Image artifacts by TCC truncation

Kenji Yamazoe[*] and Kenneth L. Ho
TSMC Technology Incorporated, San Jose, California USA

ABSTRACT

In optical lithography image simulation, sum of coherent systems (SOCS) is used for efficient computation. In SOCS, the transmission cross coefficient is decomposed into eigenvalues and eigenfunctions, with only the leading eigenpairs used to compute the optical image while the others are ignored. This process, termed TCC truncation, can induce unfavorable image artifacts. The artifacts are usually minor but can sometimes cause unwanted results, such as negative image intensity and non-periodic image from periodic mask pattern. It is shown that these artifacts can be avoided by using a smaller grid size than the Nyquist grid size of the optical image. Since the artifacts originate from TCC truncation, the error caused by TCC truncation is mathematically analyzed.

Keywords: Partially coherent image formation, transmission cross coefficient, lithography image

1. INTRODUCTION

In optical lithography simulation, we essentially need to repeat millions of image simulations to precisely apply optical proximity correction (OPC) on a full-chip domain size. In this scenario, not only accuracy but also run-time has the highest priority in image computation. Various strategies are available to improve the run-time, including using GPUs [1], but we will focus only on the basic imaging algorithm here. Among many image simulation methods, the most suitable for optical lithography is to use the eigenvalues and eigenfunctions of the transmission cross coefficient (TCC) [2,3]. This approach, often called sum of coherent systems (SOCS) in lithography simulation [4], has two main benefits which enable fast computation. First, the eigenvalue spectrum typically exhibits rapid decay, so that we can approximate the image by using only, say, less than 10% of all eigenfunctions, corresponding to the largest eigenvalues. Since this process is equivalent to approximating the TCC by truncating minor eigenfunctions, it is called TCC truncation in this paper. Second, the eigenvalues and eigenfunctions are reusable, i.e., once calculated, they can be reused to efficiently image any mask pattern.

According to the optical image formation theory, the image formed in optical lithography is band-limited [5]. The band-limit of the optical image is twice as big as the extent of the pupil function. Therefore, we can define the largest grid size, i.e., the Nyquist grid size [6], for image computation as $\lambda/(4\mathrm{NA})$, where NA is the numerical aperture of the imaging system and λ is the quasi-monochromatic exposure wavelength. Equivalently, the spatial frequency band is from $-2\mathrm{NA}/\lambda$ to $2\mathrm{NA}/\lambda$, which defines the Nyquist frequency. All information of the optical image is confined within $\pm 2\mathrm{NA}/\lambda$ in the spatial frequency domain. This is the minimum requirement to correctly compute the optical image.

Here, we define the Nyquist sampling condition in two ways: (1) the spatial grid size is less than or equal to $\lambda/(4\mathrm{NA})$, or (2) the spatial frequency band is from $-\eta$ to η, where $\eta \geq 2\mathrm{NA}/\lambda$. In this paper, when we mention the Nyquist sampling condition, it means either one of these two. As long as the Nyquist sampling condition is satisfied, we can upsample the image to a smaller spatial grid whenever required. For example, with $\lambda = 193\mathrm{nm}$ and $\mathrm{NA} = 1.35$, the image can be calculated on a $35\mathrm{nm} < \lambda/(4\mathrm{NA})$ grid then upsampled to a 2nm grid while preserving all information in the original image.

However, TCC truncation sometimes induces artifacts in the upsampled optical image even if the original image was calculated under the Nyquist sampling condition. In this paper, four artifacts are shown: (1) negative intensity, (2) non-periodic image from periodic mask, (3) sharper image slope, and (4) loss of symmetry. These artifacts can be surprising and may even pose challenges for downstream image processing and analysis. In this paper, first, image formation theory is reviewed and computation method by TCC is explained in Sec. 2. Next, two artifacts caused by TCC truncation are illustrated in Sec. 3. A solution to avoid them is presented in Sec. 4, along with two more artifacts motivated by their analysis. In Sec. 5, the image approximation error by TCC truncation is studied mathematically and its relationship to the artifacts is discussed. Lastly, Sec. 6 summarizes the paper.

---

[*] kyamazoe@tsmc.com

Optical and EUV Nanolithography XXXIX, edited by Claire van Lare, Iacopo Mochi,
Proc. of SPIE Vol. 13979, 139790F · 2026 Published by SPIE
0277-786X · doi: 10.1117/12.3091052

2. IMAGE FORMATION BY TCC

In this section, the band-limit of the image and image computation by TCC are reviewed. Figure 1 is the imaging system assumed in this paper, which is a standard 4-*f* system (if we include the illumination optics, it would be more appropriately described as a 6-*f* system) [7]. The light source $S(f,g)$ is defined at the light source plane, which is imaged onto the projection optics pupil to form the effective light source. Since the projection pupil plane and the light source plane are under a conjugate relationship, the effective light source is also denoted as $S(f,g)$. On the object plane, the mask with transmittance $m(x,y)$ is set. The optical image $I(x,y)$ is formed on the wafer plane. The mask plane and wafer plane are under a conjugate relationship. The projection optics pupil plane and wafer plane are a Fourier transform pair. Real space is denoted by (x,y) and spatial frequency is denoted by (f,g).

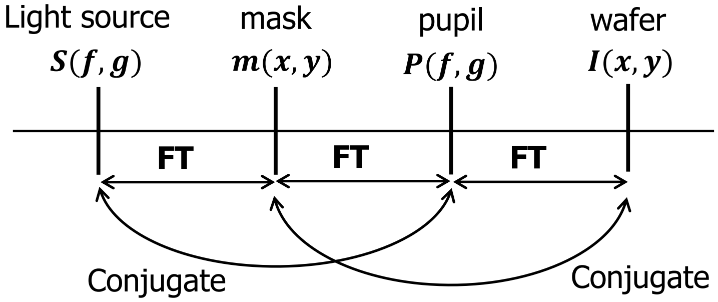


Figure 1. 4-*f* imaging system assumed in this paper. **FT** denotes the Fourier transform.

Without loss of generality, we assume scalar imaging. Then $I(x,y)$ can be written as [8]

$$I(x,y) = \iint S(f',g')\, |\mathbf{FT}^{-1}[\widetilde{m}(f-f',g-g')\, P(f,g)]|^2\, df'dg', \tag{1}$$

where **FT** denotes the Fourier transform and

$$P(f,g) = \mathrm{circ}\left(\frac{\lambda}{\mathrm{NA}}\sqrt{f^2+g^2}\right), \tag{2}$$

$$\widetilde{m}(f,g) = \mathbf{FT}[m(x,y)]. \tag{3}$$

The circ function is defined as $\mathrm{circ}(r) = 1$ for $r \le 1$ and zero otherwise. Therefore, $P(f,g) = 0$ if $\sqrt{f^2+g^2} > \mathrm{NA}/\lambda$. The autocorrelation theorem [9] imposes that $\mathbf{FT}[I(x,y)] = 0$ if $\sqrt{f^2+g^2} > 2\mathrm{NA}/\lambda$. This is the spatial frequency band-limit of the optical image. When we perform discrete Fourier transform (DFT) in computing the image, the spatial frequency extent must be at least as large as this band-limit; otherwise the image will have aliasing effect. Consequently, the largest grid size in real space is $\lambda/(4\mathrm{NA})$, which we refer to as the Nyquist grid size.

In Eq. (1), we now shift the pupil function $P(f,g)$ instead of the mask diffraction $\widetilde{m}(f,g)$ and switch the order of integration [10]. We obtain

$$I(x,y) = \iint \iint T(f',g',f'',g'')\, \widetilde{m}(f',g')\, \widetilde{m}^*(f'',g'')\, e^{i2\pi[(f'-f'')x+(g'-g'')y]}\, df'dg'\, df''dg'', \tag{4}$$

where $T(f',g',f'',g'')$ is the transmission cross coefficient (TCC) defined as

$$T(f',g',f'',g'') = \iint S(f,g)\, P(f'+f,g'+g)\, P^*(f''+f,g''+g)\, df dg. \tag{5}$$

The TCC can be decomposed by Mercer expansion [11] as

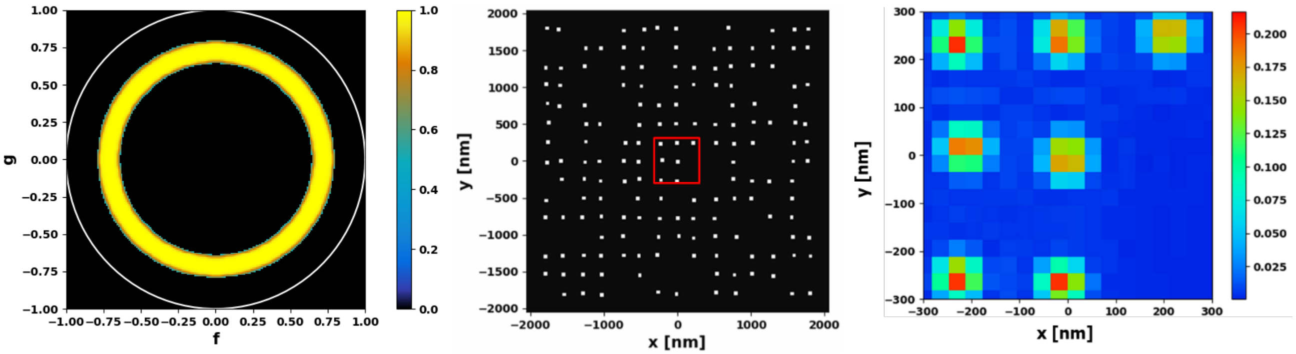
$$T(f',g',f'',g'') = \sum_i \lambda_i\, \phi_i(f',g')\, \phi_i^*(f'',g''), \tag{6}$$

where $\lambda_1 \ge \lambda_2 \ge \cdots \ge 0$ are its eigenvalues with corresponding eigenfunctions $\phi_i(f,g)$, commonly referred to as TCC kernels. Substituting Eq. (6) into Eq. (4), the final imaging equation can be written in SOCS form as

$$I(x,y) = \sum_i \lambda_i \, |\mathbf{FT}^{-1}[\phi_i(f,g)\, \tilde{m}(f,g)]|^2 \, . \qquad (7)$$

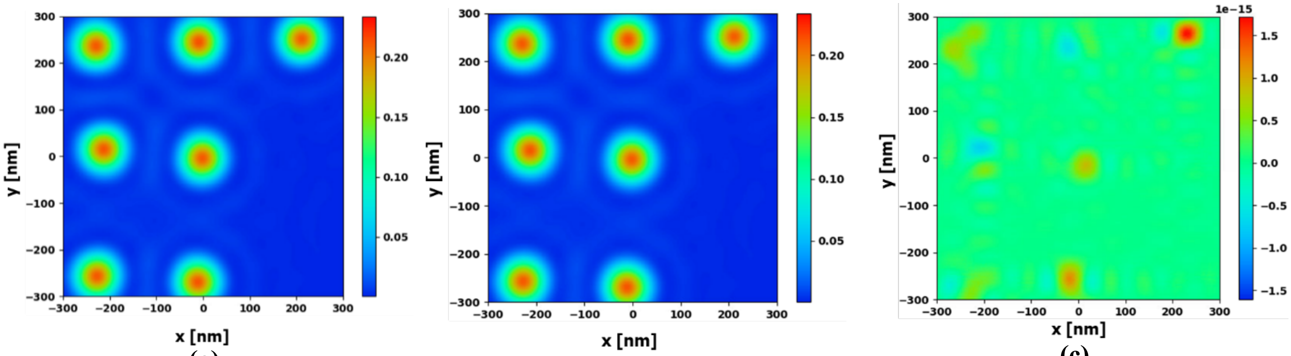
Observe that the image is guaranteed to be real and non-negative.

Equations (1) and (7) are implemented in Python and image test results are shown below. First, we assume $\lambda = 193\text{nm}$ and $\text{NA} = 1.35$ with water immersion throughout this paper. The simulation conditions and the simulated image are shown in Fig. 2. The illumination is the standard 4/5 annular illumination with outer radius $\sigma_{\max} = 0.8$, which will be used throughout this paper. The mask is a random contact hole pattern with a critical dimension (CD) of 64nm. We use a binary mask with Kirchhoff approximation [12] or so-called thin mask approximation, which will be assumed in this paper. The image is calculated by Eq. (1) with a grid size of $35.3\text{nm} < \lambda/(4\text{NA})$ as shown in Fig. 2(c). Figure 2(c) can be upsampled, for example, by Fourier transform with zero-padding. This upsampling scheme is assumed later in Sec. 5. Figure 3(a) is the upsampled image of Fig. 2(c) with a grid size of 2nm. The same process is repeated with Eq. (7) to obtain Fig. 3(b). Since we used all TCC eigenfunctions to compute Fig. 3(b), the difference between Figs. 3(a) and (b) is essentially zero as shown in Fig. 3(c). Next, let us examine the band-limit of the image. We take the Fourier transform of Fig. 3(a) to obtain the spectrum of the image, which is complex. When we plot the area where the absolute value of the image spectrum is greater than zero (actually, machine precision), we obtain Fig. 4. For visual clarity, a white circle with radius $2\text{NA}/\lambda$ is also shown. We see the image spectrum is confined inside the white circle, which confirms the theoretical band-limit.



(a) **(b)** **(c)**

Figure 2. (a) Illumination. (b) Random contact hole pattern. (c) Image inside the red box with a grid size of 35.3nm.



(a) **(b)** **(c)**

Figure 3. (a) Upsampled image by Eq. (1). (b) Upsampled image by Eq. (7). (c) Difference between (a) and (b).

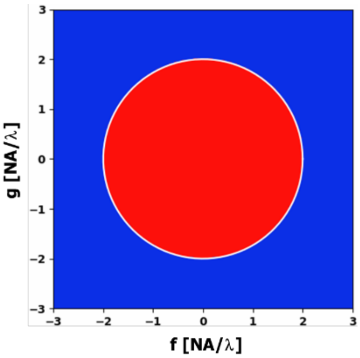
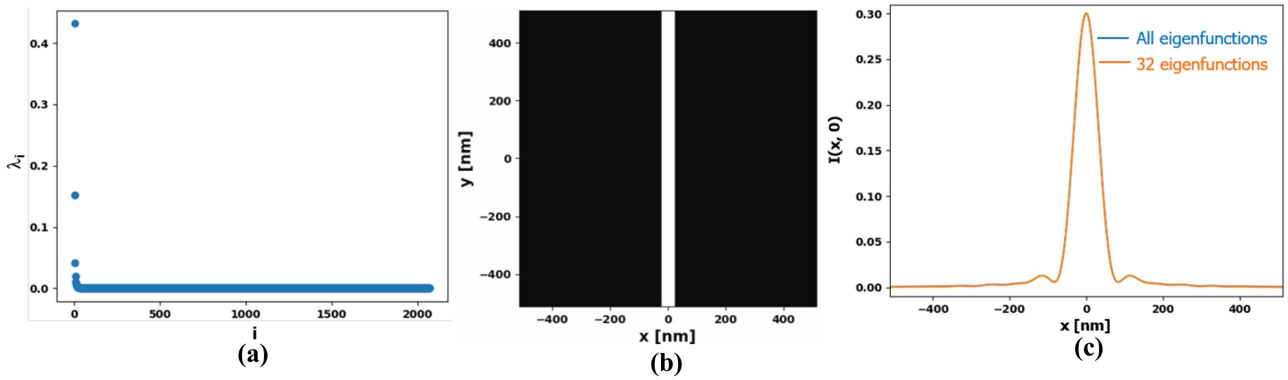


Figure 4. Image band-limit.

3. ARTIFACTS BY TCC TRUNCATION

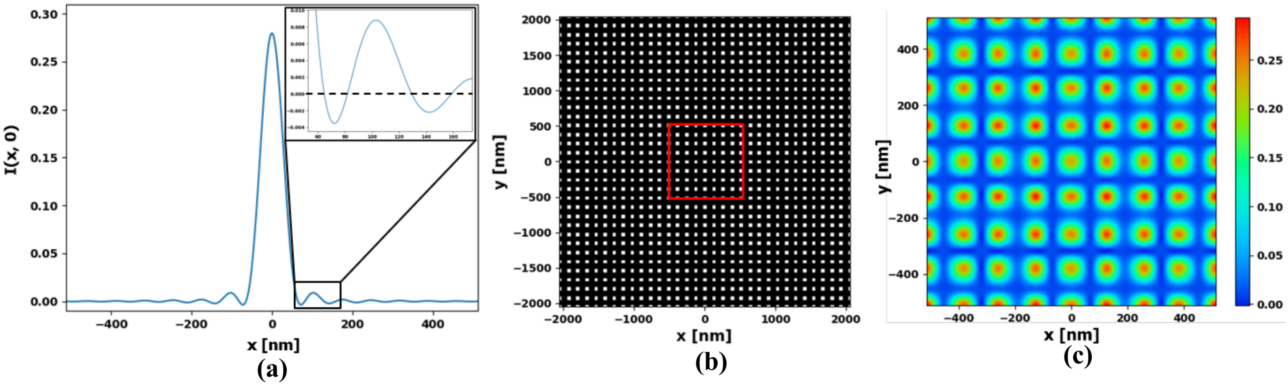
In this section, we explain the TCC truncation process and demonstrate some artifacts caused by its use. Consider again the example in Sec. 2. The associated TCC eigenvalues, normalized without loss of generality such that they sum to one, are plotted in Fig. 5(a). Since the TCC eigenfunctions are orthonormal, the magnitude of each eigenvalue decides the contribution of the corresponding eigenfunction to the full TCC. For example, the first eigenvalue $\lambda_1 \approx 0.43$ accounts for nearly 43% of the TCC (as measured in the spectral norm). The sum of the 1st through 32nd eigenvalues is 0.94. Therefore, if we use 32 eigenfunctions ($< 1.5\%$ of all 2072 in total), we can capture 94% of the TCC. To confirm this effect, a simple simulation is done. We use the mask pattern depicted in Fig. 5(b), which is an isolated line with a CD of 50nm. Images computed using all eigenfunction vs. 32 eigenfunctions are shown in Fig. 5(c). The difference is $O(10^{-4})$. Thus, TCC truncation can provide efficient computation with practical accuracy.



(a) (b) (c)

Figure 5. (a) Normalized TCC eigenvalues for the simulation condition in Sec. 2. (b) Isolated line pattern. (c) Images calculated by different numbers of TCC eigenfunctions (curves are very nearly coincident).

However, TCC truncation can introduce some unexpected artifacts. Here, we present two examples. In the following, we consider the extreme case of only one eigenfunction in order to most clearly demonstrate the phenomena. As before, we compute the image on a 35.3nm < $\lambda/(4\text{NA})$ grid and then upsample to a 2nm grid. The first artifact is negative intensity. As shown in Fig. 6(a) for the mask in Fig. 5(b), the image has non-negligible negative value, which violates a fundamental assumption and may disturb image processing at a later stage. The second artifact is non-periodic image from periodic mask pattern. The mask is set as in Fig. 6(b), which is a square contact hole array with a CD of 64nm and a pitch of 128nm. The simulation domain is set to 4096nm in each dimension. Since the mask pattern is periodic, the image should be periodic, too. However, the image is not periodic as seen in Fig. 6(c).



(a) (b) (c)

Figure 6. (a) Negative image intensity. (b) Periodic contact hole array. (c) Image inside the red box

4. CAUSE OF ARTIFACTS

In this section, we investigate the cause of the image artifacts in Sec. 3 and how to avoid them. First, we remark that artifacts do not appear if all eigenfunctions are used. Therefore, the artifacts must originate from TCC truncation. By Eq. (5), since the TCC may be non-zero for $\sqrt{(f')^2 + (g')^2} > \mathrm{NA}/\lambda$ and similarly for (f'', g'') by symmetry, we deduce that its eigenfunctions are, in general, also non-zero in $\sqrt{f^2 + g^2} > \mathrm{NA}/\lambda$. To confirm this, the first eigenfunction is plotted in Fig. 7. The black circle indicates $\sqrt{f^2 + g^2} = \mathrm{NA}/\lambda$. The eigenfunction clearly spreads outside this, up to $\sqrt{f^2 + g^2} = (1 + \sigma_{\max})\mathrm{NA}/\lambda$ according to Eq. (5). Autocorrelation theory applied to Eq. (7) then indicates that the TCC-truncated image has band-limit $\sqrt{f^2 + g^2} = 2(1 + \sigma_{\max})\mathrm{NA}/\lambda$, which is wider than the standard image Nyquist band-limit by a factor of $1 + \sigma_{\max}$. In calculating the images in Figs. 6(a) and (c), only the standard Nyquist band-limit was used. As a result, aliasing occurred, causing the artifacts observed. For reference, if we had used all eigenfunctions, then the image band-limit would reduce to $\sqrt{f^2 + g^2} = 2\mathrm{NA}/\lambda$ because Eq. (7) is algebraically equivalent to Eq. (1) in that setting. In other words, only when we use all eigenfunctions does the standard image Nyquist band-limit apply.

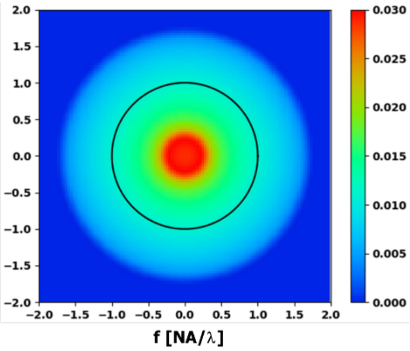


Figure 7. The first eigenfunction used to compute the images in Fig. 6.

The higher spatial frequency of the TCC-truncated image creates an additional side effect. According to Bernstein's inequality [13], the image slope satisfies

$$|\nabla I(x,y)| \leq \pi B \max[I(x,y)], \tag{8}$$

where B is the band-limit of the image, i.e., $B = 2\mathrm{NA}/\lambda$ without truncation and $B = 2(1 + \sigma_{\max})\mathrm{NA}/\lambda$ with truncation. Thus, TCC truncation can yield an artificially inflated slope. However, the effect can be subtle: while more eigenfunctions will typically decrease B (see Sec. 5), it can also increase $\max[I(x,y)]$. The end result is a balance between these two competing terms. In our tests, strong off-axis illumination tends to show higher image slope with fewer eigenfunctions. As an example, the mask in Fig. 5(b) is used again and the image is computed with either all eigenfunctions or just the first one. The image slopes are shown in Fig. 8. We see that the image with one eigenfunction has higher slope although its maximum intensity is weaker; compare Figs. 5(c) and 6(a). Notably, this effect happens even when we use the full band-limit $\sqrt{f^2 + g^2} = 2(1 + \sigma_{\max})\mathrm{NA}/\lambda$. This is an artifact of TCC truncation but not aliasing, unlike those in Sec. 3.

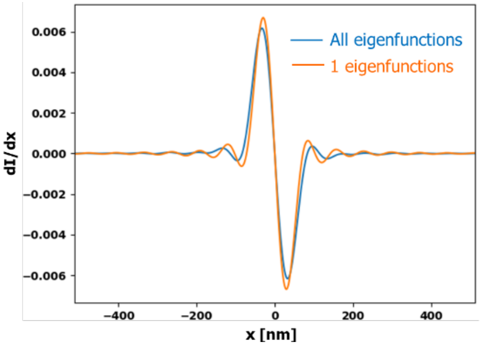


Figure 8. Image slope comparison.

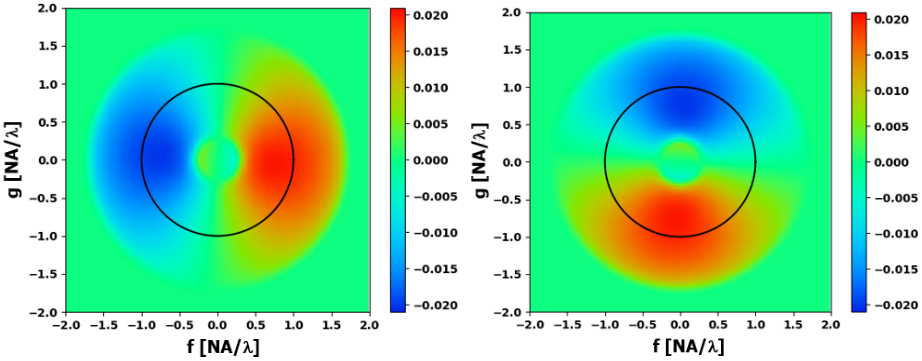


Figure 9. (a) The second eigenfunction. (b) The third eigenfunction.

Another possible non-aliasing artifact is loss of image symmetry. If the simulation setup, especially the illumination, is symmetric, then the TCC will often have repeated eigenvalues, corresponding to a degenerate subspace that reflects some aspect of that symmetry. For example, in Fig. 5(a), $\lambda_2 = \lambda_3 \approx 0.15$. The corresponding eigenfunctions, displayed in Fig. 9, are identical up to rotation. If we truncate the TCC by keeping only some, but not all, of a repeated eigenvalue set (e.g., λ_2 but not λ_3), then this symmetry will naturally be broken.

5. MATHEMATICAL ANALYSIS

In this section, we analyze the TCC truncation error and how it affects the demonstrated artifacts. Our starting point for this will be the matrix formulation [14] of Eq. (7):

$$I(x,y) = \mathbf{M}^\dagger(x,y;f,g)\,\mathbf{T}(f,g;f',g')\,\mathbf{M}(f',g';x,y), \tag{9}$$

$$\mathbf{M}(f,g;x,y) = \widetilde{m}(f,g)\,\boldsymbol{\omega}(f,g;x,y) \in \mathbb{C}^{N_f N_g}, \tag{10}$$

$$\boldsymbol{\omega}(f,g;x,y) = e^{-i2\pi(fx/L_x+gy/L_y)} \in \mathbb{C}^{N_f N_g} \tag{11}$$

$$\mathbf{T}(f,g;f',g') = \boldsymbol{\Phi}(f,g)\,\boldsymbol{\Lambda}\,\boldsymbol{\Phi}^\dagger(f',g') \in \mathbb{C}^{N_f N_g \times N_f N_g}, \tag{12}$$

$$\boldsymbol{\Phi}(f,g) = [\cdots \quad \phi_i(f,g) \quad \cdots], \tag{13}$$

$$\boldsymbol{\Lambda} = \mathrm{diag}([\cdots \quad \lambda_i \quad \cdots]), \tag{14}$$

where boldface symbols represent vectors or matrices; the notation, e.g., $\mathbf{M}(f,g;x,y)$ is sometimes used to clarify that it indexes over (f,g) along the rows and (x,y) along the columns (the latter, in this case, a singleton dimension); † is the Hermitian conjugate; (N_f, N_g) is the discretization size in the spatial frequency domain; (L_x, L_y) is the simulation domain size; and $\mathrm{diag}(v)$ is the diagonal matrix whose elements are given by v. We use the subscript k to denote the truncated versions of the above, keeping eigenvalues only up to λ_k:

$$I_k(x,y) = \mathbf{M}^\dagger \mathbf{T}_k \mathbf{M}, \tag{15}$$

$$\mathbf{T}_k = \boldsymbol{\Phi}\,\boldsymbol{\Lambda}_k \boldsymbol{\Phi}^\dagger, \tag{16}$$

$$\boldsymbol{\Lambda}_k = \mathrm{diag}([\lambda_1 \quad \cdots \quad \lambda_k \quad 0 \quad \cdots]). \tag{17}$$

Let $E = I - I_k$ be the truncation error. Then

$$|E(x,y)| \le \|\mathbf{T} - \mathbf{T}_k\|\,\|\mathbf{M}\|^2 = \|\boldsymbol{\Lambda} - \boldsymbol{\Lambda}_k\|\,\|\widetilde{m}\|^2 = \lambda_{k+1}\|m\|^2, \tag{18}$$

where we have used the unitary of both $\boldsymbol{\Phi}$ and the Fourier transform, as well as the fact that the norm of a non-negative diagonal matrix is given by its largest entry. Note that this is just Theorem 2.1 in [15]. In fact, since $I_k(x,y) \ge 0$, a slightly stronger statement holds:

$$|E(x,y)| \le \min\{\lambda_{k+1}\|m\|^2, I(x,y)\}. \tag{19}$$

The truncation error also controls the aliasing error as the exact image I is alias-free by definition. However, it is not possible, in general, to more precisely bound the amount of aliasing (i.e., as a function of the computation band-limit η) because the frequency content beyond the Nyquist band-limit can be fairly arbitrary.

Consider the first artifact from Sec. 3: negative intensity. Write the upsampled image as

$$\hat{I}_k(x,y) = \boldsymbol{\omega}^\dagger(x,y;f,g)\,\mathbf{F}(f,g;x_i,y_j)\,I_k(x_i,y_j), \tag{20}$$

where $\mathbf{F}$ is the DFT matrix acting on the computational grid $\{(x_i,y_j)\}$, appropriately normalized such that

$$\boldsymbol{\omega}^\dagger(x_i,y_j;f,g)\,\mathbf{F}(f,g;x_{i'},y_{j'}) = \delta(i,i')\,\delta(j,j') \tag{21}$$

for δ the Kronecker delta. Note that

$$\left\|\boldsymbol{\omega}^\dagger(x,y)\,\mathbf{F}\right\|_\infty \le \left\|\boldsymbol{\omega}^\dagger(x,y)\,\mathbf{F}\right\| \le \|\boldsymbol{\omega}(x,y)\|\,\|\mathbf{F}\| = 1 \tag{22}$$

for any (x,y). Analogously define the exact upsampled image $\hat{I}$ and corresponding error $\hat{E} = \hat{I} - \hat{I}_k$. Let $C \ge \max\{\hat{I}, \hat{I}_k\}$ be a sufficiently large constant. Then $\hat{I}_{\min} \equiv \min \hat{I} = C - \|\hat{I} - C\|_\infty$ and similarly for $\hat{I}_k$. Therefore,

$$\min \hat{I}_k = C - \left\|\hat{I}_k - C\right\|_\infty \ge C - \left\|\hat{I} - C\right\|_\infty - \left\|\hat{E}\right\|_\infty \ge \hat{I}_{\min} - \left\|\boldsymbol{\omega}^\dagger \mathbf{F}\right\|_\infty \|E\|_\infty \ge \hat{I}_{\min} - \|E\|_\infty. \tag{23}$$

In other words, the smallest (most negative) value that $\hat{I}_k$ can take is the minimum of the exact image, minus an $O(\lambda_{k+1})$ term. In particular, if $\hat{I}_{\min} \geq \lambda_{k+1}\|m\|^2$, then $\hat{I}_k$ is always non-negative.

Now suppose the mask is periodic with pitch (Δ_x, Δ_y) as in Fig. 6(b). Let $\Delta[u(x,y)] = u(x,y) - u(x+\Delta_x, y+\Delta_y)$ be the periodic shift difference operator and define the upsampled image periodicity error as

$$\left|\Delta[\hat{I}_k]\right| \leq \left|\Delta[\hat{I}]\right| + \left|\Delta[\hat{E}]\right| \leq \left|\Delta[\boldsymbol{\omega}]^\dagger \mathbf{F} E\right| \leq \left\|\Delta[\boldsymbol{\omega}]^\dagger \mathbf{F}\right\|_\infty \|E\|_\infty, \tag{24}$$

where $\Delta[\hat{I}] = 0$ since I has no aliasing. It can be shown that

$$\|\Delta[\boldsymbol{\omega}]\| \leq 2\left\|\boldsymbol{\omega}\left(x - \frac{\Delta_x}{2}, y - \frac{\Delta_y}{2}\right)\right\| \left\|\sin\left[\pi\left(\frac{f\Delta_x}{L_x} + \frac{g\Delta_y}{L_y}\right)\right]\right\|_\infty. \tag{25}$$

Hence,

$$\left|\Delta[\hat{I}_k]\right| \leq 2\left\|\sin\left[\pi\left(\frac{f\Delta_x}{L_x} + \frac{g\Delta_y}{L_y}\right)\right]\right\|_\infty \|E\|_\infty \tag{26}$$

by Eq. (22). Notably, the error vanishes as $(\Delta_x, \Delta_y) \to (L_x, L_y)$ as required by the DFT-induced periodicity.

Third, for the image slope,

$$|\nabla I(x,y) - \nabla I_k(x,y)| = |\nabla E(x,y)| \leq \pi B \|E\|_\infty \tag{27}$$

again by Eq. (8), which also applies in this setting; see [13].

Finally, regarding loss of symmetry, consider for concreteness the case of rotational symmetry, perhaps the most important kind in practice. Let θ be the rotational period, $R(x,y) = (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta)$ the corresponding planar rotation, and $\Delta_R[u] = u(x,y) - u(R(x,y))$ the associated rotational difference operator. Define the upsampled image rotational symmetry error as

$$\left|\Delta_R[\hat{I}_k]\right| \leq \left|\Delta_R[\hat{I}]\right| + \left|\Delta_R[\hat{E}]\right| \leq \left|\Delta_R[\boldsymbol{\omega}]^\dagger \mathbf{F} E\right| \leq \left\|\Delta_R[\boldsymbol{\omega}]^\dagger \mathbf{F}\right\|_\infty \|E\|_\infty, \tag{29}$$

where $\Delta_R[\hat{I}] = 0$ by assumption. A very similar (but more complicated) calculation as for Eq. (25) gives

$$\|\Delta_R[\boldsymbol{\omega}]\| \leq 2\left\|\boldsymbol{\omega}\left(\frac{x_\theta^+}{2}, \frac{y_\theta^+}{2}\right)\right\| \left\|\sin\left[\pi\left(\frac{f x_\theta^-}{L_x} + \frac{g y_\theta^-}{L_y}\right)\right]\right\|_\infty, \tag{30}$$

$$\begin{bmatrix} x_\theta^\pm \\ y_\theta^\pm \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} \pm \mathbf{R}\begin{bmatrix} x \\ y \end{bmatrix}, \tag{31}$$

where $\mathbf{R}$ is the 2×2 matrix representation of R. Thus,

$$\left|\Delta_R[\hat{I}_k]\right| \leq 2\left\|\sin\left[\pi\left(\frac{f x_\theta^-}{L_x} + \frac{g y_\theta^-}{L_y}\right)\right]\right\|_\infty \|E\|_\infty. \tag{32}$$

Notice that the error vanishes as $\theta \to 2\pi$ since then $\mathbf{R}$ becomes the identity.

6. SUMMARY

TCC truncation in the SOCS methodology is a standard practice for image computation in optical lithography, widely used for its efficiency and well-understood approximation theory. However, it can cause some unexpected side effects. In this paper, we have discussed four of these: negative image intensity, loss of image periodicity, sharper image slope, and loss of image symmetry. Some artifacts, such as negative intensity and loss of periodicity, are due to aliasing. These can be avoided by calculating the image on a grid size of at most $\lambda/[4(1+\sigma_{\max})\mathrm{NA}]$, which is smaller than the standard Nyquist grid size $\lambda/(4\mathrm{NA})$ by a factor of $1+\sigma_{\max}$. Others, such as sharper image slope and loss of symmetry, are due only to truncation. The latter can be avoided by ensuring that all eigenvalues of a repeated set are retained, while the former cannot be fully removed without using all eigenfunctions. Nevertheless, both types of artifacts can be effectively controlled through the truncation error, which essentially scales as $O(\lambda_{k+1})$ if k eigenfunctions are used. Thus, for any given tolerance on the magnitude of these artifacts, the truncation parameter k can always be chosen such that this tolerance is met.

ACKNOWLEDGEMENTS

We would like to thank our colleagues Jason Chou and Christopher Wong for inspiring the band-limited Bernstein's inequality. We would also like to thank Lin Zhang, Shuo-Yen Chou, H.T. Lin, and Y.C. Ku for supporting this study. We also appreciate the encouragement from Xin Zhou and Danping Peng.

REFERENCES

[1] K. Ho, H. Le, X. Zhou, and D. Peng, "GPU computing for lithography simulations in mask synthesis," presented at NVIDIA GPU Technology Conference 2020.

[2] H. H. Hopkins, "On the diffraction theory of optical image," Proc. R. Soc. London, Ser. A **217**, 408–432 (1953).

[3] M. Born and E. Wolf, *Principles of Optics*, 6th ed. (Pergamon, 1980), Chap. 10.

[4] N. B. Cobb, "Fast optical and process proximity correction algorithms for integrated circuit manufacturing," Ph.D. dissertation (Electrical Engineering and Computer Science, University of California, Berkeley, 1998).

[5] J. W. Goodman, *Introduction to Fourier optics*, ed. S. W. Director (McGraw-Hill, New York, 1996) 2nd ed. Chap. 6.

[6] J. D, Gaskill, *Linear systems, Fourier transforms, and Optics* (John Wiley & Sons, 1978) Chap. 8.

[7] W. Singer, M. Totzeck, and H. Gross, *Handbook of Optical Systems, Volume 2: Physical image formation,* ed. H. Gross, (WILRY-VCH, 2005) Chap. 21.

[8] R. Barakat, "Partially coherent imagery in the presence of aberrations," Opt. Acta **17**, 337–347 (1970).

[9] J. W. Goodman, *Introduction to Fourier optics*, ed. S. W. Director (McGraw-Hill, New York, 1996) 2nd ed. Chap. 2.

[10] K. Yamazoe, "Two models for partially coherent imaging," J. Opt. Soc. Am. A **29**, 2591-2597 (2012).

[11] E. Wolf, "New theory of partial coherence in the space–frequency domain. Part I: spectra and cross spectra of steady-state sources," J. Opt. Soc. Am. **72**, 343-351 (1982)

[12] J. W. Goodman, *Introduction to Fourier optics*, ed. S. W. Director (McGraw-Hill, New York, 1996) 2nd ed. Chap. 3.

[13] J. Chou, K. Yamazoe, C. Wong, J. Lai, and S. Chou, "Theoretical limits on pattern corner sharpness in optical lithography," presented at SPIE Advanced Lithography 2026 and the manuscript to be submitted.

[14] K. Yamazoe, "Computation theory of partially coherent imaging by stacked pupil shift matrix," J. Opt. Soc. Am. A **25**, 3111-3119 (2008).

[15] Y. C. Pati, Y.-T. Wang, J.-W. Liang, and T. Kailath, "Phase-shifting masks: automated design and mask requirements," Proc. SPIE **2197**, 314-327 (1994).