pi/N;
    angRes = 500;
    ang = pi/2 - angWidth/2 : angWidth/angRes : pi/2 + angWidth/2 - angWidth/angRes;
    delays = (0:N-1)' * pi * cos(ang);
    cleanBeamPattern = abs( sum(exp(j * delays)) ) / N;

    % First find the peak locations in the absence of phase noise.

    [a, mainPeakIdx] = max(cleanBeamPattern);
    firstDipOffset = find(cleanBeamPattern(mainPeakIdx:length(cleanBeamPattern)) < 0.05);
    firstZeroIdx = mainPeakIdx-1 + firstDipOffset(1);
    [a, sidePeakOffset] = max(cleanBeamPattern(firstZeroIdx:length(cleanBeamPattern)));
    sidePeakIdx = firstZeroIdx-1 + sidePeakOffset;

    peakAng = [ang(mainPeakIdx) ang(sidePeakIdx)];
    peakDelays = (0:N-1)' * pi * cos(peakAng);

    for sidx = 1:length(sigma)
        beamEst = zeros(size(peakAng));
        beamEst2 = zeros(size(peakAng));

        % Then use Monte Carlo to estimate parameters at the peaks.

        for idx = 1:numSamps
            noisyDelays = peakDelays + sigma(sidx) * randn(size(peakDelays));
            beamPattern = abs( sum(exp(j * noisyDelays)) ) / N;

            beamEst = beamEst + beamPattern;
            beamEst2 = beamEst2 + beamPattern.^2;
        end;

        beamEst = beamEst / numSamps;
        beamEst2 = beamEst2 / numSamps;

        mu_S(sidx, N) = beamEst(2);
        mu_M(sidx, N) = beamEst(1);
        sig_S(sidx, N) = sqrt(beamEst2(2) - beamEst(2)^2);
        sig_M(sidx, N) = sqrt(beamEst2(1) - beamEst(1)^2);

        S_SL(sidx, N) = 20*log10(beamEst(2) / beamEst(1));
    end;

end;
```

The Matlab script `computeSidelobeBreakPoints.m` computes the sidelobe breakpoint $\sigma_\phi^{\mathrm{SB}}$ as a function of array size $M$, and implements heuristics to determine the endpoints used for the linear regression.

```matlab
% computeSidelobeBreakPoints.m
%
% Fit lines to the sidelobe suppression graphs to find the sidelobe
% breakpoints.

S_SLSlope = zeros(1000, 1);
```

```
SidelobeBreakPoints  =  zeros(1000, 1);

for  N  =  2:1000
                                                                                              10
    %  Define  the  region  over  which  to  perform  the  linear  regression.

    if  (N  <  10)
        startIdx  =  min(find(S_SL(:, N)  >  −6  −  N/2));
        endIdx  =  max(find(S_SL(:, N)  <  −4));
    else
        startIdx  =  min(find(S_SL(:, N)  >  −11));
        endIdx  =  max(find(sigma  <  TiltBreakPoints(N)));
    end;
                                                                                              20
    P  =  polyfit(sigma(startIdx:endIdx)', S_SL(startIdx:endIdx, N), 1);

    %  Intersect  with  the  initial  flat  portion  to  solve  for  the  breakpoint.

    S_SLSlope(N)  =  P(1);
    SidelobeBreakPoints(N)  =  (S_SL(1, N)  −  P(2))  /  P(1);

end;
```

## C.2.3   Beam width

The Matlab script `computeNoisyBeamWidth.m` computes the noisy beam width $\theta_{\mathrm{NB}}$.

```
%  computeNoisyBeamWidth.m
%
%  Computes  metric  for  beam  width  in  the  presence  of  phase  noise,  for  a
%  range  of  array  sizes  N  and  noise  standard  deviations  below  the  tilt
%  breakpoint.

sigLen  =  20;
maxN  =  3;
numSamps  =  300;
angRes  =  500;                                                                               10

NoisyBeamWidth  =  zeros(sigLen, maxN);

for  N  =  1:maxN
    sigEnd  =  0.7  *  TiltBreakPoints(N);
    sigma  =  0  :  sigEnd/sigLen  :  sigEnd  −  sigEnd/sigLen;

    angWidth  =  pi/N;
    ang  =  pi/2  −  angWidth/2  :  angWidth/angRes  :  pi/2  +  angWidth/2  −  angWidth/angRes;
                                                                                              20
    for  sidx  =  1:length(sigma)
        delays  =  (0:N−1)'  *  pi  *  cos(ang);

        beamEst  =  zeros(size(ang));
        beamEst2  =  zeros(size(ang));

        for  idx  =  1:numSamps
            noisyDelays  =  delays  +  sigma(sidx)  *  randn(size(delays));
            beamPattern  =  abs( sum(exp(j * noisyDelays)) )  /  N;
                                                                                              30
            beamEst  =  beamEst  +  beamPattern;
            beamEst2  =  beamEst2  +  beamPattern.^2;
        end;

        beamEst  =  beamEst  /  numSamps;
        beamEst2  =  beamEst2  /  numSamps;

        sigTheta  =  beamEst2  −  beamEst.^2;
        sigTheta(find(sigTheta  <  0))  =  0;
        sigTheta  =  sqrt(sigTheta);                                                          40

        hiBeam  =  beamEst  +  sigTheta;
        lowBeam  =  beamEst  −  sigTheta;

        NoisyBeamWidth(sidx, N)  =  length(find(hiBeam/max(lowBeam)  >  sqrt(1/2)))  *  angWidth/angRes;
    end;

end;
```

## C.2.4   Phase tightening bias

The impact of Gaussian phase noise on phase tightening performance is more difficult to analyze analytically than the impact of uniform phase noise. We consequently use the Matlab function `ditherNormExpVal` to compute the expected bias of the phase estimates for Gaussian noise.

```matlab
function [E, Lam, pdist, ptDist] = ditherNormExpVal(N, gam, sig)
%
%   E = ditherNormExpVal(N, gam, sig)
%
%   Dither with Gaussian distribution, using Markov chain to compute
%     expected value.
%
%   N    : number of states in chain (spacing between 0 and 1, even)
%   gam  : true value to estimate
%   sig  : standard deviation of noise                                      10
%

if (mod(N, 2))
    error('N must be even');
end;

Del = 1 / N;
Lam = zeros(N , 1);

% ptDist is organized by bucket of length Del, aliased into the            20
% range [0, 1).          It's a discrete, aliased version of the
% continuous noise distribution.

ptDist = zeros(2*N, 1);

% Go 10 standard deviations out to approximate support.

W = 10 * sig;
leftEdge = gam - W;
                                                                           30
sweepLine = (floor(N*leftEdge) + 1) / N;
idx = round(mod(N*sweepLine, N));
if (idx == 0)
    idx = N;
end;

while (sweepLine <= gam + W)
    ptDist(idx) = ptDist(idx) + (normcdf(sweepLine, gam, sig) - normcdf(leftEdge, gam, sig));

    leftEdge = sweepLine;                                                  40
    sweepLine = sweepLine + Del;
    idx = idx + 1;
    if (idx > N)
        idx = 1;
    end;
end;

ptDist(idx) = ptDist(idx) + (normcdf(gam + W, gam, sig) - normcdf(leftEdge, gam, sig));

ptDist(N+1:2*N) = ptDist(1:N);                                             50

% Here's where the assumption that N is even comes in.
% With N even, the decision region is split nicely into whole
% buckets.        Otherwise, we would have to cut buckets in half.

for idx = 1:N
    Lam(idx) = sum(ptDist(idx:idx+N/2-1));
end;

                                                                           60
pdist = solveSteadyStateMatrix(Lam);

% now average about ML estimate

[z, midx] = maxfudge(pdist);
E = (midx-1) / N;

ks = midx;
ke = midx;
                                                                           70
for cnt=1:N/2
    ks = ks - 1;
    if (ks < 1)
        ks = N;
    end;

    ke = ke + 1;
    if (ke > N)
```

```
          ke  =  1;
        end;                                                                        80

        E  =  E  +  cnt  *  Del  *  (pdist(ke)  −  pdist(ks));
      end;

    E  =  mod(E,  1);
    if  (E  >  1  −  eps)
      E  =  0;
    end;
end                                                                                 90

%

function  pdist  =  solveSteadyStateMatrix(aLam)
% set  up  matrix  to  solve  for  steady  state  probability  dist:  Q  p  =  eN

    N  =  length(aLam);
    Q  =  spalloc(N,  N,  4*N−3);
    Q(1,1)  =  1;
    Q(1,2)  =  aLam(2)  −  1;
    Q(1,N)  =  −aLam(N);                                                            100

    for  k  =  2:N−1
        Q(k,  k−1)  =  −aLam(k−1);
        Q(k,  k)  =  1;
        Q(k,  k+1)  =  aLam(k+1)  −  1;
    end;

    Q(N,:)  =  ones(1,N);

    eN  =  zeros(N,  1);                                                            110
    eN(N,1)  =  1;

    pdist  =  Q  \  eN;

end

%

function  [a,b]  =  maxfudge(aVec)
    N  =  length(aVec);                                                             120
    b  =  1;
    a  =  aVec(1);

    for  k=1:N
        if  (aVec(k)  >  a  +  2*eps)
            b  =  k;
            a  =  aVec(k);
        end;
    end;
end                                                                                 130
```

# Bibliography

[1] N. J. Wade, *A Natural History of Vision* (The MIT Press, 1998).

[2] J. A. Jensen, *Estimation of Blood Velocities Using Ultrasound: A Signal Processing Approach* (Cambridge Univ. Press, 1996).

[3] A. E. Siegman, *Lasers* (Univ. Science Books, 1986).

[4] L. Novotny and B. Hecht, *Principles of Nano-Optics* (Cambridge Univ. Press, 2006).

[5] P. Hariharan, *Optical Holography: Principles, Techniques, and Applications*, 2nd ed. (Cambridge Univ. Press, 1996).

[6] W. Simbürger, K. Aufinger, J. Böck, S. Boguth, D. Kehrer, C. Kienmayer, H. Knapp, T. F. Meister, W. Perndl, M. Rest, C. Sandner, H. Schäfer, R. Schreiter, R. Stengl, R. Thüringer, M. Tiebout, H. D. Wohlmuth, M. Wurzer, and A. L. Scholtz, "Silicon-based RF ICs up to 100 GHz: research trends and applications," in *Proc. of ICSICT* (2004), pp. 12–19.

[7] A. Babakhani, X. Guan, A. Komijani, A. Natarajan, and A. Hajimiri, "A 77-GHz phased-array transceiver with on-chip antennas in silicon: receiver and antennas," IEEE J. Solid-St. Circ. **41**, 2795–2806 (2006).

[8] A. Natarajan, A. Komijani, X. Guan, A. Babakhani, and A. Hajimiri, "A 77-GHz phased-array transceiver with on-chip antennas in silicon: transmitter and local LO-path phase shifting," IEEE J. Solid-St. Circ. **41**, 2807–2819 (2006).

[9] M. I. Skolnik, *Introduction to Radar Systems*, 3rd ed. (McGraw-Hill, 2002).

[10] K. M. Strohm, H. Bloecher, R. Schneider, and J. Wenger, "Development of future short range radar technology," in *Proc. of EURAD* (2005), pp. 165–168.

[11] R. W. McMillan, N. C. Currie, D. D. Ferris, and M. C. Wicks, "Concealed weapon detection using microwave and millimeter wave sensors," in *Proc. of ICMMT* (1998), pp. 1–4.

[12] M. Levoy and P. Hanrahan, "Light field rendering," in *Proc. of SIGGRAPH* (1996), pp. 31–42.

[13] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen, "The lumigraph," in *Proc. of SIGGRAPH* (1996), pp. 43–54.

[14] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan, "Light field photography with a hand-held plenoptic camera," Tech. Rep. CTSR 2005-02, Stanford Univ., Calif. (2005).

[15] J. Li and P. Stoica, eds., *MIMO Radar Signal Processing* (John Wiley & Sons, 2009).

[16] R. J. Urick, *Principles of Underwater Sound*, 3rd ed. (McGraw-Hill, 1983).

[17] T. L. Szabo, *Diagnostic Ultrasound Imaging: Inside Out* (Elsevier, 2004).

[18] M. Born and E. Wolf, *Principles of Optics*, 7th ed. (Cambridge Univ. Press, 1999).

[19] A. D. Pierce, *Acoustics: An Introduction to its Physical Principles and Applications* (Acoustical Society of America, 1989).

[20] H. L. Van Trees, *Optimum Array Processing* (John Wiley & Sons, 2002).

[21] H. Krim and M. Viberg, "Two decades of array signal processing research," IEEE Signal Process. Mag. **13**, 67–94 (1996).

[22] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. (McGraw Hill, 1991).

[23] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," Proc. IEEE **57**, 1408–1418 (1969).

[24] J. R. Guerci, "Theory and application of covariance matrix tapers for robust adaptive beamforming," IEEE Trans. Signal Process. **47**, 977–985 (1999).

[25] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing* (Prentice-Hall, 1989).

[26] O. Bimber, ed., *IEEE Computer* (August 2006).

[27] A. Levin, W. T. Freeman, and F. Durand, "Understanding camera trade-offs through a Bayesian analysis of light field projections," in *Proc. of ECCV* (2008), pp. 88–101.

[28] E. Hecht, *Optics*, 2nd ed. (Addison-Wesley, 1987).

[29] E. H. Adelson and J. R. Bergen, "The plenoptic function and the elements of early vision," in *Computational Models of Visual Processing*, M. S. Landy and J. A. Movshon, eds. (The MIT Press, 1991), pp. 3–20.

[30] R. W. Boyd, *Radiometry and the Detection of Optical Radiation* (John Wiley & Sons, 1983).

[31] A. Adams and M. Levoy, "General linear cameras with finite aperture," in *Proc. of EGSR* (2007).

[32] J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, 1968).

[33] E. Kheirandish, *The Arabic Version of Euclid's Optics* (Springer, 1999).

[34] P. Flandrin, *Time-Frequency / Time-Scale Analysis* (Academic Press, 1999).

[35] B. Boashash, ed., *Time Frequency Signal Analysis and Processing: A Comprehensive Reference* (Elsevier, 2003).

[36] A. Accardi and G. Wornell, "Quasi light fields: extending the light field to coherent radiation," J. Opt. Soc. Am. A **26**, 2055–2066 (2009).

[37] W. Chun and O. S. Cossairt, "Data processing for three-dimensional displays," United States Patent 7,525,541 (April 28, 2009).

[38] R. Ziegler, S. Bucheli, L. Ahrenberg, M. Magnor, and M. Gross, "A bidirectional light field – hologram transform," Comput. Graph. Forum **26**, 435–446 (2007).

[39] M. J. Bastiaans, "Application of the Wigner distribution function in optics," in *The Wigner Distribution – Theory and Applications in Signal Processing*, W. Mecklenbräuker and F. Hlawatsch, eds. (Elsevier, 1997), pp. 375–426.

[40] A. Walther, "Radiometry and coherence," J. Opt. Soc. Am. **58**, 1256–1259 (1968).

[41] E. Wigner, "On the quantum correction for thermodynamic equilibrium," Phys. Rev. **40**, 749–759 (1932).

[42] J. Ville, "Théorie et applications de la notion de signal analytique," Cables Transm. **2A**, 61–74 (1948).

[43] K. D. Stephan, "Radiometry before World War II: measuring infrared and millimeter-wave radiation 1800–1925," IEEE Antennas Propag. **47**, 28–37 (2005).

[44] S. Chandrasekhar, *Radiative Transfer* (Dover Publications, 1960).

[45] A. T. Friberg, G. S. Agarwal, J. T. Foley, and E. Wolf, "Statistical wave-theoretical derivation of the free-space transport equation of radiometry," J. Opt. Soc. Am. B **9**, 1386–1393 (1992).

[46] P. Moon and G. Timoshenko, "The light field," J. Math. Phys. **18**, 51–151 (1939). [Translation of A. Gershun, *The Light Field* (Moscow, 1936)].

[47] A. Walther, "Radiometry and coherence," J. Opt. Soc. Am. **63**, 1622–1623 (1973).

[48] E. Wolf, "Coherence and radiometry," J. Opt. Soc. Am. **68**, 6–17 (1978).

[49] G. S. Agarwal, J. T. Foley, and E. Wolf, "The radiance and phase-space representations of the cross-spectral density operator," Opt. Commun. **62**, 67–72 (1987).

[50] A. T. Friberg, "On the existence of a radiance function for finite planar sources of arbitrary states of coherence," J. Opt. Soc. Am. **69**, 192–198 (1979).

[51] G. S. Agarwal and E. Wolf, "Calculus for functions of noncommuting operators and general phase-space methods in quantum mechanics. I. Mapping theorems and ordering of functions of noncommuting operators," Phys. Rev. D **2**, 2161–2186 (1970).

[52] J. T. Foley and E. Wolf, "Radiometry as a short-wavelength limit of statistical wave theory with globally incoherent sources," Opt. Commun. **55**, 236–241 (1985).

[53] Z. Zhang and M. Levoy, "Wigner distributions and how they relate to the light field," in *Proc. of ICCP* (2009).

[54] J. G. Kirkwood, "Quantum statistics of almost classical assemblies," Phys. Rev. **44**, 31–37 (1933).

[55] A. Rihaczek, "Signal energy distribution in time and frequency," IEEE Trans. Inf. Theory **14**, 369–374 (1968).

[56] ZEMAX Development Corporation, Bellevue, Wash., *Optical Design Program User's Guide* (2006).

[57] M. A. Alonso, "Radiometry and wide-angle wave fields. I. Coherent fields in two dimensions," J. Opt. Soc. Am. A **18**, 902–909 (2001).

[58] R. G. Littlejohn and R. Winston, "Corrections to classical radiometry," J. Opt. Soc. Am. A **10**, 2024–2037 (1993).

[59] J. R. Shewell and E. Wolf, "Inverse diffraction and a new reciprocity theorem," J. Opt. Soc. Am. **58**, 1596–1603 (1968).

[60] A. M. Zysk, P. S. Carney, and J. C. Schotland, "Eikonal method for calculation of coherence functions," Phys. Rev. Lett. **95**, 043904 (2005).

[61] R. W. Schoonover, A. M. Zysk, P. S. Carney, J. C. Schotland, and E. Wolf, "Geometrical optics limit of stochastic electromagnetic fields," Phys. Rev. A **77**, 043831 (2008).

[62] P. Royston, "Remark AS R94: a remark on algorithm AS 181: the $W$-test for normality," Appl. Statist. **44**, 547–551 (1995).

[63] A. Demir, "Computing timing jitter from phase noise spectra for oscillators and phase-locked loops with white and $1/f$ noise," IEEE Trans. Circuits Syst. **53**, 1869–1884 (2006).

[64] H. Hashemi, X. Guan, A. Komijani, and A. Hajimiri, "A 24-GHz SiGe phased-array receiver – LO phase-shifting approach," IEEE Trans. Microw. Theory **53**, 614–626 (2005).

[65] H. L. Van Trees, *Detection, Estimation, and Modulation Theory: Part II. Nonlinear Modulation Theory* (John Wiley & Sons, 2003).

[66] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," Trans. ASME – J. Basic Eng. **83**, 95–108 (1961).

[67] J. C. Candy and G. C. Temes, *Oversampling Delta-Sigma Data Converters: Theory, Design, and Simulation* (Wiley-IEEE Press, 1991).

[68] R. G. Gallager, *Discrete Stochastic Processes* (Kluwer Academic Publishers, 1996).

[69] L. Gammaitoni, "Stochastic resonance and the dithering effect in threshold physical systems," Phys. Rev. E **52**, 4691–4698 (1995).

[70] P. Debevec, "Virtual cinematography: relighting through computation," IEEE Computer **39**, 57–65 (2006).

[71] J. Udesen, F. Gran, K. L. Hansen, J. A. Jensen, C. Thomsen, and M. B. Nielsen, "High frame-rate blood vector velocity imaging using plane waves: simulations and preliminary experiments," IEEE Trans. Ultrason. Ferr. **55**, 1729–1743 (2008).

[72] C. Oestges and B. Clerckx, *MIMO Wireless Communications* (Elsevier, 2007).

[73] I. Bekkerman and J. Tabrikian, "Target detection and localization using MIMO radars and sonars," IEEE Trans. Signal Process. **54**, 3873–3883 (2006).

[74] D. W. Bliss and K. W. Forsythe, "Multiple-input multiple-output (MIMO) radar and imaging: degrees of freedom and resolution," in *Proc. of ACSSC* (2003), pp. 54–59.

[75] G. San Antonio, D. R. Fuhrmann, and F. C. Robey, "MIMO radar ambiguity functions," in *Proc. of ACSSC* (2006), pp. 36–40.

[76] J. A. Jensen, "A model for the propagation and scattering of ultrasound in tissue," J. Acoust. Soc. Am. **89**, 182–190 (1991).

[77] L. Ljung, *System Identification: Theory for the User*, 2nd ed. (Prentice-Hall, 1999).

[78] R. E. Apfel and C. K. Holland, "Gauging the likelihood of cavitation from short-pulse, low-duty cycle diagnostic ultrasound," Ultrasound Med. Biol. **17**, 179–185 (1991).

[79] T. Misaridis and J. A. Jensen, "Use of modulated excitation signals in medical ultrasound. Part I: basic concepts and expected benefits," IEEE Trans. Ultrason. Ferr. **52**, 177–191 (2005).

[80] H. L. Van Trees, *Detection, Estimation, and Modulation Theory: Part I* (John Wiley & Sons, 2001).

[81] T. Misaridis and J. A. Jensen, "Use of modulated excitation signals in medical ultrasound. Part II: design and performance for medical imaging applications," IEEE Trans. Ultrason. Ferr. **52**, 192–207 (2005).

[82] R. Y. Chiao and X. Hao, "Coded excitation for diagnostic ultrasound: a system developer's perspective," IEEE Trans. Ultrason. Ferr. **52**, 160–170 (2005).

[83] A. Pezeshki, A. R. Calderbank, W. Moran, and S. D. Howard, "Doppler resilient Golay complementary waveforms," IEEE Trans. Inf. Theory **54**, 4254–4266 (2008).

[84] R. Y. Chiao and L. J. Thomas, "Synthetic transmit aperture imaging using orthogonal Golay coded excitation," in *Proc. of ULTSYM* (2000), pp. 1677–1680.

[85] T. Misaridis and J. A. Jensen, "Use of modulated excitation signals in medical ultrasound. Part III: high frame rate imaging," IEEE Trans. Ultrason. Ferr. **52**, 208–219 (2005).

[86] J. N. Wilson, P. Gader, W. Lee, H. Frigui, and K. C. Ho, "A large-scale systematic evaluation of algorithms using ground-penetrating radar for landmine detection and discrimination," IEEE Trans. Geosci. Remote **45**, 2560–2572 (2007).

[87] R. Gross, I. Matthews, and S. Baker, "Appearance-based face recognition and light-fields," IEEE Trans. Pattern Anal. **26**, 449–465 (2004).

[88] R. Mudumbai, J. Hespanha, U. Madhow, and G. Barriac, "Distributed transmit beamforming using feedback control," IEEE Trans. Inf. Theory (in review).

[89] J. W. Goodman, *Statistical Optics* (John Wiley & Sons, 1985).

[90] H. A. Haus, *Waves and Fields in Optoelectronics* (Prentice-Hall, 1984).

[91] E. Wolf, *Introduction to the Theory of Coherence and Polarization of Light* (Cambridge Univ. Press, 2007).

[92] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge Univ. Press, 1995).

[93] D. J. Griffiths, *Introduction to Quantum Mechanics*, 2nd ed. (Pearson Education, 2005).

[94] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging* (IEEE Press, 1988).

[95] R. Ng, "Fourier slice photography," in *Proc. of SIGGRAPH* (2005), pp. 735–744.

[96] J. Bertrand and P. Bertrand, "A tomographic approach to Wigner's function," Found. Phys. **17**, 397–405 (1987).

[97] M. Petkovšek, H. S. Wilf, and D. Zeilberger, *A=B* (A K Peters, 1996).

[98] J. A. C. Veerman, J. J. Rusch, and H. P. Urbach, "Calculation of the Rayleigh-Sommerfeld diffraction integral by exact integration of the fast oscillating factor," J. Opt. Soc. Am. A **22**, 636–646 (2005).