Single Exposure Quantitative Phase Imaging with a Conventional Microscope Using Diffusion Models

Gabriel della Maggiora [*,1,2,6], **Luis Alberto Croquevielle** [*,3], **Harry Horsley**[4], **Thomas Heinis**[3], **Artur Yakimovich**[1,2,5 †]

[1] Center for Advanced Systems Understanding (CASUS), Görlitz, Germany
[2] Helmholtz-Zentrum Dresden-Rossendorf e. V. (HZDR), Dresden, Germany
[3] Department of Computing, Imperial College London, London, United Kingdom
[4] Bladder Infection and Immunity Group (BIIG), UCL Centre for Kidney and Bladder Health, Division of Medicine, University College London, Royal Free Hospital Campus, London, United Kingdom
[5] Institute of Computer Science, University of Wrocław, Wrocław, Poland
[6] School of Computation, Information and Technology, Technical University of Munich, Germany
{g.della-maggiora valdes,a.yakimovich}@hzdr.de, aac622@ic.ac.uk, h.horsley@ucl.ac.uk

Abstract

Phase imaging is gaining importance due to its applications in fields like biomedical imaging and material characterization. In biomedical applications, it can provide quantitative information missing in label-free microscopy modalities. One of the most prominent methods in phase quantification is the Transport-of-Intensity Equation (TIE). TIE often requires multiple acquisitions at different defocus distances, which is not always feasible in a clinical setting, due to hardware constraints. To address this issue, we propose the use of chromatic aberrations to induce the required through-focus images with a single exposure, effectively generating a through-focus stack. Since the defocus distance induced by the aberrations is small, conventional TIE solvers are insufficient to address the resulting artifacts. We propose Zero-Mean Diffusion, a modified version of diffusion models designed for quantitative image prediction, and train it with synthetic data to ensure robust phase retrieval. Our contributions offer an alternative TIE approach that leverages chromatic aberrations, achieving accurate single-exposure phase measurement with white light and thus improving the efficiency of phase imaging. Additionally, we present a new class of diffusion models that are well-suited for quantitative data and have a sound theoretical basis. To validate our approach, we employ a widespread brightfield microscope equipped with a commercially available color camera. We apply our model to clinical microscopy of patients' urine, obtaining accurate phase measurements.

1 Introduction

Phase imaging has emerged as a crucial technique in many applications, including biomedical imaging (Park, Depeursinge, and Popescu 2018), label-free imaging characterization (Rivenson et al. 2019), and modality imaging conversion (Wang et al. 2024a; Zuo et al. 2020). It enhances contrast for objects with little or no absorption, such as biological structures, and provides quantitative information about the morphology of microscopic objects like cells, which is typically missing in label-free microscopy (Zangle and Teitell 2014). For example, this information can be used to detect morphological changes in cells (Yakimovich et al. 2018).

Several techniques have been developed for phase quantification. Among these, the Transport-of-Intensity Equation (TIE) (Teague 1983) stands out due to its straightforward approach. TIE does not require phase unwrapping and is easy to integrate into hardware (Gupta, Mahendra, and Nishchal 2020). It enables phase retrieval under partially coherent illumination (Paganin and Nugent 1998), which is common in clinical microscopes. However, a primary challenge in TIE is the presence of low-frequency artifacts (Wu et al. 2022; Zhu et al. 2014), often requiring multiple through-focus images, which can be time-consuming and technically demanding to get (Waller, Tian, and Barbastathis 2010; Zuo et al. 2013). Alternatively, regularization techniques (Lustig, Donoho, and Pauly 2007; Metzler et al. 2018) can be employed, but they are tailored to specific acquisition scenarios and may not capture all image features (Zuo et al. 2020).

In this work, we propose to perform phase imaging by leveraging chromatic aberrations produced in systems with broad-spectrum sources, enabling precise phase estimation with a single exposure. This approach can be implemented with a conventional microscope equipped with a commercially available polychromatic (RGB) camera, making it highly practical for clinical use. However, the use of white light introduces two challenges: a fixed effective defocus distance (Waller et al. 2010), and blurring in the captured diffraction patterns due to source incoherence (Zuo et al. 2020, 2015a). To address these issues, we propose a data-driven approach using a deep diffusion model (Ho, Jain, and Abbeel 2020; Saharia et al. 2022) trained on synthetic data (Yang et al. 2023a), reducing the reliance on real-world acquisitions (della Maggiora et al. 2022; Yang et al. 2023b). This model provides a robust and practical solution for phase retrieval with polychromatic sources. Our main contribu-

[*]These authors contributed equally.
[†]Corresponding author.

tions are:

- We propose a novel training paradigm for phase retrieval based on physics-based synthetic data generation (Yang et al. 2023a).
- We present a flexible and practical method for clinical phase imaging.
- We introduce Zero-Mean Diffusion, a variation of diffusion models, along with theoretical insights that explain its improved performance for phase retrieval tasks.

2 Related Work

The induced phase of an object's incident wave φ can be determined by analyzing its measured diffraction pattern I. For a monochromatic plane wave with wavelength λ and wavenumber $k = \frac{2\pi}{\lambda}$, the Transport-of-Intensity Equation (TIE), establishes a connection between the rate of change of the image in the propagation direction and the lateral phase gradient (Zuo et al. 2020):

$$-k\frac{\partial I(x,y;z)}{\partial z} = \nabla_{(x,y)}\cdot\left[I(x,y;z)\nabla_{(x,y)}\varphi(x,y;z)\right] \quad (1)$$

where (x, y) are lateral spatial coordinates and z is the defocus distance, and $z = 0$ is the center of the object. Equation (1) can be solved for I within the near field of a propagated distance z by the Fresnel diffraction integral. For a wavelength λ and denoting $\mathbf{x} = (x, y)$ this gives

$$I(\mathbf{x};z,\lambda) = \left|A(\mathbf{x})e^{i\varphi(\mathbf{x})} * \frac{e^{ikz}}{i\lambda z}e^{i\frac{k}{2z}\|\mathbf{x}\|_2^2}\right|^2, \quad (2)$$

where $A(\mathbf{x})$ denotes the amplitude. To recover the phase φ from (1), two approaches exist. First, assuming the object is predominantly non-absorptive allows for treating it as a phase object, such that $I(\mathbf{x}; z = 0) = I_0$ and (1) becomes:

$$-k\frac{\partial I(\mathbf{x};z)}{\partial z} = I_0\nabla^2_{\mathbf{x}}\varphi(\mathbf{x};z).$$

On the other hand, for an absorptive object, Teague's assumption (Teague 1983) yields an alternative formulation. In both cases, we can solve for the phase φ by using the Fourier transform in the lateral spatial coordinates and differential operators acting on $I(x, y; z)$ (see details in Appendix A).

Important numerical challenges render phase recovery an ill-posed problem. Techniques such as Tikhonov regularization (Sixou et al. 2013), Total Variation (Cheng et al. 2009), or Gaussian smoothing (Nakajima 1998) can mitigate instabilities, but they complicate the retrieval of low spatial frequency components. Additionally, approximating $\frac{\partial I(x,y;z)}{\partial z}$ through finite differences is susceptible to noise: a small distance yields better detail recovery at the cost of worse Signal-to-Noise Ratio (SNR), while larger distances worsen the paraxial approximation, under which the TIE is valid. (Zhu et al. 2014; Zuo et al. 2020). These issues can be addressed by increasing the number of acquisitions. While effective, this adds overhead to the method, making it less practical for real-world applications.

Another solution is to employ polychromatic illumination, such as white light. This configuration is useful for phase quantification because the intensity observed at the defocus plane can be interpreted as imaging at multiple defocus distances (Zuo et al. 2020), as we show in Figure 1. Nevertheless, a naive application of this idea results in a blurred diffraction pattern due to the superposition of the different defocus point spread functions at varying propagation distances. There is also blurring induced by temporal coherence arising from the superposition of diffuse spots at different axial propagation distances, resulting in an image with fewer details. Nonetheless, this configuration can work by using the wavelength-dependent characteristics of Fresnel diffraction and an RGB camera (Waller et al. 2010; Zuo et al. 2015b). This results in a single-exposure phase quantification method that effectively alleviates the problem of mechanical defocusing in TIE phase retrieval.

Deep learning has recently emerged as a powerful tool for addressing image reconstruction challenges (Wang et al. 2024b), including super-resolution (Li et al. 2021), inpainting (Lugmayr et al. 2022), and the reconstruction of undersampled MRI (Ahishakiye et al. 2021) and CT images (Li et al. 2022). In the field of Quantitative Phase Imaging (QPI) (Wang et al. 2024a), deep learning methods have demonstrated superior phase estimation compared to traditional 2-shot and single-image QPI techniques (Zhang et al. 2021; Wu et al. 2022). These image reconstruction problems often focus on minimizing the L_1 and L_2 norms, or other distance metrics. Despite their effectiveness, these methods usually provide blurry solutions and lack a way of modeling the uncertainty in the model (Lehtinen et al. 2018).

Generative models have gained popularity because they can produce more realistic reconstructions than regression methods. Additionally, they model the uncertainty in ill-posed inverse problems, providing a way to sample the probability distribution associated with the problem. Diffusion models are generative likelihood-based models that have been successfully applied to problems like under-sampled MRI image reconstruction (Chung and Ye 2022), deblurring (Li et al. 2023), image denoising (Chung et al. 2023), and phase retrieval (della Maggiora et al. 2023; Zhang et al. 2021), showing great stability during training and providing state-of-the-art sample quality.

To obtain the phase of an image, we propose using white light as the image source with a polychromatic (RGB) camera-based acquisition. Conveniently, microscopes with such setups are commercially available and widespread in clinical imaging. Since white light is polychromatic, we can quantify phase using the induced defocus stack. Solving the Transport-of-Intensity Equation (TIE) with these measurements is possible in principle because different wavelengths can be translated to different defocus distances under the same wavelength (Zuo et al. 2020; Waller et al. 2010). Mathematically, the TIE model given by (1) can be reformulated using an auxiliary variable $\xi = \lambda z$, since both z and λ are related in equation (1). Consequently, the TIE can be expressed in terms of ξ as follows (Waller et al. 2010):

$$-2\pi\frac{\partial I(\mathbf{x};\xi)}{\partial \xi} = \nabla_{\mathbf{x}}\cdot\big[I(\mathbf{x};\xi)\nabla_{\mathbf{x}}\varphi(\mathbf{x};\xi)\big]. \quad (3)$$

The main advantage of this formulation is that it allows to

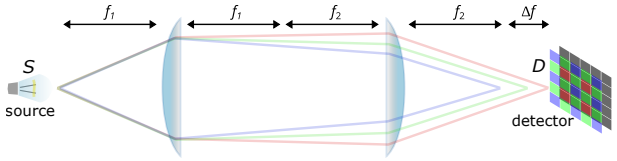

Figure 1: Example of optical system. The defocus distance Δf induced by chromatic aberration depends on the focal lengths f_1 and f_2. The resulting Δf for a specific wavelength λ corresponds to variable z in (2).

perform phase recovery while maintaining a constant defocus plane z (Waller et al. 2010). It also maintains the advantages of the previous TIE formulation, such as only requiring partial temporal coherence. Finally, to solve the equation for φ, we can use the color filtering that any digital color camera provides and obtain the required differentially defocused intensity images without the problems introduced by capturing multiple defocus planes.

To circumvent the issues present with using chromatic aberrations as the defocus stack, we propose the use of a diffusion model to solve this problem as an instance of conditioned probabilistic sampling and quantify the phase using polychromatic image acquisition with a white light-based source, achieving high accuracy.

3 Methods

Sections 3.1 and 3.2 describe our formulation for TIE and the design of a sample dataset for learning phase recovery. Sections 3.3 and 3.4 describe our use of diffusion models to solve phase recovery as an instance of conditioned sampling.

3.1 Forward Model

Consider a thin object $T(\mathbf{x}) = A(\mathbf{x})e^{i\phi(\mathbf{x})}$, measured in $z = 0$. The equation for the intensity image under partially incoherent illumination (Zuo et al. 2020) is

$$\begin{aligned} &I(\mathbf{x}; z, \lambda) \\ &\quad = \int S(\mathbf{u}) \left| \int T(\mathbf{x}') h_{z,\lambda}(\mathbf{x} - \mathbf{x}') e^{i2\pi \mathbf{u}\mathbf{x}'} \mathbf{dx}' \right|^2 \mathbf{du} \\ &\quad = \int S(\mathbf{u}) I_{\mathbf{u}}(\mathbf{x}) \mathbf{du}, \end{aligned}$$

where $S(\mathbf{u})$ and $I_{\mathbf{u}}$ correspond to the Fourier transforms of the source and the image, respectively, and h is the system's impulse response. This suggests that the observed intensity image is a superposition of all the coherent images obtained from the source. Under the paraxial approximation and assuming that the source is a coherent plane wave, that is, $S(\mathbf{u}) = \delta(\mathbf{u})$, the image can be obtained by the Fresnel integral given by equation (2).

In polychromatic acquisitions, the sensor is usually sensitive to a specific bandwidth. Generally speaking, a polychromatic sensor is built as an array of monochromatic sensors sensitive to three main components: red, green, and blue. Each has a different sensitivity to each wavelength. The sensor's quantum efficiency determines how much energy each frequency transforms into a signal. Given this model, the system can be expressed as a function of the power spectral distribution (PSD) $S_\lambda(\mathbf{u})$ (Zuo et al. 2020). Then, the intensity image is given by

$$\begin{aligned} &I(\mathbf{x}; z) \\ &\quad = \iint S_\lambda(\mathbf{u}) \left| \int T(\mathbf{x}') h_{z,\lambda}(\mathbf{x} - \mathbf{x}') e^{i2\pi \mathbf{u}\mathbf{x}'} \mathbf{dx}' \right|^2 d\mathbf{u}d\lambda \\ &\quad = \iint S_\lambda(\mathbf{u}) I_{\mathbf{u},\lambda}(\mathbf{x}) \mathbf{du} d\lambda. \end{aligned} \tag{4}$$

This means that the resulting intensity image is equal to the superposition of the different power densities within the defined bandwidth, i.e., the average of all images produced by the different coherent wavelength components.

3.2 Dataset Simulation

To simulate polychromatic acquisitions, we use the ImageNet dataset (Deng et al. 2009). Each image undergoes a grayscale conversion and is treated as a pure phase object. Then, we simulate the image under a fixed defocus distance z randomly selected between 0.1μm and 3μm. For the simulation, we use the Fresnel propagator for each wavelength ranging from 400nm to 700nm with equal contribution, discretizing with 6nm increments. For each increment, the source is viewed as a coherent plane wave. Since the quantum efficiency of the sensor influences the signal captured by the detector, we simulate a Gaussian distribution for each channel centered at $\lambda_r = 630$nm, $\lambda_g = 550$nm, and $\lambda_b = 450$nm respectively. We model the quantum efficiency of the sensor using the following function:

$$Q_c(\lambda) = e^{-\frac{(\lambda - \lambda_c)^2}{2\sigma_c^2}}$$

where σ_c is a parameter representing the sensitivity of the sensor to the corresponding wavelength. To model the PSD for a specific channel c we assume the model $S_\lambda(\mathbf{u}) = \delta_\lambda(\mathbf{u}) Q_c(\lambda)$. Replacing in (4), we get a single integral with respect to the λ variable, which is approximated as:

$$I_c(\mathbf{x}; z) = \frac{1}{W} \sum_{i=1}^{W} Q_c(\lambda_i) I(\mathbf{x}; z, \lambda_i),$$

where c denotes the respective channel. The wavelength range is divided into 6nm-sized intervals, and the left endpoint λ_i and size $1/W$ of each interval are used for a Riemann sum approximation of the integral in (4). Finally, white Gaussian noise is added to the image to simulate the noise in real acquisitions. This procedure allows us to construct a dataset of polychromatic acquisitions along with the correspondent phase of the images, which we use to train a deep learning model for phase reconstruction. Our methodology is based on diffusion models and explained in Sections 3.3 and 3.4. See Figure 12 in the appendix for more details.

3.3 Conditional Denoising Diffusion Models

Phase recovery is considered ill-posed, meaning that a given acquisition may be consistent with multiple different phases. Hence, we frame the problem as an instance of conditioned sampling, which explicitly accounts for the uncertainty of

the recovery. Suppose we have access to a dataset of input-output pairs, denoted as $\mathcal{D} = \{(X_i, Y_i)\}_{i=1}^N$. We model these samples as coming from an unknown joint distribution $q_{\text{data}}(X, Y)$. In our context, X corresponds to a sample drawn from polychromatic image acquisition, and Y corresponds to the phase of the image.

Our objective then becomes to learn an approximate way of sampling $Y \sim q_{\text{data}}(\cdot|X)$, for any given X. To do this we propose the use of Conditional Variational Diffusion Models (CVDM) (della Maggiora et al. 2023). In this framework, the basic setup is the existence of a *forward process* $\{Y_t\}_{t\in[0,1]}$, where $Y_0 \sim q_{\text{data}}(\cdot|X)$ and the evolution of the process is given by the following stochastic differential equation (SDE) (Song et al. 2021):

$$dY_t = -\frac{1}{2}\beta(t, X)Y_t dt + \sqrt{\beta(t, X)}dW_t, \tag{5}$$

where $\beta(t, X)$ is a *variance schedule* function. The intuitive idea is that noise is gradually added to an initial sample $Y_0 \sim q_{\text{data}}(\cdot|X)$, such that Y_1 at the end of the process is close to a $\mathcal{N}(0, I)$ variable. For this process, we denote the density at time t by $q_t(\cdot|X)$. The interesting part is learning to (approximately) reverse this forward process. This gives rise to a *reverse process* $\{Z_t\}_{t\in[0,1]}$, whose density at time t we denote by $p_t(\cdot|X)$, where we aim to have $p_t \approx q_{1-t}$.

Specifically, the reverse process starts with $p_0(\cdot|X) = q_1(\cdot|X)$. It is known that q_1 approximates a $\mathcal{N}(0, I)$ distribution, so this step is usually implemented as $Z_0 \sim \mathcal{N}(0, I)$. After sampling Z_0, the reverse dynamics (given by a different SDE) are simulated until we get $Z_1 \sim p_1(\cdot|X) \approx q_{\text{data}}(\cdot|X)$. This can be done in different ways (Ho, Jain, and Abbeel 2020; Song et al. 2021), usually via a score or noise-prediction model. In CVDM, a noise-prediction model is learned, and the reverse process is simulated via ancestral sampling (della Maggiora et al. 2023).

We now explain the way in which CVDM learns and simulates the reverse process. Given equation (5) it can be shown (della Maggiora et al. 2023) that the process $\{Y_t\}$ conditioned on Y_0 evolves such that

$$q_t(\cdot|Y_0 = Y, X) = \mathcal{N}\left(\sqrt{\gamma(t, X)}Y, (1 - \gamma(t, X))I\right) \tag{6}$$

where $\gamma(t, X) = e^{-\int_0^t \beta(s,X)ds}$. Here, γ is also referred to as a *variance schedule* function. It captures the longer-term dynamics of the forward process, while β controls the instantaneous rate of change according to (5). Interestingly, for any $s < t$ we have that $q_t(\cdot|Y_s = y_s, X)$ is also normally distributed, and Bayes' theorem implies that the posterior distribution is a Gaussian too (Kingma et al. 2021; della Maggiora et al. 2023):

$$\begin{aligned} &q_s(\cdot|Y_t = y_t, Y_0 = Y, X) \\ &\quad = \mathcal{N}\Big(\alpha(y_t, Y, s, t, X), \sigma(s, t, X)I\Big) \end{aligned} \tag{7}$$

for $s < t$. Here, the functions α and σ have an explicit form in terms of the schedule functions β and γ. In CVDM, this motivates the modeling of the reverse process such that

$$\begin{aligned} &p_{1-s}(\cdot|Z_{1-t} = y_t, X) \\ &\quad = \mathcal{N}\Big(\alpha(y_t, Y_\nu, s, t, X), \sigma(s, t, X)I\Big) \end{aligned} \tag{8}$$

for $s < t$. The only difference between (7) and (8) is that Y is unavailable in (8) and hence replaced by a learned model $Y_\nu(y_t, t, X)$. Equation (8) is the key tool used in CVDM to simulate the reverse process. Concretely, first a partition $(t_0, \ldots, t_N)$ of $[0, 1]$ is chosen. Then, $Z_0 \sim \mathcal{N}(0, I)$ is sampled. Finally, equation (8) is used to iteratively sample $Z_{t_{i+1}}$ given Z_{t_i}, until $Z_1 = Z_{t_N}$ is obtained.

The main thing left to explain is how the predictor Y_ν is trained. First, notice that equation (6) leads to the following parameterization. Conditioned on $Y_0 = Y$, the variable Y_t can be written as

$$Y_t(Y, \varepsilon) = \sqrt{\gamma(t, X)}Y + \sqrt{1 - \gamma(t, X)}\varepsilon \tag{9}$$

where $\varepsilon \sim \mathcal{N}(0, I)$. A natural approach to learn Y_ν is to maximize the log-likelihood term $\mathbb{E}_{(X,Y)\sim q_{\text{data}}}[p_1(Y|X)]$. CVDM and other diffusion models use a variational lower bound to get a more tractable expression. In the end, a diffusion loss term is obtained, which after some simplification corresponds to (della Maggiora et al. 2023):

$$\mathcal{L}_{\text{diff}}(X, Y) = \frac{1}{2}\mathbb{E}_{\varepsilon,t}\left[\|Y - Y_\nu(Y_t(Y, \varepsilon), t, X)\|_2^2\right].$$

This measures how well we can guess Y given the value of Y_t, when the process starts from $Y_0 = Y$. The goal then becomes to minimize $\mathbb{E}_{(X,Y)\sim q_{\text{data}}}[\mathcal{L}_{\text{diff}}(X, Y)]$. Experimental results (Ho, Jain, and Abbeel 2020) show that reparameterizing the model to guess the Gaussian noise ε improves accuracy. Hence, CVDM learns a noise-prediction model ε_ν, via a loss term $\mathcal{L}_{\text{noise}}$ equivalent to $\mathcal{L}_{\text{diff}}$.

A key advantage of CVDM in comparison to SR3 (Saharia et al. 2021) and other conditioning diffusion methods (Dhariwal and Nichol 2021) is that it allows for automatically learning the schedule functions. This avoids the need to fine-tune β and γ by hand, which is costly. However, when learning these functions it is necessary to ensure some basic properties, which are encoded in two loss terms $\mathcal{L}_\beta(X)$ and $\mathcal{L}_{\text{prior}}(X, Y)$. Moreover, a regularization term $\mathcal{L}_\gamma(X)$ is needed to avoid pathological schedules. The final loss function used in CVDM is hence given by

$$\begin{aligned} \mathcal{L}_{\text{CVDM}} = \mathbb{E}_{(X,Y)\sim q_{\text{data}}}\big[\\ \mathcal{L}_\beta(X) + \mathcal{L}_{\text{prior}}(X, Y) + \mathcal{L}_{\text{noise}}(Y, X) + a\mathcal{L}_\gamma(X)\big], \end{aligned}$$

where a is the weight of the regularization term $\mathcal{L}_\gamma$. In practice, a Monte Carlo estimator of $\mathcal{L}_{\text{CVDM}}$ is optimized by using the available $\{(X_i, Y_i)\}_{i=1}^N$ samples. For more details, see Appendix B.

3.4 Zero-Mean Diffusion

A standard practice in diffusion models is to normalize all data to the $[-1, 1]$ interval. This improves numerical stability but becomes an issue for quantitative imaging because the reconstructed data loses direct physical interpretation. In this work, we propose an alternative to normalization called Zero-Mean Diffusion (ZMD), which avoids this problem and has an interesting theoretical basis.

Concretely, we use a simple regression model to learn the expected value of the sample data, in conjunction with a diffusion model which predicts the difference between the regression model and the actual sample. This type of strategy

has been used before, with several works showing that modeling residuals instead of the full objective can be simpler and lead to better results (Li et al. 2021; Yang, Srivastava, and Mandt 2022). In addition, theoretical analysis suggests a reason for the high reconstruction quality achieved by ZMD.

As explained in Section 3.3, we are interested in sampling from a conditional distribution $q_{\text{data}}(\cdot|X)$. The approach via diffusion models is to define a stochastic forward process $\{Y_t\}_{t\in[0,1]}$, which evolves according to (5) and such that $q_0(\cdot|X) = q_{\text{data}}(\cdot|X)$. The forward process can be approximately reversed by starting at $Z_0 \sim p_0(\cdot|X)$ with $p_0 = \mathcal{N}(0, I) \approx q_1$, and then simulating the reverse dynamics until we get $Z_1 \sim p_1(\cdot|X)$ with $p_1 \approx q_0 = q_{\text{data}}$, which is the distribution of interest.

We now formally describe ZMD. We define a new stochastic (forward) process $\left\{\tilde{Y}_t\right\}$ given by $\tilde{Y}_t = Y_t - \mu_t$ for $t \in [0, 1]$ where $\mu_t = \mathbb{E}_{q_t(\cdot|X)}[Y_t]$ is the expected value of Y_t. Notice that, similar to our convention with Y_t and Z_t, we omit explicit reference to X when talking about $\tilde{Y}_t$ and μ_t. We do this to simplify notation since the conditioning data X remains fixed. An important property of the centered process $\left\{\tilde{Y}_t\right\}$ is that it also evolves according to (5).

Proposition 1. *The process $\left\{\tilde{Y}_t\right\}$ given by $\tilde{Y}_t = Y_t - \mu_t$ evolves according to the following SDE:*

$$d\tilde{Y}_t = -\frac{1}{2}\beta(t, X)\tilde{Y}_t dt + \sqrt{\beta(t, X)}dW_t. \tag{10}$$

The proof can be found in Appendix C. Proposition 1 provides an alternative sampling scheme for $q_0(\cdot|X)$ via diffusion models. The idea is to use CVDM as outlined in the previous section but applied now to the $\left\{\tilde{Y}_t\right\}$ process. We use an analogous notation for this method version so that $\tilde{q}_t(\cdot|X)$ denotes the density for this process at time t. Our main goal is learning to simulate a reverse process $\left\{\tilde{Z}_t\right\}$, characterized at time t by a density $\tilde{p}_t(\cdot|X)$ and where the key property is $\tilde{p}_t \approx \tilde{q}_{1-t}$.

Proposition 1 guarantees that the dynamics of $\left\{\tilde{Y}_t\right\}$ are exactly the same as those of $\{Y_t\}$, which means CVDM can be implemented in the same way as before, that is, by using the same parameterization and minimizing the loss function $\mathcal{L}_{\text{CVDM}}$. Nonetheless, there is a caveat. As described in the previous section, in practice, a Monte Carlo estimator of $\mathcal{L}_{\text{CVDM}}$ is minimized by using the available $\{(X_i, Y_i)\}_{i=1}^N$ samples. For ZMD, the Monte Carlo estimator would need access to $\{(X_i, Y_i - \mu_0)\}_{i=1}^N$ samples. Hence, ZMD needs the value of μ_0 or at least a good estimate.

Our solution is to define a learnable function $\mu_\theta(X)$ to approximate μ_0. To make sure the approximation is accurate, we use the following loss term

$$\mathcal{L}_{\text{mean}} = \mathbb{E}_{(X,Y)\sim q_{\text{data}}}\left[\|Y - \mu_\theta(X)\|_2^2\right],$$

which is now included in the full ZMD loss function:

$$\mathcal{L}_{\text{ZMD}} = \mathcal{L}_{\text{CVDM}} + \omega\mathcal{L}_{\text{mean}}, \tag{11}$$

where ω controls the relative weights of the terms. For $\mathcal{L}_{\text{mean}}$, the expected value is calculated with respect to $(X, Y) \sim q_{\text{data}}$. On the other hand, for $\mathcal{L}_{\text{CVDM}}$ the expected value is with respect to $(X, \tilde{Y}) \sim \tilde{q}_{\text{data}}$, where $\tilde{q}_{\text{data}}$ represents the distribution of $(X, Y - \mu_0)$. This is because CVDM is now learning to reverse the centered process $\left\{\tilde{Y}_t\right\}$. In practice, we optimize over many iterations. For each one, we simulate $(X, \tilde{Y}) \sim \tilde{q}_{\text{data}}$ by sampling (X, Y) from our dataset, and setting $\tilde{Y} = Y - \mu_\theta(X)$. As μ_θ improves, this procedure becomes increasingly better at approximating the sampling we need.

Once all models have been trained, we can use ZMD to approximately sample from $q_0(\cdot|X) = q_{\text{data}}(\cdot|X)$ by following an almost identical procedure as that of standard CVDM:

1. Equation (10) implies $\tilde{q}_1 \approx \mathcal{N}(0, I)$. Hence, a good approximation for $\tilde{p}_0$ is $\mathcal{N}(0, I)$.
2. After choosing a time discretization $(t_0, \ldots, t_N)$ of the $[0, 1]$ interval, equation (8) can be used used to iteratively sample $\tilde{Z}_{t_{i+1}}$ given $\tilde{Z}_{t_i}$, until getting $\tilde{Z}_1 = \tilde{Z}_{t_N}$. If the noise-prediction model $\tilde{\varepsilon}_\nu$ is accurate, then it holds that $\tilde{Z}_1 \sim \tilde{p}_1$ with $\tilde{p}_1 \approx \tilde{q}_0$.

The previous two steps produce a sample $\tilde{Z}_1 \mathrel{\dot\sim} \tilde{q}_0$. To get an (approximate) sample from the distribution of interest $q_0(\cdot|X)$, we compute $\tilde{Z}_1 + \mu_\theta(X)$. This can be understood intuitively. Since μ_0 (or its estimate μ_θ) is subtracted at the start of the forward process, it must be added back at the end of the reverse process. For the mathematical formalization of this idea, please see Appendix C.

As shown in Section 4, ZMD works very well as an alternative to the standard practice of normalizing data to the $[-1, 1]$ interval. Hence, it provides a robust way of using diffusion models for quantitative data. Besides these strong empirical results, there is a deeper theoretical consideration underlying the performance of ZMD. To explain this, we will use $\tilde{p}_1^{\mu_0}$ to denote the distribution of the random variable $\tilde{Z}_1 + \mu_0$ where $\tilde{Z}_1 \sim \tilde{p}_1$.

Standard CVDM (Section 3.3) can be used to sample from a distribution p_1, while ZMD can be used to sample from a distribution $\tilde{p}_1^{\mu_0}$. Both p_1 and $\tilde{p}_1^{\mu_0}$ should approximate $q_0(\cdot|X) = q_{\text{data}}(\cdot|X)$, so it is natural to ask which method leads to a better approximation. To answer this question, we could compare $D_{\text{KL}}(q_0||p_1)$ with $D_{\text{KL}}(q_0||\tilde{p}_1^{\mu_0})$. While these divergences cannot be computed analytically, there are known upper bounds for them, which we denote $M(p_1)$ and $M(\tilde{p}_1^{\mu_0})$, respectively. Our main theoretical result is that the upper bound for ZMD is always better.

Theorem 1. *For the stochastic processes as defined in Sections 3.3 and 3.4, and for any distribution q_0, it holds that $M(\tilde{p}_1^{\mu_0}) \leq M(p_1)$.*

The proof can be found in Appendix C. We conjecture that this theoretical advantage of using a centered forward process may be one of the reasons for the good performance we get when using ZMD. Notice that this only holds if we have access to an accurate model μ_θ of μ_0.

4 Results

To evaluate our model, we conduct three distinct experiments. First, we generate a set of synthetic samples analogous to those in the training set. Second, we acquire four

through-focus stacks (which allow us to compute a ground truth) and subsequently test our method on these images. Finally, we use an uncurated clinical dataset comprised of real-world clinical images. While ground truth data is unavailable for this dataset, we compare our results with reference values reported in the literature. For the first two experiments, we include 2-shot (using two images to measure the phase) acquisition methods as additional reference values. In all reported results, "2s" stands for the 2-shot modality, "p" for the polychromatic (single exposure) modality, and "MP" refers to the metrics obtained by the regression model μ_θ, which serves as a baseline.

4.1 Synthetic Images

To validate the model's performance, we use the HCOCO dataset (Cong et al. 2019) to simulate synthetic examples. Table 1 shows the results, and Figure 2 shows a qualitative comparison of the methods. We also measured in the polychromatic modality the difference of the regression (MP) baseline of ZMD against the full method. The regression baseline achieved a MS-SSIM of 0.95 $\pm$ 0.02 and a MAE of 0.06 $\pm$ 0.02.

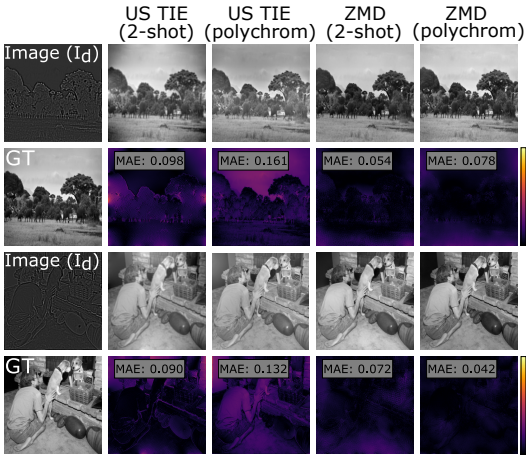

Figure 2: QPI methods assessed using synthetic images from HCOCO. The colorbar goes from 0 to 1

Metric / Model	US TIE (2s)	US TIE (p)	ZMD (2s)	ZMD (p)
MS-SSIM (↑)	0.91 $\pm$ 0.05	0.81 $\pm$ 0.09	0.94 $\pm$ 0.03	**0.97 $\pm$ 0.02**
MAE (↓)	0.11 $\pm$ 0.04	0.14 $\pm$ 0.05	0.08 $\pm$ 0.02	**0.06 $\pm$ 0.02**

Table 1: Performance metrics for Synthetic HCOCO (average value and standard deviation over dataset).

4.2 Through-focus Images

We assessed our approach using four stacks of clinical urine microscopy samples, each acquired as a through-focus series. The ground truth phase was obtained by solving the TIE using US TIE as solver (Zhang et al. 2020). To obtain an accurate ground-truth we estimated $\partial I(x, y; z)/\partial z$ using a 20th-degree polynomial fitted on each pixel of the volume (Waller, Tian, and Barbastathis 2010). The optimization was done considering the images at $d = \pm 2k\mu$m, taking k from 1 to 20. We compared our polychromatic method, utilizing a defocus distance of $z = 2\mu$m, against the 2-shot ZMD and US TIE methods, which were evaluated at $z = \pm 2\mu$m. Diffusion models were trained only on synthetic data. Figures 3 and 9 (Supplementary Material) display the reconstructions for all stacks, while Tables 2 and 3 present the MS-SSIM and MAE scores for each sample and method, corresponding to the rows in the figures. The first two samples are in Figure 3, and the next two are in Figure 9.

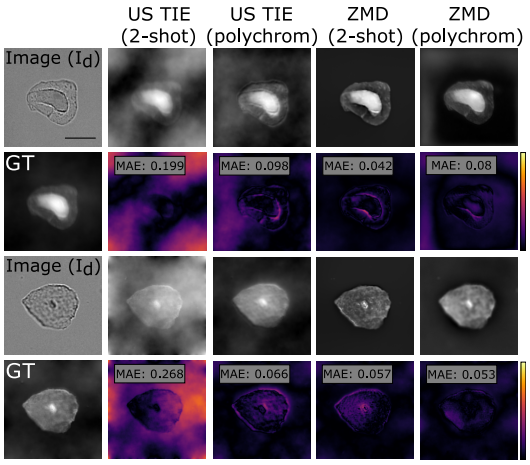

Figure 3: QPI methods assessed using clinical microscopy images depicting two overlapping (two upper rows) and an individual (two lower rows) epithelial cells. Scale bar (upper leftmost image) is 50 μm. The colorbar goes from 0 to 1

Sample	US TIE (2s)	US TIE (p)	MP (p)	ZMD (2s)	ZMD (p)
1	0.61	0.65	0.78	**0.91**	0.82
2	0.63	0.78	0.88	0.86	**0.89**
3	0.74	0.68	0.83	0.81	**0.88**
4	0.78	0.74	0.84	0.75	**0.86**

Table 2: MS-SSIM on Through-Focus Brightfield Images.

Sample	US TIE (2s)	US TIE (p)	MP (p)	ZMD (2s)	ZMD (p)
1	0.20	0.10	0.09	**0.04**	0.08
2	0.27	0.07	0.06	0.06	**0.05**
3	0.18	0.13	**0.10**	0.11	**0.10**
4	0.18	0.10	**0.09**	0.10	**0.09**

Table 3: MAE on Through-Focus Brightfield Images.

4.3 Clinical microscopy dataset

To evaluate the applicability of our proposed methodology to a real-world clinical context, we conduct experiments using the Urinary Tract Infection (UTI) dataset as outlined in (Liou et al. 2024). The images in this dataset were acquired for UTI patient condition evaluation in a point-of-care clin-

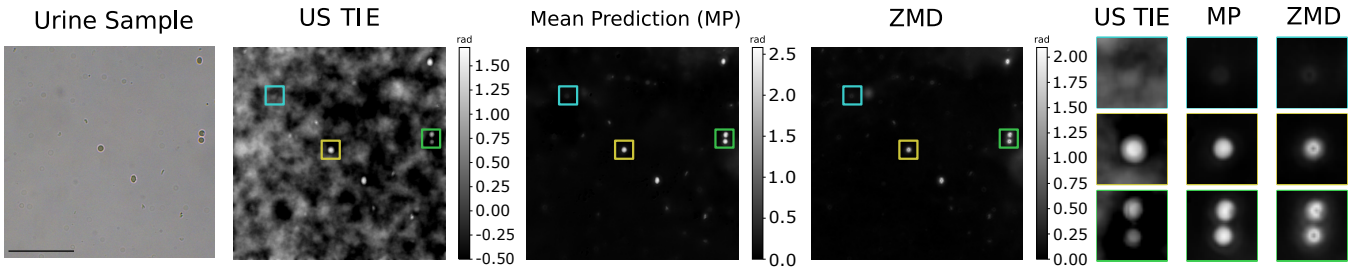

Figure 4: Urine sample with red blood cells. From left to right: US TIE phase reconstruction, Mean Prediction, and Zero-Mean Diffusion prediction. Regions of the image are enlarged to better show reconstruction details. Phase reconstructions are quantitative, with values between 0 and 2.5 radians. Scale bar is 200 μm.

ical setting. Image acquisition was performed in a through-focus manner with 200 defocused images separated by 1μm each using an Olympus BX41F microscope frame, U-5RE quintuple nosepiece, U-LS30 LED illuminator, U-AC Abbe condenser, and two types of objectives: an Olympus Plan 20x/0.40 Infinity/0.17 (used in Figure 10) and a UPlanFL N 20x/0.5 UIS2 Infinity/0.17/OFN26.5 (used in Figures 3, 4, 5, 6 and 9). Polychromatic images were taken with a commercially available digital scientific 16-bit color camera (Infinity 3S-1UR, Teledyne Lumenera) connected to the frame of the microscope using a 0.5x C-mount adapter. We calculate the phase on images with both a small and a high number of objects. Figure 5 shows the phase quantification performed by our model. Figure 4 shows the reconstruction for red blood cells using three different polychromatic methods: US TIE, Mean Prediction (that is, the regression model μ_θ of ZMD without the residual prediction), and ZMD. Crucially, our quantitative results are consistent with the values reported in the literature, that is, between 1.8 and 2.3 rad (Park et al. 2019; Nguyen et al. 2017).

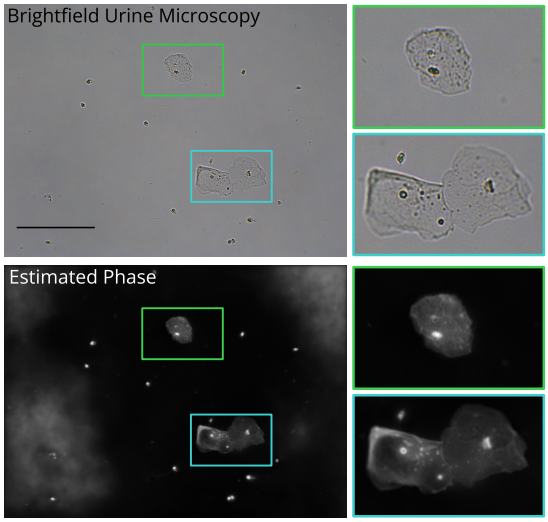

Figure 5: Phase quantification performed on clinical images. Regions of the images are enlarged to better show the objects in the image (epithelial cells). Scale bar (upper leftmost image) is 200 μm.

5 Discussion

Our method was tested in 2-shot and polychromatic modalities, with ZMD showing significantly better performance than US TIE in the latter (see Tables 1, 3, 4a, 4b). This highlights ZMD's clinical potential, capturing fine morphometric details of both flake-shaped epithelial cells and disk-shaped red blood cells in urine, including their biconcave morphology and overlapping cytoplasm. Results were validated using two microscopes with different objectives: a UPlanFL N 20x/0.5 UIS2 Infinity/0.17/OFN26.5 objective (Figures 3, 4, 5, 6, and 9) and an Olympus Plan 20x/0.40 Infinity/0.17 objective (Figure 10).

The improvement over the traditional 2-shot modality stems from the model's use of three RGB measurements, enabling better noise handling and more reliable estimation of the through-focus gradient and enhancing image quality.

In our experiments, ZMD recovered more structure than US TIE and Mean Prediction, uniquely revealing the ring cell morphology of red blood cells. As shown in Figures 3 and 4, US TIE produced cloud artifacts. Polychromatic acquisitions inherently exhibit spatial and temporal incoherence, causing some blurring of diffraction patterns and the quantified phase (e.g., Figure 3). Despite this trade-off, ZMD's improved noise handling and robust performance highlight its potential for medical imaging applications.

6 Conclusions

We present a novel method for reliably quantifying phase using chromatic aberrations rather than a purposely acquired through-focus stack. To solve this problem, we apply a denoising diffusion model designed for quantitative tasks. We show that this technique achieves high accuracy using a standard microscope frame with an RGB (color) camera, and our method's applicability is validated on clinical samples. As all necessary hardware is commercially available and commonly present in clinical microscopy setups, our approach holds significant promise for real-world clinical adoption.

Future work should focus on leveraging Physics-Informed Neural Networks (PINNs) to solve the Helmholtz equation, enabling more accurate wave modeling without the paraxial approximation. This would address current limitations and improve simulation realism. Additionally, creating a standardized QPI dataset is crucial for accurate evaluation and benchmarking of different methods.

Acknowledgments

This work was partially funded by the Center for Advanced Systems Understanding (CASUS), which is financed by Germany's Federal Ministry of Education and Research (BMBF) and by the Saxon Ministry for Science, Culture, and Tourism (SMWK) with tax funds on the basis of the budget approved by the Saxon State Parliament. The authors acknowledge the support of SMWK: ScaDS.AI and Chile's National Agency for Research and Development (ANID) through their Scholarship Program (DOCTORADO BECAS CHILE/2023 - 72230222).

References

Abadi, M.; Agarwal, A.; and et al., P. B. 2015. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. Software available from tensorflow.org.

Ahishakiye, E.; Van Gijzen, M. B.; Tumwiine, J.; Wario, R.; and Obungoloch, J. 2021. A survey on deep learning in medical image reconstruction. *Intelligent Medicine*, 1(3): 118–127.

Benton, J.; Bortoli, V.; Doucet, A.; and Deligiannidis, G. 2024. Nearly d-linear convergence bounds for diffusion models via stochastic localization. In *The Twelfth International Conference on Learning Representations*. OpenReview.

Cheng, H.; Liu, H.; Zhang, Q.; and Wei, S. 2009. Phase Retrieval Using the Transport-of-Intensity Equation. In *2009 Fifth International Conference on Image and Graphics*, 417–421.

Chung, H.; Kim, J.; Mccann, M. T.; Klasky, M. L.; and Ye, J. C. 2023. Diffusion Posterior Sampling for General Noisy Inverse Problems. arXiv:2209.14687.

Chung, H.; and Ye, J. C. 2022. Score-based diffusion models for accelerated MRI. arXiv:2110.05243.

Cong, W.; Zhang, J.; Niu, L.; Liu, L.; Ling, Z.; Li, W.; and Zhang, L. 2019. Image Harmonization Dataset iHarmony4: HCOCO, HAdobe5k, HFlickr, and Hday2night.

della Maggiora, G.; Castillo-Passi, C.; Qiu, W.; Liu, S.; Milovic, C.; Sekino, M.; Tejos, C.; Uribe, S.; and Irarrazaval, P. 2022. DeepSPIO: Super Paramagnetic Iron Oxide Particle Quantification Using Deep Learning in Magnetic Resonance Imaging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(1): 143–153.

della Maggiora, G.; Croquevielle, L. A.; Deshpande, N.; Horsley, H.; Heinis, T.; and Yakimovich, A. 2023. Conditional Variational Diffusion Models. In *The Twelfth International Conference on Learning Representations*.

Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, K.; and Fei-Fei, L. 2009. ImageNet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE.

Dhariwal, P.; and Nichol, A. 2021. Diffusion Models Beat GANs on Image Synthesis. ArXiv:2105.05233 [cs, stat].

Franzese, G.; Rossi, S.; Yang, L.; Finamore, A.; Rossi, D.; Filippone, M.; and Michiardi, P. 2023. How Much Is Enough? A Study on Diffusion Times in Score-Based Generative Models. *Entropy*, 25(4).

Gupta, A. K.; Mahendra, R.; and Nishchal, N. K. 2020. Single-shot phase imaging based on transport of intensity equation. *Optics Communications*, 477: 126347.

Ho, J.; Jain, A.; and Abbeel, P. 2020. Denoising Diffusion Probabilistic Models. In *Advances in Neural Information Processing Systems*, volume 33, 6840–6851. Curran Associates, Inc.

Karatzas, I.; and Shreve, S. 1991. *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics (113) (Book 113). Springer New York. ISBN 9780387976556.

Kingma, D.; Salimans, T.; Poole, B.; and Ho, J. 2021. Variational diffusion models. *Advances in neural information processing systems*, 34: 21696–21707.

Lehtinen, J.; Munkberg, J.; Hasselgren, J.; Laine, S.; Karras, T.; Aittala, M.; and Aila, T. 2018. Noise2Noise: Learning Image Restoration without Clean Data. arXiv:1803.04189.

Li, D.; Ma, L.; Li, J.; Qi, S.; Yao, Y.; and Teng, Y. 2022. A comprehensive survey on deep learning techniques in CT image quality improvement. *Medical & Biological Engineering & Computing*, 60(10): 2757–2770.

Li, H.; Yang, Y.; Chang, M.; Feng, H.; Xu, Z.; Li, Q.; and Chen, Y. 2021. SRDiff: Single Image Super-Resolution with Diffusion Probabilistic Models. *arXiv:2104.14951 [cs]*. ArXiv: 2104.14951.

Li, X.; Ren, Y.; Jin, X.; Lan, C.; Wang, X.; Zeng, W.; Wang, X.; and Chen, Z. 2023. Diffusion Models for Image Restoration and Enhancement – A Comprehensive Survey. arXiv:2308.09388.

Liou, N.; De, T.; Urbanski, A.; Chieng, C.; Kong, Q.; David, A. L.; Khasriya, R.; Yakimovich, A.; and Horsley, H. 2024. A clinical microscopy dataset to develop a deep learning diagnostic test for urinary tract infection. *Scientific Data*, 11: 155.

Lugmayr, A.; Danelljan, M.; Romero, A.; Yu, F.; Timofte, R.; and Van Gool, L. 2022. RePaint: Inpainting using Denoising Diffusion Probabilistic Models. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE.

Lustig, M.; Donoho, D.; and Pauly, J. M. 2007. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine*, 58(6): 1182–1195.

Metzler, C. A.; Schniter, P.; Veeraraghavan, A.; and Baraniuk, R. G. 2018. prDeep: Robust Phase Retrieval with a Flexible Deep Network. arXiv:1803.00212.

Nakajima, N. 1998. Phase-retrieval system using a shifted Gaussian filter. *Journal of the Optical Society of America A*, 15(2): 402.

Nguyen, T. H.; Kandel, M.; Shakir, H. M.; Best-Popescu, C.; Arikkath, J.; Do, M. N.; and Popescu, G. 2017. Halo-free Phase Contrast Microscopy. *Scientific Reports*, 7(1).

Nielsen, F. 2019. On the Kullback-Leibler divergence between location-scale densities. *arXiv preprint arXiv:1904.10428*.

Paganin, D.; and Nugent, K. A. 1998. Noninterferometric Phase Imaging with Partially Coherent Light. *Phys. Rev. Lett.*, 80: 2586–2589.

Park, H. S.; Eldridge, W. J.; Yang, W.-H.; Crose, M.; Ceballos, S.; Roback, J. D.; Chi, J.-T. A.; and Wax, A. 2019. Quantitative phase imaging of erythrocytes under microfluidic constriction in a high refractive index medium reveals water content changes. *Microsystems & Nanoengineering*, 5(1).

Park, Y.; Depeursinge, C.; and Popescu, G. 2018. Quantitative phase imaging in biomedicine. *Nature Photonics*, 12(10): 578–589. Number: 10 Publisher: Nature Publishing Group.

Rivenson, Y.; Liu, T.; Wei, Z.; Zhang, Y.; de Haan, K.; and Ozcan, A. 2019. PhaseStain: the digital staining of label-free quantitative phase microscopy images using deep learning. *Light: Science & Applications*, 8(1): 23. Number: 1 Publisher: Nature Publishing Group.

Ronneberger, O.; Fischer, P.; and Brox, T. 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. arXiv:1505.04597.

Saharia, C.; Chan, W.; Chang, H.; Lee, C.; Ho, J.; Salimans, T.; Fleet, D.; and Norouzi, M. 2022. Palette: Image-to-Image Diffusion Models. In *Special Interest Group on Computer Graphics*, 1–10. Vancouver BC Canada: ACM. ISBN 978-1-4503-9337-9.

Saharia, C.; Ho, J.; Chan, W.; Salimans, T.; Fleet, D. J.; and Norouzi, M. 2021. Image Super-Resolution via Iterative Refinement. arXiv:2104.07636.

Sixou, B.; Davidoiu, V.; Langer, M.; and Peyrin, F. 2013. Absorption and phase retrieval with Tikhonov and joint sparsity regularizations. *Inverse Problems &; Imaging*, 7(1): 267–282.

Song, Y.; Sohl-Dickstein, J.; Kingma, D. P.; Kumar, A.; Ermon, S.; and Poole, B. 2021. Score-Based Generative Modeling through Stochastic Differential Equations. *arXiv:2011.13456 [cs, stat]*. ArXiv: 2011.13456.

Teague, M. R. 1983. Deterministic phase retrieval: a Green's function solution. *J. Opt. Soc. Am.*, 73(11): 1434–1441.

Waller, L.; Kou, S. S.; Sheppard, C. J. R.; and Barbastathis, G. 2010. Phase from chromatic aberrations. *Optics Express*, 18(22): 22817–22825. Publisher: Optica Publishing Group.

Waller, L.; Tian, L.; and Barbastathis, G. 2010. Transport of Intensity phase-amplitude imaging with higher order intensity derivatives. *Opt. Express*, 18(12): 12552–12561.

Wang, K.; Song, L.; Wang, C.; Ren, Z.; Zhao, G.; Dou, J.; Di, J.; Barbastathis, G.; Zhou, R.; Zhao, J.; and Lam, E. Y. 2024a. On the use of deep learning for phase recovery. *Light: Science & Applications*, 13(1): 4. Number: 1 Publisher: Nature Publishing Group.

Wang, Z.; Yu, X.; Wang, C.; Chen, W.; Wang, J.; Chu, Y.-H.; Sun, H.; Li, R.; Li, P.; Yang, F.; Han, H.; Kang, T.; Lin, J.; Yang, C.; Chang, S.; Shi, Z.; Hua, S.; Li, Y.; Hu, J.; Zhu, L.; Zhou, J.; Lin, M.; Guo, J.; Cai, C.; Chen, Z.; Guo, D.; Yang, G.; and Qu, X. 2024b. One for Multiple: Physics-informed Synthetic Data Boosts Generalizable Deep Learning for Fast MRI Reconstruction. arXiv:2307.13220.

Wu, X.; Wu, Z.; Shanmugavel, S. C.; Yu, H. Z.; and Zhu, Y. 2022. Physics-informed neural network for phase imaging based on transport of intensity equation. *Optics Express*, 30(24): 43398–43416. Publisher: Optica Publishing Group.

Yakimovich, A.; Witte, R.; Andriasyan, V.; Georgi, F.; and Greber, U. F. 2018. Label-free digital holo-tomographic microscopy reveals virus-induced cytopathic effects in live cells. *MSphere*, 3(6): 10–1128.

Yang, Q.; Wang, Z.; Guo, K.; Cai, C.; and Qu, X. 2023a. Physics-Driven Synthetic Data Learning for Biomedical Magnetic Resonance: The imaging physics-based data synthesis paradigm for artificial intelligence. *IEEE Signal Processing Magazine*, 40(2): 129–140.

Yang, Q.; Wang, Z.; Guo, K.; Cai, C.; and Qu, X. 2023b. Physics-Driven Synthetic Data Learning for Biomedical Magnetic Resonance: The imaging physics-based data synthesis paradigm for artificial intelligence. *IEEE Signal Processing Magazine*, 40(2): 129–140.

Yang, R.; Srivastava, P.; and Mandt, S. 2022. Diffusion Probabilistic Modeling for Video Generation. ArXiv:2203.09481 [cs, stat].

Zangle, T. A.; and Teitell, M. A. 2014. Live-cell mass profiling: an emerging approach in quantitative biophysics. *Nature methods*, 11(12): 1221–1228.

Zhang, J.; Chen, Q.; Sun, J.; Tian, L.; and Zuo, C. 2020. On a universal solution to the transport-of-intensity equation. *Optics Letters*, 45(13): 3649–3652. Publisher: Optica Publishing Group.

Zhang, Y.; Noack, M. A.; Vagovic, P.; Fezzaa, K.; Garcia-Moreno, F.; Ritschel, T.; and Villanueva-Perez, P. 2021. PhaseGAN: a deep-learning phase-retrieval approach for unpaired datasets. *Optics Express*, 29(13): 19593–19604. Publisher: Optica Publishing Group.

Zhu, Y.; Shanker, A.; Tian, L.; Waller, L.; and Barbastathis, G. 2014. Low-noise phase imaging by hybrid uniform and structured illumination transport of intensity equation. *Optics Express*, 22(22): 26696–26711. Publisher: Optica Publishing Group.

Zuo, C.; Chen, Q.; Tian, L.; Waller, L.; and Asundi, A. 2015a. Transport of intensity phase retrieval and computational imaging for partially coherent fields: The phase space perspective. *Optics and Lasers in Engineering*, 71: 20–32.

Zuo, C.; Chen, Q.; Yu, Y.; and Asundi, A. 2013. Transport-of-intensity phase imaging using Savitzky-Golay differentiation filter - theory and applications. *Optics Express*, 21(5): 5346.

Zuo, C.; Li, J.; Sun, J.; Fan, Y.; Zhang, J.; Lu, L.; Zhang, R.; Wang, B.; Huang, L.; and Chen, Q. 2020. Transport of intensity equation: a tutorial. *Optics and Lasers in Engineering*, 135: 106187.

Zuo, C.; Sun, J.; Zhang, J.; Hu, Y.; and Chen, Q. 2015b. Lensless phase microscopy and diffraction tomography with multi-angle and multi-wavelength illuminations using a LED matrix. *Optics Express*, 23(11): 14314.