NON-LINE-OF-SIGHT PHASOR FIELD CAMERA

by

Xiaochun Liu

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Electrical and Computer Engineering)

at the

UNIVERSITY OF WISCONSIN–MADISON

2022

Date of final oral examination: 04/29/2022

The dissertation is approved by the following members of the Final Oral Committee:
Andreas Velten, Assistant Professor, Electrical and Computer Engineering
William A. Sethares, Professor, Electrical and Computer Engineering
Chu Ma, Assistant Professor, Electrical and Computer Engineering
Yin Li, Assistant Professor, Computer Sciences

"If you are working on something exciting that you really care about, you don't have to be pushed. The vision pulls you."

- Steve Jobs

ACKNOWLEDGMENTS

This thesis would not have been possible without the support of my advisor, collaborators, friends, and my family members. First of all, I cannot thank my advisor Andreas Velten enough. Without his guidance, support, and encouragement, this thesis would not exist. One thing I most appreciate is how he is supportive at all times. I still remember the excitement when Andreas introduced the NIAC corner camera project on the first day I met him. I am sincerely grateful for all the freedom and support Andreas provides. I feel tremendously lucky and could not have asked for better advisors.

I also feel lucky to work with great colleagues in the Computational Optics Group at the University of Wisconsin-Madison. Thanks to Toan Huu Le for being my office mate who is sitting back with me and spending time discussing both scientific and non-sense topics. Thank you, Ji Hyun, for being a collaborative and supportive lab mate. Thank you, Sebastian Bauer, for all suggestions through paper collaborations that help me become a better listener. Thank Fiona Kine, Marco La Manna, Syed Azer Reza, and Atul Ingle for all the help during my very first year in the lab. Somehow, after Talha Sultan, Trevor Seets, Felipe Gutierrez Barragan, Yizhou Lu, Ehsan Ahmadi, Alex Bocchieri and Khadijeh Masumnia-Bisheh join the group, I am the "senior" PhD student. Thank you all for the support and memorable time we spent together.

It is my pleasure to work with collaborators at both multi-disciplinary and worldwide levels. I would like to thank Ibon Guillen, Miguel Galindo, Miguel Angel Cosculluela, Julio Marco, Adrian Jarabo, and Prof. Diego Gutierrez at Universidad de Zaragoza, I3A for fruitful discussions and multiple research projects about Non-Line-of-Sight imaging. Many thanks to Deyang Jiang and Xin Lou at the ShanghaiTech University for their collaborations and works regarding Non-Line-of-Sight imaging on a limited resource platform. I also would like to thank Simone Riccardo, Marco Renna, and Alberto Tosi at the Politecnico di Milano for their professional knowledge about Single-photon avalanche diode. I am grateful to work with Eric Brandt and Eftychios Sifakis at the Computer Science Department

at the University of Wisconsin-Madison for their in-depth knowledge and work in a low latency Non-Line-of-Sight imaging pipeline.

Furthermore, I feel grateful to Ming Zhou and Prof. Zongfu Yu, who helped and guided me as a fresh PhD student during my first several months. A big shout out to the Double Mint Hot Pot for being a great host to help me reduce my stress and anxiety level. There is lots of fun and joy for me with my friends and wife there in Madison.

Finally, I would like to thank my parents, Yan, Jinhong, and my wife, Renwen. They always support me and encourage me to pursue what I dream of. Dad and Mom's regular reminders of my mental and physical health and Wife's delicious cooking and the time we spend together have always been my support. This thesis is dedicated to them.

This thesis and my PhD study are supported by DARPA through the DARPA REVEAL project (HR0011-16-C0025), the NASA Innovative Advanced Concepts (NIAC) Program (NNX15AQ29G), the Air Force Office of Scientific Research (AFOSR) Young Investigator Program (FA9550-15-1-0208), the Office of Naval Research (ONR, N00014-15-1-2652), the DURIP program (FA9550-18-1-0409), the National Science Foundation grants NSF (IIS-2008584, CCF-1812944, IIS-1763638, IIS-2106768) and a grant from UW–Madison's Draper Technology Innovation Fund. I am truly grateful to have this support which enables me to focus on my research.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Non-line-of-sight (NLOS) imaging recovers objects from diffusely reflected indirect light using transient illumination in combination with a computational inverse method. Despite its many potential applications, existing methods lack practical usability due to several shared limitations, including the assumption of single scattering only, lack of occlusions, and Lambertian reflectance. Line-of-sight (LOS) imaging systems, on the other hand, can address these and other imaging challenges despite relying on the mathematically simple processes of linear diffractive wave propagation.

The goal of this dissertation is to design computational imaging solvers for NLOS imaging using wave diffraction theory. To achieve this goal, we show that the NLOS imaging problem can also be formulated as a diffractive wave propagation problem. This allows to image NLOS scenes from raw time-of-flight data by applying the mathematical operators that model wave propagation inside a conventional line-of-sight imaging system. By doing this, we have developed a method that yields a new class of imaging algorithms mimicking the various capabilities of LOS cameras. This new class of NLOS imaging algorithms relies on solving a wave diffraction integral, namely the Rayleigh-Sommerfeld Diffraction (RSD) integral. We have developed a Fast Fourier Transform based reconstruction algorithm based on RSD, for the first time, NLOS imaging of complex scenes with strong multiple scattering and ambient light, arbitrary materials, large depth range, and occlusions. Our new RSD based method presented in this thesis enables the reconstruction of room-sized scenes from non-confocal, parallel multi-pixel captured measurements in seconds with less memory usage.

1 INTRODUCTION

Time resolved Non-Line-of-Sight (NLOS) imaging uses measured signals from an occluded scene with computational algorithms to see around a corner. A conceptual image is shown in Fig. 1.1. This figure shows a conceptual understanding of Non-Line-of-Sight imaging. People perform measurements indirectly from a scene. The goal is to use collected measurements to form an image of the hidden scene. As a growing topic in the Computational Imaging area, there have been rich developments for both hardware and software algorithms. After a conceptual exploration of time resolved NLOS imaging [5, 6], Velten *et al.* propose and demonstrate NLOS imaging under a lab condition. On the hardware side, Velten *et al.* use a femtosecond laser with a streak camera to capture signals, then use a filtered backprojection (FBP) algorithm to perform reconstructions. This FBP method is similar to the solution of a computed tomography problem [7, 8]. Instead of using a costly streak camera, Buttafava *et al.* [9] shows a gated Single-Photon Avalanche Diode (SPAD) can be used to capture fast NLOS signals with a lower cost than a streak camera. Up until now, a picosecond laser and SPAD hardware setup is commonly used in time resolved NLOS imaging [4, 2, 3, 10, 11, 12]. Based on acquisition schemes, a measured time resolved signal can be divided into two branches which is confocal [4] or non-confocal [7]. The confocal acquisition requires SPAD and laser spots on a relay wall to be co-located as opposed to the non-confocal. This non-confocal acquisition shows a great advantage in terms of real-time NLOS imaging applications regarding signal to noise ratio [12]. Galindo *et al.* provide a simulated dataset library about different acquisition schemes by using computer graphic rendering technics [13]. A conceptual understanding of Non-Line-of-Sight imaging is shown in Fig. 1.1.

NLOS imaging can be studied and solved by a wave propagation model, which is called a phasor field NLOS imaging model [3, 14]. With this phasor field model, one can study time resolved NLOS imaging problem from an optical diffraction perspective. For example, this idea of diffraction can be used to perform hidden scene reconstruction from measurements [1].

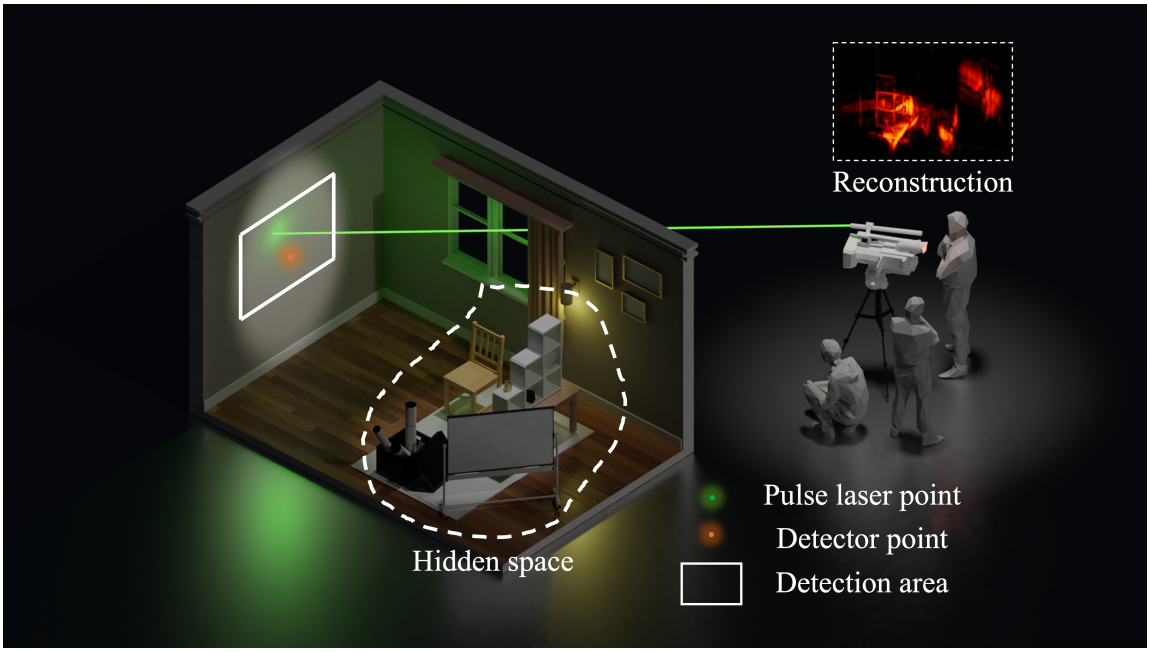

Figure 1.1: This figure shows a conceptual understanding of Non-Line-of-Sight imaging. People perform measurements indirectly from a scene. The goal is to use collected measurements to form an image of the hidden space. For example, as shown in the figure, people can only shoot light using a pulse laser through this window shown in Figure. Light will bounce back and forth between the wall and the hidden space. These bouncing signals can be measured via a single photon detection device which is shown at the detector point. Then smart algorithms are used to form a three-dimensional image using captured signals which is shown as a reconstruction in this figure.

1.1 Related work

In this section, we summarize developments in NLOS imaging additional to the works mentioned above. Based on the signals used in measurements, we categorize works into two categories: picosecond time resolved NLOS imaging and others without a use of temporal resolution.

Here is a list of developments for the former category. There are several insightful developments for the NLOS phasor field model as well as experimental

demonstration [15, 16, 17, 1, 18, 19]. These results can be used to study NLOS image reconstructions, and new hardware designs. A geometrical based solution is used to solve NLOS imaging problem [20, 10]. A joint albedo-normal efficient method is applied to recover NLOS geometry [21]. Algebraic based iterative solutions [22, 23, 24] which is used in the Computed tomography achieve a similar reconstruction quality as the FBP method [7, 8]. A faster GPU implementation of backprojection solver [25] can be used to decrease the computational cost for iterative solutions. An modified Laplacian of Gaussian (LOG) filter [26] is presented and can be added to the FBP based solution. A modified FBP solution based on a two weighting factors has shown robustness to noisy scenarios [27]. Making use of occlusions can improve quality of reconstructions [28, 29, 30, 31]. Hardware time delay can be used to co-design temporal focusing with reconstructions [32]. Circular, confocal NLOS imaging acquired data looks similar to a sinogram in tomography which is different from the confocal and the non-confocal acquisition method mentioned above [33]. To recover a room, one can estimate it by fitting planes from temporal measurements [34]. Tracking people around the corner can use a less computational resources than to recover an image around the corner [35]. Bayesian statistics reconstruction account for random errors in measurements. Benchmarks and simulated dataset are provided [36, 13]. Color NLOS imaging uses signal-pixel SPAD sensors [37, 38]. Gated array SPAD sensor as well as a dynamic scene reconstruction have been demonstrated [39, 40]. Deep learning based method can be used for a human pose estimation [41] and reconstructions [42]. Unlike measurements are scanned on a large relay wall, a small amount of measurements collected from a pinhole on the door can be used to perform a hidden target tracking [43].

There are some works focusing on the analysis to this problem: feature visibility (missing cone problem) [44], a justification for the FBP method [11], using the Wigner Distribution model to explain NLOS imaging [45], a study related to hidden object's BRDF behavior in reconstructions [46], a calibration procedure for the current NLOS imaging setups [47], a discussion related to some limitations of current NLOS imaging methods [48]. There are some recent review papers which

summarize developments in NLOS imaging [49, 50, 51].

Methods used with a nanosecond temporal resolution, or without any temporal resolution, or non-optical signals are listed below. Nanosecond NLOS imaging scene reconstructions with a modified hardware leads to a cheap design [52]. Some passive approaches such as optical speckle correlations method [53], optical speckle tracking [54], image based tracking [55], localization [56] can be applied to an active NLOS imaging environment. Superheterodyne synethetic wave method is similar to the NLOS phasor field method which is interesting for the passive NLOS imaging scenario [57, 58, 59, 60]. Computational periscopy uses a shadow from a nature light to perform hidden scene reconstructions [61, 62, 63, 64]. Some works explore NLOS imaging using other physical signals, such as acoustic [65], mid-infrared speckle [66] and long-wave infrared [67] signals, optical polarization [68], doppler radar signals [69].

1.2 Contributions

The goal of this thesis is to design of computational imaging solvers for NLOS imaging using wave diffraction theory. In order to achieve this goal, we make the following contributions in this thesis:

- **Non-Line-of-Sight Phasor Field Diffraction Model** [3, 1](Chapter 2): In this work, despite its many potential applications in Non-line-of-sight imaging (NLOS), existing methods lack practical usability due to several shared limitations, including the assumption of single scattering only, lack of occlusions, and Lambertian reflectance. Line-of-sight (LOS) imaging systems, on the other hand, can address these and other imaging challenges despite relying on the mathematically simple processes of linear diffractive wave propagation. In this work we show that the NLOS imaging problem can also be formulated as a diffractive wave propagation problem. This allows to image NLOS scenes from raw time-of-flight data by applying the mathematical operators that model wave propagation inside a conventional line-of-sight imaging

system. By doing this, we have developed a method that yields a new class of imaging algorithms mimicking the various capabilities of LOS cameras. To demonstrate our method, we derive three imaging algorithms, each with its own unique novel capabilities, modeled after three different LOS imaging systems. These algorithms rely on solving a wave diffraction integral, namely the Rayleigh-Sommerfeld Diffraction (RSD) integral. Fast solutions to RSD and its approximations are readily available, directly benefiting our method. We demonstrate, for the first time, NLOS imaging of complex scenes with strong multiple scattering and ambient light, arbitrary materials, large depth range, and occlusions. Our method handles these challenging cases without explicitly developing a light transport model. We believe that our approach will help unlock the potential of NLOS imaging, and the development of novel applications not restricted to laboratory conditions, as shown in our results.

- **Fast Phasor Field Diffraction Algorithm:** [1](Chapter 3): We introduce an NLOS reconstruction method using the phasor field formalism along with a convolutional Fast Fourier Transform (FFT) based Rayleigh Sommerfeld Diffraction (RSD) algorithm to provide fast non-approximative scene reconstructions for general capture setups, in particular including non-confocal setups using a single laser and a sensor array. Our hardware prototype includes a single-photon avalanche diode detector and a picosecond pulse laser which will be mentioned specifically later. When used in the confocal scenario, this new method performs at speed similar to LCT and FK Migration, while requiring significantly less memory. In addition to applying our new algorithm to open source data [3, 2], we also perform several additional experiments.

- **Wigner Distribution Model for Phasor Field Diffraction:** [45](Chapter 4): In phasor field NLOS imaging, the Rayleigh-Sommerfeld Diffraction (RSD) model is shown to be a key solution to Non-Line-of-Sight imaging problems [3]. All existing applications of the Wigner Distribution Function are used when Fresnel approximation is valid. However, in NLOS imaging ap-

plication, the diffraction happens close to the relay wall where only the RSD holds as an exact solution. This RSD also gives an exact solution to the wave propagation as opposed to the approximations such as Fresnel or Fraunhofer diffraction which are commonly known in classical optics [70, 71, 72]. It is shown that the RSD can be used to solve scanning free, real-time, three-dimensional NLOS reconstruction problem [1]. Dove *et al.* [15] present a two dimensional spatial Wigner Distribution Function in a paraxial region with the approximated Fresnel diffraction for NLOS phasor field model.

The RSD with the Wigner Distribution Function has never been discussed in the context of real-world NLOS measurements. In this paper, we will study the RSD in the Wigner Distribution Function domain and compare it with the Fresnel diffraction under real-world parameters like finite relay wall size, discrete spatial sampling, and different acquisition schemes such as confocal and non-confocal measurements. Another angle to describe our work is to use the Wigner Distribution Function to explain Non-Line-of-Sight imaging and clarify when approximations are useful and meaningful in practice.

The key contributions of this chapter are listed below: 1. We study Rayleigh-Sommerfeld Diffraction in the Wigner Distribution domain and show that the exact Rayleigh-Sommerfeld Diffraction solution does not have any geometrical interpretations as opposed to the Fresnel approximation which has a shear mapping interpretation in the Wigner Distribution Function domain. 2. We derive a lateral resolution limit from the exact Rayleigh-Sommerfeld Diffraction solution for NLOS reconstructions. 3. We provide an understanding of spatial sampling for phasor field wavefronts on a relay wall. 4. We characterize errors from the Fresnel diffraction. We show that this error is less in the confocal acquisition, which makes it an applicable candidate for reconstruction algorithms. In addition, errors from the confocal acquisition and the non-confocal acquisition are visualized in the Wigner Distribution Function domain.

- **Feature Visibility Analysis:** [44](Chapter 5): In this work, we aim to provide

a generic description for direct bounce (3rd bounce) NLOS measurements and show how much information they encode and how this affects practical NLOS imaging problems. Consider the NLOS reconstruction shown in Figure 1.2. It contains three very similar patches that only vary slightly in orientation and have different positions in the reconstruction space. Yet while two of the patches are reconstructed clearly and accurately, the third is completely missing from the reconstruction. A closer inspection of prior published results reveals that similar artifacts are seen in reconstructions using a variety of diverse reconstruction methods. Surfaces with certain normal vectors are missing in the reconstruction or scenes with simple surfaces are chosen to avoid the problem. The main purpose of this work is to explain this phenomenon. As we show in this chapter, any NLOS measurement can be expressed as an integral operator known in the literature as elliptical Radon integral. We analyze this measurement function in the Fourier domain and show that a significant part of the measurement space is not accessed by the NLOS measurement and thus represents a null space for NLOS reconstruction. Finally, we investigate the Fourier domain representations of common scenes and scene features to identify features that fall into the null space and cannot be reconstructed. Because our analysis involves a generic description of the NLOS measurement, it is independent of the reconstruction algorithm used. We expect our findings will inform inverse solution design and future NLOS reconstruction methods.

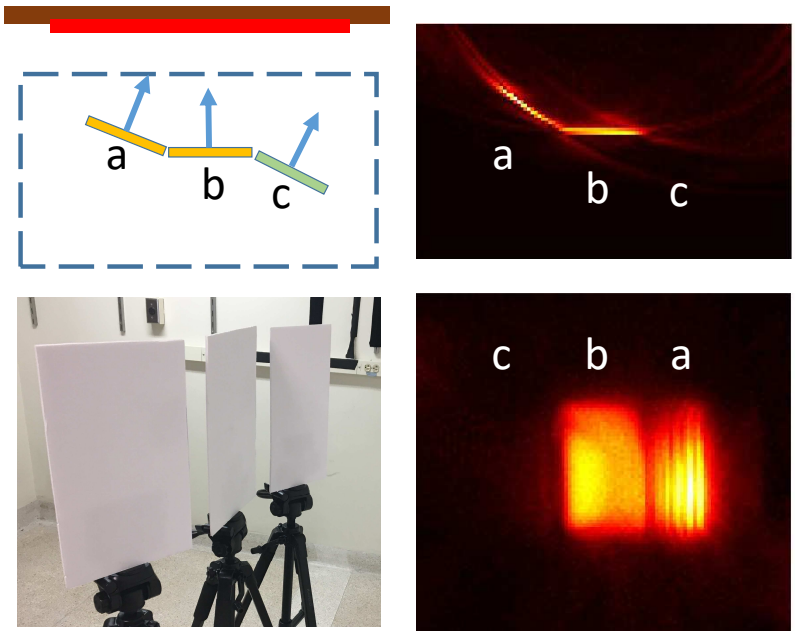

Figure 1.2: **Three patches rotation example.** Experimental reconstruction of a scene containing three patches denoted by **a, b, c**. Patches **a** and **c** are parallel but the latter one's surface normal vector does not point towards the NLOS relay wall and the patch does not appear in the reconstruction shown on the right. The first row stands for the top view and the second row for the front view. The explanation for this astonishing effect will be developed in the main text, see Sec. 5.5, and also Fig. 5.8 for a graphical explanation.

Liu *et al.*[3, 1] and Reza *et al.*[14] introduce a virtual wave phasor field formalism that is the basis of this chapter. Using the phasor field method, the NLOS imaging problem can be stated as a line of sight optical imaging problem based on diffraction and solved using existing diffraction theory methods.

2.1 A formal definition

Time-of-flight LOS imaging has used a phasor formalism together with Fourier domain ranging[73] to describe the emitted modulated light signal. Kadambi et al.[74] extended this concept to reconstruct NLOS scenes using phasors to describe hardware intensity modulation. We show that a similar description can be used to model the physics of light transport through the scene. The key insight of our method is that propagation through a scene of intensity-modulated light can be modeled using a Rayleigh-Sommerfeld diffraction (RSD) operator acting on a quantity we term the *phasor field*. This allows us to formulate any NLOS imaging problem as a wave imaging problem (Figure 1), and to transfer well-established insights and techniques from classic optics into the NLOS domain. Given a captured time-resolved dataset of light transport through a NLOS scene, and a choice of a *template* LOS imaging system, our method provides a recipe that results in a NLOS imaging algorithm mimicking the capabilities of the corresponding LOS system. This template system can be any real or hypothetical wave imaging system that includes a set of light sources and detectors. The resulting algorithms can then be efficiently solved using diffraction integrals like the RSD. For now, we focus on the phasor field framework and computational RSD solver will be introduced in Ch. 3.

We start by mathematically defining our phasor field $\mathcal{P}(\vec{x}, t)$. Let $\mathcal{E}(\vec{x}, t)$ $[\sqrt{\mathrm{Wm}^{-2}}]$ be a quasi-monochromatic scalar field at position $\vec{x} \in \mathcal{S}$ and time t, incident on (or reflected from) a Lambertian surface $\mathcal{S}$, with center frequency Ω_0 and bandwidth

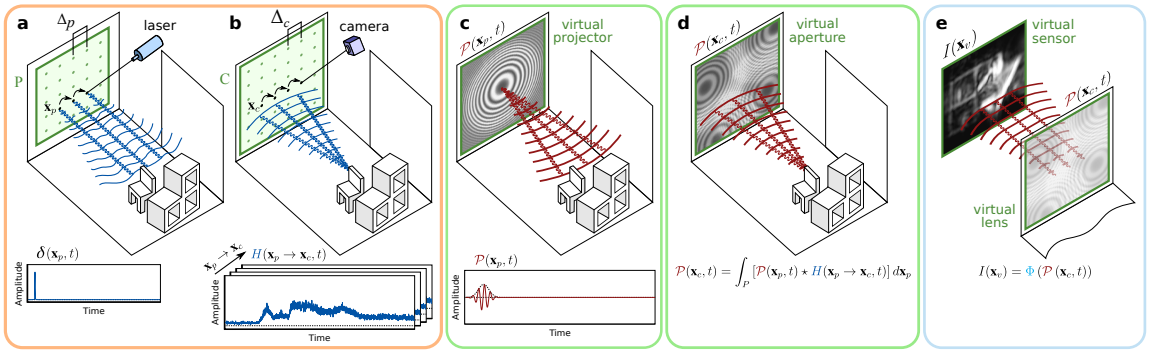

Figure 2.1: **NLOS as a virtual LOS imaging system.** Capturing scene data: **a**, A pulsed laser sequentially scans a relay wall; **b**, the light reflected back from the scene onto the wall is recorded at the sensor yielding an impulse response H of the scene. **c**, Virtual light source: The phasor field wave of a virtual light source $\mathcal{P}(x_p, t)$ is modeled after the wavefront of the light source of the template LOS system. **d**, The scene response to this virtual illumination $\mathcal{P}(x_c, t)$ is computed using H. **e**, The scene is reconstructed from the wavefront $\mathcal{P}(x_c, t)$ using wave diffraction theory. The function $\Phi(\cdot)$ is also taken from the template LOS system.

$\Delta\Omega \ll \Omega_0$. We can then define

$$\mathcal{P}(\vec{x}, t) \equiv \left\langle \frac{1}{\tau} \int_{t-\tau/2}^{t+\tau/2} \left|\mathcal{E}(\vec{x}, t')\right|^2 \, dt' \right\rangle - \left\langle \frac{1}{T} \int_{t-T/2}^{t+T/2} \left|\mathcal{E}(\vec{x}, t')\right|^2 \, dt' \right\rangle \tag{2.1}$$

as the mean subtracted irradiance $[\mathrm{Wm}^{-2}]$ at point $\vec{x}$ and time t. The $\langle\cdot\rangle$ operator denotes spatial speckle averaging (for the reflected case) accounting for laser illumination, and τ represents the averaging of the intensity at a fast detector, with $\tau \ll 1/\Delta\Omega \ll T$. The second integral in the equation above is a long-term average intensity over an interval $T \gg \tau$ of the signal as seen by a conventional non-transient photodetector.

Now, let us define the Fourier component of $\mathcal{P}(\vec{x}, t)$ for frequency ω as

$$\mathcal{P}_{0,\omega}(\vec{x}) \equiv \int_{-\infty}^{+\infty} \mathcal{P}(\vec{x}, t)\, e^{-i\omega t} dt, \tag{2.2}$$

from which we can define a monochromatic component of the phasor field $\mathcal{P}_\omega(\vec{x}, t)$ as

$$\mathcal{P}_\omega(\vec{x}, t) \equiv \mathcal{P}_{0,\omega}(\vec{x})\, e^{i\omega t}. \tag{2.3}$$

Using the above, our phasor field $\mathcal{P}(\vec{x}, t)$ can be expressed as a superposition of monochromatic plane waves as $\mathcal{P}(\vec{x}, t) = \int_{-\infty}^{+\infty} \mathcal{P}_\omega(\vec{x}, t)\, d\omega/2\pi$. Since $\mathcal{P}(\vec{x}, t)$ is a real quantity, the Fourier components $\mathcal{P}_{0,\omega}(\vec{x})$ are complex and symmetric about $\omega = 0$. Note that in many places of this manuscript we assign $\mathcal{P}(\vec{x}, t)$ an explicitly complex value; in these cases it is implied that the correct real representation is $\frac{1}{2}(\mathcal{P}(\vec{x}, t) + \mathcal{P}^*(\vec{x}, t))$. In practice the complex conjugate can be safely ignored in our calculations. As can be seen in Section B in the supplemental, given an isotropic source plane S and a destination plane D, and assuming that the electric field at S is incoherent, the propagation of its monochromatic component $\mathcal{P}_\omega(\vec{x}, t)$ is defined by an RSD-like propagation integral:

$$\mathcal{P}_\omega(\vec{x}_d, t) = \gamma \int_S \mathcal{P}_\omega(\vec{x}_s, t) \frac{e^{ik|\vec{x}_d - \vec{x}_s|}}{|\vec{x}_d - \vec{x}_s|} d\vec{x}_s, \tag{2.4}$$

where γ is an attenuation factor, and $k = 2\pi/\lambda$ is the wave number for wavelength $\lambda = 2\pi/\omega$, $\vec{x}_s \in S$ and $\vec{x}_d \in D$. Note that, as described in Section B in the supplemental, we approximate γ as a constant over the plane S as $\gamma \approx 1/|\langle S \rangle - \vec{x}_d|$; this approximation has a minor effect on the signal amplitude at the sensor, but does not change the phase of our phasor field. While Equation 2.4 is defined for monochromatic signals, it can be used to propagate broadband signals by propagating each monochromatic component independently; this can be efficiently done by time-shifting the phasor field (more details are provided in Section B.1 of the supplemental).

The key insight of Equation 2.4 is that, given the assumption of constant γ, the propagation of our phasor field is defined by the same RSD operator as any other physical wave. Therefore, in order to image a scene from a *virtual camera* with aperture at plane C, we can apply the image formation model of any wave-based LOS imaging system directly over the phasor field $\mathcal{P}(\vec{x}_c, t)$ at the aperture, with

$\vec{x}_c \in C$. The challenge is how to compute $\mathcal{P}(\vec{x}_c, t)$ from an illuminating input phasor field $\mathcal{P}(\vec{x}_p, t)$, where $\vec{x}_p$ is a point in the *virtual projector* aperture P, given a particular NLOS scene (see Fig. 2.1).

Since light transport is linear in space and time-invariant[75, 76], we can characterize light transport through the scene as an impulse response function $H(\vec{x}_p \rightarrow \vec{x}_c, t)$, where $\vec{x}_p$ and $\vec{x}_c$ are the positions of the emitter and detector, respectively. The phasor field at the virtual aperture $\mathcal{P}(\vec{x}_c, t)$ can thus be expressed as a function of the input phasor field $\mathcal{P}(\vec{x}_p, t)$ and $H(x_p \rightarrow \vec{x}_c, t)$ as

$$\mathcal{P}(\vec{x}_c, t) = \int_P [\mathcal{P}(\vec{x}_p, t) \star H(\vec{x}_p \rightarrow \vec{x}_c, t)] \, d\vec{x}_p, \tag{2.5}$$

where $\star$ denotes the convolution operator. Any imaging system can be characterized by its image formation function $\Phi(\cdot)$, which transduces the incoming field into an image

$$I(\vec{x}_v) = \Phi\left(\mathcal{P}\left(\vec{x}_c, t\right)\right), \tag{2.6}$$

where $\vec{x}_v$ is the point being imaged (i.e., the point at the *virtual sensor*). This in turn can be formulated as an RSD propagator, requiring to solve a diffraction integral in order to generate the final image.

In an NLOS scenario, $H(\vec{x}_p \rightarrow \vec{x}_c, t)$ usually corresponds to 5D transients acquired via an ultrafast sensor focused on $\vec{x}_c$, and sequentially illuminating the relay wall with short pulses at different points $\vec{x}_p$ (see Fig. 2.1). Points $\vec{x}_p$ and $\vec{x}_c$ correspond to a virtual LOS imaging system projected onto the relay wall. Once $H(\vec{x}_p \rightarrow \vec{x}_c, t)$ has been captured, both the wavefront $\mathcal{P}(\vec{x}_p, t)$ and the imaging operator $\Phi(\cdot)$ can be implemented computationally, so they are not bounded by hardware limitations. We can leverage this to employ different $\mathcal{P}(\vec{x}_p, t)$ functions from any existing LOS imaging system[71] to emulate its characteristics in an NLOS setting.

Overall speaking, phasor field NLOS framework enables us to design and implement a virtual camera to satisfy different NLOS imaging tasks. Fig. 2.2 shows the result on a complex scene with a virtual confocal camera. The scene contains

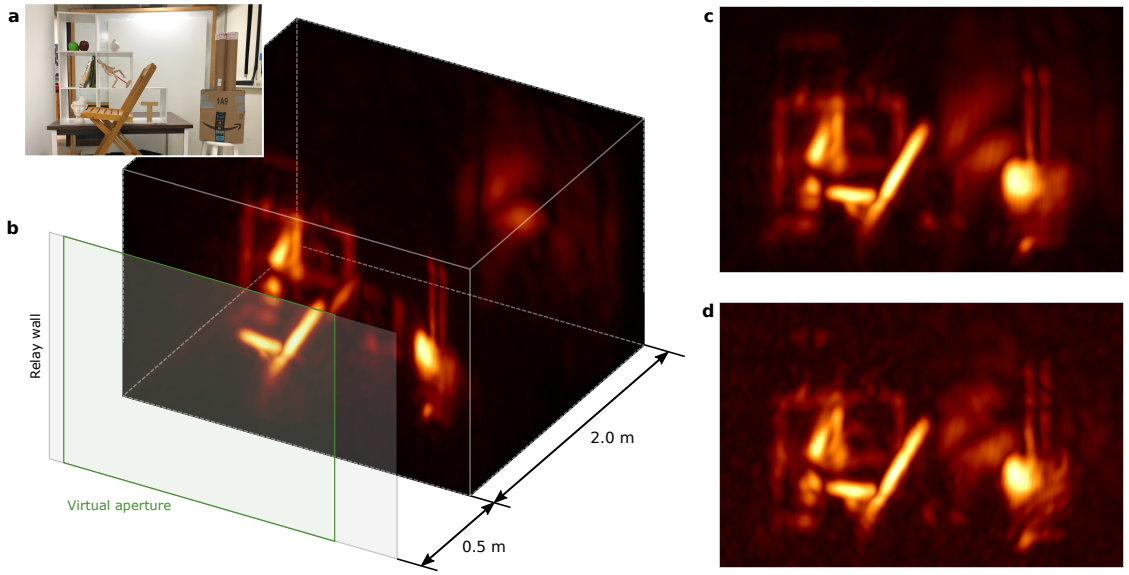

Figure 2.2: **Reconstructions of a complex NLOS scene. a**, Photograph of the scene as seen from the relay wall. The scene contains occluding geometries, multiple anisotropic surface reflectances, large depth, and strong ambient and multiply scattered light. **b**, 3D visualization of the reconstruction using phasor fields ($\lambda = 6$ cm). We include the relay wall location and the coverage of the virtual aperture for illustration purposes. **c**, Frontal view of the scene, captured with an exposure time of 10 ms per laser position. **d**, Frontal view captured with just a 1 ms exposure time (24 seconds for the complete scan).

multiple objects with occlusions distributed over a large depth, a wide range of surface reflectances and albedos, and strong interreflections. Despite this challenging scenario, phasor field method is able to image many details of the scene, at the correct depths, even with an ultra-short, 1 ms exposure. More analysis on the robustness of our method to capture noise can be found in Liu *et al.*[3].

2.2 A signal processing framework

In this section, we apply phasor field framework introduced in Ch. 2.1 to NLOS imaging reconstruction problem and describe it within a signal processing frame-

work. The concept of phasor field NLOS imaging is shown in Fig. 2.3. Data from the scene is collected by illuminating a set of points $\vec{x}_p$ on a relay surface P and collecting the light returned at points $\vec{x}_c$ on a relay surface C. This data set represents impulse responses $H(\vec{x}_p \to \vec{x}_c, t)$ of the scene. Using such an impulse response we can compute the scene response at points $\vec{x}_c$ to an input signal $\mathcal{P}(\vec{x}_p, t)$ as

$$\mathcal{P}(\vec{x}_c, t) = \int_P [\mathcal{P}(\vec{x}_p, t) \underset{t}{*} H(\vec{x}_p \to \vec{x}_c, t)] d\vec{x}_p \tag{2.7}$$

where the $\underset{t}{*}$ operator indicates a convolution in time. We call the quantities $\mathcal{P}(\vec{x}_p, t)$ and $\mathcal{P}(\vec{x}_c, t)$ phasor field wavefronts. $\mathcal{P}(\vec{x}_c, t)$ describes the wavefront that would be returned from the scene if it were illuminated by a illumination wave $\mathcal{P}(\vec{x}_p, t)$. Reconstructing an image from the wave front of a reflected wave is the fundamental problem solved by a line of sight imaging system. The reconstruction operation

$$I(\vec{x}_v, t) = \Phi(\mathcal{P}(\vec{x}_c, t)) \tag{2.8}$$

resulting in a 3D image $I(\vec{x}_v)$ of the scene amounts to propagation of the wavefront at C back into the scene into the points $\vec{x}_v$ where it has the shape of the scene objects. The Fourier domain version $\Phi_{\mathcal{F}}(\cdot)$ of the wave propagation operator $\Phi(\cdot)$ is known as the Rayleigh-Sommerfeld Diffraction (RSD) integral:

$$\Phi(\mathcal{P}_{\mathcal{F}}(\vec{x}_c, \Omega)) = \left|\mathcal{R}_{\vec{x}_v}(\mathcal{P}_{\mathcal{F}}(\vec{x}_c, \Omega))\right|^2. \tag{2.9}$$

The RSD in the considered context is calculated by

$$\mathcal{R}_{\vec{x}_v}(\mathcal{P}_{\mathcal{F}}(\vec{x}_c, \Omega)) = \alpha(\vec{x}_v) \int_C \mathcal{P}_{\mathcal{F}}(\vec{x}_c, \Omega) \underbrace{\frac{e^{-ik|\vec{x}_c - \vec{x}_v|}}{|\vec{x}_c - \vec{x}_v|}}_{\text{RSD diffraction kernel}} d\vec{x}_c\,. \tag{2.10}$$

In this equation, $k = \Omega/c$ denotes the wavenumber and c across our paper refers to the speed of light. The conventional RSD propagates the electric field, but in this context propagation of an intensity modulation is required. The phasor field RSD differs from the conventional version by the amplitude correction factor $\alpha(\vec{x}_v)$ [3].

This factor depends on the location $\vec{x}_v$ of the reconstruction point and could be precomputed once the geometry of the relay surface is known. Alternatively, it can be disregarded, as it only causes a slowly varying error in brightness of reconstructed points, but not their location. The RSD in Eq. (2.10) is a function of each individual monochromatic phasor field component. For this reason, the wavefront $\mathcal{P}(\vec{x}_c, t)$ received at the aperture has been replaced by its Fourier domain representation $\mathcal{P}_{\mathcal{F}}(\vec{x}_c, \Omega)$. Throughout this paper, frequency domain quantities are denoted by the same variable as the respective time domain quantities, but with the subscript $\mathcal{F}$ and the argument angular frequency Ω instead of t. For instance, $\mathcal{F}_t\left(\mathcal{P}(\vec{x}_p, t)\right) = \mathcal{P}_{\mathcal{F}}(\vec{x}_p, \Omega)$ and $\mathcal{F}_t\left(H(\vec{x}_p \to \vec{x}_c, t)\right) = H_{\mathcal{F}}(\vec{x}_p \to \vec{x}_c, \Omega)$, where $\mathcal{F}_t(\cdot)$ denotes the Fourier transform with respect to time. Note that in this paper, the RSD propagation direction is from the camera aperture (i.e., relay surface C) into the reconstruction volume.

It is important to note that both illumination $\mathcal{P}(\vec{x}_p, t)$ and image formation $\Phi(\cdot)$ are implemented virtually on a computer. For this reason, they can be chosen to mimic any LOS imaging system. For the purpose of NLOS 3D image reconstruction, one option is to choose a transient camera sending a virtual phasor field pulse

$$\mathcal{P}(\vec{x}_p, t) = e^{i\Omega_C t}\delta(\vec{x}_p - \vec{x}_{ls})e^{-\frac{(t-t_0)^2}{2\sigma^2}} \tag{2.11}$$

from the virtual light source position $\vec{x}_{ls}$ into the scene. The center frequency Ω_C has to be chosen according to the spatial relay wall sampling. The smallest achievable wavelength should be larger than twice the largest distance between neighboring points $\vec{x}_p$ and $\vec{x}_c$ and larger than the temporal resolution of the imaging hardware [3]. For example, given a spatial sampling of 1 cm, the smallest possible modulation wavelength is larger than 2 cm. For the following, we set $t_0 = 0$. The illumination pulse as a function of time needs to be converted into the frequency domain, so that each corresponding frequency is then propagated separately by the RSD in Eq. (2.10). The temporal Fourier transform of the illumination phasor

field yields

$$\mathcal{P}_{\mathcal{F}}(\vec{x}_p, \Omega) = \mathcal{F}_t\left(\mathcal{P}(\vec{x}_p, t)\right) = \delta(\vec{x}_p - \vec{x}_{ls}) \left(2\pi\delta(\Omega - \Omega_C) \underset{f}{*} \sigma\sqrt{2\pi}e^{-\frac{\sigma^2\Omega^2}{2}}\right). \quad (2.12)$$

The result $\mathcal{P}_{\mathcal{F}}(\vec{x}_p, \Omega)$ in the frequency domain is a Gaussian centered around the central frequency Ω_C as it is shown in Fig.1. Figuratively, the RSD propagates the light wave arriving at the aperture (i.e., relay surface C) back into the scene, thereby reconstructing it. Equivalently, one can think of it as a virtual imaging system that forms the image acquired by a virtual sensor behind the relay wall.

After processing all frequency components through space with the RSD, the result at $\vec{x}_v$ needs to be converted to the time domain again by applying the inverse Fourier transform. The overall reconstruction is therefore calculated by

$$I(\vec{x}_v, t) = \left| \int_{-\infty}^{+\infty} e^{i\Omega t} \mathcal{R}_{\vec{x}_v}\Big(\underbrace{\underbrace{\mathcal{P}_{\mathcal{F}}(\vec{x}_p, \Omega)}_{\text{Illumination phasor field}} \cdot H_{\mathcal{F}}(\vec{x}_p \to \vec{x}_c, \Omega)}_{\text{Phasor field at the camera aperture (relay surface C)}} \Big) \frac{d\Omega}{2\pi} \right|^2, \quad (2.13)$$

where the integral over P has vanished as there is only one virtual illumination point $\vec{x}_{ls}$. Calculating the square is omitted in the actual reconstruction implementation, as it only affects the scene contrast.

The RSD as defined in Eq. (2.10) can propagate the wave from an arbitrary surface to any arbitrary point $\vec{x}_v$. Consider a NLOS imaging measurement is captured on a planar relay wall, then we can use two parallel planes setup for the RSD calculation. Then multiple convolutional RSD methods have been introduced in the literature [77, 78, 79]. We introduce the scalar coordinates $\vec{x}_c = (x_c, y_c, 0)$ and $\vec{x}_v = (x_v, y_v, z_v)$ and rewrite the RSD in Eq. (2.10) as follows:

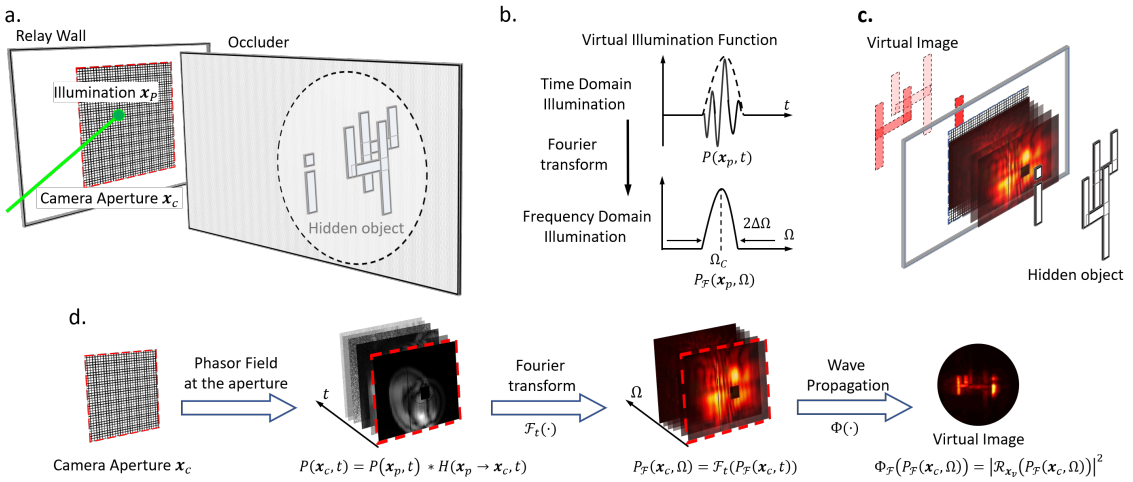

Figure 2.3: Illustration of proposed phasor field fast NLOS imaging method. **a.** refers to the NLOS imaging scenarios, including relay wall, occluder, and hidden object. Measurements are performed on the relay wall, including illumination point $\vec{x}_p$ and camera aperture $\vec{x}_c$. **b.** shows the virtual illumination in the reconstruction in time and frequency domain. **d.** shows the entire reconstruction pipeline. The wave propagation model is described in the following. Overall, our proposed method can be thought of as building a virtual lens as shown in **c**, which creates the corresponding virtual image of hidden objects from the captured phasor field.

$$
\begin{aligned}
\mathcal{P}_{\mathcal{F}}(\vec{x}_v, \Omega) &= \mathcal{R}_{z_v}\Big(\mathcal{P}_{\mathcal{F}}(\vec{x}_p, \Omega) \cdot H_{\mathcal{F}}(\vec{x}_p \to \vec{x}_c, \Omega)\Big) \\
&= \mathcal{R}_{z_v}\Big(\mathcal{P}_{\mathcal{F}}(\vec{x}_c, \Omega)\Big) \\
\mathcal{P}_{\mathcal{F}}(x_v, y_v, z_v, \Omega) &= \mathcal{R}_{z_v}\Big(\mathcal{P}_{\mathcal{F}}(x_c, y_c, 0, \Omega)\Big) \\
&= \int_{-\infty}^{+\infty}\!\!\!\int \mathcal{P}_{\mathcal{F}}(x_c, y_c, 0, \Omega)\, \underbrace{\frac{\alpha(x_v, y_v, z_v) e^{-i\frac{\Omega}{c}\sqrt{(x_c - x_v)^2 + (y_c - y_v)^2 + z_v^2}}}{\sqrt{(x_c - x_v)^2 + (y_c - y_v)^2 + z_v^2}}}_{\text{RSD diffraction kernel}} \, dx_c \, dy_c \\
&= \int_{-\infty}^{+\infty}\!\!\!\int \mathcal{P}_{\mathcal{F}}(x_c, y_c, 0, \Omega) \cdot \underbrace{G(x_v - x_c, y_v - y_c, z_v, \Omega)}_{\text{2D convolution kernel}} \, dx_c \, dy_c \\
&= \underbrace{\mathcal{P}_{\mathcal{F}}(x_c, y_c, 0, \Omega) * G(x_c, y_c, z_v, \Omega)}_{\text{Spatial 2D convolution}}, \qquad (2.14)
\end{aligned}
$$

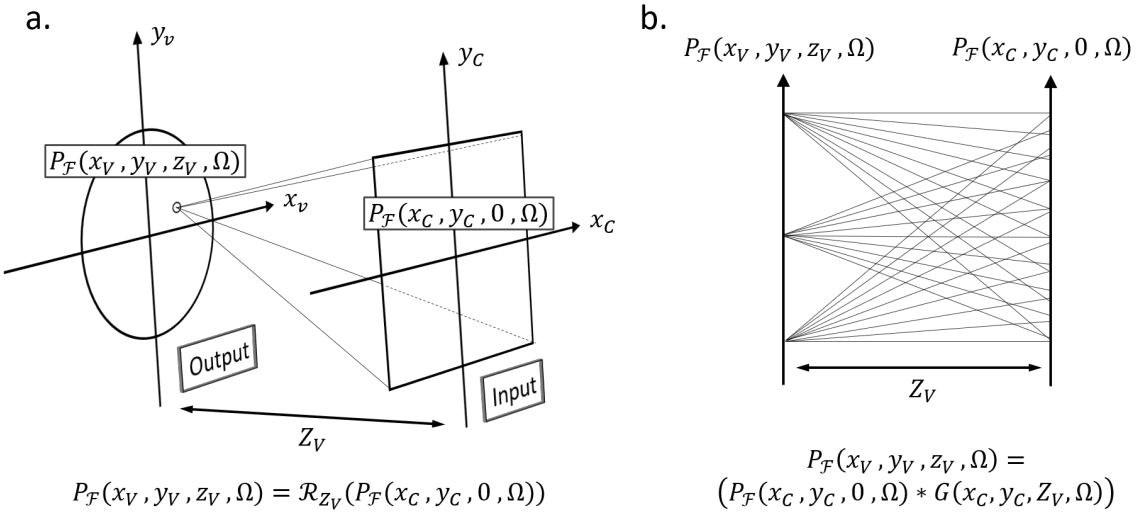

Figure 2.4: Rayleigh Sommerfeld Diffraction (RSD) calculation. **a.** shows the two parallel planes geometrical setup for the reconstruction. The input and output planes are space with z_{v}. **b.** refers to the side view for **a.**

where the geometrical setup is illustrated in Fig. 2.4.

Equation (2.14) considers two parallel planes with spacing z_{v} in a Cartesian coordinate system. For this reason, the RSD notation changed from $\mathcal{R}_{\vec{x}_{\mathrm{v}}}(\cdot)$ for the point $\vec{x}_{\mathrm{v}}$ to $\mathcal{R}_{z_{\mathrm{v}}}(\cdot)$ to indicate that the propagation holds for all points in the plane at distance z_{v} from the relay wall. For a single frequency component Ω, the relation between the wavefront $\mathcal{P}_{\mathcal{F}}(\mathrm{x_c}, \mathrm{y_c}, 0, \Omega)$ at the camera aperture plane and the wavefront $\mathcal{P}_{\mathcal{F}}(\mathrm{x_v}, \mathrm{y_v}, z_{\mathrm{v}}, \Omega)$ at the virtual image plane is a two-dimensional spatial convolution with the 2D convolution kernel defined by $\mathrm{G}(\mathrm{x_c}, \mathrm{y_c}, z_{\mathrm{v}}, \Omega) = \frac{\alpha(\mathrm{x_v}, \mathrm{y_v}, z_{\mathrm{v}}) \cdot \exp(-\mathrm{i}\frac{\Omega}{\mathrm{c}}\sqrt{\mathrm{x_c^2}+\mathrm{y_c^2}+z_{\mathrm{v}}^2})}{\sqrt{\mathrm{x_c^2}+\mathrm{y_c^2}+z_{\mathrm{v}}^2}}$ where the factor $\alpha(\mathrm{x_v}, \mathrm{y_v}, z_{\mathrm{v}})$ will be ignored during reconstruction. Note that the RSD in Eq. (2.14) needs to be calculated for each individual frequency component $\mathcal{P}_{\mathcal{F}}(\mathrm{x_c}, \mathrm{y_c}, 0, \Omega)$. Considering the virtual pulse illumination in Eq. (2.11), the wavefront $\mathcal{P}_{\mathcal{F}}(\mathrm{x_c}, \mathrm{y_c}, 0, \Omega)$ is a broad-band signal; its spectrum is a Gaussian centered around Ω_{C} as shown in Eq. (2.12). For this reason, it is sufficient to consider the frequency range $\Omega \in [\Omega_{\mathrm{C}} - \Delta\Omega, \Omega_{\mathrm{C}} + \Delta\Omega]$. Although the magnitude is not completely zero outside this interval, it is very small and

can be neglected. The chosen range $\Delta\Omega$ depends on the virtual illumination pulse bandwidth and thus on the pulse width parameter σ. Thus, applying Eq. (2.14) for the frequencies $\Omega \in [\Omega_C - \Delta\Omega, \Omega_C + \Delta\Omega]$ and subsequent inverse Fourier transform with respect to time is equivalent to sending the designed modulated virtual illumination pulse wavefront $\mathcal{P}(\vec{x}_p, t) = e^{i\Omega_C t} e^{-\frac{t^2}{2\sigma^2}}$ into the hidden scene, capturing its reflection at the visible relay wall, and propagating it back into the scene or imaging it onto a virtual imaging sensor using a virtual lens. The relay wall functions as a virtual aperture. Thus, with a finite frequency interval, a phasor field propagation model is shown below,

$$\mathcal{P}(x_v, y_v, z_v, t) = \int_{\Omega_C - \Delta\Omega}^{\Omega_C + \Delta\Omega} e^{j\Omega t} \cdot \underbrace{\mathcal{R}_{z_v}\Big(\mathcal{P}_{\mathcal{F}}(x_c, y_c, 0, \Omega)\Big)}_{\text{Monochromatic wavefront at depth } z_v} \frac{d\Omega}{2\pi} \tag{2.15}$$

The output $\mathcal{P}(x_v, y_v, z_v, t)$ in Eq. (2.15) depends on the time t, as each reconstruction point is illuminated only for a short period of time. Taking the absolute value of $\mathcal{P}(x_v, y_v, z_v, t)$ in Eq. (2.15) and squaring it makes us arrive at a 4D reconstruction (cf. Eq. (2.13)). We can understand this reconstruction as a movie of a virtual pulse travelling through the hidden scene, as shown in Fig. 2.5. In this figure, a patch shaped as a 4 is being illuminated by a spherical wavefront coming from the illumination point on the relay surface. This $\mathcal{P}(x_v, y_v, z_v, t)$ in the output as a four-dimensional function contains direct and indirect signals from a unknown scene. This direct and indirect effect is shown in Liu *et al.*[3] with a movie in the supplementary material. Typically, for a three-dimensional reconstruction, we are interested the direct (3rd) bounce signal from a unknown scene. This can be performed by calculating the spherical geometry as a function of point source illumination position $(x_{ls}, y_{ls}, 0)$ and replacing t at each voxel (x_v, y_v, z_v):

$$t := \frac{1}{c}\sqrt{(x_v - x_{ls})^2 + (y_v - y_{ls})^2 + z_v^2}. \tag{2.16}$$

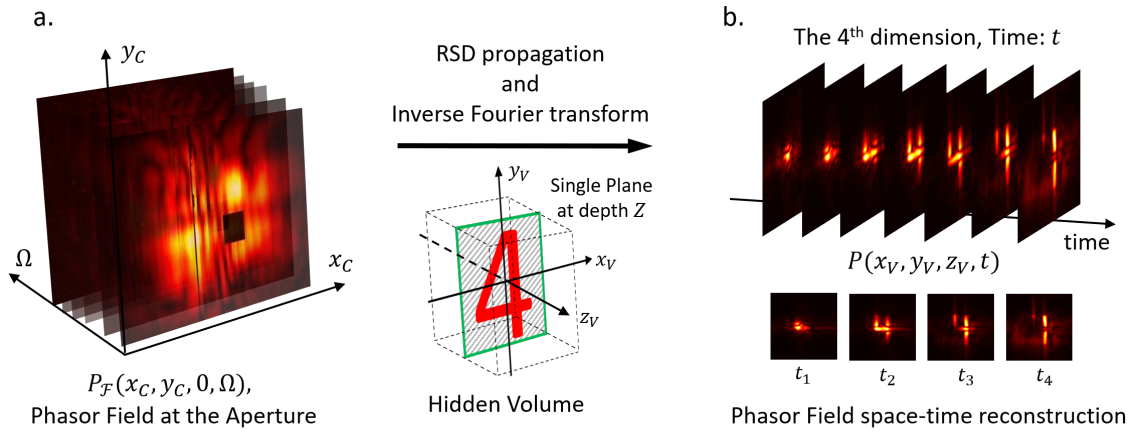

Figure 2.5: Space-time wave propagation using RSD. **a.** The phasor field collected at the aperture forms a spatial frequency cube. Given the output plane, by using the RSD propagation model, we can recover the hidden wavefront at any time instances. **b.** shows this space-time wave propagation method where one can reveal a spherical wavefront that moves into the hidden scene. Even though **b** only shows reconstruction at a single depth plane, our proposed method can be generalized into the three-dimensional volume as well, which leads to a four-dimensional reconstruction space-time volume.

2.3 Appendix: Results and additional notes

Phasor field camera for confocal NLOS data

The RSD reconstruction method for NLOS data presented so far only deals with the non-confocal case, which means that the illumination point $\vec{x}_p$ and the camera point $\vec{x}_c$ on the relay wall are different. However, a confocal dataset $H^c(\vec{x}_p \rightarrow \vec{x}_c, t)$ as used in LCT and FK migration algorithms [4, 2] only contains data with $\vec{x}_p = \vec{x}_c$:

$$H^c(\vec{x}_p \rightarrow \vec{x}_c, t) = H(\vec{x}_p \rightarrow \vec{x}_c, t)\delta(\vec{x}_p - \vec{x}_c)\,. \tag{2.17}$$

Such a dataset is not suitable for implementing the virtual point light source described in Eq. (2.11). Instead, we can model an illumination wavefront that is

focused on $\vec{x}_v$:

$$\mathcal{P}(\vec{x}_p, t) = e^{j\Omega(t-\frac{1}{c}|\vec{x}_v-\vec{x}_p|)} e^{-\frac{(t-t_0-\frac{1}{c}|\vec{x}_v-\vec{x}_p|)^2}{2\sigma^2}}. \tag{2.18}$$

Setting t_0 to 0 and applying the Fourier transform leads to

$$\mathcal{P}_{\mathcal{F}}(\vec{x}_p, \Omega) = \left(2\pi\delta(\Omega-\Omega_C) \underset{f}{*} \sqrt{2\pi}\sigma e^{-\frac{\sigma^2\Omega^2}{2}}\right) e^{-j\frac{\Omega}{c}|\vec{x}_v-\vec{x}_p|}. \tag{2.19}$$

Inserting into Eq. (2.13) yields

$$\begin{aligned}
&I(\vec{x}_v, t) = \\
&\left|\int_{-\infty}^{+\infty} e^{j\Omega t}\mathcal{R}_{\vec{x}_v}\Big(\int_P (2\pi)^{\frac{3}{2}}\sigma e^{-\frac{\sigma^2(\Omega-\Omega_C)^2}{2}} e^{-j\frac{\Omega}{c}|\vec{x}_v-\vec{x}_p|}\delta(\vec{x}_p-\vec{x}_c)\cdot H_{\mathcal{F}}(\vec{x}_p\to\vec{x}_c,\Omega)d\vec{x}_p\Big)\frac{d\Omega}{2\pi}\right|^2 \\
&=\left|\int_{-\infty}^{+\infty} e^{j\Omega t}\mathcal{R}_{\vec{x}_v}\Big((2\pi)^{\frac{3}{2}}\sigma e^{-\frac{\sigma^2(\Omega-\Omega_C)^2}{2}} e^{-j\frac{\Omega}{c}|\vec{x}_v-\vec{x}_c|}\cdot H_{\mathcal{F}}(\vec{x}_c\to\vec{x}_c,\Omega)\Big)\frac{d\Omega}{2\pi}\right|^2 \\
&=\left|\int_{-\infty}^{+\infty} e^{j\Omega t}\int_C (2\pi)^{\frac{3}{2}}\sigma e^{-\frac{\sigma^2(\Omega-\Omega_C)^2}{2}} e^{-j\frac{\Omega}{c}|\vec{x}_v-\vec{x}_c|}\cdot H_{\mathcal{F}}(\vec{x}_c\to\vec{x}_c,\Omega)e^{-jk|\vec{x}_v-\vec{x}_c|}d\vec{x}_c\frac{d\Omega}{2\pi}\right|^2 \\
&=\left|\int_{-\infty}^{+\infty} e^{j\Omega t}\int_C (2\pi)^{\frac{3}{2}}\sigma e^{-\frac{\sigma^2(\Omega-\Omega_C)^2}{2}}\cdot H_{\mathcal{F}}(\vec{x}_c\to\vec{x}_c,\Omega)e^{-jk2|\vec{x}_v-\vec{x}_c|}d\vec{x}_c\frac{d\Omega}{2\pi}\right|^2.
\end{aligned} \tag{2.20}$$

The reconstruction thus uses an RSD operator with an additional factor of two doubling all distances. We use our fast RSD operator to evaluate this RSD integral.

Fourier domain histogram

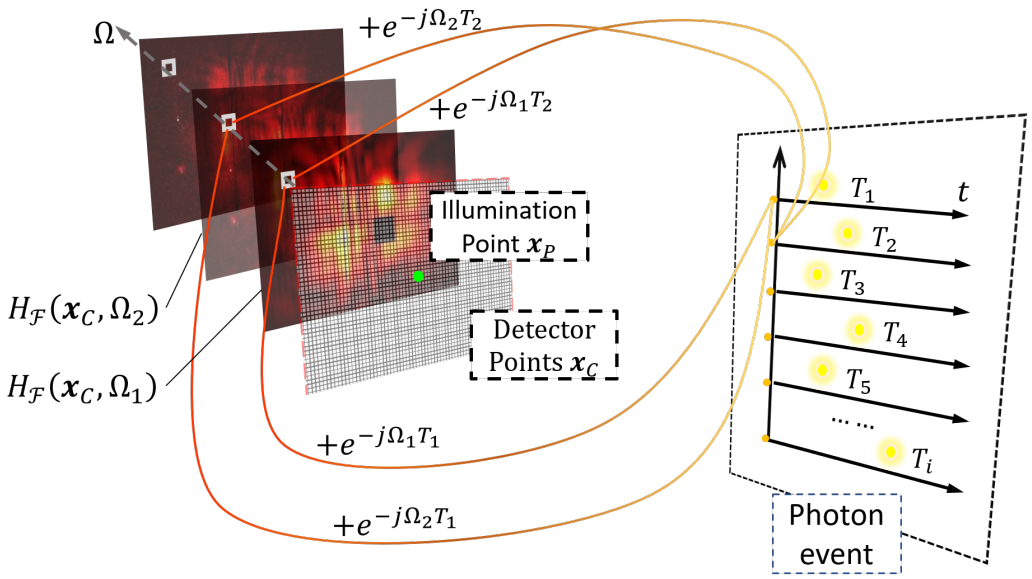

Figure 2.6: Illustration of Fourier Domain Histogram. Instead of binning the photon event in time, we propose doing the binning in the frequency domain. This allows us directly to sample the phasor field wavefront $H_{\mathcal{F}}(\vec{x}_c, \Omega)$ used for reconstructions. Ω stands for the frequency range for the phasor field wavefront. The equation for the Fourier Domain Histogram can be applied during measurements, which is a summation of complex phasors (or a separated real and imaginary part).

According to Eq. (2.13), the virtual wave acquired at the virtual aperture is calculated by $\mathcal{P}_{\mathcal{F}}(\vec{x}_c, \Omega) = \mathcal{P}_{\mathcal{F}}(\vec{x}_p, \Omega) \cdot H_{\mathcal{F}}(\vec{x}_p \rightarrow \vec{x}_c, \Omega)$. This requires the Fourier domain representation of the impulse response $H_{\mathcal{F}}(\vec{x}_p \rightarrow \vec{x}_c, \Omega)$ from $\vec{x}_p$ to $\vec{x}_c$. A new memory efficient direct acquisition method for $H_{\mathcal{F}}(\vec{x}_p \rightarrow \vec{x}_c, \Omega)$ is presented in the following.

The SPAD detector uses Time-Correlated Single Photon Counting (TCSPC) to generate the transient responses $H(\vec{x}_p \rightarrow \vec{x}_c, t)$. After the emission of a laser pulse, a SPAD pixel receives one photon and an electronic signal is transmitted to the TCPSC unit that encodes the time between the emission of the laser pulse and the detection of an associated returning photon. The arrival times of all photons

during a measurement interval are transferred to a computer and are arranged in a histogram to obtain the transient scene response $H(\vec{x}_p \to \vec{x}_c, t)$ for a given $\vec{x}_p$ and $\vec{x}_c$. To obtain $H_{\mathcal{F}}(\vec{x}_p \to \vec{x}_p, \Omega)$ we could collect and store these TCSPC histograms and perform the Fourier transform on it. A more memory efficient way is to build the frequency spectrum directly from the timing data obtained from the hardware. We call this new capturing method a Fourier Domain Histogram (FDH) and its creation process is shown in Fig. 2.6. It can be written as

$$\begin{aligned} H_{\mathcal{F}}(\vec{x}_p \to \vec{x}_c, \Omega) &= \int_{-\infty}^{+\infty} H(\vec{x}_p \to \vec{x}_c, t) \cdot e^{-j\Omega t}\, dt \\ &= \int_{-\infty}^{+\infty} \left(\sum_{i=1}^{I} \delta(t - T_i) \right) \cdot e^{-j\Omega t}\, dt \\ &= \sum_{i=1}^{I} e^{-j\Omega T_i}\,. \end{aligned} \tag{2.21}$$

The travel times T_i are discrete; the time resolution is determined by the acquisition hardware (in the context of NLOS imaging typically a few to tens of picoseconds). Equation (2.21) means that the Fourier domain histogram $H_{\mathcal{F}}(\vec{x}_p \to \vec{x}_c, \Omega)$ is acquired by multiplying each of the I photon travel times T_i, $i = 1, \ldots, I$, by a phase term depending on the considered frequency Ω and adding the result to the previous value for that frequency. As a consequence, instead of a large number of time bins (on the order of thousands), only one value for each Ω needs to be stored and processed. Fig. 2.6 illustrates the generation of the FDH. Similar to the time domain histogram binning, this FDH performs binning for each captured photon.

We want to remark that the travel times T_i in Eq. (2.21) are measured from the respective illumination position on the relay wall into the scene and back to the relay wall at the detector focus position. The travel times from the laser setup to the illumination on the relay wall and from the detector focus point on the relay wall to the detector setup have been subtracted and are not part of H. Alternatively,

the total travel time from laser to detector can be incorporated and the travel times from laser to wall and wall to detector are combined into Δt. The final result from Eq. (2.21) is then multiplied by $e^{j\Omega\Delta t}$ to correct for this constant time offset.

Results

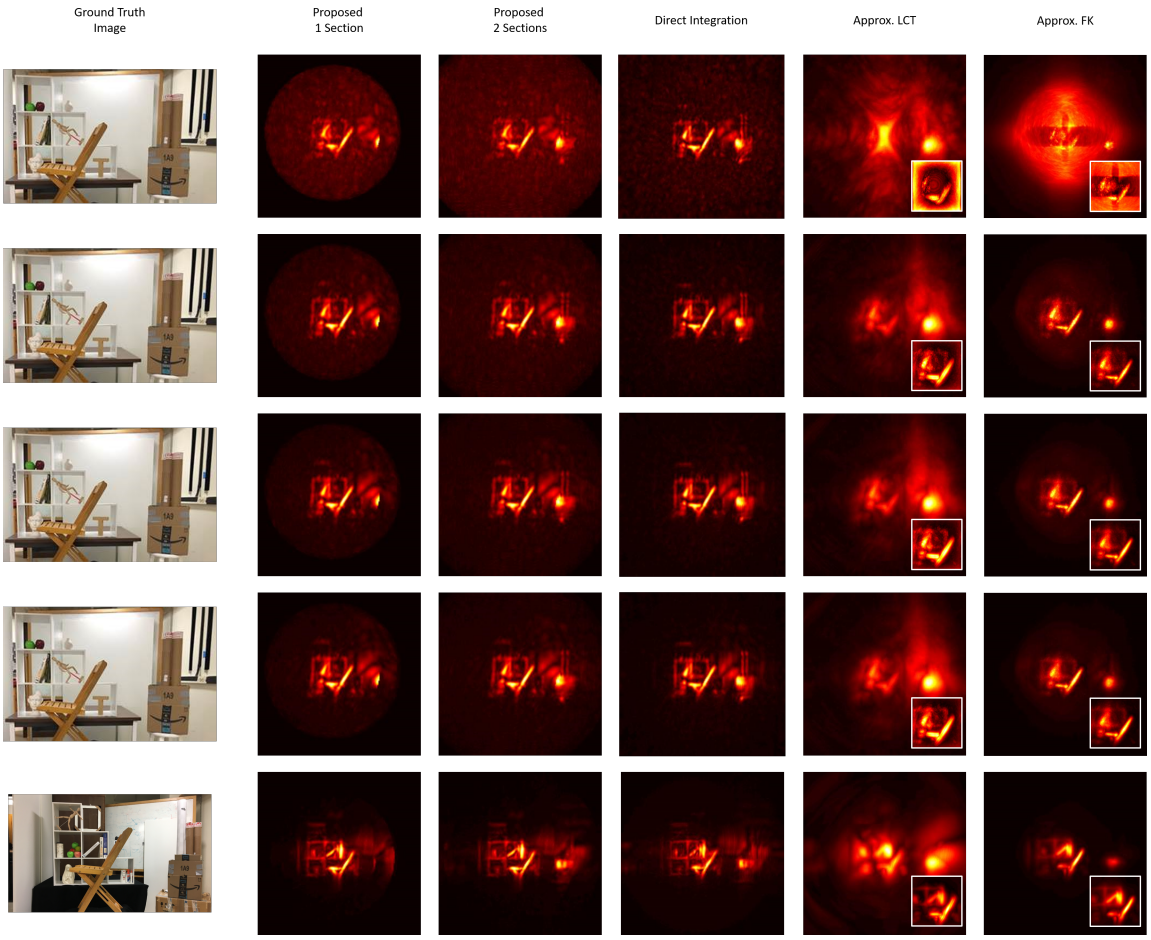

Figure 2.7: Methods comparison on Office Scene: Exposure time per each pixel measurement from first row to last row is 1 ms, 5 ms, 10 ms, 20 ms, 1000 ms (note that the 1000 ms Office Scene dataset was acquired with slight differences in the object location). The total acquisition time from first row to last row is 23 s, 117 s, 4 min, 8 min, 390 min. The width of result in each dimension is 3 m as details provided in [1]. Each column shows the reconstruction with different methods. The first two columns stand for our proposed RSD [1] based solver with one or two spatial sections. The circle in the first column is actually the size of the farthest reconstruction plane which is the one with the largest region that is calculated with the same distance shift B_1. All planes in front of this one have a smaller reconstruction area; due to the maximum operation along the depth dimension, the circle size is defined by the largest one in the back. The third column is the Direct Integration (back-projection solver) as a comparison for the first two columns. The last two columns refer to the approximation method [2] which approximate non-confocal by confocal data and solve it through the scanning-based solver (LCT: forth column, FK-migration: fifth column). For the last two columns, each small image shows the results from midpoint approximation[2] in order to approximate confocal data from non-confocal measurements. The respective larger image results from zero-padding applied to the input data to show the same reconstruction volume as the first three columns.

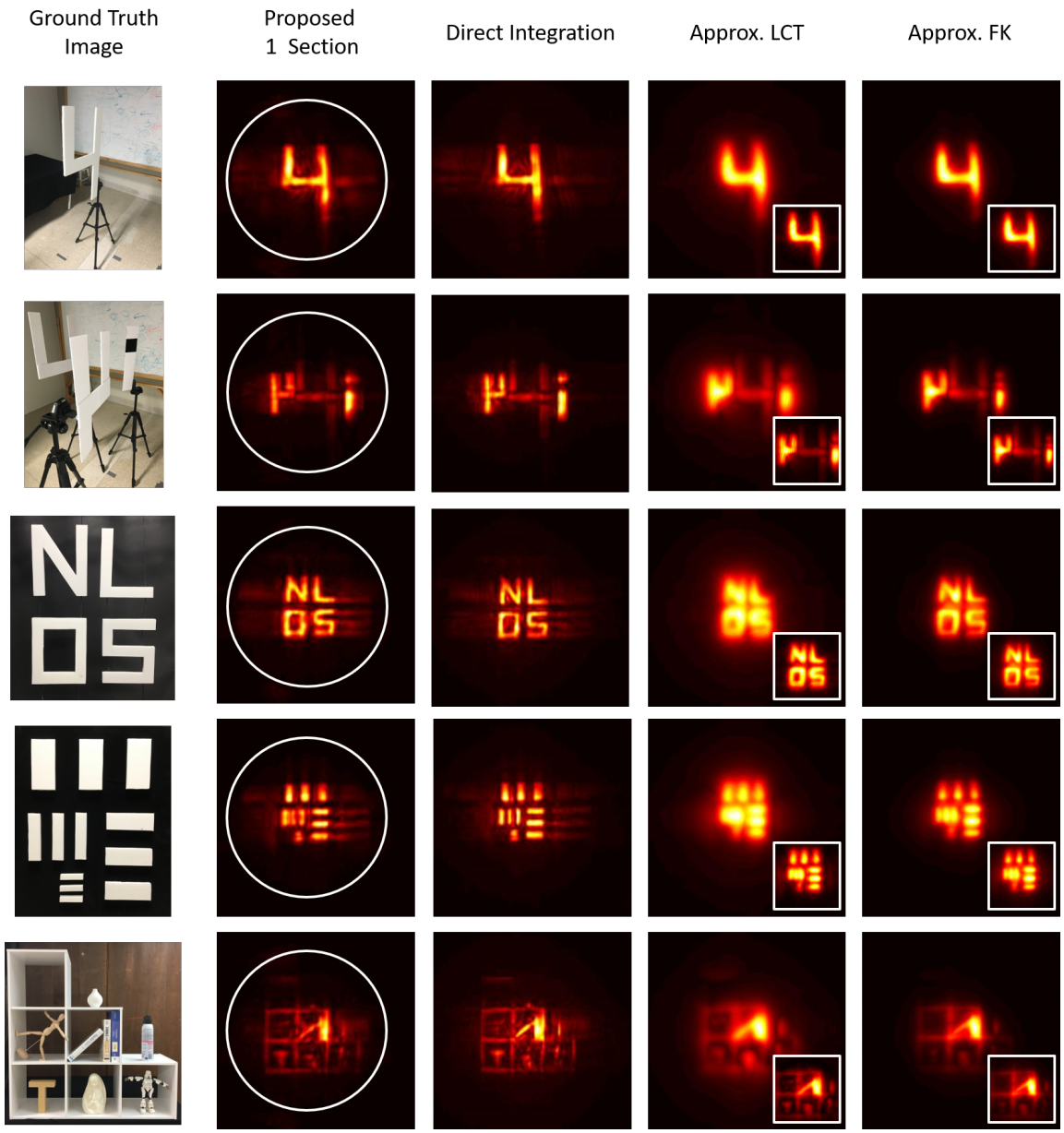

Figure 2.8: Methods comparison on simple targets: Exposure time for these scenes are all 1000 ms per each pixel measurement. The total acquisition time for each target is 390 min. The width of result in each dimension is 2 m as details provided in [1]. Each row shows a different simple target, each column the reconstruction from different methods. The first column stands for our proposed RSD [1] based solver with one spatial section (inside white circle) corresponding to one spatial section presented in [1] . The second column is from Direct Integration (back-projection solver) for comparison with the first column. The last two columns show the approximate method [2] which approximate non-confocal as a confocal datasets and reconstruct through confocal solvers (LCT: forth column, FK-migration: fifth column). For the last two columns, each small image shows the results from mid-point approximation[2] in order to approximate confocal data from non-confocal measurements. The respective larger image results from zero-padding applied to the input data to show the same reconstruction volume as the first three columns.

Office Scene Acquisition Time	Proposed 1 section	Proposed 2 sections	Direct Integration	Approx. LCT (low res)	Approx. FK (low res)
1 ms	19.9 s	30.1 s	8248 s	15.6 s (3.80 s)	22.4 s (5.52 s)
5 ms	15.2 s	24.3 s	8685 s	15.7 s (3.76 s)	22.5 s (5.52 s)
10 ms	15.7 s	24.7 s	8534 s	15.8 s (3.80 s)	22.9 s (5.5 s)
20 ms	16 s	23.9 s	8667 s	15.5 s (3.79 s)	22.3 s (5.51 s)
1000 ms	18.7 s	37 s	5776 s	15.5 s (3.74 s)	22.4 s (5.5 s)

Table 2.1: Office Scene run time comparison: This table shows the actual run time for generating the results in Fig. 2.7. Our proposed method starts from the captured wavefront and has the same volume size as the Direct Integration method (150 x 150 x 125 voxels). For showing the best reconstruction quality of the approx LCT and approx FK methods, we use a voxel grid of 256 x 256 x 512 with 1 cm sampling resolution on the relay wall. Approx LCT and approx FK can be much faster when down-sampling the spatial dimensions as shown in brackets (128 x 128 x 512), but the results are even more blurry than the ones shown in Fig. 2.7. Note that down-sampling the spatial domain is not possible, as the number of spatial voxels has to equal the number of time bins and lower time resolution would lead to even worse results (but faster run time). The flexibility of adapting the full 3D voxel grid is an advantage of our RSD algorithm.

Dataset	Proposed	Direct Integration	Approx. LCT (low res)	Approx. FK (low res)
4	2.9 s	1298 s	15.5 s (3.7 s)	21.8 s (5.5 s)
44i	2.8 s	1316 s	15.4 s (3.72 s)	22.1 s (5.5 s)
NLOS	2.9 s	1292 s	15.6 s (3.69 s)	23 s (5.47 s)
Resolution Bar	2.9 s	1290 s	15.4 s (3.71 s)	25 s (5.49 s)
Shelf Light On	2.7 s	1302 s	15.3 s (3.67 s)	22.3 s (5.55 s)

Table 2.2: Simple scenes run time comparison: This table shows the actual run time for generating the results in Fig. 2.8. Our proposed method starts from the captured wavefront and has the same volume size as the Direct Integration method. For showing the best reconstruction quality of the approx LCT and approx FK methods, we use a voxel grid of 256 x 256 x 512 with 1 cm sampling resolution on the relay wall. Approx LCT and approx FK can be much faster when down-sampling the spatial dimensions as shown in brackets (128 x 128 x 512).

Dataset	Scene Depth	Material
Officescene 1 ms, 5 ms, 10 ms, 20 ms	0.5 m - 2.5 m	Wooden chair, white shelf, cardboard, books, plastic, white board, statue ...
Officescene 1000 ms	0.5 m - 2.5 m	Wooden chair, white shelf, cardboard, books, plastic, white board, statue ...
4	1 m	White styrofoam
44i	0.5 m - 1.3 m	White styrofoam
NLOS	0.75 m	White styrofoam
Resolution Bar	0.75 m	White styrofoam
Shelf Light On	0.8 m	White shelf, cardboard, books, plastic ...

Table 2.3: Target scene parameters: scene depth complexity (distance away from the relay wall), targets material.

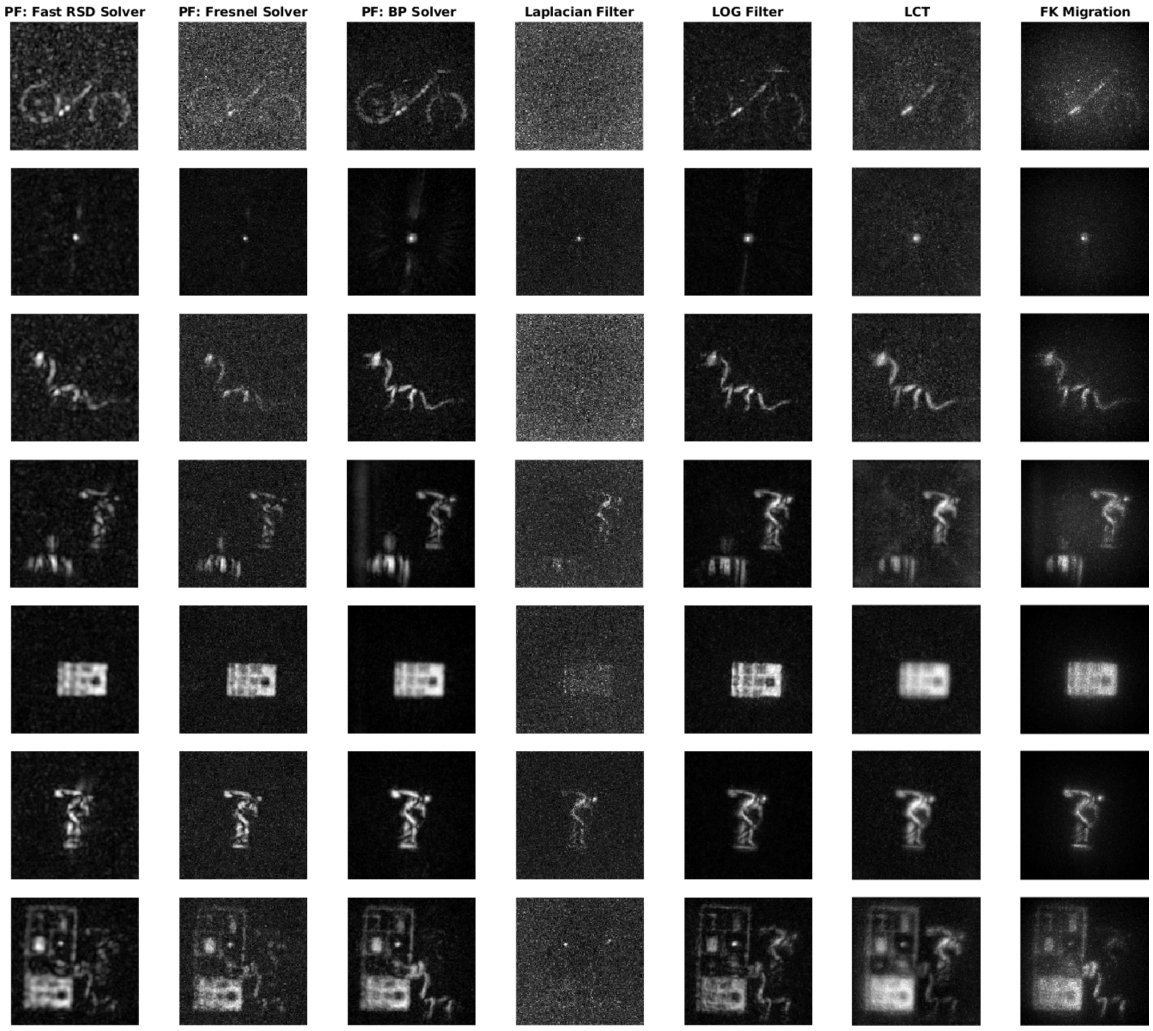

Figure 2.9: Comparisons on confocal scanning shortest exposure datasets [2]: The first three rows correspond to the phasor field (PF) NLOS method [3] out of which the first two rows present our fast implementation [1] (one with RSD, one with Fresnel diffraction kernel[3]) and the third row shows the results using the convolution backprojection kernel calculated from the LCT [4]. The fourth (Laplacian filter) and fifth (LOG: Laplacian of Gaussian) rows are filtered backprojection with filter implementation from [3] and the backprojection step is calculated from the convolution provided by LCT. The last two rows show LCT and FK-Migration [2]. For the shortest exposure dataset, we interestingly find out that LOG is quite robust. The Fresnel diffraction solver seems suited for confocal data, although it does not perform well on non-confocal data [3].

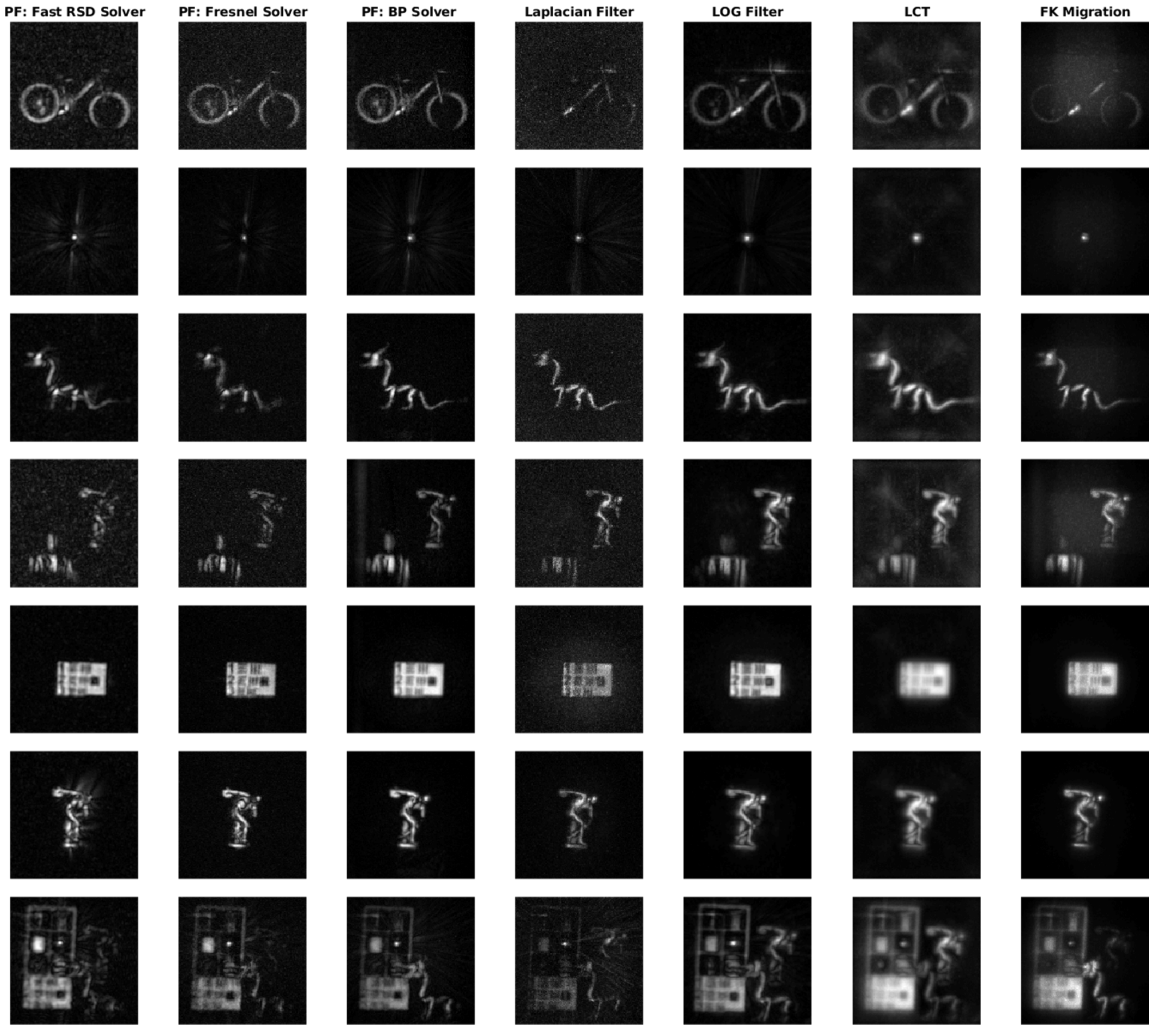

Figure 2.10: Comparisons on confocal scanning longest exposure datasets [2]: The first three rows correspond to the phasor field (PF) NLOS method [3] out of which the first two rows present our fast implementation [1] (one with RSD, one with Fresnel diffraction kernel [3]) and the third row shows the results using the convolution backprojection kernel calculated from the LCT [4]. The fourth (Laplacian filter) and fifth (LOG: Laplacian of Gaussian) rows are filtered backprojection with filter implementation from [3] and the backprojection step is calculated from the convolution provided by LCT. The last two rows show LCT and FK-Migration [2]. For the longest exposure dataset, almost all methods perform well. The Fresnel diffraction solver seems suited for confocal data, although it does not perform well on non-confocal data [3].

3 COMPUTATIONAL SOLVER FOR PHASOR FIELD NON-LINE-OF-SIGHT IMAGING

In this section, we present a computational solver based on the Phasor field NLOS imaging framework mentioned in Ch. 2. Using this computational solver, we can perform a reconstruction using a captured NLOS data. To do that, we start from a continuous model and its discrete version first. Then we carry out its step-by-step recipe for its numerical implementation. Unless otherwise specific mention the input/output geometry, we assume input wavefront is captured on a plane in the Cartesian coordinate. As shown in Ch. 2 Eq. (2.15), solving this Rayleigh Sommerfeld Diffraction (RSD) efficiently is a key to a reconstruction step. A side from notation used in the previous chapter, we use all scalar notation in this chapter which makes it converges to a self-contain programming recipe.

3.1 Continuous model

Starting with the RSD equation in an analytical form shown in Eq. (3.1). The output wavefront $u_2(\mu, \nu, k)$ is calculated through a spatial convolution of input wavefront $u_1(x, y, k)$ and a convolution kernel $h(x, y, k, Z)$. $(x, y), (\mu, \nu)$ stand for coordinate variables on input and output plane. Since input and output planes are parallel to each other, Z denotes the spacing between them. Equation only describes a single monochromatic wave's behavior, thus it is a function of wavenumber $k = \frac{2\pi}{\lambda} = \frac{\omega}{c}$ which λ, ω, c refers to wavelength, angular frequency and speed of light in air.

$$\begin{aligned} u_2(\mu, \nu, k) &= \iint_{-\infty}^{+\infty} u_1(x, y, k) \frac{e^{-jk\sqrt{(x-\mu)^2+(y-\nu)^2+Z^2}}}{\sqrt{(x-\mu)^2+(y-\nu)^2+Z^2}} \, dx \, dy \\ &= \Big(u_1(x, y, k) \underset{x-y}{*} h(x, y, k, Z) \Big) (\mu, \nu) \end{aligned} \tag{3.1}$$

$h(x, y, k, Z)$ refers to the RSD propagation kernel $h^{RSD}(x, y, z)$ in Eq.(4.12).

$h(x, y, k, Z)$ in Eq. (3.1) is as below,

$$h(x, y, k, Z) = \frac{e^{-jk\sqrt{x^2+y^2+Z^2}}}{\sqrt{x^2+y^2+Z^2}} \tag{3.2}$$

From Eq. (3.1), we can study the RSD propagation in the spatial frequency Fourier domain using Fourier transform convolution property,

$$U_2(f_\mu, f_\nu, k) = U_1(f_x, f_y, k) \cdot H(f_x, f_y, k, Z) \tag{3.3}$$

A short explanation for symbols used in Eq. (3.3) is provided here: $U_1(f_x, f_y, k)$ and $U_2(f_\mu, f_\nu, k)$ stands for the two dimensional Fourier transform $\mathcal{F}(u_1(x, y, k))$ and $\mathcal{F}(u_2(\mu, \nu, k))$ along $(x, y), (\mu, \nu)$ dimension. k, Z stands for the wavenumber and distance. Next, we are going to discretize this RSD model in order to build its numerical solver.

3.2 Discrete model

To carry out a numerical procedure, we have to discretize the continuous RSD model shown in the Eq. (3.1).

There are several key parameters related to a discrete sampling:

- Input, output aperture $S_{in} - S_{in}$ $S_{out} - S_{out}$ refer to a physical dimension, unit in meter, same side length in both dimension.
- Input, output discrete matrix in size $N - N$
- Input, output aperture spatial sampling δ_{in} δ_{out}, unit meter. It also means a spacing between sampling grids.

In practice, S_{in} and δ_{in} refers to the physical size of the scanning area on the visible wall and sampling spacing of the discrete aperture array. S_{out} governs the physical size of the reconstruction area for each depth slice, and δ_{out} stands for the spacing of reconstruction grid.

We use additional symbols $[n\mu, n\nu]$ and $[nx, ny]$ to represent discrete index. Consider an input and output square window of N by N, which their origin are at the center of the aperture. We choose N as an even number for the simplification. Moreover, we consider a spatial sampling interval for any plane is the same in both horizontal and vertical direction.

Thus, the discrete model for the Eq.(3.1) is defined as follows:

$$u_2[n\mu, n\nu] = \delta_{in}^2 \sum_{nx,ny=-N/2}^{N/2-1\,N/2-1} u_1[nx, ny] \cdot \frac{exp\Big[-j \cdot k \cdot r[n\mu, n\nu, nx, ny]\Big]}{r[n\mu, n\nu, nx, ny]}$$

$$r[n\mu, n\nu, nx, ny] = \Big[(nx \cdot \delta_{in} - n\mu \cdot \delta_{out})^2 + (ny \cdot \delta_{in} - n\nu \cdot \delta_{out})^2 + Z^2\Big]^{1/2}$$

$$\delta_{in} = S_{in}/N$$

$$\delta_{out} = S_{out}/N \tag{3.4}$$

Now, there are two cases based on whether an input and output grid sampling spacing is equal to each other or not. When it is equal to each other, it falls down to the common used scenario shown before [1].

Case 1 - equal sampling

When $\delta_{in} = \delta_{out} = \delta$, Eq. (3.4) turns into a standard discrete linear convolution as follows:

$$u_2[n\mu, n\nu] = \sum_{nx,ny=-N/2}^{N/2-1\,N/2-1} u_1[nx, ny] \cdot h[n\mu - nx, n\nu - ny, Z]$$

$$h[nx, ny] = \delta^2 \cdot \frac{exp[-jk\sqrt{nx^2\delta^2 + ny^2\delta^2 + Z^2}]}{\sqrt{nx^2\delta^2 + ny^2\delta^2 + Z^2}} \tag{3.5}$$

With equally sampling case above in Eq. (3.5), final reconstructions can be calculated using the convolution-multiplication theorem,

$$u_2(\mu,\nu) = \mathcal{F}^{-1}\Big(\mathcal{F}(u_1(x,y)) \cdot \mathcal{F}(h(x,y))\Big) \tag{3.6}$$

This type of equal sampling case for reconstructions is shown in Liu *et al.*[3].

Case 2 - fractional sampling

As show in Fig. 3.1, when $\delta_{in} \neq \delta_{out}$, input and output spatial sampling spacing is connected by a fractional number $\alpha = \delta_{in}/\delta_{out}$. Thus, $\delta_{out} \geqslant \delta_{in}$ in this case to account for the loss of resolution in the reconstruction domain. α also means the reconstructed image resolution drops linearly as a function of distance. Remember that we did not change the size of a numerical matrix, so by changing the output grid spacing δ_{out}, it allows us to cover a larger unknown area with the same number of numerical matrix. By doing this fractional sampling mentioned above, it gives us a trapezoidal grip as opposed to a cube grid using a same number of numerical elements for reconstruction which is shown in Fig. 3.1.

We define $\alpha = \delta_{in}/\delta_{out}$ which is a fractional number, $\delta = \delta_{out}$ and $\delta_{in} = \alpha \cdot \delta$. Thus, Eq. (3.4) turns into

$$u_2[n\mu, n\nu] = \alpha^2\delta^2 \sum_{nx,ny=-N/2}^{N/2-1}\sum^{N/2-1} u_1[nx, ny] \cdot \frac{exp\Big[-j \cdot k \cdot r[n\mu, n\nu, nx, ny]\Big]}{r[n\mu, n\nu, nx, ny]}$$

$$r[n\mu, n\nu, nx, ny] = \Big[(nx \cdot \alpha \cdot \delta - n\mu \cdot \delta)^2 + (ny \cdot \alpha \cdot \delta - n\nu \cdot \delta)^2 + Z^2\Big]^{1/2} \tag{3.7}$$

Unlike Eq. (3.5) turns in a standard linear convolution which uses the convolution-multiplication property in Eq. (3.6), in this fractional sampling case, Eq. (3.7) can not performed directly using a stand Fourier transform. To solve this fractional situation, a new type of fractional scale convolution need to be considered.

To see how this new type of solver works, first, let's define new variables $nx' = nx \cdot \alpha$ and $ny' = ny \cdot \alpha$, rewrite Eq. (3.7) as follows,

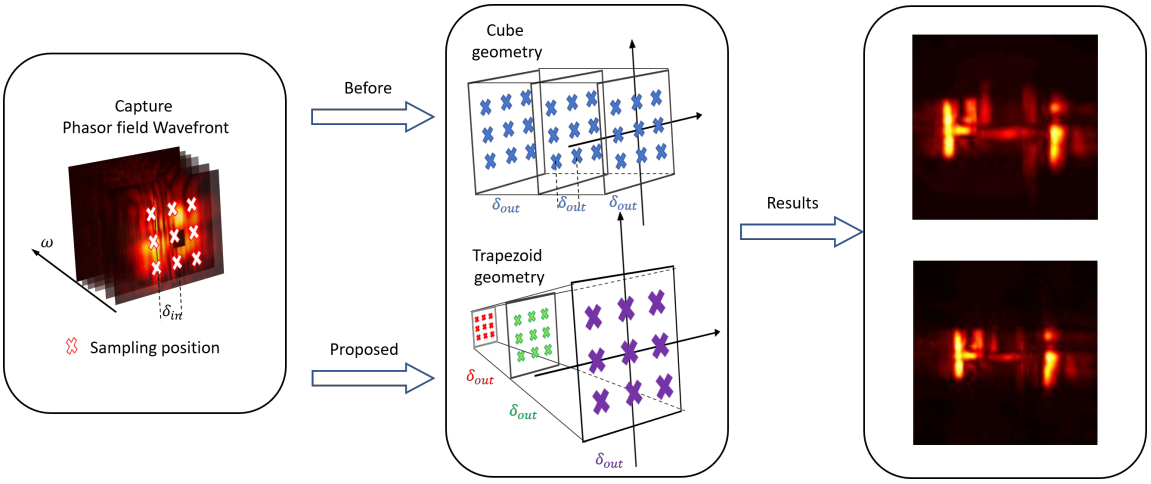

Figure 3.1: shows proposed reconstruction comparing to past algorithms. Input captured Phasor field wavefront is sampled uniformly on the relay wall by δ_{in}. Past fast algorithms can reconstruct hidden images defined on the same uniform grid in the output domain which means $\delta_{out} = \delta_{in}$. We consider a fraction rate solver allows us to reconstruct hidden images where input and output grid are defined by a fractional number which means $\alpha = \frac{\delta_{in}}{\delta_{out}}$. To construct a real world camera perspective projection, our solver allows δ_{out} varies as a function of distance. This is illustrated as red, green and purple color δ_{out} in the figure above. With a same number of matrix element, but changing δ_{out}, it covers larger field of view of hidden space as shown on the right.

$$u_2[n\mu, n\nu] = \alpha^2\delta^2 \sum_{nx,ny=-N/2}^{N/2-1\ N/2-1} u_1[\frac{nx'}{\alpha}, \frac{ny'}{\alpha}] \cdot \frac{exp\Big[-j\cdot k\cdot r[n\mu, n\nu, nx, ny]\Big]}{r[n\mu, n\nu, nx, ny]}$$

$$r[n\mu, n\nu, nx, ny] = \Big[(nx'\cdot\delta - n\mu\cdot\delta)^2 + (ny'\cdot\delta - n\nu\cdot\delta)^2 + Z^2\Big]^{1/2} \tag{3.8}$$

Then in order to perform Eq. (3.8) using a similar convolution-multiplication property, the Fourier transform of $u_1[\frac{nx'}{\alpha}, \frac{ny'}{\alpha}]$ which it is on a scaled grid has to be calculated. So we define $U_1^{\alpha}(nx', ny') = \mathcal{F}^{\alpha}(u_1(nx, ny))$ as a scaled version of Fourier transform as follows. Both nx, nx', ny, ny' are variables with a range

from $N/2$ to $N/2-1$. Term $(-j\frac{2\pi}{N})$ inside the complex exponential is the same as a standard discrete Fourier transform which it defines a fundamental frequency.

Thus, this scaled Fourier transform solver is shown below in Eq. (3.9),

$$\begin{aligned} U_1^{\alpha}(nx', ny') &= \mathcal{F}^{\alpha}(u_1(nx, ny)) \\ &= \sum_{nx,ny=-N/2}^{N/2-1\ N/2-1} u_1(nx, ny) \cdot exp\Big[(-j\frac{2\pi}{N}) \cdot -\alpha(nx \cdot nx' + ny \cdot ny')\Big] \end{aligned} \tag{3.9}$$

Then, with the defined scaled Fourier transform in Eq. (3.9), the case 2 fractional sampling scenario in Eq. (3.7) can be calculated as follows,

$$u_2(\mu, \nu) = \mathcal{F}^{-1}\Big(\mathcal{F}^{\alpha}(u_1(x, y)) \cdot \mathcal{F}(h(x, y))\Big) \tag{3.10}$$

There are three numerical procedures in Eq. (3.10), one is a solver to perform the scaled Fourier transform $\mathcal{F}^{\alpha}(\cdot)$ shown in Eq. (3.9), one is a solver to construct the RSD kernel $h(x, y)$ from Eq. (3.7), the last solver is a standard fast Fourier transform $\mathcal{F}(\cdot)$ which is available directly. In the next section, based on equations above, we derive close-formed numerical procedures for Eq. (3.10).

3.3 Numerical procedure

Two numerical solvers are shown in this section which are a scaled Fourier transform solver and a RSD kernel solver. For the equal sampling scenario shown in Eq. (3.6) Sec. 3.2, it only needs the RSD kernel solver. For the fractional sampling scenario shown in Eq. (3.10) Sec. 3.2, it needs both the scaled Fourier transform solver as well as the RSD kernel solver.

Scaled Fourier transform solver

We derive a numerical procedure to implement the scaled Fourier transform $\mathcal{F}^{\alpha}(\cdot)$ shown in Eq. (3.9). As we show later in this section that a multi-dimensional scaled Fourier transform solver can be treated as a one-dimensional version in multiple times. Thus, we start from a one-dimensional scaled Fourier transform and then prove how we extend this one-dimensional solver into multi-dimensional cases.

First, we consider a one-dimensional example as following: $U_1^{\alpha}(nx') = \mathcal{F}^{\alpha}(u_1(nx))$ which it calculates $\mathcal{F}^{\alpha}(u_1(nx))$ at the fractional rate α. Eq. (3.9) can be further simplified as following,

$$
\begin{aligned}
U_1^{\alpha}(nx') &= \mathcal{F}^{\alpha}(u_1(nx)) \\
&= \sum_{nx=-N/2}^{N/2-1} u_1(nx) \cdot k_N^{-\alpha \cdot nx \cdot nx'} \quad \text{where} \quad k_N = exp\big[-j\frac{2\pi}{N}\big] \\
\textit{since} \quad & \alpha \cdot nx \cdot nx' = \frac{\alpha}{2}nx^2 + \frac{\alpha}{2}nx'^2 - \frac{\alpha}{2}(nx' - nx)^2 \\
&:= \sum_{nx=-N/2}^{N/2-1} u_1(nx) \cdot k_N^{\frac{\alpha}{2}(nx'-nx)^2} \cdot k_N^{-\frac{\alpha}{2}nx^2} \cdot k_N^{-\frac{\alpha}{2}nx'^2} \\
&= k_N^{-\frac{\alpha}{2}nx'^2} \cdot \sum_{nx=-N/2}^{N/2-1} \underbrace{u_1(nx) \cdot k_N^{-\frac{\alpha}{2}nx^2}}_{u_1'(nx)} \cdot \underbrace{k_N^{\frac{\alpha}{2}(nx'-nx)^2}}_{k_N'(nx'-nx)} \\
&= k_N^{-\frac{\alpha}{2}nx'^2} \cdot \sum_{nx=-N/2}^{N/2-1} u_1'(nx) \cdot k_N'(nx' - nx) \qquad (3.11) \\
&= k_N^{-\frac{\alpha}{2}nx'^2} \cdot A(nx') \qquad (3.12)
\end{aligned}
$$

Moreover, Eq. (3.11) can be treated directly as a standard convolution. This means $A(nx')$ in Eq. (3.12) can be calculated from a standard fast Fourier transform algorithm as shown below in Eq. (3.13). **FFT** and **IFFT** below refer to the fast Fourier transform and its inverse. $\bullet$ refers to the point-wise multiplication.

$$A(nx') = \sum_{nx=-N/2}^{N/2-1} u_1'(nx) \cdot k_N'(nx' - nx)$$

$$\textit{define} \quad u_1'(nx) = u_1(nx) \cdot k_N^{-\frac{\alpha}{2}nx^2} \text{ and } k_N'(nx) = k_N^{\frac{\alpha}{2}nx^2}$$

$$:= \mathbf{IFFT}\Big\{\mathbf{FFT}\big\{u_1'(nx)\big\} \bullet \mathbf{FFT}\big\{k_N'(nx)\big\}\Big\} \tag{3.13}$$

In Eq. (3.13), $u_1'(nx)$ is the product between input $u_1(nx)$ and $k_N^{-\frac{\alpha}{2}nx^2}$ which is given from the grid vector $nx = -N/2$ to $N/2-1$. Then, once $A(nx')$ is calculated from the fast Fourier transform algorithm in Eq. (3.13), the last step as shown in Eq. (3.12) is to do a multiplication between $k_N^{-\frac{\alpha}{2}nx'^2}$ and $A(nx')$. This quadratic phase kernel $k_N^{-\frac{\alpha}{2}nx'^2}$ is also a function of the grid vector $nx' = -N/2$ to $N/2-1$.

Thus, a three steps illustration of this scaled Fourier transform solver in a one-dimensional case is shown as below.

$$\begin{aligned} U_1^{\alpha}(nx') &= \mathcal{F}^{\alpha}(u_1(nx)) \\ &= \sum_{nx=-N/2}^{N/2-1} u_1(nx) \cdot k_N^{-\alpha \cdot nx \cdot nx'} \quad \text{where} \quad k_N = exp\big[-j\frac{2\pi}{N}\big] \\ &= k_N^{-\frac{\alpha}{2}nx'^2} \bullet \mathbf{IFFT}\Big\{\mathbf{FFT}\big\{u_1(nx) \bullet k_N^{-\frac{\alpha}{2}nx^2}\big\} \bullet \mathbf{FFT}\big\{k_N^{\frac{\alpha}{2}nx^2}\big\}\Big\} \\ &= k_N^{-\frac{\alpha}{2}nx'^2} \underbrace{\bullet}_{\text{step 3}} \underbrace{\mathbf{IFFT}\Big\{\mathbf{FFT}\big\{\underbrace{u_1(nx) \bullet k_N^{-\frac{\alpha}{2}nx^2}}_{\text{step 1 multiplication}}\big\} \bullet \mathbf{FFT}\big\{k_N^{\frac{\alpha}{2}nx^2}\big\}\Big\}}_{\text{step 2 convolution}} \end{aligned} \tag{3.14}$$

A multi-dimensional scaled Fourier transform solver can be constructed by using the one-dimensional solver shown in Eq. (3.14) multiple times. Since we deal with a wavefront captured on a planar relay wall, we use a two dimensional case as an example in the following.

Eq. (3.15) shows a procedure to evaluate a two-dimensional scaled Fourier

transform using the one-dimensional solver mentioned above. $U^{\alpha}(u,v)$ stands for a scaled Fourier transform of $u(x,y)$ and the fractional rate α is the same for both x and y dimension.

$$
\begin{aligned}
U^{\alpha}(u,v) &= \mathcal{F}^{\alpha}(u_1(x,y)) \\
&= \sum_{y=-N/2}^{N/2-1} \sum_{x=-N/2}^{N/2-1} u(x,y)\cdot k_N^{-\alpha(xu+yv)} \quad \text{where} \quad k_N = exp\big[-j\frac{2\pi}{N}\big] \\
&= \sum_{y=-N/2}^{N/2-1} \sum_{x=-N/2}^{N/2-1} u(x,y)\cdot k_N^{-\alpha xu}\cdot k_N^{-\alpha yv} \\
&= \sum_{y=-N/2}^{N/2-1} \underbrace{\Big\{ \sum_{x=-N/2}^{N/2-1} u(x,y)\cdot k_N^{-\alpha xu} \Big\}}_{\text{Use } \textbf{1d solver} \text{ in Eq. (3.14) on } u(x,y) \text{ along y-dim}} \cdot k_N^{-\alpha yv} \\
&= \underbrace{\sum_{y=-N/2}^{N/2-1} u^{'}(x,y)\cdot k_N^{-\alpha yv}}_{\text{Use } \textbf{1d solver} \text{ in Eq. (3.14) on } u^{'}(x,y) \text{ along x-dim}}
\end{aligned}
\tag{3.15}
$$

RSD kernel solver

A numerical step for the RSD kernel implementation is presented in [1]. Notice that no matter a equal or a fractional sampling scenario, this RSD kernel solver is just a function of spatial sampling δ. Then it can be calculated as follows. Notations follows the same as in Eq.(3.4).

Step 1: Discretize depth

$$\hat{z} = \frac{Z}{\delta}$$

Step 2: Symmetry zero padding and precision control

$$N' = \frac{\lambda Z}{\delta^2}$$
$$pad = \alpha \cdot \frac{N' - N}{2} \quad \text{where } (\alpha \geqslant 1)$$
$$u_1[nx, ny] = padarray\Big(u_1[nx, ny], [pad, pad], 0\Big)$$
$$\text{Update discrete size: } N = N'$$

Step 3: Uncertainty parameter

$$\eta^2 = \frac{\lambda Z}{N\delta^2} = \frac{\lambda \hat{z}}{N\delta}$$

Step 4: Compute Convolution kernel

$$r = \sqrt{nx^2/\hat{z}^2 + ny^2/\hat{z}^2 + 1}$$
$$h[nx, ny] = \frac{exp\big[-j2\pi \cdot \hat{z}^2/(\eta^2 N) \cdot r\big]}{r}$$

Thus $h[nx, ny]$ refers to the final RSD convolutional kernel.

3.4 Appendix: Result

We use publicly available datasets to verify our results [1]. As it is shown in Fig. 3.2, we verify our proposed solvers in two cases with a different choice of fractional sampling rate parameters. The running time for the first row (letter 44i) is 4.5 second, and the seconds row (office scene) is 35 seconds. As it can be seen from the results, as the fractional rate changes in distance, targets are getting smaller. Even though we test our solvers with a set of fractional sampling rate α, this is just a verification for us to test the tolerance of numerical errors. Since there are some works can be found regarding NLOS imaging resolution statement [7, 4, 3, 45], we believe α can be chosen optimally but this is beyond what we aim for this project for

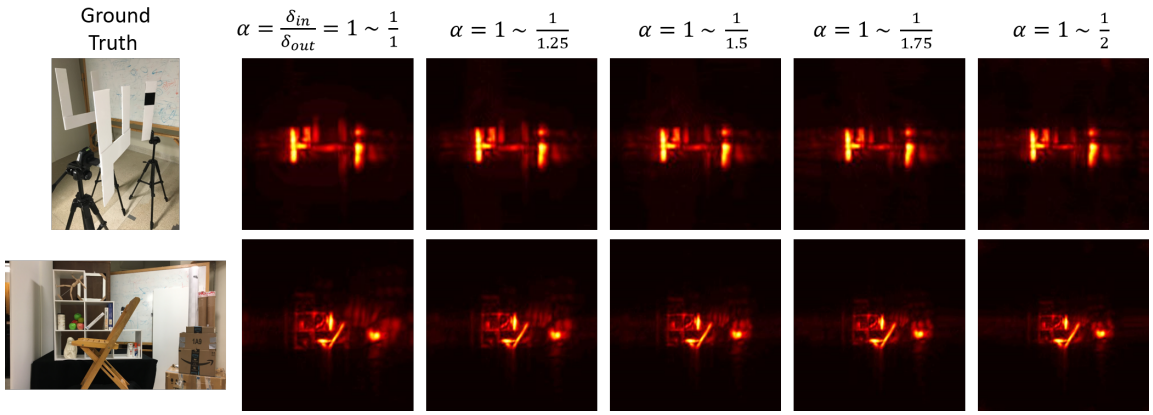

Figure 3.2: We apply our proposed perspective projection solver to existing datasets. The fractional rate $\alpha = \frac{\delta_{in}}{\delta_{out}}$ in Eq. (3.10) is chosen such that it varies in distance. As results shows above, the first column refers to the ground truth images where experiments are taken. Second column α is a constant and not varies in distance. So results in the second column refer to the equal sampling case in Eq. (3.5). The rest columns refers to the fractional sampling case in Eq. (3.7) and solved by our proposed solver in Eq. (3.14). From the third column to the last one, we test our solvers with a different choices of α. In the third column, α varies from 1 to $\frac{1}{1.25}$, and the last one is from 1 to $\frac{1}{2}$. This is just a verification of solvers and α can be chosen optimally as a function of a physical aperture size.

now. Some future works can also focus on simplifying the proposed solver more, or approximating some steps for a faster computation [80, 81].

4 WIGNER DISTRIBUTION DESCRIPTION FOR PHASOR FIELD NON-LINE-OF-SIGHT IMAGING

In classical optics, the Wigner Distribution Function (WDF) [82] and its Fourier transform pair the Ambiguity Function [83] are powerful analytical tools to study optical diffraction, phase space [84], and partially coherent light [85]. An optical imaging process (such as propagation through a lens or free-space propagations) can be interpreted as a simple geometrical transformation in the WDF domain [86, 87, 88]. This conceptual understanding is also useful in light field imaging. Several works draw connections between light fields and the Wigner Distribution Function [89, 90, 91, 92]. In this section, we describe the connections between Wigner Distribution Function and Non-Line-of-Sight imaging within the Phasor field framework [45].

In phasor field NLOS imaging, the Rayleigh-Sommerfeld Diffraction (RSD) model is shown to be a key solution to Non-Line-of-Sight imaging problems [3]. All existing applications of the Wigner Distribution Function are used when Fresnel approximation is valid. However, in NLOS imaging application, the diffraction happens close to the relay wall where only the RSD holds as an exact solution. This RSD also gives an exact solution to the wave propagation as opposed to the approximations such as Fresnel or Fraunhofer diffraction which are commonly known in classical optics [70, 71, 72]. It is shown that the RSD can be used to solve scanning free, real-time, three-dimensional NLOS reconstruction problem [1]. Dove *et al.* [15] present a two dimensional spatial Wigner Distribution Function in a paraxial region with the approximated Fresnel diffraction for NLOS phasor field model.

The RSD with the Wigner Distribution Function has never been discussed in the context of real-world NLOS measurements. In this section, we will study the RSD in the Wigner Distribution Function domain and compare it with the Fresnel diffraction under real-world parameters like finite relay wall size, discrete spatial sampling, and different acquisition schemes such as confocal and non-confocal

measurements. Another angle to describe our work is to use the Wigner Distribution Function to explain Non-Line-of-Sight imaging and clarify when approximations are useful and meaningful in practice.

The first part of this section contains a short review of the Wigner Distribution Function and its description of a linear system in Ch. 4.1. Then using this linear system Wigner Distribution Function description, we show how the Wigner Distribution Function is being used to model and solve problems in Ch. 4.2.

Notation Setup

Throughout this section, we use notations as follows. $\mathcal{F}(\cdot)$ stands for the Fourier transform, $W_f(\cdot,\cdot)$ for the Wigner Distribution Function (WDF) where the footnote f refers to an input function. We use integral(s) to describe a linear operator (for example, diffraction throughout this paper).

For example, to describe a standard linear operator in space or frequency (input, output spatial/frequency representation), we express the linear integral as follows:

$$\begin{aligned} f_o(x_o) &= \int h_{xx}(x_o, x_i) f_i(x_i)\, dx_i \\ F_o(\mu_o) &= \frac{1}{2\pi} \int h_{\mu\mu}(\mu_o, \mu_i) F_i(\mu_i)\, d\mu_i\,, \end{aligned} \tag{4.1}$$

In Eq. (4.1), an input $f_i(x_i)$ and a output $f_o(x_o)$ are denoted by footnotes (the same for their frequency representations $F_i(\mu_i) = \mathcal{F}(f_i(x_i))$, $F_o(\mu_o) = \mathcal{F}(f_o(x_o))$). For a one dimensional linear operator above, the kernel $h_{xx}(x_o, x_i)$ in its primary domain x or $h_{\mu\mu}(\mu_o, \mu_i)$ in its frequency domain μ can be used to describe the relationship between input $f_i(x_i)$ and output $f_o(x_o)$ functions (signals). In later sections, $h_{xx}(x_o, x_i)$ is used to describe a wave propagation which is tied to a physical diffraction process.

We list the most frequently used notations below:

- $\mathcal{F}(\cdot)$: Fourier transform
- $f_i(x_i)/F_i(\mu_i)$: Input spatial/frequency representation

- $f_o(x_o)/F_o(\mu_o)$: Output spatial/frequency representation
- $f^*(x)/F^*(\mu)$: Complex conjugate of spatial/frequency representation
- $h_{xx}(x_o, x_i)$: Linear operator spatial integral kernel
- $h_{xx}(x)$: Linear operator spatial convolution kernel
- $h_{\mu\mu}(\mu_o, \mu_i)$: Linear operator frequency integral kernel
- $\mathcal{W}_{f_i}(x_i, \mu_i)$: Input Wigner Distribution Function for object f_i.
- $\mathcal{W}_{f_o}(x_o, \mu_o)$: Output Wigner Distribution Function for object f_o.
- $\underset{x-y}{*}$: A convolution along both $x - y$ dimension.

4.1 Wigner distribution in classical imaging

Most imaging phenomena can be (approximately) described by the linearity of coherent wave or its intensity and formulated as linear operators. In this section, we review the linear operator in the WDF domain in a formula cookbook fashion [93, 87, 88]. Then we apply this WDF framework to show the RSD and Fresnel diffraction (RSD, Fresnel propagators) in the WDF domain.

What is Wigner distribution

To describe a physical object, its spatial $f(x)$ (x refers to the spatial coordinate) and spatial frequency $F(\mu)$ signal representations can be converted through the Fourier transform. For example, $f(x)$ can be a image on the x coordinate and $F(\mu)$ refers to its Fourier transform. In a word, a standard way to present this object is either in the space **or** in the spatial frequency domain. However, WDF $\mathcal{W}_f(x, \mu)$ gives us both space **and** spatial frequency representation for this object which is different from the Fourier transform.

The WDF $\mathcal{W}_f(x, \mu)$ of this object f can be calculated through its $f(x)$ or $F(\mu)$ is given below:

$$\mathcal{W}_f(x,\mu) = \int_{-\infty}^{+\infty} \underbrace{f(x+\frac{\tau}{2})\, f^*(x-\frac{\tau}{2})}_{\text{Spatial representation}}\, e^{-j2\pi\mu\tau} d\tau$$
$$= \int_{-\infty}^{+\infty} \underbrace{F(\mu+\frac{\xi}{2})\, F^*(\mu-\frac{\xi}{2})}_{\text{Spatial frequency representation}}\, e^{j2\pi x\xi} d\xi\,, \tag{4.2}$$

$\mathcal{W}_f(x,\mu)$ in Eq. (4.2) is a function of both space x and spatial frequency μ for a one dimensional object f. Overall, the Wigner Distribution Function representation always doubles the dimension for the notation of an object (1d signals have 2d WDF, 2d signals have 4d WDF).

Linear operators in the Wigner distribution domain

Given an integral expression in space or frequency for a linear operator, the skeleton for this linear operator in the WDF domain can be derived immediately [93, 87, 88]. For example, a linear system can be described as a linear operator to describe wave propagation in the context of wave optics. For simplicity, we use a one-dimensional linear operator as an example in this section.

The linear system WDF description aims to build a relationship between an input WDF $\mathcal{W}_{f_i}(x_i,\mu_i)$ and a output WDF $\mathcal{W}_{f_o}(x_o,\mu_o)$ by using a four dimensional kernel $K(x_o,\mu_o,x_i,\mu_i)$. Intuitively speaking, it tries to capture how a transformation effects in both space and spatial frequency domains.

$$\mathcal{W}_{f_o}(x_o,\mu_o) = \frac{1}{2\pi}\iint K(x_o,\mu_o,x_i,\mu_i)\mathcal{W}_{f_i}(x_i,\mu_i)\, dx_i d\mu_i\,, \tag{4.3}$$

$K(x_o,\mu_o,x_i,\mu_i)$ in Eq. (4.3) links an input object $\mathcal{W}_{f_i}(x_i,\mu_i)$ and a output object $\mathcal{W}_{f_o}(x_o,\mu_o)$ in the WDF domain. Once we know a spatial $h_{xx}(x_o,x_i)$ or a frequency description $h_{\mu\mu}(\mu_o,\mu_i)$ (notations are stated in Eq. (4.1)) for a linear system, we can derive the associated $K(x_o,\mu_o,x_i,\mu_i)$ in Eq. (4.3) as follows,

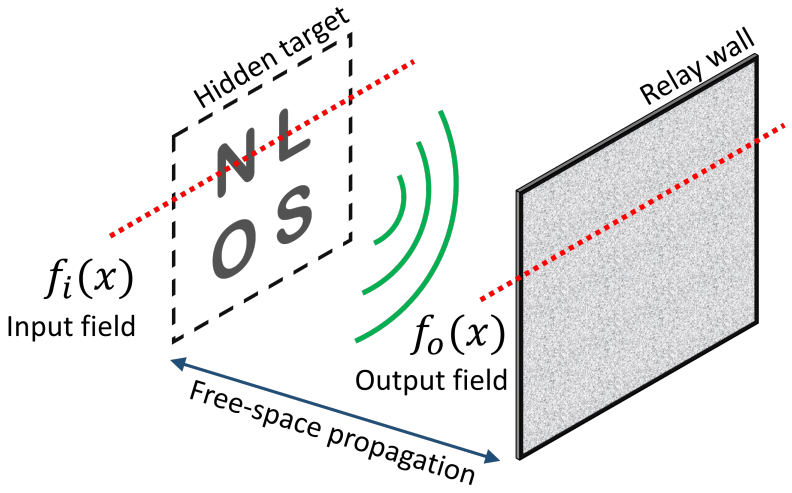

Figure 4.1: Two parallel planes (lines) setup geometry in Sec. 4.1. $f_i(x)$ and $f_o(x)$ represent line slices of the field used in Eq. (4.5).

$$
\begin{aligned}
&K(x_o, \mu_o, x_i, \mu_i) = \\
&:= \iint h_{xx}(x_o + \frac{x_o^{'}}{2}, x_i + \frac{x_i^{'}}{2})\, h_{xx}^{*}(x_o - \frac{x_o^{'}}{2}, x_i - \frac{x_i^{'}}{2})\, \exp\left[-j\mu_o x_o^{'} + j\mu_i x_i^{'}\right] dx_o^{'} dx_i^{'} \\
&:= (\frac{1}{2\pi})^2 \iint h_{\mu\mu}(\mu_o + \frac{\mu_o^{'}}{2}, \mu_i + \frac{\mu_i^{'}}{2})\, h_{\mu\mu}^{*}(\mu_o - \frac{\mu_o^{'}}{2}, \mu_i - \frac{\mu_i^{'}}{2})\, \exp\left[j\mu_o^{'} x_o - j\mu_i^{'} x_i\right] d\mu_o^{'} d\mu_i^{'},
\end{aligned}
\tag{4.4}
$$

Eq. (4.3) and Eq. (4.4) can be adapted to any linear operators with $h_{xx}(x_o, x_i)$ (or $h_{\mu\mu}(\mu_o, \mu_i)$) in analytical forms. Notice that $K(x_o, \mu_o, x_i, \mu_i)$ is completely described the linear physical process. Constraints on variables (x_o, μ_o, x_i, μ_i) can be made to model this linear physical process even more (such as an energy constraint, a frequency bandwidth, and a spatial truncation). Some special linear systems directly have closed-form expressions without deriving by definitions from Eq. (4.3) and Eq. (4.4).

Diffraction in the Wigner distribution domain

In this section, we describe the RSD and Fresnel diffraction in the WDF domain by using formulas in Sec.4.1. These two diffraction propagators are related to the NLOS imaging problem.

Before deriving the RSD and Fresnel diffraction in the WDF domain, we need to introduce a notation setup for this section. For simplicity, in the one-dimensional Cartesian coordinate system, we consider two parallel lines with a spacing z (distance between lines). The geometrical setup is shown in Fig. 4.1. Spatial representations for $h_{xx}^{RSD}(x_o, x_i)$ (RSD) and $h_{xx}^{Fre}(x_o, x_i)$ (Fresnel) are used to describe the propagation from an input field $f_i(x_i)$ to a output field $f_o(x_o)$. In order to derive the RSD and Fresnel diffraction in the WDF domain, the first step is to write down $h_{xx}(x_o, x_i)$ in an analytical form, then plugging $h_{xx}(x_o, x_i)$ into Eq. (4.3) and Eq. (4.4) leads to WDF descriptions for the RSD and Fresnel diffraction.

The first step is to review a standard way of describing the RSD and Fresnel diffraction in the space domain. Both the RSD and the Fresnel diffraction in this geometrical parallel plane setup case, can be treated as a spatial convolution [77, 78, 71]. Thus, it reduces one variable $x = x_o - x_i$ for the kernel from $h_{xx}(x_o, x_i)$ to $h_{xx}(x)$ in Eq. (4.5). We use notation $\underset{x}{*}$ to represent the convolution along x dimension.

$$f_o(x_o) = f_i(x_i) * h_{xx}(x_o, x_i) := f_i(x_i) \underset{x}{*} h_{xx}(x)\,, \tag{4.5}$$

For the RSD, $h_{xx}(x)$ refers to:

$$h_{xx}(x) = h_{xx}^{RSD}(x) = \frac{e^{jk\sqrt{x^2+z^2}}}{\sqrt{x^2+z^2}} \tag{4.6}$$

For the Fresnel diffraction, $h_{xx}(x)$ refers to:

$$h_{xx}(x) = h_{xx}^{Fre}(x) = \frac{e^{jkz}}{j\lambda z} e^{\frac{jk}{2z}x^2} = \alpha(z) e^{\frac{jk}{2z}x^2}\,, \tag{4.7}$$

$k = \omega/c$ in both cases stand for the wavenumber of a monochromatic wave, which

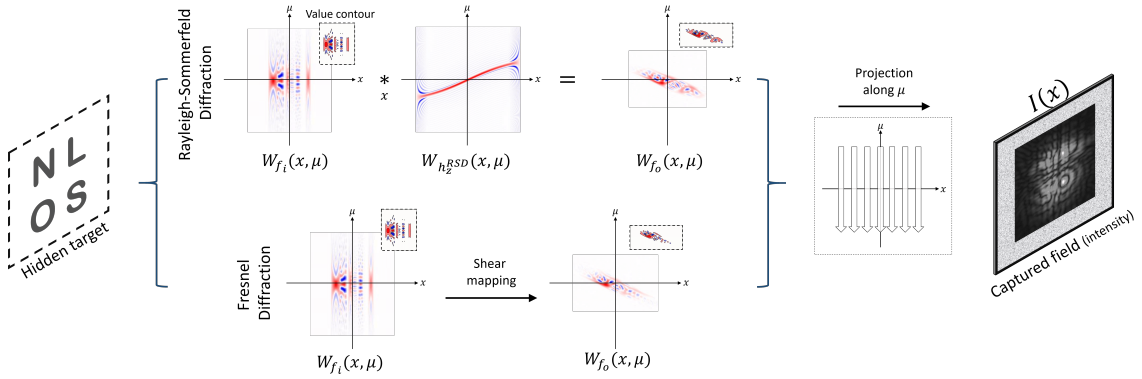

Figure 4.2: The RSD and Fresnel diffraction in the WDF domain Eq. (4.10). $\mathcal{W}_{f_i}(x,\mu)$ and $\mathcal{W}_{f_o}(x,\mu)$ are WDF for the input field $f_i(x)$ and the output field $f_o(x)$ in Fig. 4.1. The RSD in the WDF domain is shown in the first row and the Fresnel is in the second row. Both propagators starts from a same target field WDF $\mathcal{W}_{f_i}(x,\mu)$, the differences lie in transformations in the WDF domain. The RSD refers to a convolution along spatial coordinate x with WDF of the RSD kernel $\mathcal{W}_{h_z^{RSD}}(x_i,\mu)$ in Eq. (4.8). The Fresnel diffraction refers to a shear mapping in Eq. (4.9). We also plot the corresponding value contours (level) for each WDF plot which are shown in dash windows.

ω refers to the angular frequency and c means the speed of light travelling in air. z refers to the propagation distance (The spacing between an input and a output plane).

The next step is to show the RSD and the Fresnel diffraction in the WDF domain. Given a spatial kernel description $h_{xx}(x_o,x_i)$ either from the RSD or the Fresnel diffraction, we can plug Eq. (4.6 & 4.7) into Eq. (4.4) to calculate the corresponding kernel $K(x_o,\mu_o,x_i,\mu_i)$. Then, we use Eq. (4.3) to link input WDF $\mathcal{W}_{f_i}(x_i,\mu)$ with output WDF $\mathcal{W}_{f_o}(x_o,\mu)$ by $K(x_o,\mu_o,x_i,\mu_i)$. Thus, we achieve WDF descriptions for the RSD and the Fresnel diffraction. We skip algebraic steps here, calculations are provided in Sec. 4.3

The RSD and the Fresnel diffraction in the WDF domain are given below. $\mathcal{W}_{f_i}(x_i,\mu)$ and $\mathcal{W}_{f_o}(x_o,\mu)$ stand for the WDF of an input and a output wavefront.

- **RSD in the WDF domain** in Eq. (4.8) refers to a convolution along the spatial

x direction. $\mathcal{W}_{h_z^{RSD}}(x_i, \mu)$ stands for the WDF of the RSD convolution kernel $h_z^{RSD}(x) = \frac{e^{jk\sqrt{x^2+z^2}}}{\sqrt{x^2+z^2}}$ where a footnote z denotes a propagating distance.

$$\begin{aligned}\mathcal{W}_{f_o}(x_o, \mu) &= \mathcal{W}_{h_z^{RSD}}(x_i, \mu) \underset{x}{*} \mathcal{W}_{f_i}(x_i, \mu) \\ &:= \int \mathcal{W}_{h_z^{RSD}}(x_o - x_i, \mu)\mathcal{W}_{f_i}(x_i, \mu)\, dx_i\,, \end{aligned} \tag{4.8}$$

- **Fresnel diffraction in the WDF domain** in Eq. (4.9) refers to a shear mapping (coordinate transformation) as a function of a propagating distance z.

$$\mathcal{W}_{f_o}(x, \mu) = \mathcal{W}_{f_i}(x - \frac{z}{k}\mu, \mu)\,, \tag{4.9}$$

Then, an output intensity $I_o(x_o) = |f_o(x_o)|^2$ of the wavefront $f_o(x_o)$ can be calculated from the marginal distribution of the output WDF $\mathcal{W}_{f_o}(x_o, \mu)$ (projection along frequency coordinate μ),

$$\begin{aligned} I_o(x_o) &= |f_o(x_o)|^2 \\ &= \int \underbrace{\mathcal{W}_{f_o}(x_o, \mu)}_{\text{Eq.(4.8 or 4.9)}}\, d\mu \quad \text{- projection along } \mu \\ I_o^{RSD}(x) &:= \iint \mathcal{W}_{h_z^{RSD}}(x - x_i, \mu)\mathcal{W}_{f_i}(x_i, \mu)\, dx_i\, d\mu \\ I_o^{Fre}(x) &:= \int \mathcal{W}_{f_i}(x - \frac{z}{k}\mu, \mu)\, d\mu\,, \end{aligned} \tag{4.10}$$

Fig. 4.2 illustrates calculation steps in Eq. (4.10). Propagation using the Fresnel diffraction results in shearing of the WDF. Propagation using the exact RSD propagator, however, does not have a simple geometrical interpretation in the WDF domain. In the next section, more details are discussed in the context of NLOS imaging.

4.2 Wigner distribution in Non-Line-of-Sight imaging

In this section, we discuss NLOS imaging within the phasor field virtual wave optics and the WDF framework. To understand how the NLOS imaging problem is related to wave optics, we review the phasor field method [3, 14, 15, 16, 17, 1]. Then combining the phasor field framework with WDF descriptions, we derive a spatial lateral resolution limit using the exact RSD solution. We explore differences between confocal and non-confocal measurements, and errors from the Fresnel approximation. Notice that, previous chapter Ch. 2 presents an extensive discussion regarding to the Phasor field model. This short review in this section is only for the sake of clarifications. Readers with enough background in the Phasor field model are encouraged to skip this short review section.

Phasor field model review

We need to introduce some additional variables to illustrate captured NLOS signals. $g(x, y, t)$ represents a captured time response [a], coordinate (x, y) refers to a spatial location of a detector pixel on a relay wall, t refers to a time index. We assume all time responses $g(x, y, t)$ are captured from a plane relay wall.

First, the phasor field $p(x, y)$ is defined to be a single frequency component of $G(x, y, \omega)$ which $G(x, y, \omega)$ stands for the temporal Fourier transform of the captured time response $g(x, y, t)$. In Eq. (4.11), $p(x, y)$ can be calculated through the Fourier transform of a convolution in time with a temporal harmonic function $e^{j\omega t}$, or product with a shifted delta $\delta(\xi - \omega)$ in the Fourier domain,

$$\begin{aligned} p(x, y) &= \mathcal{F}(g(x, y, t) \underset{t}{*} e^{j\omega t}) \\ &= G(x, y, \xi) \cdot \delta(\xi - \omega) \end{aligned} \tag{4.11}$$

[a] Across the entire paper, captured time responses (signal) refer to shifted version of raw temporal measurements. This shifting process could be done during the acquisition by calculating a line-of-sight time delay respect to a distance between physical hardware and focused points on the relay wall. More descriptions please refer to [7].

The angular frequency variable ω and its associated wavelength λ are connected through the wavenumber $k = \omega/c = 2\pi/\lambda$. Overall, the phasor field $p(x, y)$ is defined to be a single frequency content of each captured time response $g(x, y, t)$ in the Fourier domain.

Then, a NLOS imaging process can be understood as follows: An unknown phasor field $p_i(x, y)$ (input) carrying the object's information propagates to a relay wall, where the phasor field $p_o(x, y)$ is captured (output). The goal for reconstructions is to invert this diffraction process from the captured field $p_o(x, y)$ to have a virtual image representation which ideally is the same as $p_i(x, y)$. This diffraction process from $p_i(x, y)$ to $p_o(x, y)$ can be modeled as the RSD propagator which is shown in [3]. $\alpha(x, y)$ refers to an additional amplitude correction factor.

$$\begin{aligned} p_o(x, y) &= \alpha(x, y)\Big(p_i(x, y) \underset{x-y}{*} h_{xx}^{RSD}(x, y, z)\Big) \\ &\propto p_i(x, y) \underset{x-y}{*} h_{xx}^{RSD}(x, y, z) \end{aligned} \tag{4.12}$$

$h_{xx}^{RSD}(x, y, z)$ in Eq.(4.12) refers to the RSD convolution kernel in two dimensional case as following (one dimension in Eq. (4.6)),

$$h_{xx}^{RSD}(x, y, z) = \frac{e^{jk\sqrt{x^2+y^2+z^2}}}{\sqrt{x^2 + y^2 + z^2}} \tag{4.13}$$

Next, since Eq. (4.12) describes the phasor field diffraction at each frequency ω, by using Eq. (4.11) with different choices of ω, $p_i(x, y, \omega)$ has another frequency dimension. Then, one can extend the model into a space-time broadband propagator [1] for describing the captured space-time signals as following,

$$p_o(x, y, t) = \int_{\omega\in\Omega} e^{j\omega t}\Big(\underbrace{p_i(x, y, \omega) \underset{x-y}{*} h_{xx}(x, y, z)}_{\text{Diffraction function at } \omega}\Big)\, d\omega \tag{4.14}$$

Here are some additional explanations for Eq.(4.14):

- Eq. (4.14) can be used to describe a propagation either from a hidden target to captured signals (forward propagation), or from captured signals to a virtual image (reconstruction).

- $\omega \in \Omega$ stands for a closed frequency interval that is chosen to match the captured system spatial and temporal resolution (resolvable wavefronts). A NLOS picosecond time-resolved system usually has 60 to 70 picosecond temporal resolution, but phasor field wavefronts are limited by the spatial sampling resolution on a relay wall. Given a discrete spatial sampling grid with a spacing Δ on the relay wall (for example 1 cm), phasor field wave components need to be satisfied the half wavelength condition $\lambda_\omega > \Delta/2, \omega = 2\pi c/\lambda_\omega$ [94].

- In Eq.(4.11), one can also use a temporal illumination convolution kernel function $e^{j\omega_c t}e^{-\frac{t^2}{2\sigma^2}}$ (Gaussian-modulated sinusoidal pulse) to generate the phaosr field $p_i(x, y, \omega) = \mathcal{F}(g(x, y, t) \underset{t}{*} e^{j\omega_c t}e^{-\frac{t^2}{2\sigma^2}})$, then uses Eq.(4.14) for reconstruction. This Gaussian-modulated sinusoidal pulse models an object that reflects a temporally changing phasor field wavefront which is shown in [3].

- The benefit of using Eq. (4.14) is that we can study the space-time NLOS signals by decomposing them into diffraction processes at each individual frequency. For example, the diffraction inside Eq. (4.14), one can approximate $h_{xx}(x, y, z)$ in Eq. (4.14) by the Fresnel propagator $h_{xx}^{Fre}(x, y, z)$ instead of the RSD kernel $h_{xx}^{RSD}(x, y, z)$. So that each frequency component has a geometrical shear mapping transformation in the WDF domain as shown in Eq. (4.10).

Resolution analysis

To understand the achievable lateral resolution in NLOS reconstructions, one have to understand the central frequency for the phasor field. Here we provide an example of the phasor field central frequency in a reconstruction pipeline. As

we discussed ealier in the previous section for Eq.(4.14), considering the phasor field coming from a gaussian-modulated sinusoidal pulse $e^{j\omega_c t}e^{-\frac{t^2}{2\sigma^2}}$, ω_c defines the central frequency for the captured phasor field. Lateral resolution is bounded by the diffraction limits at the central frequency ω_c without exploiting optical occlusions in the hidden scene [15]. Next, we want to show the lateral resolution at a central frequency when using the exact RSD propagator given a finite size relay wall.

First, we model a limited size relay wall using an aperture function $T(x, y)$. Generally speaking, this aperture function can be modeled as real or complex functions which is used in optical coded imaging. However, to derive lateral resolution limits, we consider the aperture function as an binary function $T(x, y) \in \{0, 1\}$ in Eq. (4.15). N stands for the aperture half side length (for example $N = 1$ as a 2 m by 2 m scanning wall).

$$T(x, y) = \begin{cases} 1, & \text{if } |x| \leqslant N, |y| \leqslant N \\ 0, & \text{otherwise} \end{cases} \tag{4.15}$$

Second, the achievable lateral resolution can be characterized by the point spread function $PSF(x_t, y_t, z, \omega)$ from a point object in the hidden scene at (x_t, y_t, z) shown in Eq. (4.16). $h(x - x_t, y - y_t, z)$ stands for the scattered phasor field wavefront from a point object with distance z away from the aperture. The multiplication between $T(x, y)$ and $h(x - x_t, y - y_t, z)$ stands for the captured phasor field on the relay wall. $h^*(x, y, z)$ stands for the complex conjugate of the RSD propagation convolution kernel $h(x, y, z) = h_{xx}^{RSD}(x, y, z)$ shown in Eq. (4.13).

$$\begin{aligned} &PSF(x_t, y_t, z, \omega) \\ &= \left| \Big(\underbrace{\overbrace{T(x, y)}^{\text{Aperture function}} \quad \overbrace{h(x - x_t, y - y_t, z)}^{\text{Point object wavefront}}}_{\text{Captured wavefront}} \Big) \underset{x-y}{*} \underbrace{h^*(x, y, z)}_{\text{RSD propagation kernel}} \right|^2 \end{aligned} \tag{4.16}$$

Eq. (4.16) can be used to calculate the lateral resolution limit as a function of

an aperture function $T(x, y)$, the central frequency $\omega_c = 2\pi c/\lambda_c$, the hidden point object position (x_t, y_t, z). The choice of central wavelength λ_c depends on the spatial sampling spacing Δ which is $\lambda_c > \alpha \cdot 2\Delta$ for $\alpha > 1$ in the discussion for Eq. (4.14) and a system's temporal resolution. Decreasing spatial sampling spacing Δ for a fixed aperture function $T(x, y)$ would lead more spatial sampling points N, but lower the achievable central wavelength λ_c which leads a higher lateral resolution. This central wavelength λ_c is chosen to be at the scale of $4\,\text{cm} \sim 6\,\text{cm}$ with $\Delta = 1\,\text{cm}$ in the previous phasor field experiments [3]. Fig. 4.3 shows the reconstructed image of multiple point targets that lie in different lateral positions. Fig. 4.3 plots $\text{PSF}(x_t, y_t, z, \omega)$ in Eq. (4.16) by using a central frequency $\omega_c = 2\pi c/(\lambda_c = 4\,\text{cm})$ with different target depth z settings. We can also study the lateral resolution in the frequency space by applying the Fourier transform to the phasor field $\mathcal{F}\big((T(x, y)h(x - x_t, y - y_t, z)) \underset{x-y}{*} h^*(x, y, z)\big)$ as shown in Fig. 4.3.

In the WDF domain, when using the Fresnel approximation instead of the RSD to model diffraction, the resolution loss has a more straightforward geometrical explanation. Given an input field WDF $\mathcal{W}_{p_i}(x, \mu)$ and an aperture WDF $\mathcal{W}_T(x, \mu)$, which $T(x) = \text{rect}[x] = 1\,(|x| < 1/2), 0\,(\text{otherwise})$ is a one dimensional version of Eq. (4.15), then the output field $\mathcal{W}_{p_o}(x, \mu)$ in the WDF domain is as below,

$$
\begin{aligned}
\mathcal{W}_{p_o}(x, \mu) &= \underbrace{\mathcal{W}_T(x, \mu)}_{\text{WDF of } T(x)} \underset{\mu}{*} \mathcal{W}_{p_i}(x - \frac{z}{k}\mu, \mu) \\
&= \Big\{ \overbrace{\underbrace{2(1 - |2x|)\,\text{rect}[x]}_{\text{spatial truncation}}\ \text{sinc}[2(1 - |2x|\mu)]}^{\mathcal{W}_T(x,\mu)} \Big\} \\
&\quad \underset{\mu}{*}\ \mathcal{W}_{p_i}(x - \frac{z}{k}\mu, \mu)
\end{aligned}
\tag{4.17}
$$

Eq. (4.17) uses the WDF Multiplication theorem: the multiplication of $T(x)$ and $p_i(x)$ in the space domain corresponds to the convolution in the Fourier domain which is corresponding to a convolution along the frequency coordinate $\underset{\mu}{*}$ in the WDF domain. The loss of resolution comes from WDF of the aperture function

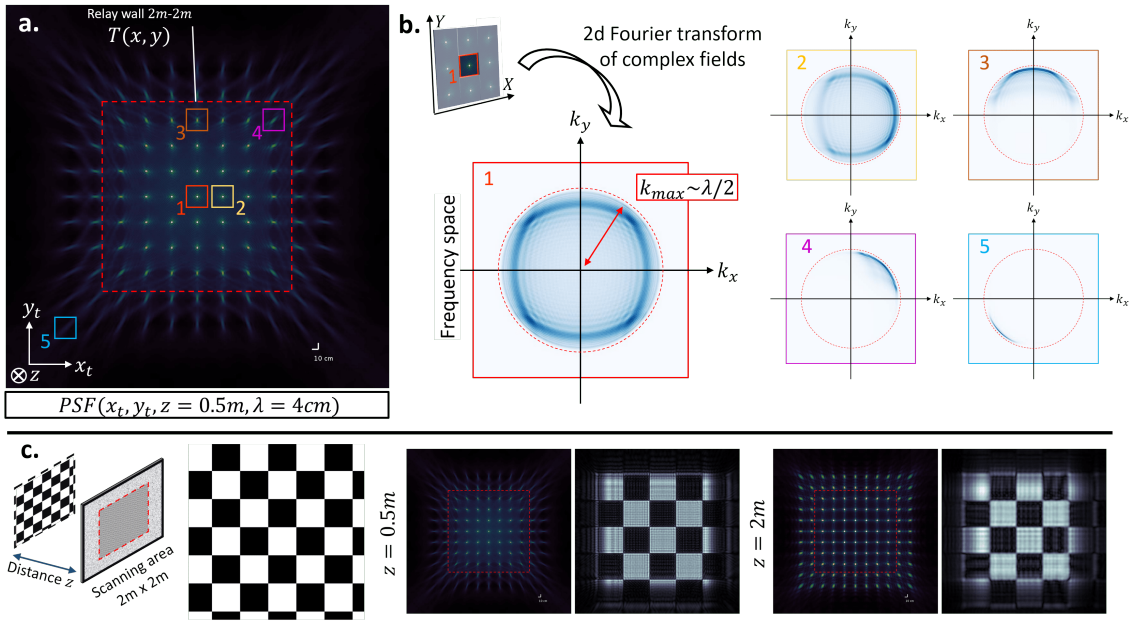

Figure 4.3: Achievable lateral resolution from RSD discussed in Sec. 4.2 Eq. (4.16). **a.** we show the point spread function $\mathrm{PSF}(x_t, y_t, z, \omega)$ from multiple point targets in the hidden with $T(x, y)$ (2 m by 2 m, red dash box), a target depth $z = 0.5$ m away from a relay wall, a central wavelength $\lambda = 4$ cm. Point spread function varies at each lateral location. **b.** we pick five positions (color boxes from 1-5 in **b**) from **a** to illustrate the frequency bandwidth (2d Fourier transform on the complex field). Point position at the center of aperture (number 1, red box) achieves almost maximum bandwidth corresponding to $\lambda/2$. The further away from the center of the aperture, the worse distortion, and the smaller region is covered in the frequency domain. **c.** we show a PSF plot and a reconstructed checkerboard pattern for two depth $z = 0.5$ m - 2 m.

$\mathcal{W}_T(x, \mu)$ in Eq. (4.17). Since $\mathcal{W}_T(x, \mu)$ has a truncation term $2(1 - |2x|)\,\mathrm{rect}[x]$ along the spatial x dimension. Multiplication with $\mathrm{rect}[x]$ would result in zero everywhere in the WDF for $x > |\frac{1}{2}|$. For a fixed size scanning aperture, the output field shears more in the WDF domain as the distance z is increasing which causes more loss of information in the frequency domain for the output field WDF $\mathcal{W}_{p_o}(x, \mu)$.

Confocal and Non-confocal measurements analysis

The confocal NLOS measurement requires co-locating a SPAD and a laser focused point on a relay wall and sequentially scanning the co-located focused points to measure time responses [4]. Otherwise, general measurement setups are referred to as non-confocal measurements [7]. In this section, we add an illumination wavefront function into the phasor field forward diffraction model, which characterize the differences between the confocal and the non-confocal measurement. We model at the central wavelength λ_c (with angular frequency $\omega_c = 2\pi c/\lambda_c$), but the analysis can be applied to each frequency component.

Here are notations would be used in the following. The geometrical setup is shown in Fig. 4.4 on the left side. We consider a hidden object $f(x, y)$ as an amplitude object at a depth z away from a relay wall (on the $x - y$ plane at $z = 0$). For a non-confocal measurement, we have to assign variables (x_i, y_i) to represent a single illumination point source on the relay wall. Collected phasor field wavefronts at the central frequency are denoted by $p_o^{con}(x, y, \omega_c)$ for confocal and $p_o^{n\text{-}con}(x, y, \omega_c)$ for non-confocal. The aperture function $T(x, y)$ follows the same notation as in the previous section.

We modify the phasor field forward diffraction model to include two steps [b]:
1. Propagation from a virtual illumination on the relay wall to a target plane.
2. Propagation from the target plane back to the relay wall. For the confocal configuration, the illumination aperture is as the same size as the aperture function. Thus the illumination field can be modeled as a wavefront starting from $T(x, y)$. For the non-confocal configuration, the illumination field is a spherical wavefront starting from a point illumination source at (x_i, y_i) on $x - y$ plane.

For the first step, Eq. (4.18) calculates the illumination wavefront $u(x, y)$ for confocal $u^{con}(x, y)$ and non-confocal $u^{n\text{-}con}(x, y)$ at the target plane z. $h(x, y, z) = h_{xx}^{RSD}(x, y, z)$ refers to Eq. (4.13),

[b] As for the confocal measurement, two-way propagation (from illumination to object and object to relay wall) can also be modeled as one-way propagation by thinking the object emits light at the same time but traveling at the half-speed [2].

$$u^{\text{con}}(x,y) = \left(T(x,y)e^{j\phi(x,y)}\right) \underset{x-y}{*} h(x,y,z)$$
$$u^{\text{n-con}}(x,y) = h(x-x_i, y-y_i, z) \tag{4.18}$$

For the second step, the received phasor field wavefront $p_o(x,y,\omega_c)$ from the hidden object $f(x,y)$ is as follows,

$$\begin{aligned} p_o(x,y,\omega_c) &= \iint \underbrace{\overbrace{u(x,y)}^{\text{Illumination}} \overbrace{f(x,y)}^{\text{hidden object}}}_{\text{Target wavefront } p_i(x,y,\omega_c)} h^*(x-\nu, y-\mu, z)\, d\nu\, d\mu \\ &= p_i(x,y,\omega_c) \underset{x-y}{*} h^*(x,y,z) \end{aligned} \tag{4.19}$$

For the same hidden object $f(x,y)$, confocal and non-confocal measurements *"see"* different unknown target wavefronts $p_i(x,y,\omega_c)$ because of the illumination wavefront $u(x,y)$ as shown in Eq. (4.19). More importantly, Eq. (4.19) also refers that one can probe hidden object's $f(x,y)$ different frequency components by creating illuminating $u(x,y)$ from different spatial points on the relay wall. The reason is as follows. $p_i(x,y,\omega_c) = u(x,y)f(x,y)$ comes from the hidden object $f(x,y)$ with a spatial modulation from the illumination wavefront $u(x,y)$. $u(x,y)$ is different between $u^{\text{con}}(x,y)$ in confocal measurements and $u^{\text{n-con}}(x,y)$ in non-confocal measurements in Eq. (4.18). Even for non-confocal measurements from different single illumination points (x_i, y_i), the unknown target field $p_i(x,y,\omega_c)$ would be different. Based on the Fourier transform multiplication and convolution properties, multiplication in the space $x-y$ domain corresponding to a convolution in the spatial frequency space. Different illumination wavefronts result in different frequency convolution samples on the hidden object $f(x,y)$.

Fig. 4.4 uses Eq. (4.19) to show differences between the non-confocal and the confocal measurement in the WDF domain. In Fig. 4.4, $|\mathcal{W}_{p_o}^{\text{RSD}} - \mathcal{W}_{p_o}^{\text{Fre}}|(x,\mu)$ shows error maps between the RSD and Fresnel diffraction in the WDF domain. The

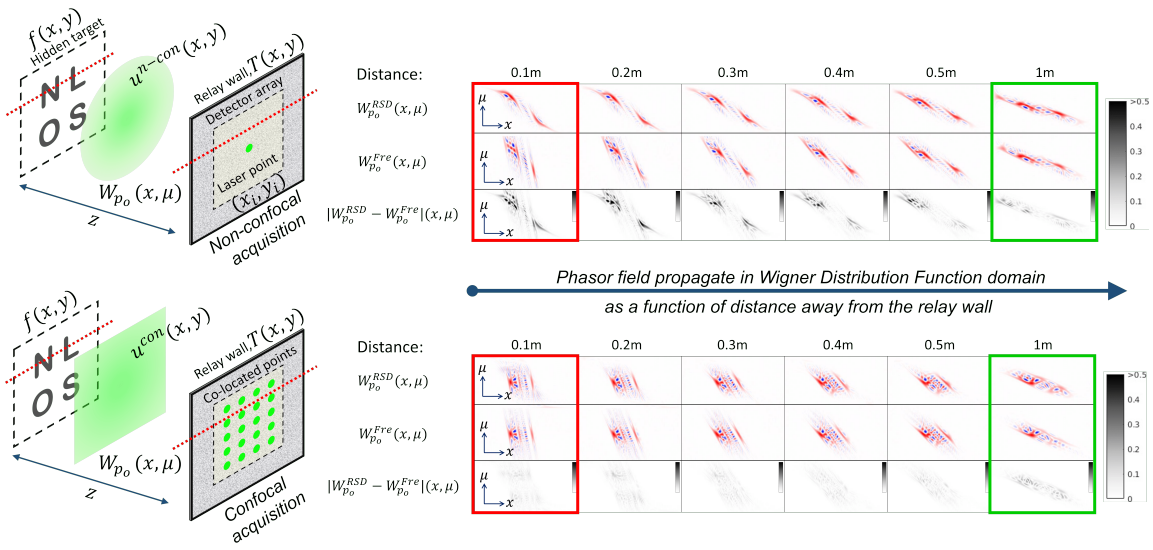

Figure 4.4: Difference between the RSD and the Fresnel propagation in the WDF domain with Non-confocal, confocal acquisitionsin Eq. (4.18 & 4.19) Sec. 4.2. This numerical simulation use the same aperture function $T(x, y)$ and a input hidden target $f(x, y)$ for both non-confocal and confocal acquisitions. The non-confocal single illumination point (x_i, y_i) is at the center. Illumination functions $u^{n\text{-}con}(x, y)$ and $u^{con}(x, y)$ are shown in Eq. (4.18). Each row shows phasor field distributions in the WDF domain as a function of distance z. $\mathcal{W}^{RSD}_{p_o}(x, \mu)$, $\mathcal{W}^{Fre}_{p_o}(x, \mu)$ refers to the phasor field WDF distribution from the RSD or the Fresnel diffraction. For each depth, we plot the absolute difference between normalized WDF for RSD and Fresnel $|\mathcal{W}^{RSD}_{p_o} - \mathcal{W}^{Fre}_{p_o}|(x, \mu)$ (Normalized WDF's value between 0-1). The Fresnel approximation for the non-confocal and the confocal show different errors by the absolute difference map in the WDF domain (from red to green box).

Fresnel diffraction (shear mapping transform in the WDF domain) works as a better approximation for the confocal acquisition than the non-confocal acquisition. A confocal measurement contains more frequency components of the hidden target than a single illumination non-confocal measurement. This means that reconstructions from confocal data should always look *"sharper"* than reconstructions from single illumination non-confocal data even under the same lab condition. With multiple illumination points, non-confocal measurements can increase frequency

components of the hidden target.

Error metric for the Fresnel diffraction

Using Fresnel diffraction for reconstructions can be understood as choosing a poor lens with "aberrations" as opposed to use RSD to model a perfect imaging system. The formula to describe the errors made by the Fresnel diffraction can be obtained by replacing the kernel $h^*(x, y, z)$ in Eq. (4.16) by the Fresnel propagation kernel $h_{\text{Fre}}(x, y, z) = \alpha(z) e^{\frac{jk}{2z}(x^2+y^2)}$. This focusing error $E(x_t, y_t, z_t, x, y, z)$ is as follows,

$$
\begin{aligned}
&E(x_t, y_t, z_t, x, y, z) \\
&= \iint \overbrace{T(x, y)}^{\text{Aperture function}} \; \overbrace{h(x - x_t, y - y_t, z_t)}^{\text{Point object wavefront}} \; \overbrace{u(x_t, y_t, z_t)}^{\text{Illumination wavefront}} \; \underbrace{h^*_{\text{Fre}}(x - \nu, y - \mu, z)}_{\text{Fresnel propagation kernel}} \, d\nu \, d\mu
\end{aligned}
\tag{4.20}
$$

As shown in Eq. (4.20), the focusing error $E(x_t, y_t, z_t, x, y, z)$ is a six dimensional function. The first three arguments (x_t, y_t, z_t) are from a location of hidden point object and the remaining (x, y, z) are from the Fresnel propagator $h_{\text{Fre}}(x, y, z)$. Since $E(x_t, y_t, z_t, x, y, z)$ depends on the illumination wavefront $u(x_t, y_t, z_t)$ at the location of hidden target shown in Eq. (4.18), this error in reconstructions made by the Fresnel diffraction for confocal measurements and non-confocal measurements are different. The behavior of Fresnel diffraction operator varies depending on acquisitions is also provided in Fig. 4.4 in the WDF domain.

Overall, this error $E(x_t, y_t, z_t, x, y, z) \in \mathbb{C}$ leads to both magnitude and phase error in the reconstruction domain. We give an illustration of this error in magnitude in Fig. 4.5 using the illumination function $u^{\text{con}}(x, y)$. One can use Eq. (4.20) to evaluate more general situations with different error metrics depending on desired applications.

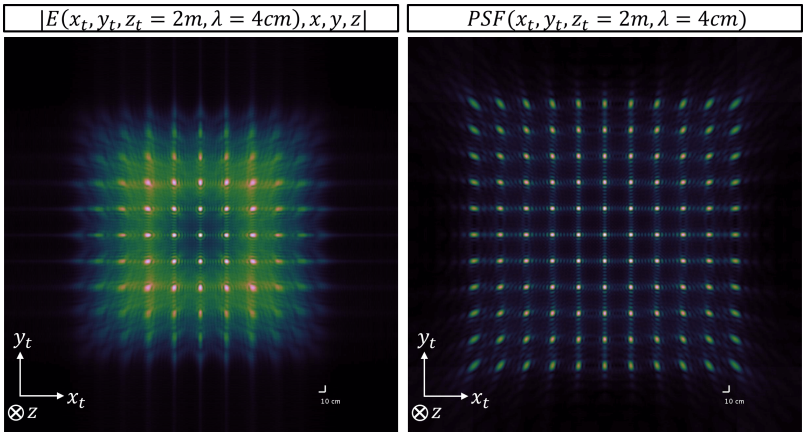

Figure 4.5: Error plot for Eq. (4.20). $|E(x_t, y_t, z_t, x, y, z)|$ refers to complex error field magnitude. $PSF(x_t, y_t, z = 2m, \lambda = 4cm)$ stands for the ideal PSF plot from RSD propagator for referencing.

4.3 Appendix: Algebraic steps for two Phasor field propagators in the Wigner distribution domain

In this section, we provide an algebraic steps to shown the RSD and the Fresnel diffraction in the WDF domain. Consider two parallel planes with spacing z, both Rayleigh-Sommerfeld Diffraction (RSD) and Fresnel diffraction link an input wavefront $f_i(x_i)$ to a output wavefront $f_o(x_o)$ by a spatial convolution with a convolution kernel $h_{xx}(x_o, x_i)$. Because of the convolution kernel is a special case of an integral kernel, $h_{xx}(x_o, x_i)$ becomes $h_{xx}(x_o - x_i)$.

$$f_o(x_o) = f_i(x_i) * h_{xx}(x_o, x_i) = f_i(x_i) * h_{xx}(x_o - x_i)\,, \tag{4.21}$$

Using the Wigner Distribution Function (WDF) convolution theorem, convolution along x in spatial domain applies convolution along x but in the WDF domain. This infers Eq. (4.21) has a equivalent transformation in the WDF domain as follows,

$$\mathcal{W}_{f_o}(x_o, \mu_o) = \mathcal{W}_h(x, \mu) \underset{x}{*} \mathcal{W}_{f_i}(x, \mu) \tag{4.22}$$

Overall, there are two approaches to derive the RSD, Fresnel diffraction in the WDF domain. One is to apply WDF's convolution theorem, which means that the only thing that has to be done is to simplify $\mathcal{W}_h(x, \mu)$ given the RSD and Fresnel convolution kernel expressions. Another is to use Eq. (4.3,4.4) to derive by definitions. We show the challenge of deriving the RSD in the WDF using the first approach and derive the Fresnel diffraction in the WDF based on definitions.

For the RSD, to best of our knowledge, its kernel expression $h_{xx}(x_o - x_i) = \frac{e^{jk\sqrt{(x_o-x_i)^2+z^2}}}{\sqrt{(x_o-x_i)^2+z^2}}$ does not have a simplified analytical form in the WDF domain. In this work, we rely on numerical implementations to evaluate the RSD in the WDF domain.

As for the Fresnel diffraction, given approximations coming from a binomial expansion on the RSD kernel, it is way more easier to treat analytically. In the following, we show the Fresnel diffraction in the WDF domain by definitions following algebraic procedures in Eq. (4.3,4.4). The Fresnel spatial convolution kernel $h_{xx}(x_o - x_i)$ can be written down as below and k stands for the wavenumber,

$$\begin{aligned} h_{xx}(x_o - x_i) &= \frac{e^{jkz}}{j\lambda z} \exp\left[\frac{jk}{2z}(x_o - x_i)^2\right] \\ &= \alpha(z) \exp\left[\frac{jk}{2z}(x_o - x_i)^2\right], \end{aligned} \tag{4.23}$$

Eq. (4.23) describes the Fresnel diffraction in its spatial representation. Use Eq. (4.24) to derive the integral kernel $K(x_o, \mu_o, x_i, \mu_i)$ in the WDF domain as follows:

$$K(x_o, \mu_o, x_i, \mu_i) = \iint \underbrace{h_{xx}(x_o + \frac{x_o'}{2}, x_i + \frac{x_i'}{2})}_{\text{term 1}} \underbrace{h_{xx}^*(x_o - \frac{x_o'}{2}, x_i - \frac{x_i'}{2})}_{\text{term 2}} \exp\left[-j\mu_o x_o' + j\mu_i x_i'\right] dx_o' dx_i', \tag{4.24}$$

Plug $h_{xx}(x_o, x_i)$ from Eq. (4.23) into Eq. (4.24), simplify term 1 and term 2,

$$\text{term 1} = \alpha(z)\exp\Big[\frac{jk}{2z}(x_o - x_i + \frac{x_o^{'}}{2} - \frac{x_i^{'}}{2})^2\Big]$$
$$\text{term 2} = \exp\Big[-\frac{jk}{2z}(x_o - x_i - \frac{x_o^{'}}{2} + \frac{x_i^{'}}{2})^2\Big]\alpha^*(z)\,, \tag{4.25}$$

Plugging in terms from Eq. (4.25), Eq. (4.24) becomes,

$$\begin{aligned}
&K(x_o, \mu_o, x_i, \mu_i)\\
&= \iint |\alpha(z)|^2 \exp\Big[\frac{jk}{z}\underbrace{(x_o - x_i)(x_o^{'} - x_i^{'})}_{\text{use } a^2-b^2=(a+b)(a-b)}\Big]\exp\Big[-j\mu_o x_o^{'} + j\mu_i x_i^{'}\Big]\,dx_o^{'}dx_i^{'}\\
&= |\alpha(z)|^2 \int \underbrace{\int \exp\Big[\frac{jk}{z}(x_o - x_i)x_o^{'}\Big]\exp\Big[-j\mu_o x_o^{'}\Big]\,dx_o^{'}}_{\text{term 3}} \exp\Big[-\frac{jk}{z}(x_o - x_i)x_i^{'}\Big]\exp\Big[j\mu_i x_i^{'}\Big]\,dx_i^{'}\,,
\end{aligned} \tag{4.26}$$

Simplify term 3 in Eq. (4.26) by using the Fourier transform property,

$$\begin{aligned}
\text{term 3} &= \int \exp\Big[\frac{jk}{z}(x_o - x_i)x_o^{'}\Big]\exp\Big[-j\mu_o x_o^{'}\Big]\,dx_o^{'}\\
&= 2\pi\delta(\mu_o - \frac{k(x_o - x_i)}{z})\\
&= 2\pi\delta(x_i - x_o + \frac{z}{k}\mu_o)\,,
\end{aligned} \tag{4.27}$$

Replace term 3 in Eq. (4.26) by Eq. (4.27),

$$
\begin{aligned}
&K(x_o, \mu_o, x_i, \mu_i) \\
&= |\alpha(z)|^2 \cdot 2\pi\delta(x_i - x_o + \frac{z}{k}\mu_o) \cdot \int \mathbf{1} \cdot \exp\Big[-j\big(\frac{k(x_o - x_i)}{z} - \mu_i\big)x_i'\Big]\, dx_i' \\
&= |\alpha(z)|^2 \cdot 2\pi\delta(x_i - x_o + \frac{z}{k}\mu_o) \cdot 2\pi\delta\big(\frac{k}{z}(x_o - x_i) - \mu_i\big)\,, \qquad (4.28)
\end{aligned}
$$

Since Eq. (4.28) consists of a multiplication between two delta functions, we can use it as a constraint to simplify variables,

$$
\begin{cases}
x_i - x_o + \frac{z}{k}\mu_o = 0 \\
\frac{k}{z}(x_o - x_i) - \mu_i = 0\,,
\end{cases} \qquad (4.29)
$$

which leads to constraints $x_i = x_o - \frac{z}{k}\mu_o$ and $\mu_i = \mu_o$.

Above all, in the WDF domain, the Fresnel diffraction integral kernel $K(x_o, \mu_o, x_i, \mu_i)$ is as follows,

$$
\begin{aligned}
&K(x_o, \mu_o, x_i, \mu_i) = (2\pi)^2 |\alpha(z)|^2\, \delta(x_i - x_o + \frac{z}{k}\mu_o)\, \delta(\mu_i - \mu_o) \\
&K(x_o, \mu_o, x_i, \mu_i) \overbrace{=}^{\text{normalize}} \delta(x_i - x_o + \frac{z}{k}\mu_o)\, \delta(\mu_i - \mu_o)\,, \qquad (4.30)
\end{aligned}
$$

Using $K(x_o, \mu_o, x_i, \mu_i)$ in Eq. (4.30) and plugging it in Eq. (4.3), then the corresponding input, output field WDF transformation using the Fresnel diffraction is as follows,

$$
\begin{aligned}
&\mathcal{W}_{f_o}(x_o, \mu_o) \\
&= \frac{1}{2\pi} \iint K(x_o, \mu_o, x_i, \mu_i) \mathcal{W}_{f_i}(x_i, \mu_i)\, dx_i d\mu_i \\
&\propto \iint \delta(x_i - x_o + \frac{z}{k}\mu_o)\delta(\mu_i - \mu_o)\mathcal{W}_{f_i}(x_i, \mu_i)\, dx_i d\mu_i\,, \qquad (4.31)
\end{aligned}
$$

Finally, applying constraints on variables for Eq. (4.31) leads to a shear mapping in the WDF domain,

$$\mathcal{W}_{f_o}(x, \mu) = \mathcal{W}_{f_i}(x - \frac{z}{k}\mu, \mu)\,, \tag{4.32}$$

5 ANALYSIS OF FEATURE VISIBILITY IN NON-LINE-OF-SIGHT MEASUREMENTS

In this chapter, we formulate an equation describing a general Non-line-of-sight (NLOS) imaging measurement and analyze the properties of the measurement in the Fourier domain regarding the spatial frequencies of the scene it encodes. We conclude that for a relay wall with finite size, certain scene configurations and features are not detectable in an NLOS measurement. We then provide experimental examples of invisible scene features and their reconstructions, as well as a set of example scenes that lead to an ill-posed NLOS imaging problem.

5.1 Introduction

In this chapter, we aim to provide a generic description for direct bounce (3rd bounce) NLOS measurements and show how much information they encode and how this affects practical NLOS imaging problems. Consider the NLOS reconstruction shown in Figure 5.1. It contains three very similar patches that only vary slightly in orientation and have different positions in the reconstruction space. Yet while two of the patches are reconstructed clearly and accurately, the third is completely missing from the reconstruction. A closer inspection of prior published results reveals that similar artifacts are seen in reconstructions using a variety of diverse reconstruction methods. Surfaces with certain normal vectors are missing in the reconstruction or scenes with simple surfaces are chosen to avoid the problem. The main purpose of this work is to explain this phenomenon.

As we show below, any NLOS measurement can be expressed as an integral operator known in the literature as elliptical Radon integral. We analyze this measurement function in the Fourier domain and show that a significant part of the measurement space is not accessed by the NLOS measurement and thus represents a null space for NLOS reconstruction. Finally, we investigate the Fourier domain representations of common scenes and scene features to identify features

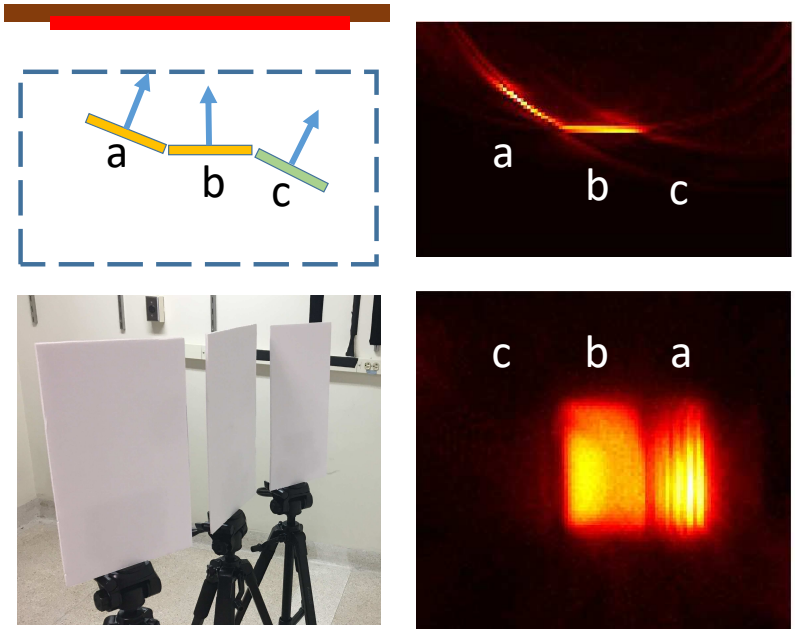

Figure 5.1: **Three patches rotation example.** Experimental reconstruction of a scene containing three patches denoted by **a, b, c**. Patches **a** and **c** are parallel but the latter one's surface normal vector does not point towards the NLOS relay wall and the patch does not appear in the reconstruction shown on the right. The first row stands for the top view and the second row for the front view. The explanation for this astonishing effect will be developed in the main text, see Sec. 5.5, and also Fig. 5.8 for a graphical explanation.

that fall into the null space and cannot be reconstructed. Because our analysis involves a generic description of the NLOS measurement, it is independent of the reconstruction algorithm used. We expect our findings will inform inverse solution design and future NLOS reconstruction methods.

5.2 Related Work

Statements of the properties of NLOS reconstructions are sometimes included with a presentation of reconstruction algorithms. For example, the available resolution

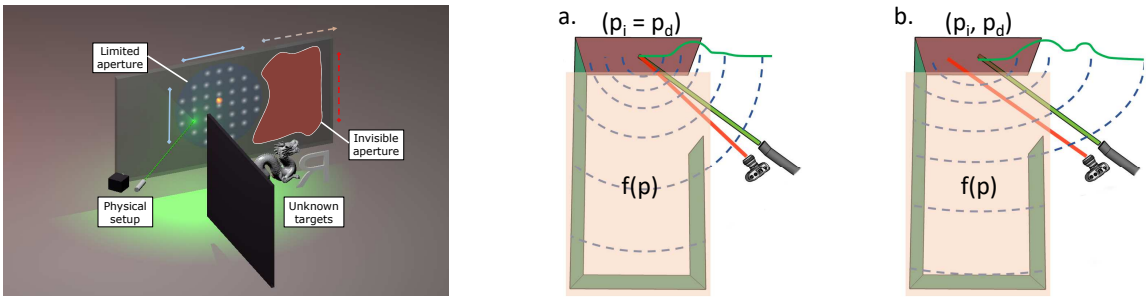

Figure 5.2: **Illustration of a NLOS measurement:** The panels a and b show two popular NLOS measurement setups. Subfigure **a** shows the confocal measurement which means the illumination and detection point are co-located, while **b** shows the non-confocal measurement. The green curves in **a** and **b** sketch the acquired time responses, i.e., the integration of reflecting scene features along each circle (confocal case) or ellipse (non-confocal case).

has been analyzed [9, 4]. Since the underlying mathematical forward model shares a similar root as the computed tomography problem (Radon integral), much more extensive work on this problem is available through related problems. Radon integrals have been studied for their applications in ultrasound imaging and medical computed tomographies [95, 96, 97]. The effect of a finite-sized sampling aperture is similar to the missing cones problem that has been the subject of extensive research in medical computed tomography (CT) imaging [98, 99].

Radon integrals can be categorized into classical Radon integral (planar), spherical Radon transform (SRT) and elliptical (ellipsoid) Radon transform (ERT). The planar Radon transform with its high dimensional model [100] is well known in the CT field. The spherical Radon transform (SRT) with its inverse solution design can be found in [95, 96]. A similar inverse formula for the confocal measurement setup applies for the SRT model used in ultrasound [96] and NLOS imaging [4]. The elliptical Radon transform (ERT) has been studied by Moon et al. [97]. An inverse to the ERT was proposed in [101] and is similar to (but not the same as) the NLOS non-confocal filtered backprojection method [7, 8, 9].

5.3 NLOS imaging problem

In this section, we are going to provide basic mathematical tools for modeling the NLOS imaging problem in a general way by introducing the NLOS measurement function and its properties. Then we introduce the concept of limited aperture NLOS imaging which results in an incomplete measurement space. This incomplete measurement space is essential for the next section dealing with our measurement analysis in the spatial Fourier domain in Sec. 5.4.

NLOS Measurement Function

The NLOS measurement scenarios are illustrated in Figure 5.2 on the left, and two measurement setups are shown in the subfigures **a** and **b**. In an NLOS measurement, the scene is illuminated from a point $\vec{p}_i$ (illumination position) and light returning from the scene is recorded at a detection point $\vec{p}_d$ (detection position) after a certain time interval t. Both $\vec{p}_i$ and $\vec{p}_d$ are within the finite area of the relay wall that we call the NLOS sampling aperture. In practice, this sampling aperture is bounded by the limited field of view because of the remote detection.

We use $f(\vec{p})$ to represent the unknown 3D scene we would like to recover. Here $f(\vec{p})$ is a function of the vector $\vec{p} = (x, y, z)$ storing the reflectance values in space. We assume uniform scattering, thus the reflectance value is angle independent. The function g represents one single time response measurement at illumination position $\vec{p}_i$ and detection position $\vec{p}_d$. Thus, each time resolved measurement $g(\vec{p}_i, \vec{p}_d, t)$ is a function of illumination position $\vec{p}_i$, detection position $\vec{p}_d$ and time t. The NLOS measurement is made up of a set of detection positions (detection grid) from a single or multiple subsequent point illuminations on the wall. We use $G = \{g_1, g_2, g_3, ..., g_n\}$ to represent all measurements. We focus on illustrating the transformation from $f(\vec{p})$ to one single measurement $g(\vec{p}_i, \vec{p}_d, t)$ first. Then it is straightforward to understand the transformation from $f(\vec{p})$ to the entire measurement set G. Subfigures **a** and **b** of Figure 5.2 represent the confocal and non-confocal measurement setups. Since the non-confocal case is the more general version of

NLOS measurement, we start our forward modeling from the non-confocal measurement.

We can use an integral operator $A : f(\vec{p}) \to g(\vec{p}_i, \vec{p}_d, t)$ to represent the linear transform from the unknown function $f(\vec{p})$ to a temporal measurement at given detection and illumination positions and time $g(\vec{p}_i, \vec{p}_d, t)$. This results in the NLOS measurement equation:

$$g(\vec{p}_i, \vec{p}_d, t) = \int_{\mathbb{R}^3} \Upsilon(d_i, d_d) \cdot \delta(d_i + d_d - t \cdot c) \cdot f(\vec{p})\, d\vec{p}\,. \tag{5.1}$$

In the literature, this integral is ofter referred to as the Elliptical Radon Transform (ERT). The delta function kernel in this equation describes the geometry of the integration, and c denotes the speed of light. To simplify notation, we use the distance terms $d_i = |\vec{p} - \vec{p}_i|$ and $d_d = |\vec{p} - \vec{p}_d|$ which represent the distances between the integral variable $\vec{p}$ and illumination/detection position, respectively. The term $\Upsilon(d_i, d_d)$ stands for the intensity drop-off associated with distances traveled by the light. This intensity term is normally not included in the ERT, however we add it in this treatment as it is needed to correctly model the physical measurement process. The integration (5.1) can be performed for an entire family of thin ellipsoid surfaces having different foci $\vec{p}_i$ and $\vec{p}_d$.

Any set of NLOS measurements made from locations within the NLOS aperture area can be expressed as a set of these measurement functions. The integral in Equation (5.1) is difficult to treat analytically because of non-constant curvatures. Therefore we also consider two simplified scenarios. If we allow only measurements where $\vec{p}_i = \vec{p}_d$, we obtain the confocal NLOS measurement (Figure 5.2 **a**)

$$g(\vec{p}_i, \vec{p}_i, t) = \int_{\mathbb{R}^3} \Upsilon(d_i) \cdot \delta(2 \cdot d_i - t \cdot c) \cdot f(\vec{p})\, d\vec{p}\,. \tag{5.2}$$

This integral without intensity term $\Upsilon(d_i)$ is also known in the literature as the spherical Radon transform (SRT) [95, 96].

Another useful tool to approximate the measurement function is to locally replace the elliptical integral by integrals over planes that are tangential to the

ellipsoids. This refers to a zoom-in version of the integration at a local volume. Since the integral is a linear operator, the measurement can be represented as superposition of all individual inputs (linearity). Then for each individual input, the local integral version can be approximated by the planar integral. This means:

We linearize the unknown 3D function $f(\vec{p})$ by a summation of local functions $f_1, f_2, ..., f_n$ which store the features representing the 3D image. Each local function represents a local pattern within cube volumes at different positions such that each subspace contains only small sections of the ellipsoids. Each ellipsoid $g(\vec{p}_i, \vec{p}_i, t)$ can within that subspace be approximated by planes $g'(t, \theta, \phi)$ where the planar angles θ and ϕ are such that the plane normal vector points to the center of the ellipsoid and t is proportional to twice the distance between the ellipsoid center and the plane. In polar coordinates, this yields

$$g'_{(\theta,\phi)}(t) = \iiint_{-\infty}^{\infty} f(x, y, z) \cdot \delta(\sin\theta\cos\phi x + \sin\theta\sin\phi y + \cos\theta z - t \cdot c)\, dx\, dy\, dz\,.$$

This planar approximation has been described before in [7, 8] to approximate the NLOS imaging result from non-confocal streak camera measurements. It is also similar in nature to the approximation made when using a piecewise definition of a locally varying point spread function as is often done in point spread function deconvolution problems.

Above all, we provide the basic tools for describing the NLOS forward model which we are going to use in Section 5. There, we provide our main tools to analyze the incomplete measurement space effect. Our idea is to mimic the imaging system by illustrating the information content in the Fourier domain by the modulation transfer function (MTF) dependent on a given limited aperture.

5.4 NLOS Measurements in the Fourier domain

A complete description of a measurement of a section of $f(\vec{p})$ can be described as the set of all $g(\vec{p}_i, \vec{p}_d, t)$ for which $\vec{p}_i$ and $\vec{p}_d$ are in the NLOS aperture plane and t is such that the measurement ellipsoid goes through $f(\vec{p})$. To obtain insight into the

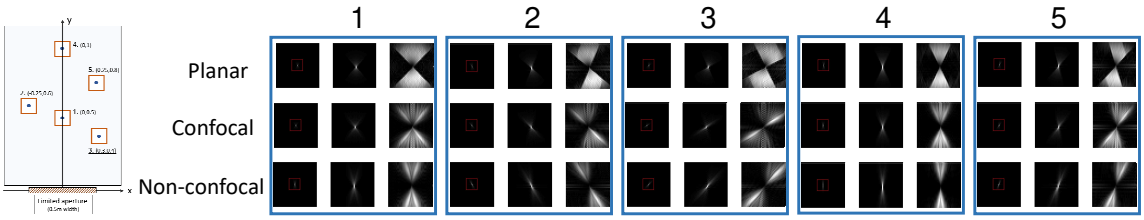

Figure 5.3: **Local Measurement MTF**: Fourier domain representations of the NLOS Measurement Function for five points in the unknown geometry. The left graphic represents the geometric setup. We construct a half meter limited aperture with five volumes of interest varying in depth and horizontal offset. The panels show the measurement function for all points for the different models. Planar corresponds to the PRT, confocal to the SRT, and non-confocal to the SRT. For each point, the first column shows the computed patterns, the middle column shows the same pattern zoomed in, and the right column shows the zoomed in pattern after correcting for the lower values at higher frequencies to enhance visualization.

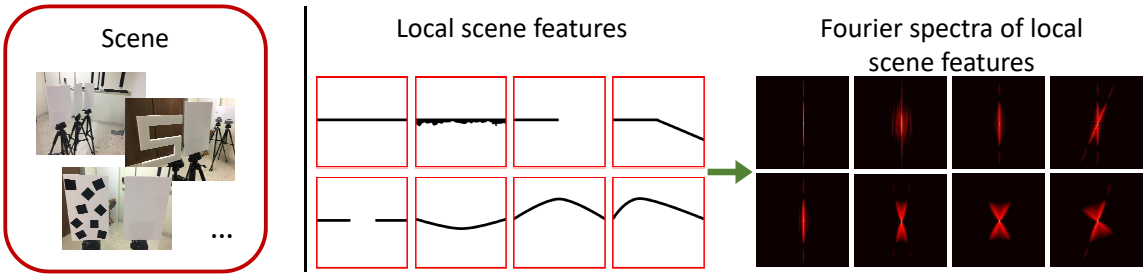

Figure 5.4: **Local scene features:** This figure shows a set of common NLOS scene features in the red boxes and their Fourier transforms. Rotating of the features simply corresponds to rotating by the same angle in the Fourier domain. The patterns are (top left to bottom right) a smooth planar surface, a rough planar surface, the edge of a planar surface, a corner between two surfaces, a gap in a planar surface, a convex curved surface, and two concave curved surfaces. The spectrum of a planar surface is a line. Roughness, curvature, and edges result in spectra that also cover other regions of the Fourier space.

patterns in the scene that are sampled by this measurement, we want to analyze them in the Fourier domain. The Fourier domain NLOS measurement function can be understood as a Modulation Transfer Function (MTF) of NLOS imaging.

For the planar Radon transform (PRT, Equation (5.3)), the Fourier transform of g' can be computed analytically using the projection-slice theorem.

Projection slice theorem

The projection slice theorem (PST) is well-known and widely used in the area of computed tomography and other fields. We will give a short explanation of the 2D version here; the 3D version is shown in the supplemental document. The PST provides an elegant tool for using projections (i.e., integrals) along parallel lines of an unknown 2D scene: it shows that the 1D Fourier transform of such projections actually represents one line through the origin of the 2D Fourier spectrum of the unknown scene. Repeating the process for different projection angles provides the missing lines; scene reconstruction is then easily achieved by inverse 2D FT.

The unknown 2D function is denoted by $f(x, y)$. Let us first assume that the projection angle θ is zero. The projection dependent on the displacement u with respect to the origin and vertically to the projection direction is then given by [102]

$$\begin{aligned} f_{\text{proj}}(u, \theta = 0) &= \iint f(x, y)\delta(x - u)\,dx\,dy \\ &= \int f(u, y)\,dy\,. \end{aligned}$$

Performing the 1D Fourier transform with respect to u yields

$$\begin{aligned} \mathcal{F}_u\{f_{\text{proj}}(u, 0)\} &= \iint f(u, y)e^{-j2\pi f_u u}\,du\,dy \\ &= \iint f(x, y)e^{-j2\pi f_x x}\,dx\,dy \\ &= \iint f(x, y)e^{-j2\pi(f_x x + f_y y)}\,dx\,dy\Big|_{f_y=0}\,. \end{aligned} \tag{5.3}$$

This corresponds to the 2D Fourier transform of the unknown function $f(x, y)$ at the line with $f_y = 0$. Repeating the projection with different angles θ results in

the corresponding line of the 2D Fourier transform; note that rotation in the spatial domain corresponds to a rotation by the same angle in the frequency domain. By considering all angles θ from 0 to 2π and adding all Fourier spectrum lines, the full function $f(x, y)$ can then be reconstructed by inverse 2D Fourier transform. It is advisable to apply a high pass filter before the inverse Fourier transform, as high frequencies are underrepresented because of the spectral lines meeting in the origin, but diverging for higher frequencies, which means that there is less information available for these frequencies. This is called *filtered backprojection* [100].

In the 3D case, the projection along parallel planes and subsequent 1D Fourier transform of the calculated (scalar) projection values for each plane provides one line of the 3D spectrum of $f(\vec{p})$. Rotating the projection planes about the origin provides the spectrum along all lines through the origin. See the supplementary document for details.

Local Fourier cone

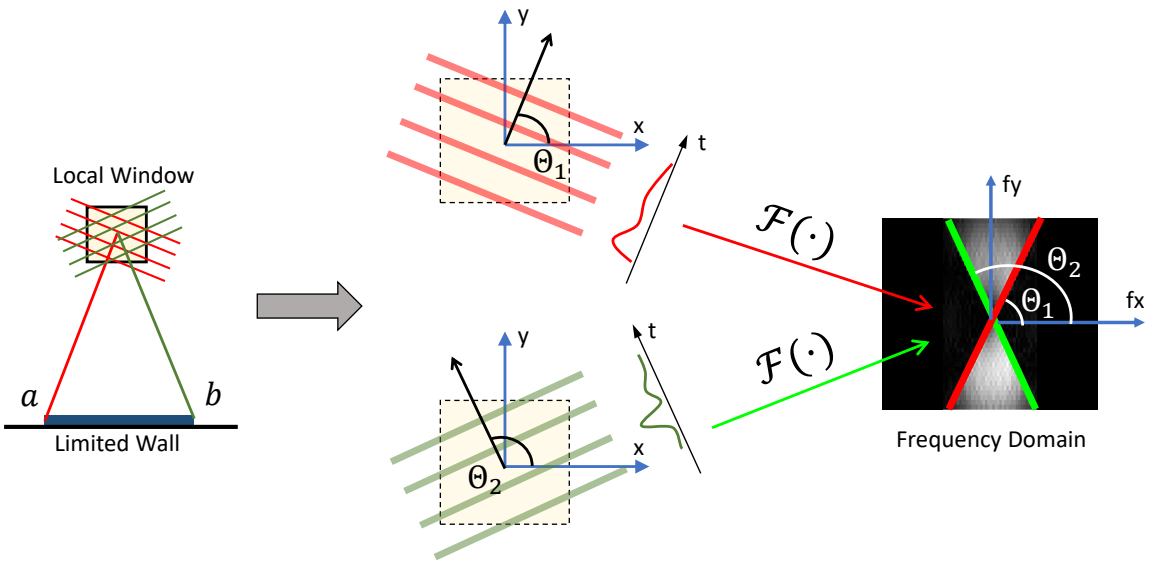

Figure 5.5: Fourier slice theorem and cone generation.

The left part of Figure 5.5 shows the measurement for two cases where the

illumination and detection points are co-located at a or b at either side of the relay wall. In the middle of that figure, we can see that the collected time response for those two cases are projections along directions with angles Θ_1 or Θ_2 with respect to the relay wall. The exact projections are performed along ellipsoids; however, we approximate these as planes. As it is shown on the right of Figure 5.5, the Fourier slice theorem states that the 1D Fourier transform of the projections of the scene along parallel planes at angle Θ in the primal domain is equal to a slice at angle Θ in the 2D Fourier spectrum of the scene. To create a cone we then simply draw the corresponding lines for all accessible points on the relay wall. By using a limited size relay wall, the black part of the spectrum is not contained in any measurements.

Null space and cone variation: As shown in Fig. 5.3, given the same size of the limited aperture, the closer the considered volume is to the sampling aperture, the larger the visible angle range of this cone is (compared with position 1 and position 4 in Fig. 5.3). Moreover, offset with regard to the center of the aperture also tilts this sampling cone in the Fourier domain (refer to positions 2, 3 and 5 in Figure 5.3). Overall, things outside this cone will never be sampled in the Fourier domain by the limited aperture unless the aperture size is increased. This means that, e.g., a wall placed perpendicularly to the relay wall has a Fourier transform which, apart from the DC value at frequency 0, will fully vanish in the missing cone and cannot be seen in any reconstruction not accounting for the missing cone.

Model validity: To verify the model validity, we perform a computation for planar, confocal and non-confocal measurement (fixed illumination as one focus at the center) with the same limited aperture. With an acceptable error in the discrete model, we can see from Fig. 5.3 that the local planar model gives a good boundary estimation for the confocal and non-confocal cases. Notice that we fixed one focus $\vec{p}_i$ in the non-confocal case resulting in a more narrow MTF pattern. By moving the non-confocal illumination spot, this narrow cone slightly rotates which could achieve a similar angle coverage in the confocal measurement.

5.5 Measurement of scene features

Local scene features: To use the proposed analysis, we need to decompose any complicated NLOS imaging scenes into simple features such as planes etc. This is shown in Fig. 5.4, which also displays their Fourier spectra which defines their reconstructability, cf. Fig. 5.5 and the following paragraph. Most of the simple features can be represented as a thin surface with different roughness level as well as edge like discontinuity patterns and curvatures. Rotation of the patterns in space simply corresponds to the same rotation in the Fourier domain.

MTF with scene features: To assess visibility of the targets we have to consider the overlap between the target spectra and the MTF sampling cone. This will be shown illustratively in the results, see Section 5.5. Note that the most valuable information is encoded in the high frequency components of the cone. The center of the cone is at the origin and thus samples the low spatial frequencies. The high frequency components far away from the origin are necessary to create high resolution reconstructions.

Limited aperture ill-posed example: Limited aperture NLOS imaging with arbitrary targets should be viewed as an ill-posed problem. Consider a simple patch with different rotation angles with respect to the limited aperture. Once it is facing the sampling aperture and located at its center, its local Fourier spectrum is optimally covered by the MTF Fourier cone. However, when the patch pattern rotates, its Fourier pattern also rotates by the same angle. Once it rotates outside the angle covered by each local cone, the recorded information only includes the origin in the Fourier plane, corresponding to a constant or zero spatial frequency. See the experiments in Sec. 5.5 for an illustration of this effect. Overall, this means we can still see that a structure is there (after all, there is light coming back) but we cannot actually uniquely reconstruct it. On the other hand, a local pattern containing high spatial frequency contents like edges or surface roughness has a much broader spectrum and at least part of it always overlaps with the measurement cone making it at least partially reconstructible.

Rule-of-thumb criterion: From this, we can also derive a simple rule for visibil-

ity of scene patches. Since a planar scene segment is represented by a line through the origin in the Fourier domain, the segment will be completely visible if the line lies within the sampled cone, and invisible if it is at an angle outside the cone. **In the primal domain this means that a segment is only visible when its normal vector points toward the relay wall.**

Overall, for NLOS imaging scenarios, more higher spatial frequency components of unknown targets subsequently result in more fluctuations in the temporal measurements. For this reason, they lead to a higher chance to actually see the target by using all deterministic linear inverse solutions from a limited aperture.

Completely invisible feature example

To provide further evidence for the existence of a universal, reconstruction independent null-space, we provide a set of example scenes that are simple enough to be treated analytically for a confocal measurement set. To do this, we replace $f(\vec{p})$ by specific functions to represent the scene features inside the integral measurement equations (5.1) and (5.2). To simplify the calculation, we consider only a two dimensional scenario of a confocal NLOS measurement.

Smooth wall example (Challenge scene): As it is shown in Figure 5.6, the unknown target is a smooth wall at an angle to the relay wall described by $\delta(\cos(\theta)x - \sin(\theta)y)$ for $y \geqslant 0$. For confocal NLOS measurements, our illumination and detection are co-located at the x-axis meaning that $\vec{p}_i = \vec{p}_d = (x_0, 0),\ x_0 \geqslant 0$. Since the distance correction simply amplifies the signal at each given time index i, we use the notation $C(i)$ to represent this correction term. We use the term $g(i)$ to represent the reflectance integration at each time index i. By plugging this special function into the Radon integral, we obtain the following equation:

$$g(i) = C(i) \cdot \int_0^\infty \int_{-\infty}^\infty \delta(\cos(\theta)x - \sin(\theta)y) \cdot \delta((x - x_0)^2 + y^2 - (\frac{c \cdot i}{2})^2)\, dx\, dy\,. \tag{5.4}$$

By performing the distance correction $C(i)$ to account for the distortion from the

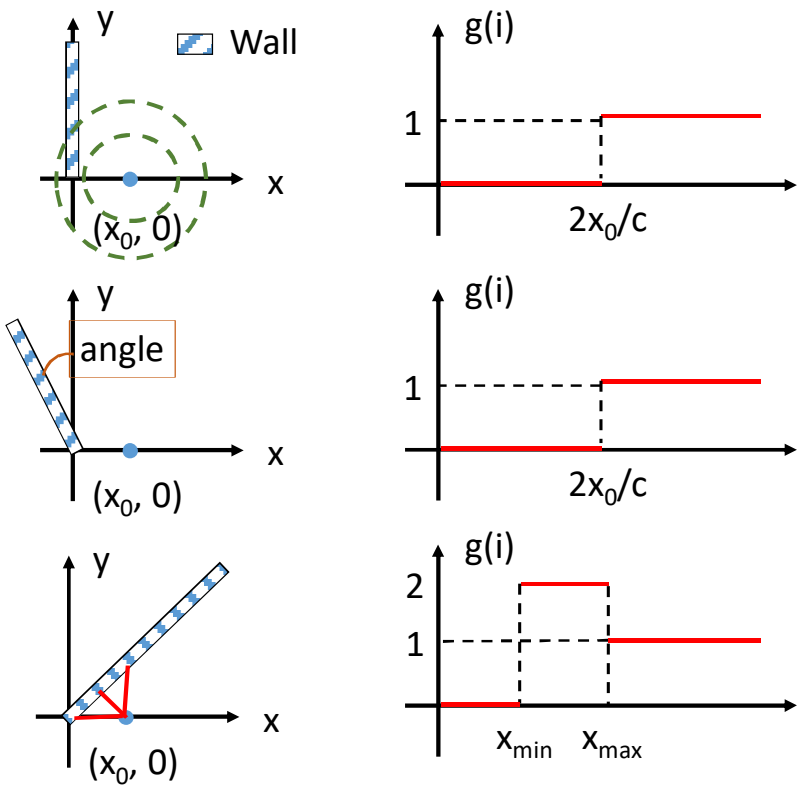

Figure 5.6: Three wall examples. The first column represents the geometry, the second column stands for the measurement $g(i)$ after the intensity correction from $t(i)$. The scenes in the first two rows lead to identical captured data.

intensity drop off, the measurement $g(i)$ simply represents a set of reflectance integrals in the unknown space.

Using the fact that $\delta(\cos(\theta)x - \sin(\theta)y)$ is only 1 for $\cos(\theta)x = \sin(\theta)y$, we can replace x by $y\tan(\theta)$ and further simplify the equation by getting rid of integral variable x as follows:

$$g(i) = C(i) \cdot \int_0^{\infty} \delta((\tan(\theta)y - x_0)^2 + y^2 - (\frac{c \cdot i}{2})^2)\, dy \tag{5.5}$$

From this equation, it is clear to see that the final measurement can be viewed

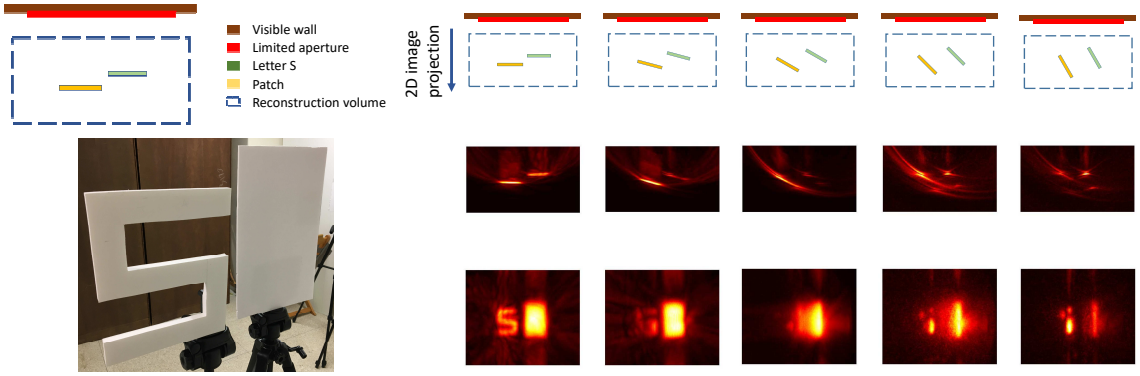

Figure 5.7: **Simple letter S and rectangle patch rotation experiments** The first row represents the schematic of the setup including the entire visible wall, limited aperture, targets (letter S and patch) as well as reconstruction volume. The next two rows show the maximum projection along the depth dimension, thus a 2D bird view and 2D front view are provided. For clear illustration, we present results using two-color (bright and dark view). The thickness of the letter S and the rectangular patch approximately equals 5 cm and 0.5 cm. As the angle increases, the ill-posed effect becomes more obvious, and certain features are missing in the measurement space and therefore cannot be resolved by the reconstruction

as a sum of 1 or 0 at any given time index $\hat{i}$ when the kernel inside the integral $(\tan(\theta)y - x_0)^2 + y^2 - (\frac{c\cdot\hat{i}}{2})^2 = 0$. This parabolic equation may have zero, one or two solutions with fixed θ, x_0. This means the final measurement g may be 0, 1 or 2 at any given time index. A simple observation is that, once the wall reaches angle $\theta = \pi/2$, the value of the integral is the same for all $\vec{p}_i = (x_0, 0)$, $x_0 \geqslant 0$ on the aperture line. The signal g with different wall rotation angles is shown in Figure 5.6. Thus, all scenes with walls at angles above $\theta = \pi/2$ result in the same NLOS data and are thus not distinguishable without any prior constraints. Extending this treatment to a 3D space results in more complex equations and is therefore less instructive. It is subject of further study.

Experiments

We perform two experiments of ill-posed scenarios which are first predicted by our proposed local cone model. For the reconstruction process, we choose a filtered backprojection [7, 8, 9] without thresholding process. For visualization, we choose the maximum projection along the depth dimension in the front view. Also, we provide a maximum projection for the bird view.

In Figure 5.7 we show a rotating letter S and planar patch example. As shown in the result, the simple patch pattern is completely resolvable when facing directly to the limited aperture. In this situation all patch normals are pointing at the relay wall aperture. As the targets rotate, normals start pointing past the aperture and the planes start to disappear. Not all target patches disappear simultaneously due to the local variations in Fourier cone discussed above. Eventually, all patch normals point outside the aperture and the entire targets are outside the Fourier cone. We can only resolve the high spatial frequency pattern at edges in the final reconstruction.

Another three patches example was already shown in Figure 5.1. Having now the theoretical tools at hand, we see in Fig. 5.8 that both patches **a** and **c** have the same magnitude spectrum, but because of the spatially dependent measurement cone, only the offset of patch **c** is captured. In the reconstruction, this patch is therefore almost entirely invisible.

5.6 Appendix: Results and additional notes

Local MTF pattern with/without intensity drop off

As mentioned in **main text section 4.1**, the generalized NLOS forward model without the intensity drop off term $\Upsilon(d_i, d_d)$ can be described as a Radon integral. All previous works approximate the intensity drop off by the squared traveling

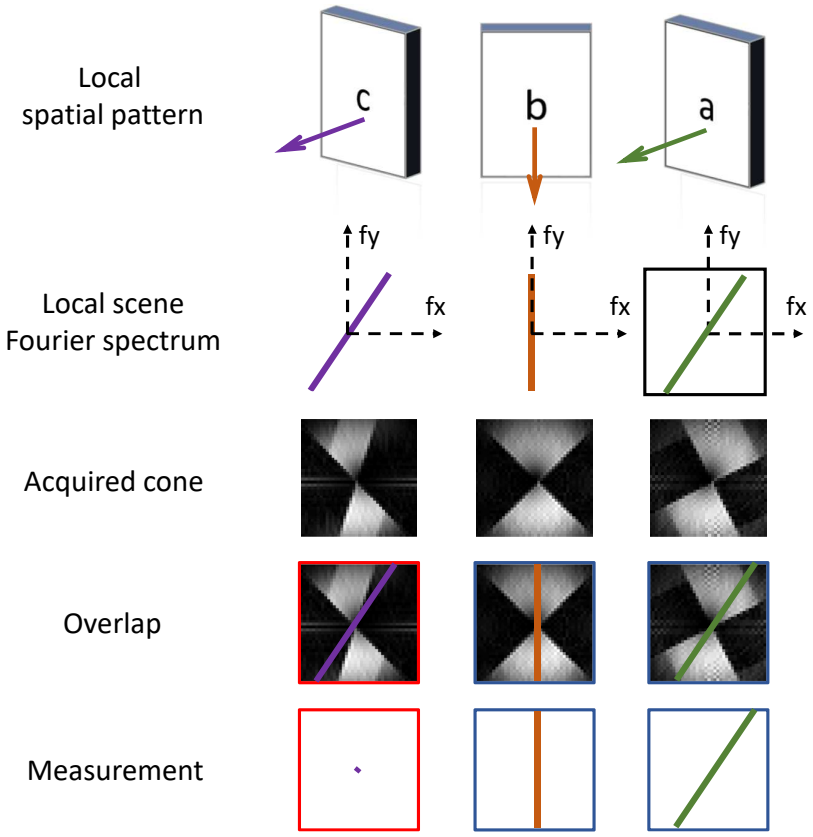

Figure 5.8: **Explanation of the missing feature in Fig. 5.1** The measurement cone is shown as gray area and illustrates the part of the Fourier spectrum that is actually acquired by NLOS measurements. It varies with the position in the 3D space that is to be reconstructed. Both patches **a** and **c** of the scene are oriented the same way; the spatial shift just corresponds to a phase shift in the Fourier domain, but not in a change of the magnitude spectrum. This means both patches have the same magnitude spectrum, but due to the fact that the measurement cone is different at their respective positions, the measurement of patch **c** only captures the offset and not the rest of the spectrum. For this reason, this patch cannot be reconstructed.

distance:

$$g(\vec{p}_i, \vec{p}_d, t) = \int_{\mathcal{D}} \Upsilon(d_i, d_d) \cdot \delta(d_i + d_d - t \cdot c) \cdot f(\vec{p})\, d\vec{p} \tag{5.6}$$

$$\approx \int_{\mathcal{D}} \frac{1}{d_i^2} \cdot \frac{1}{d_d^2} \cdot \delta(d_i + d_d - t \cdot c) \cdot f(\vec{p})\, d\vec{p}\,. \tag{5.7}$$

In the main text, we provide the MTF pattern without the intensity drop off. Here we provide the MTF pattern at the same five positions but considering the intensity drop off effect. As it is shown in Fig. 5.9, this intensity drop off is negligible for the spatial frequency analysis.

An intuition behind Equation (5.6) is that the modeling process starting from the unknown reflectance $f(\vec{p})$ instead of the field amplitude should be observed. Furthermore, the kernel $\delta(d_i + d_d - t \cdot c)$ governs the integral geometry, whereas the intensity drop-off $\Upsilon(d_i, d_d)$ reduces the value by a specific weight. This effect is similar to the amplitude error made in the diffraction integral approximation.

Non-confocal illumination extended MTF cone

There are two main types of the NLOS forward sampling method: confocal and non-confocal setup. As it can be seen in **main text Figure 2**, confocal illumination and detection positions are co-located whereas for the non-confocal setup, both positions are different. To allow for easier comparison, we consider the same detection grid sampling at a finite relay wall. Then for the non-confocal measurement, we analyze the different Fourier cone behaviors along with varying illumination source position. This is shown in Figure 5.10. We consider a fixed local window (symmetry at the center, 0.5 m in depth) and three illumination positions within the limited sampling area. Different illumination positions also provide rotation in the Fourier cone for the non-confocal measurement. Even though the Fourier cone varies, as long as the illumination positions only cover the limited relay area, the cone stays within the boundary estimated from the planar model.

We show the generalized extended MTF cone by changing the illumination position at the limited sampling area. Also, we show the consistency with our proposed model.

An intuition for why the non-confocal measurement cone varies depending on the illumination position is provided in the following. We know that the non-confocal forward model can be described as the integration along a thin ellipsoid surface in space having two foci (one is the illumination, the other the detection

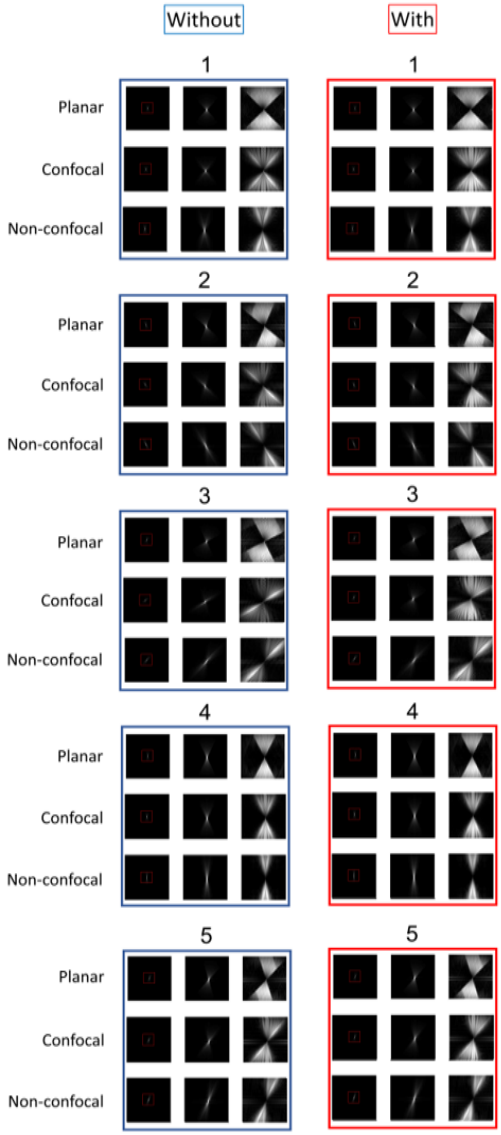

Figure 5.9: **MTF pattern without/with intensity drop off term:** This shows the MTF pattern differences without and with distance drop-off in the forward model. The first column represents no distance drop off term. The second one incorporates the distance drop-off. Each box contains the planar, confocal and non-confocal MTF patterns. Number 1-5 represent the different local positions which are the same as in **main text Figure 3**.

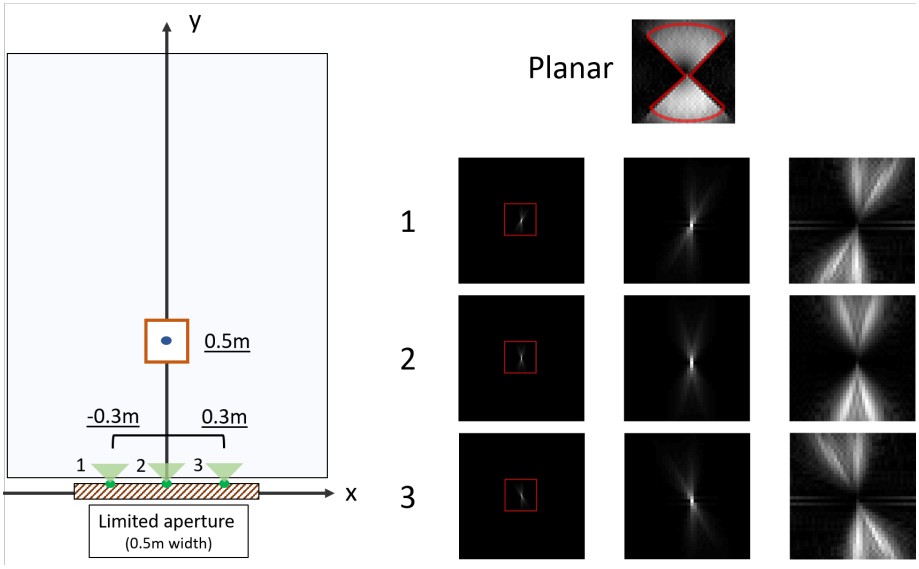

Figure 5.10: **Non-confocal illumination variation:** The non-confocal acquisition setup has two degrees of freedom for focused detection and illumination positions. In the figure, we fixed the array of detection positions at the limited aperture (width 0.5 m). We pick three illumination positions (1-3) within the limited aperture. We consider the same local volume position, and each local MTF pattern is shown on the right. Even though each non-confocal Fourier cone is narrower than the planar cone on the top, by using the multiple illumination source positions, the same Fourier cone coverage can be achieved.

position). For a fixed position window, the ellipsoid curvature is changed by moving one of the foci. In this case, the tangent line with its normal vector changes also. Thus each local cone also rotates when moving the illumination source along the hyperplane.

Here we list a summary considering two types of NLOS measurement in both mathematical and practical sense. Mathematically speaking, the non-confocal NLOS measurement is a more generalized version of confocal measurement. However, it can be seen in **main text Figure 3** that using the same limited size relay wall, the confocal setup covers a slightly wider area in the Fourier domain compared to one single illumination for the non-confocal measurement. As we show in this

section, this downside of the non-confocal measurement can be reduced in practice by simply adding some extra illumination source positions.

NLOS Radon integral modeling error

Some modeling error still constrains the model used for the NLOS problem. Notice that this *modeling error* is the error in the mathematical sense. Researchers avoid conflicting this error by adjusting the scene targets to fulfill the modeling gap. Here we provide a specific description for the modeling error may cause the *inverse crime* (Inverse crime means the forward model is not accurate enough causing the corresponding inverse solutions to fail in the real-world measurements).

Basic notation: We use the symbol f to represent the unknown function we would like to recover, A for a linear operator as the forward model, g for one measurement, G for a finite set of measurements, $\vec{p}$ represent the space variables (x, y, z). $\vec{p}_i$ and $\vec{p}_d$ stands for illumination point and detector point on the visible relay wall.

1. Lambertian approximation This is the approximation for the unknown function. By assuming the hidden scene reflects light isotropically, the unknown function f can be described as $f(\vec{p}) = f(x, y, z)$ meaning that the reflectance is a constant value in space. However, in reality, surface reflectance should at least based on incident and observation which leads to a higher dimension. This is well known in the computer graphic society.

2. Direct bounce modeling approximation With Approximation 1, one can model direct bounce in the integral equation as described in Eq. (5.8). However, this approximation ignores the indirect bounces (multi-bounces) signal within the invisible area. To make a distinction, we use B to represent the direct bounce model:

$$g(\vec{p}_i, \vec{p}_d, t) \approx Bf(\vec{p}) = \int \Upsilon(|\vec{p} - \vec{p}_i|, |\vec{p} - \vec{p}_d|) \cdot f(\vec{p}) \cdot \delta(|\vec{p} - \vec{p}_i| + |\vec{p} - \vec{p}_d| - t \cdot c)\, d\vec{p}\,. \tag{5.8}$$

In this equation, $g(\vec{p}_i, \vec{p}_d, t)$ stands for a single time response captured from given

illumination and detector positions. $\Upsilon(|\vec{p}-\vec{p}_i|,|\vec{p}-\vec{p}_d|)$ stands for the intensity drop-off term. The kernel function $\delta(|\vec{p}-\vec{p}_i|+|\vec{p}-\vec{p}_d|-t\cdot c)$ mainly describes the geometry of the integral. Based on whether illumination position $\vec{p}_i$ and detector position $\vec{p}_d$ are co-located or not, the geometry looks either spherically or ellipsoidally.

However, the actual forward model A is not equal to B only if the multibounce light is negligible. In fact, we can model the time response measurement $g(\vec{p})$ into two parts $g'(\vec{p})_{direct}$ and $g'(f(\vec{p}))_{multi}$.

The actual physically accurate model results to

$$g(\vec{p}_i,\vec{p}_d,t)=Af(\vec{p})=g'(\vec{p}_i,\vec{p}_d,t)_{direct}+g'(f(\vec{p}))_{multi}=\underbrace{Bf(\vec{p})}_{\textbf{Direct}}+\underbrace{g'(f(\vec{p}))}_{\textbf{MPI}} \quad (5.9)$$

It is obvious to see that the multibounce light depends on the unknown function f. Solving the entire inverse problem accounting for the multibounce signal is a bi-convex problem since A and f are both unknown. Most of the methods presented so far are trying to approximate $B^{-1}(g(\vec{p}_i,p_d,t))$ by numerical optimization or a linear combination of backprojection method. Both of them fail mathematically with the term $g'(f(p))$ in Eq. (5.9) causing some multibounce signal artifacts in the reconstruction volume.

3. Delta kernel model error Even though the model used before gives a well-approximated result (Eq. (5.8)), the direct bounce modeling still has an internal error in reality.

The kernel function $\delta(|\vec{p}-\vec{p}_i|+|\vec{p}-\vec{p}_d|-t\cdot c)$ in the integral (5.8) refers to infinite temporal resolution. Since inverse methods trying to solve the model refers to the infinite temporal resolution, based on our observation, this inaccurate modeling may harm the inverse solution more and more when the temporal resolution from the measurement becomes lower. This means the *inverse crime* happens.

In reality, the actual light pulse coming from the physical setup may have a specific temporal degradation or shape distortion resulting from the complex relay wall surface. Previous methods try to ignore the first returning pulse reflected from

the visible surface. We argue that it would be useful to use it to correct this delta modeling error.

The good thing for the approximation is that it makes the inverse process tractable by only adding extra physical constraints, such as looking at a simple patch or reducing the scene complexity. On the other hand, modeling a proper NLOS formation may result in the inverse becoming mathematically intractable. We want to address those difficulties and gaps for the readers for future research.

Three-dimensional Fourier slice sampling

In our **main text Section 4.1**, we mention the planar model for the NLOS forward measurement function. Here we provide a detailed model to represent the 3D Fourier slice sampling (**main text Section 5.1 refers to 2D case**).

Basic notation: $I(x, y, z)$ represents the unknown volume function in the cartesian coordinates x, y, z. $F(u, v, h)$ refers to its Fourier transform. $t(x, y, z)$ refers to the measurement (integral) data.

Cartesian to spherical coordinates: If we regard the center of the 3D unknown volume as the origin, ρ can be regarded as the distance of each plane integral and θ and ϕ directly refer to the direction of each projection integral.

For the measurement data:

$$g(\rho, \theta, \phi) = t(x, y, z)\,. \tag{5.10}$$

For the unknown volume function, $g(\rho, \theta, \phi)$ refers to the measurement data in spherical coordinates. To agree with the coordinate, we translate the coordinate for $F(u, v, h)$ to $G(\omega, \theta, \phi)$ as well.

3D Plane Radon integral: Based on the 3D Radon transform, the transformation between the original data $I(x, y, z)$ and measurement data in spherical coordinates $g(\rho, \theta, \phi)$ can be interpreted as follows:

$$g(\rho, \theta, \phi) = \iiint_{-\infty}^{\infty} I(x, y, z)\delta(\sin\theta\cos\phi x + \sin\theta\sin\phi y + \cos\theta z - \rho)\,dx\,dy\,dz\,. \tag{5.11}$$

Then, we calculate the 1D Fourier transform of the measurement $g(\rho, \theta, \phi)$ along ϕ as follows:

$$G(\omega, \theta, \phi) = \int_{-\infty}^{\infty} g(\rho, \theta, \phi)e^{-j2\pi\omega\rho}\,d\rho\,. \tag{5.12}$$

Replacing $g(\rho, \theta, \phi)$ with equation in 5.11 yields full expression.

$$G(\omega, \theta, \phi) = \int_{-\infty}^{\infty} \iiint_{-\infty}^{\infty} I(x, y, z) \\ \delta(\sin\theta\cos\phi x + \sin\theta\sin\phi y + \cos\theta z - \rho)\, dx\, dy\, dz e^{-j2\pi\omega\rho}\, d\rho\,. \quad (5.13)$$

Then the to make delta kernel $\delta(\sin\theta\cos\phi x + \sin\theta\sin\phi y + \cos\theta z - \rho)$ inside the equation become one, we can get rid of the ρ variable by using $\rho = \sin\theta\cos\phi x + \sin\theta\sin\phi y + \cos\theta z$. Then we simplify the equation:

$$G(\omega, \theta, \phi) = \iiint_{-\infty}^{\infty} I(x, y, z) e^{-j2\pi(\omega\sin\theta\cos\phi x + \omega\sin\theta\sin\phi y + \omega\cos\theta z)}\, dx\, dy\, dz\,, \quad (5.14)$$

whereas the Fourier transform of the unknown volume function $f(x, y, z)$ in Cartesian coordinate results as follows:

$$F(u, v, h) = \iiint_{-\infty}^{\infty} I(x, y, z) e^{-j2\pi(ux+vy+hz)}\, dx\, dy\, dz\,. \quad (5.15)$$

It is easy to see comparing Eqs. (5.14) and (5.15) that the 3D Fourier slice can be achieved as follows:

$$G(\omega, \theta, \phi) = F(\omega\sin\theta\cos\phi, \omega\sin\theta\sin\phi, \omega\cos\theta)\,. \quad (5.16)$$

From Eq. (5.16) we can see that, with a given projection direction, the Fourier transform of the spatial-temporal measurement $G(\omega, \theta, \phi)$ is a slice of the Fourier transform of the original data $F(u, v, h)$.

Planar model filtered backprojection method

The previous section mentioned the 3D Fourier slice form of the 3D planar Radon integral. Here we list the filter backprojection method for this type of integration. We keep the same notation as in the previous section.

To recover $I(x, y, z)$ from its Fourier spectrum $F(u, v, h)$, we can simply apply

the inverse Fourier transform as follows:

$$I(x,y,z) = \iiint_{-\infty}^{\infty} F(u,v,h) e^{j2\pi(ux+vy+hz)}\, du\, dv\, dh\,. \tag{5.17}$$

Rewriting Equation (5.17) by replacing the Cartesian coordinate with spherical coordinates yields

$$\begin{aligned} I(x,y,z) = \iiint_{-\infty}^{\infty} & F(\omega\sin\theta\cos\phi, \omega\sin\theta\sin\phi, \omega\cos\theta) \\ & e^{j2\pi(\omega\sin\theta\cos\phi x+\omega\sin\theta\sin\phi y+\omega\cos\theta z)}\omega^2\sin\theta\, d\omega\, d\theta\, d\phi\,. \end{aligned} \tag{5.18}$$

As we know from Equation (5.16), we could replace the term $F(\omega\sin\theta\cos\phi, \omega\sin\theta\sin\phi, \omega\cos\theta)$ in Equation (5.18) as follows:

$$I(x,y,z) = \iiint_{-\infty}^{\infty} G(\omega,\theta,\phi) e^{j2\pi(\omega\sin\theta\cos\phi x+\omega\sin\theta\sin\phi y+\omega\cos\theta z)}\omega^2\sin\theta\, d\omega\, d\theta\, d\phi\,. \tag{5.19}$$

We reuse the defined $\rho = \sin\theta\cos\phi x + \sin\theta\sin\phi y + \cos\theta z$ again to simplify the equation and rewrite Equation (5.18) as:

$$\begin{aligned} I(x,y,z) &= \iiint_{-\infty}^{\infty} G(\omega,\theta,\phi) e^{j2\pi\omega\rho}\omega^2\sin\theta\, d\omega\, d\theta\, d\phi \\ &= \int_0^{2\pi}\int_0^{\pi}[\int_{-\infty}^{\infty} \omega^2 \cdot G(\omega,\theta,\phi) e^{j2\pi\omega\rho}\, d\rho]\sin\theta\, d\theta\, d\phi \\ &= \int_0^{2\pi}\int_0^{\pi}[kernel * g(\rho,\theta,\phi)]\sin\theta\, d\theta\, d\phi \end{aligned} \tag{5.20}$$

Notice that in this equation, the final line $*$ represents the convolution operator which its frequency response is ω^2. Based on the definition, it is Laplacian filter.

6 CONCLUSIONS AND FUTURE WORK

Throughout this thesis, we design computational imaging solvers for Non-line-of-sight (NLOS) imaging using wave diffraction theory. By doing this, we have developed a method that yields a new class of imaging algorithms mimicking the various capabilities of line of sight cameras.

In Chapter 2, we drive new class of NLOS imaging algorithms rely on solving wave diffraction integral, namely the Rayleigh-Sommerfeld Diffraction (RSD) integral. We illustrate the robustness and versatility of our method by implementing three virtual NLOS imaging systems based on common LOS techniques: a conventional photography camera capable of imaging NLOS scenes without knowledge of the timing or location of the illumination source; a transient photography system capable of capturing transient videos of the hidden scene revealing higher-order interreflections beyond 3rd bounce; and a confocal time-gated imaging system robust to interreflections. The computational cost of our algorithm is bounded by the RSD solver computing the image formation model with complexity $O(N^3 \log(N))$. The examples shown highlight the primary benefit of our approach: By turning NLOS into a virtual LOS system, the intrinsic limitations of previous approaches no longer apply, enabling a new class of NLOS imaging methods that leverage existing wave-based imaging methods. Formulating NLOS light propagation as a wave does not impose limitations on the types of problems that can be addressed, nor the datasets that can be used. Any signal can be represented as a superposition of phasor field waves; our formulation thus can be viewed as a choice of basis to represent any kind of NLOS data. Expressing the NLOS problem this way allows to create a direct analogy to LOS imaging, which can be exploited to derive suitable imaging algorithms, and to implement them efficiently. In the future, we anticipate its application to other LOS imaging systems to, for instance, separate light transport into direct and global components, or utilize the phase of phasor field for enhanced depth resolution. Our virtual imaging system could also be used to create a *second* virtual imaging system to see around two corners, assuming

the presence of a secondary relay Lambertian surface in the hidden scene, or to select and manipulate individual light paths to isolate specific aspects of the light transport in different NLOS scenes. In that context, combining our theory with light transport inversions, via, e.g., an iterative approach, could potentially lead to better results, and is an interesting avenue of future work.

In Chapter 3, we introduce an NLOS reconstruction method using the phasor field formalism along with a convolutional Fast Fourier Transform (FFT) based Rayleigh Sommerfeld Diffraction (RSD) algorithm to provide fast non-approximative scene reconstructions for general capture setups, in particular including non-confocal setups using a single laser and a sensor array. Our hardware prototype includes a single-photon avalanche diode detector and a picosecond pulse laser which will be mentioned specifically later. When used in the confocal scenario, this new method performs at speed similar to LCT and FK Migration, while requiring significantly less memory. In addition to applying our new algorithm to open source data [3, 2], we also perform several additional experiments. To the best of our knowledge, our proposed method is the first to solve the general non-confocal NLOS imaging scenario with a similar time requirement and computational complexity as the fastest existing algorithms. In contrast to them, however, our method has much lower memory requirements. This allows us to reconstruct larger scenes and will enable implementation on embedded systems and GPU units where memory is limited. Knowledge of the relative layouts of the two grids may reduce or eliminate the requirement for working memory in the interpolation. One can also fine tune the trade-off between Fourier domain oversampling, more sophisticated interpolation methods, and reconstruction quality. It might also be possible to perform further down-sampling along the temporal dimension and use single instead of double precision variables to require less memory. These approaches are interesting topics for future research and can draw from considerable prior work on this problem in related FK Migration application areas. At present, however, the method takes several hundred times more memory than our proposed method. The LCT includes a similar re-sampling step that creates large memory requirements. Re-sampling and interpolation problems in this domain are studied in the literature covering planar

and spherical inverse radon transforms. We believe our method will enable real time NLOS imaging and reconstruction of large room scale scenes at full resolution.

In Chapter 4, we provide the Wigner Distribution Model for Phasor field NLOS imaging. Our work also provides an understanding of the connections between wave optics, Wigner Distribution Function, and NLOS imaging problems. We show that the phasor field NLOS imaging method with an exact Rayleigh-Sommerfeld Diffraction operator does not have any standard geometrical interpretations in the Wigner Distribution Function domain. The achievable theoretical lateral resolution from the Rayleigh-Sommerfeld Diffraction operator is given, and it can be evaluated numerically for different settings. For analytical purposes, it is possible to use the Fresnel diffraction as an approximation to model the phasor field propagation as a shear mapping in the Wigner Distribution Function domain if errors are considered. This error is different among different acquisition schemes, and it is less for confocal than for non-confocal measurements. Thus, the Fresnel approximations can be considered on confocal measurements. The differences between the confocal and non-confocal measurements are described by adding the illumination wavefront function into the phasor field model. This means that, theoretically, one can probe a hidden object's spatial frequency contents under non-confocal acquisitions by illuminating virtual point source from different positions on the relay wall. One can also apply concepts introduced in Sec. 4.1 to other linear NLOS imaging formation models in the Wigner Distribution Function domain. Our reliance on the Rayleigh-Sommerfeld Diffraction operator offers both advantages and disadvantages. The Rayleigh-Sommerfeld Diffraction operator is known to be hard to treat analytically. This is why past treatments of Wigner Distribution Function imaging rely on the Fresnel diffraction operator to have a simplified analytical equation. Unfortunately, this Fresnel approximation is not valid in most Non-Line-of-Sight reconstructions. Knowing theoretical limits of the exact solution is helpful, such as knowing the lateral resolution limit. Our work may stimulate further researches, such as the compressed Non-Line-of-Sight sampling, co-designing measurements and reconstruction methods, and other Non-Line-of-Sight imaging related areas.

In Chapter 5, we introduce a feature visibility model. Our statements apply to

any NLOS measurement that involves only the direct (3rd bounce) light from the hidden scene. Missing features may be reconstructed by algorithms by making use of higher order bounces, or missing information may be filled in using priors. This hole filling solution is different from deblurring and de-noising tasks common in imaging. The modulation transfer functions we derive are zero outside the Fourier cone (as opposed to just being very small like a Gaussian). This means deblurring methods based on deconvolution will fail, even for noiseless data. The limited aperture NLOS scenario is essential to be understood for future inverse method design. All the deterministic linear inverse methods can only recover the scene features which are contained in the limited measurement space. Overall, based on our local MTF sampling cone model, it is easy to see this limited aperture problem. By allowing any targets around the corner, some scene features may not be well represented, or completely missing in the incomplete measurement space. Future inverse method design should go beyond deterministic inverse methods by adding prior constraints to the inverse model to specifically account for this ill-posed limited aperture problem. In this work we provide only an approximate analytical model that we back up with numerical computations of the exact functions. An exact analytical expression for the Fourier cone is subject of further research.

BIBLIOGRAPHY

[1] Xiaochun Liu, Sebastian Bauer, and Andreas Velten. Phasor field diffraction based reconstruction for fast non-line-of-sight imaging systems. *Nature Communications*, 2020.

[2] David B. Lindell, Gordon Wetzstein, and Matthew O'Toole. Wave-based non-line-of-sight imaging using fast f-k migration. *ACM Trans. Graph. (SIGGRAPH)*, 38(4):116, 2019.

[3] Xiaochun Liu, Ibón Guillén, Marco La Manna, Ji Hyun Nam, Syed Azer Reza, Toan Huu Le, Adrian Jarabo, Diego Gutierrez, and Andreas Velten. Non-line-of-sight imaging using phasor-field virtual wave optics. *Nature*, 572(7771):620–623, 2019.

[4] Matthew O'Toole, David B Lindell, and Gordon Wetzstein. Confocal non-line-of-sight imaging based on the light-cone transform. *Nature*, 555(7696):338, 2018.

[5] Ramesh Raskar and James Davis. 5D time-light transport matrix: What can we reason about scene properties? *MIT Technical Report*, 2008.

[6] Ahmed Kirmani, Tyler Hutchison, James Davis, and Ramesh Raskar. Looking around the corner using ultrafast transient imaging. *International journal of computer vision*, 95(1):13–28, 2011.

[7] Andreas Velten, Thomas Willwacher, Otkrist Gupta, Ashok Veeraraghavan, Moungi G Bawendi, and Ramesh Raskar. Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging. *Nature Communications*, 3:745, 2012.

[8] Otkrist Gupta, Thomas Willwacher, Andreas Velten, Ashok Veeraraghavan, and Ramesh Raskar. Reconstruction of hidden 3D shapes using diffuse reflections. *Optics Express*, 20(17):19096–108, Aug 2012.

[9] Mauro Buttafava, Jessica Zeman, Alberto Tosi, Kevin Eliceiri, and Andreas Velten. Non-line-of-sight imaging using a time-gated single photon avalanche diode. *Optics Express*, 23(16):20997–21011, 2015.

[10] Shumian Xin, Sotiris Nousias, Kyros Kutulakos, Aswin Sankaranarayanan, Srinivasa Narasimhan, and Ioannis Gkioulekas. A theory of Fermat paths for non-line-of-sight shape reconstruction. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6800–6809, 2019.

[11] Byeongjoo Ahn, Akshat Dave, Ashok Veeraraghavan, Ioannis Gkioulekas, and Aswin C. Sankaranarayanan. Convolutional approximations to the general non-line-of-sight imaging operator. In *The IEEE International Conference on Computer Vision (ICCV)*, October 2019.

[12] Ji Hyun Nam, Eric Brandt, Sebastian Bauer, Xiaochun Liu, Eftychios Sifakis, and Andreas Velten. Real-time non-line-of-sight imaging of dynamic scenes. *arXiv preprint arXiv:2010.12737*, 2020.

[13] Miguel Galindo, Julio Marco, Matthew O'Toole, Gordon Wetzstein, Diego Gutierrez, and Adrian Jarabo. A dataset for benchmarking time-resolved non-line-of-sight imaging. In *International Conference on Computational Photography (ICCP)*. IEEE, 2019.

[14] Syed Azer Reza, Marco La Manna, Sebastian Bauer, and Andreas Velten. Phasor field waves: A huygens-like light transport model for non-line-of-sight imaging applications. *Optics Express*, 27(20):29380–29400, 2019.

[15] Justin Dove and Jeffrey H. Shapiro. Paraxial theory of phasor-field imaging. *Opt. Express*, 27(13):18016–18037, Jun 2019.

[16] Jeremy A Teichman. Phasor field waves: a mathematical treatment. *Optics Express*, 27(20):27500–27506, 2019.

[17] Syed Azer Reza, Marco La Manna, Sebastian Bauer, and Andreas Velten. Phasor field waves: experimental demonstrations of wave-like properties. *Optics Express*, 27(22):32587–32608, 2019.

[18] Justin Dove and Jeffrey H Shapiro. Nonparaxial phasor-field propagation. *Optics Express*, 28(20):29212–29229, 2020.

[19] Justin Justin Michael Dove. *Theory of phasor-field imaging*. PhD thesis, Massachusetts Institute of Technology, 2020.

[20] Chia-Yin Tsai, Kiriakos N Kutulakos, Srinivasa G Narasimhan, and Aswin C Sankaranarayanan. The geometry of first-returning photons for non-line-of-sight imaging. In *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 7216–7224, 2017.

[21] Sean I Young, David B Lindell, Bernd Girod, David Taubman, and Gordon Wetzstein. Non-line-of-sight surface reconstruction using the directional light-cone transform. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1407–1416, 2020.

[22] Marco La Manna, Fiona Kine, Eric Breitbach, Jonathan Jackson, Talha Sultan, and Andreas Velten. Error backprojection algorithms for non-line-of-sight imaging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(7):1615–1626, 2018.

[23] Chia-Yin Tsai, Aswin Sankaranarayanan, and Ioannis Gkioulekas. Beyond volumetric albedo – a surface optimization framework for non-line-of-sight imaging. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1545–1555, 2019.

[24] Julian Iseringhausen and Matthias B Hullin. Non-line-of-sight reconstruction using efficient transient rendering. *arXiv preprint arXiv:1809.08044*, 2018.

[25] Victor Arellano, Diego Gutierrez, and Adrian Jarabo. Fast back-projection for non-line of sight reconstruction. *Optics Express*, 25(10):11574–11583, 2017.

[26] Martin Laurenzis and Andreas Velten. Feature selection and back-projection algorithms for nonline-of-sight laser–gated viewing. *Journal of Electronic Imaging*, 23(6):063003, 2014.

[27] Xiaohua Feng and Liang Gao. Improving non-line-of-sight image reconstruction with weighting factors. *Optics letters*, 45(14):3921–3924, 2020.

[28] Christos Thrampoulidis, Gal Shulkind, Feihu Xu, William T Freeman, Jeffrey H Shapiro, Antonio Torralba, Franco NC Wong, and Gregory W Wornell. Exploiting occlusion in non-line-of-sight active imaging. *IEEE Transactions on Computational Imaging*, 4(3):419–431, 2018.

[29] Feihu Xu, Gal Shulkind, Christos Thrampoulidis, Jeffrey H Shapiro, Antonio Torralba, Franco NC Wong, and Gregory W Wornell. Revealing hidden scenes by photon-efficient occlusion-based opportunistic active imaging. *Optics express*, 26(8):9945–9962, 2018.

[30] Felix Heide, Matthew O'Toole, Kai Zang, David Lindell, Steven Diamond, and Gordon Wetzstein. Non-line-of-sight imaging with partial occluders and surface normals. *ACM Trans. Graph.*, 2019.

[31] Sheila W Seidel, John Murray-Bruce, Yanting Ma, Christopher Yu, William T Freeman, and Vivek K Goyal. Two-dimensional non-line-of-sight scene estimation from a single edge occluder. *arXiv preprint arXiv:2006.09241*, 2020.

[32] Adithya Pediredla, Akshat Dave, and Ashok Veeraraghavan. SNLOS: Non-line-of-sight scanning through temporal focusing. In *International Conference on Computational Photography*, volume 1, 2019.

[33] Mariko Isogawa, Dorian Chan, Ye Yuan, Kris Kitani, and Matthew O'Toole. Efficient non-line-of-sight imaging from transient sinograms. *arXiv preprint arXiv:2008.02787*, 2020.

[34] Adithya Kumar Pediredla, Mauro Buttafava, Alberto Tosi, Oliver Cossairt, and Ashok Veeraraghavan. Reconstructing rooms using photon echoes: A plane based model and reconstruction algorithm for looking around the corner. In *2017 IEEE International Conference on Computational Photography (ICCP)*, pages 1–12. IEEE, 2017.

[35] James Brooks and Daniele Faccio. A single-shot non-line-of-sight range-finder. *Sensors*, 19(21):4820, 2019.

[36] Jonathan Klein, Martin Laurenzis, Dominik L. Michels, and Matthias B. Hullin. A quantitative platform for non-line-of-sight imaging problems. In *British Machine Vision Conference 2018, BMVC 2018, Northumbria University, Newcastle, UK, September 3-6, 2018*, page 104, 2018.

[37] G Musarra, A Lyons, E Conca, Y Altmann, F Villa, F Zappa, MJ Padgett, and D Faccio. Non-line-of-sight 3D imaging with a single-pixel camera. *arXiv preprint arXiv:1903.04812*, 2019.

[38] Gabriella Musarra, Ashley Lyons, Enrico Conca, Federica Villa, Franco Zappa, Yoann Altmann, and Daniele Faccio. 3D RGB non-line-of-sight single-pixel imaging. In *Imaging and Applied Optics 2019 (COSI, IS, MATH, pcAOP)*, page IM2B.5. Optical Society of America, 2019.

[39] Marco Renna, Ji Hyun Nam, Mauro Buttafava, Federica Villa, Andreas Velten, and Alberto Tosi. Fast-gated 16-1 spad array for non-line-of-sight imaging applications. *Instruments*, 4(2):14, 2020.

[40] Marco La Manna, Ji-Hyun Nam, Syed Azer Reza, and Andreas Velten. Non-line-of-sight-imaging using dynamic relay surfaces. *Optics Express*, 28(4):5331–5339, 2020.

[41] Mariko Isogawa, Ye Yuan, Matthew O'Toole, and Kris M Kitani. Optical non-line-of-sight physics-based 3d human pose estimation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7013–7022, 2020.

[42] Javier Grau Chopite, Matthias B Hullin, Michael Wand, and Julian Iseringhausen. Deep non-line-of-sight reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 960–969, 2020.

[43] Christopher A Metzler, David B Lindell, and Gordon Wetzstein. Keyhole imaging: Non-line-of-sight imaging and tracking of moving objects along a single optical path at long standoff distances. *arXiv preprint arXiv:1912.06727*, 2019.

[44] Xiaochun Liu, Sebastian Bauer, and Andreas Velten. Analysis of feature visibility in non-line-of-sight measurements. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 10140–10148, 2019.

[45] Xiaochun Liu and Andreas Velten. The role of wigner distribution function in non-line-of-sight imaging. In *2020 IEEE International Conference on Computational Photography (ICCP)*, pages 1–12. IEEE, 2020.

[46] Ibón Guillén, Xiaochun Liu, Andreas Velten, Diego Gutierrez, and Adrian Jarabo. On the effect of reflectance on phasor field non-line-of-sight imaging. In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 9269–9273. IEEE, 2020.

[47] Jonathan Klein, Martin Laurenzis, Matthias B Hullin, and Julian Iseringhausen. Calibration scheme for non-line-of-sight imaging setups. *Optics Express*, 28(19):28324–28342, 2020.

[48] Martin Laurenzis. Investigation of some limitations of non-line-of-sight scene reconstruction. In *Emerging Imaging and Sensing Technologies for Security and Defence V; and Advanced Manufacturing Technologies for Micro-and Nanosystems in Security and Defence III*, volume 11540, page 115400U. International Society for Optics and Photonics, 2020.

[49] Daniele Faccio, Gordon G Wetzstein, and Andreas Velten. Non-line-of-sight imaging. *Nature reviews physics*, 2020.

[50] Martin Laurenzis. Shadows used to peer around corners, 2019.

[51] Tomohiro Maeda, Guy Satat, Tristan Swedish, Lagnojita Sinha, and Ramesh Raskar. Recent advances in imaging around corners. *arXiv preprint arXiv:1910.05613*, 2019.

[52] Felix Heide, Lei Xiao, Wolfgang Heidrich, and Matthias B Hullin. Diffuse mirrors: 3d reconstruction from diffuse indirect illumination using inexpensive time-of-flight sensors. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3222–3229, 2014.

[53] Ori Katz, Pierre Heidmann, Mathias Fink, and Sylvain Gigan. Non-invasive single-shot imaging through scattering layers and around corners via speckle correlations. *Nature photonics*, 8(10):784, 2014.

[54] Brandon M Smith, Matthew O'Toole, and Mohit Gupta. Tracking multiple objects outside the line of sight using speckle imaging. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6258–6266, 2018.

[55] Jonathan Klein, Christoph Peters, Jaime Martin, Martin Laurenzis, and Matthias B Hullin. Tracking objects outside the line of sight using 2d intensity images. *Scientific reports*, 6:32491, 2016.

[56] Jeremy Boger-Lombard and Ori Katz. Passive optical time-of-flight for non line-of-sight localization. *Nature communications*, 10, 2019.

[57] Florian Willomitzer, Fengqiang Li, Muralidhar Madabhushi Balaji, Prasanna Rangarajan, and Oliver Cossairt. High resolution non-line-of-sight imaging with superheterodyne remote digital holography. In *Computational Optical Sensing and Imaging*, pages CM2A–2. Optical Society of America, 2019.

[58] Prasanna Rangarajan, Florian Willomitzer, Oliver Cossairt, and Marc P Christensen. Spatially resolved indirect imaging of objects beyond the line of sight. In *Unconventional and Indirect Imaging, Image Reconstruction, and Wavefront Sensing 2019*, volume 11135, page 111350I. International Society for Optics and Photonics, 2019.

[59] Florian Willomitzer, Prasanna V Rangarajan, Fengqiang Li, Muralidhar M Balaji, Marc P Christensen, and Oliver Cossairt. Synthetic wavelength holography: An extension of gabor's holographic principle to imaging with scattered wavefronts. *arXiv preprint arXiv:1912.11438*, 2019.

[60] Christopher A Metzler, Felix Heide, Prasana Rangarajan, Muralidhar Madabhushi Balaji, Aparna Viswanath, Ashok Veeraraghavan, and Richard G Baraniuk. Deep-inverse correlography: towards real-time high-resolution non-line-of-sight imaging. *Optica*, 7(1):63–71, 2020.

[61] Charles Saunders, John Murray-Bruce, and Vivek K Goyal. Computational periscopy with an ordinary digital camera. *Nature*, 565(7740):472–475, 2019.

[62] Adam B Yedidia, Manel Baradad, Christos Thrampoulidis, William T Freeman, and Gregory W Wornell. Using unknown occluders to recover hidden scenes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 12231–12239, 2019.

[63] Sheila W Seidel, Yanting Ma, John Murray-Bruce, Charles Saunders, William T Freeman, C Yu Christopher, and Vivek K Goyal. Corner occluder computational periscopy: Estimating a hidden scene from a single photograph. In *2019 IEEE International Conference on Computational Photography (ICCP)*, pages 1–9. IEEE, 2019.

[64] Miika Aittala, Prafull Sharma, Lukas Murmann, Adam Yedidia, Gregory Wornell, Bill Freeman, and Frédo Durand. Computational mirrors: Blind inverse light transport by deep matrix factorization. In *Advances in Neural Information Processing Systems*, pages 14311–14321, 2019.

[65] David B. Lindell, Gordon Wetzstein, and Vladlen Koltun. Acoustic non-line-of-sight imaging. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6780–6789, 2019.

[66] Shawn Divitt, Dennis F Gardner, and Abbie T Watnik. Seeing around corners in the mid-infrared using speckle imaging. In *Frontiers in Optics*, pages FTh3F–3. Optical Society of America, 2019.

[67] Tomohiro Maeda, Yiqin Wang, Ramesh Raskar, and Achuta Kadambi. Thermal non-line-of-sight imaging. In *2019 IEEE International Conference on Computational Photography (ICCP)*, pages 1–11. IEEE, 2019.

[68] Kenichiro Tanaka, Yasuhiro Mukaigawa, and Achuta Kadambi. Polarized non-line-of-sight imaging. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2136–2145, 2020.

[69] Nicolas Scheiner, Florian Kraus, Fangyin Wei, Buu Phan, Fahim Mannan, Nils Appenrodt, Werner Ritter, Jurgen Dickmann, Klaus Dietmayer, Bernhard Sick, et al. Seeing around street corners: Non-line-of-sight detection and tracking in-the-wild using doppler radar. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2068–2077, 2020.

[70] Arnold Sommerfeld. *Electrodynamics: lectures on theoretical physics*, volume 3. Academic Press, 2013.

[71] Joseph W Goodman. *Introduction to Fourier optics*. Roberts and Company Publishers, 2005.

[72] Max Born and Emil Wolf. *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Elsevier, 2013.

[73] Mohit Gupta, Shree K Nayar, Matthias B Hullin, and Jaime Martin. Phasor imaging: A generalization of correlation-based time-of-flight imaging. *ACM Transactions on Graphics (ToG)*, 34(5):1–18, 2015.

[74] Achuta Kadambi, Hang Zhao, Boxin Shi, and Ramesh Raskar. Occluded imaging with time-of-flight sensors. *ACM Transactions on Graphics (ToG)*, 35(2):1–12, 2016.

[75] Pradeep Sen, Billy Chen, Gaurav Garg, Stephen R Marschner, Mark Horowitz, Marc Levoy, and Hendrik PA Lensch. Dual photography. In *ACM SIGGRAPH 2005 Papers*, pages 745–755. 2005.

[76] Matthew O'Toole, Felix Heide, Lei Xiao, Matthias B Hullin, Wolfgang Heidrich, and Kiriakos N Kutulakos. Temporal frequency probing for 5d transient analysis of global light transport. *ACM Transactions on Graphics (ToG)*, 33(4):1–11, 2014.

[77] Fabin Shen and Anbo Wang. Fast-Fourier-transform based numerical integration method for the rayleigh-sommerfeld diffraction formula. *Applied optics*, 45(6):1102–1110, 2006.

[78] Victor Nascov and Petre Cătălin Logofătu. Fast computation algorithm for the Rayleigh-Sommerfeld diffraction formula using a type of scaled convolution. *Applied optics*, 48(22):4310–4319, 2009.

[79] Jaakko Astola and Leonid Yaroslavsky. *Advances in signal transforms: theory and applications*, volume 7. Hindawi Publishing Corporation, 2007.

[80] Deyang Jiang, Xiaochun Liu, Jianwen Luo, Zhengpeng Liao, Andreas Velten, and Xin Lou. Ring and radius sampling based phasor field diffraction algorithm for non-line-of-sight reconstruction. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (01):1–1, 2021.

[81] Zhengpeng Liao, Deyang Jiang, Xiaochun Liu, Andreas Velten, Yajun Ha, and Xin Lou. Fpga accelerator for real-time non-line-of-sight imaging. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2021.

[82] Eugene P Wigner. On the quantum correction for thermodynamic equilibrium. In *Part I: Physical Chemistry. Part II: Solid State Physics*, pages 110–120. Springer, 1997.

[83] Philip Mayne Woodward. *Probability and Information Theory, with Applications to Radar: International Series of Monographs on Electronics and Instrumentation*, volume 3. Elsevier, 2014.

[84] Markus Testorf, Bryan Hennelly, and Jorge Ojeda-Castaneda. *Phase-Space Optics: Fundamentals and Applications: Fundamentals and Applications*. McGraw Hill Professional, 2009.

[85] Martin J Bastiaans. Application of the wigner distribution function to partially coherent light. *JOSA A*, 3(8):1227–1238, 1986.

[86] Martin J Bastiaans. The wigner distribution function applied to optical signals and systems. *Optics communications*, 25(1):26–30, 1978.

[87] Martin J Bastiaans. Application of the wigner distribution function in optics. *The Wigner Distribution—Theory and Applications in Signal Processing*, pages 375–426, 1997.

[88] Haldun M. Ozaktas, Zeev Zalevsky, and M. Alper Kutay. *The Fractional Fourier Transform with applications in optics and signal processing*. Wiley, 2001.

[89] Walter D Furlan, Manuel Martinez-Corral, Bahram Javidi, and Genaro Saavedra. Analysis of 3-d integral imaging displays using the wigner distribution. *Journal of Display Technology*, 2(2):180–185, 2006.

[90] Remo Ziegler, Simon Bucheli, Lukas Ahrenberg, Marcus Magnor, and Markus Gross. A bidirectional light field-hologram transform. In *Computer Graphics Forum*, volume 26, pages 435–446. Wiley Online Library, 2007.

[91] Zhengyun Zhang and Marc Levoy. Wigner distributions and how they relate to the light field. In *2009 IEEE International Conference on Computational Photography (ICCP)*, pages 1–10. IEEE, 2009.

[92] Roarke Horstmeyer, Se Baek Oh, and Ramesh Raskar. Iterative aperture mask design in phase space using a rank constraint. *Optics express*, 18(21):22545–22555, 2010.

[93] Martin J Bastiaans. The wigner distribution function and its applications to optics. In *AIP conference proceedings*, volume 65, pages 292–312. AIP, 1980.

[94] Mathias Fink. Time-reversed acoustics. *Scientific American*, 281(5):91–97, 1999.

[95] Minghua Xu and Lihong V Wang. Universal back-projection algorithm for photoacoustic computed tomography. *Physical Review E*, 71(1):016706, 2005.

[96] Jurij Tasinkevych and Ihor Trots. Circular Radon transform inversion technique in synthetic aperture ultrasound imaging: an ultrasound phantom evaluation. *Archives of Acoustics*, 39(4):569–582, 2015.

[97] Sunghwan Moon. On the determination of a function from an elliptical Radon transform. *Journal of Mathematical Analysis and Applications*, 416(2):724–734, 2014.

[98] Alexander H Delaney and Yoram Bresler. Globally convergent edge-preserving regularized reconstruction: an application to limited-angle tomography. *IEEE Transactions on Image Processing*, 7(2):204–221, 1998.

[99] Martin Benning, Christoph Brune, Marinus Jan Lagerwerf, and Carola-Bibliane Schönlieb. TGV sinogram inpainting for limited angle tomography. In *Technical Report*. University of Cambridge, 2015.

[100] Stanley R Deans. *The Radon transform and some of its applications*. Courier Corporation, 2007.

[101] Rim Gouia-Zarrad and Gaik Ambartsoumian. Approximate inversion algorithm of the elliptical Radon transform. In *Mechatronics and its Applications (ISMA), 2012 8th International Symposium on*, pages 1–4. IEEE, 2012.

[102] Jürgen Beyerer, Fernando Puente León, and Christian Frese. *Machine vision: Automated visual inspection: Theory, practice and applications*. Springer, 2015.