Understanding Extreme Stochastic Events in EUV Resists

Patrick Naulleau, Suchit Bhattarai, Andrew Neureuther

Center for X-ray Optics, Lawrence Berkeley National Laboratory, Berkeley, CA
EECS, University of California, Berkeley, CA

The problem of stochastics in photoresist patterning is of significant concern in the commercialization of extreme ultraviolet lithography, especially in patterning of contacts or vias. Traditionally, contact hole critical dimension (CD) variability is characterized as a Gaussian process, however, recent experimental results have demonstrated significant deviations from Gaussian statistics, especially on the small CD side. Modeling results show that this non-Gaussian variation is expected and can be attributed to the non-linear behavior of the contact hole exposure latitude as a function of dose. It is shown that if we consider the noise in the dose or sensitivity domain, it can still be treated as a Gaussian process. The CD statistics are then determined by mapping the Gaussian dose/sensitivity noise through the non-linear and deterministic CD versus dose function.

1. Introduction The problem of stochastics in photoresist patterning is of significant concern in the commercialization of extreme ultraviolet lithography, especially in the patterning of contacts or vias. Traditionally, contact hole critical dimension (CD) variability is characterized as a Gaussian process, however, recent experimental results have demonstrated significant deviations from Gaussian statistics, especially on the small CD side [1]. Here we use the stochastic Multivariate Poisson Propagation Model (MPPM) [2-6] to understand this behavior. The stochastic modeling results exhibit the same non-Gaussian variations as observed experimentally. Decomposing the model into deterministic and stochastic components demonstrates that the non-Gaussian behavior can be attributed to the non-linear characteristics of the contact hole exposure latitude as a function of dose. It is shown that if we consider the noise in the dose or sensitivity domain, it can still be treated as a Gaussian process. The CD statistics are then determined by mapping the Gaussian dose/sensitivity noise through the non-linear and deterministic CD versus dose function.

2. Stochastic modeling To capture both photon and material stochastic effects in the modeling, we use the MPPM described in detail in Ref. 6. In this model, all relevant counting terms are treated as Poisson random variables. Note that this is not the same as a Monte Carlo model and in principle can be treated as an analytic model. In the implementation used here [7] the model is realized numerically. The modeling process begins with an arbitrary aerial image representing the time averaged photon density as a function of position (pixel). In the description below we assume a chemically amplified resist, however, the model is readily adjusted to treat the non-chemically- amplified case. In the MPPM process, the first random variable represents the absorbed photons per pixel thus capturing the photon noise effect. Other key random variables in the process include the chemical yield (or quantum efficiency) and the local chemically active component density which for a chemically amplified resist is the photo-acid generator (PAG) density, the local quencher density, and the local protecting group density. The end result of the model is a random variable

representing the deprotection ratio at which point a dissolution model is applied which itself may also be treated as a statistical process. We note that in the results presented here, non-stochastic development is assumed. As described above, the model consists of a series of Poisson random variables acted on and combined through a series of functions thus the

problem can be viewed as a classical error propagation problem. Note that since the MPPM is implemented numerically, one is not restricted to using Poisson distributions; any desired distribution can be used for each of the random variables. Because the MPPM is essentially an error propagation model, deterministic variables will still have an impact on the stochastic output since those variables affect the functions acting on the input random variables. Determinist variables used in the MPPM include the aerial image, absorptivity, electron mean free path, chemical diffusion range, quencher diffusion range, and chemical reaction rates. 3. Stochastic requirements The increasing concern over stochastics with shrinking CD comes from two fundamental sources. The first source is related to the smaller dimensions leading smaller counting statistics and the second source comes from the fact that smaller CD implies an increasing number of critical features per exposure. Thus, we have both an increase in noise and an increase in the total number of features we are trying to print. The second concern determines the maximum failure rate and thus the required signal to noise ratio. As described in Ref. 1, a reasonable estimate for the via feature count can be determined based on the minimum half pitch (HPmin), which we define as

$$HP_{min} = \frac{\lambda}{2 \cdot \sqrt{} } $$ (1)

(1) where NA is the numerical aperture, λ is the imaging wavelength, and σ_{pole} is the partial coherence of a single pole from a quadrupole illumination setting such as shown in Fig. 1. For a wavelength of 13.5 nm, and pole size of 0.2, numerical apertures of 0.33 and 0.55 yield minimum contact pattern half pitches of 18 nm and 11 nm, respectively.
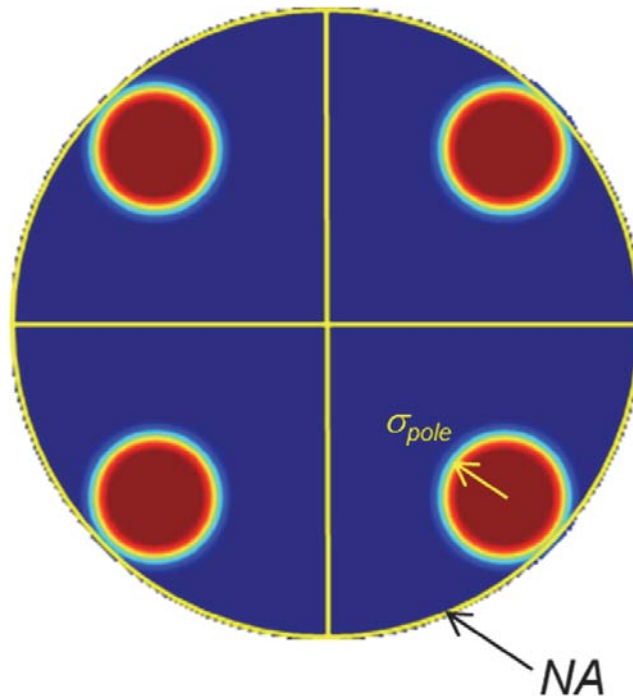

Fig. 1. Pupil fill geometry used to determine minimum half pitch for contact features.

Taking the minimum half pitch and assuming a fill factor and die size, we can then determine the critical feature count. Assuming a fill factor of 1.5% and a die size of 26 mm x 32 mm yields the results in column 2 from Table 1. The acceptable error rate is then determined based on the

target die yield where it is assumed that a yielding die requires all the features to print. Assuming a yield requirement of 99% implies the required failure rates shown in column 3 of Table 1. Table 1.

 Estimated critical feature count and acceptable failure rate as a function of minimum half pitch.

HP_{min} (nm)	Feature count	Failure rate
26	0.18×10^{11}	55×10^{14}
18	0.39×10^{11}	26×10^{14}
11	1.03×10^{11}	10×10^{14}
8	1.95×10^{11}	5×10^{14}

The estimates above show failure rate requirements implying signal to noise ratio magnitudes of 7 and 8 standard deviations assuming a Gaussian process.

4. Modeled failure rates

Next, we use the MPPM to model expected failure rates. For the aerial image we assume a dense square grid of 16-nm half pitch contacts, a numerical aperture of 0.33, 13.5-nm wavelength, and quadrupole illumination with a σ_{pole} of 0.1. The pole offset is selected such that the first diffracted orders from one illumination pole fully overlap the opposite illumination poles. For the resist parameters, we assume a chemically amplified resist process with a dose to clear value of 15.7 mJ/cm_2, a quantum efficiency of 3, a PAG density of 0.2/nm_3, quencher density of 0.085/nm_3, protecting group density of 2/nm_3, absorptivity of 4.2/μm, thickness of 35 nm, acid diffusion range of 12 nm, electron diffusion range of 4 nm, and quencher diffusion range of 3 nm. We also assume a threshold development model. Figure 2 shows a representative set 81 contacts produced by the model.
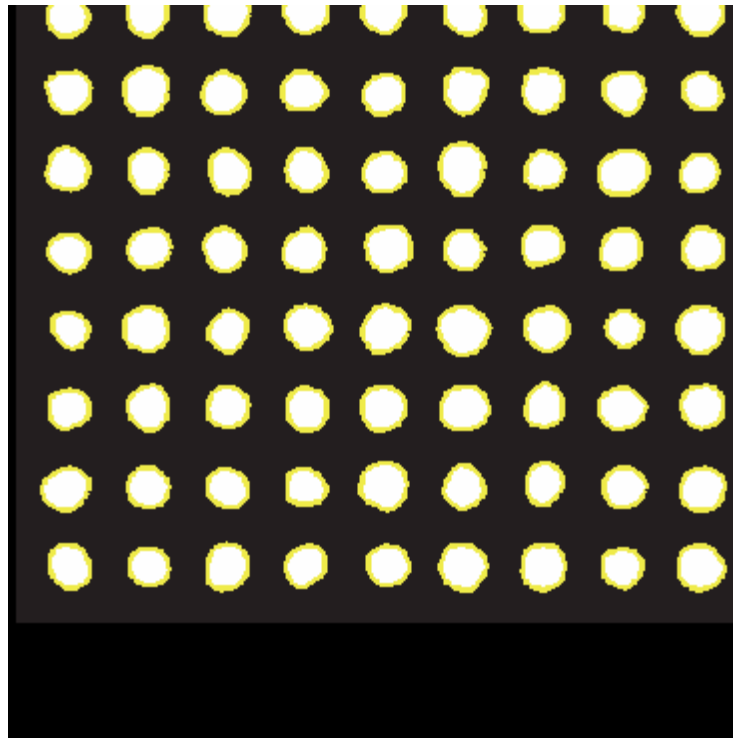

Fig. 2. Set of 81 representative contacts in resist produced by the MPPM model.

Figure 3 shows the modeled CD probability distribution function (PDF) considering approximately 2.6M contacts. The skewed distribution is evident with non-Gaussian tail on the small CD side. The non-Gaussian behavior can be better visualized with a logarithmic y axis (Fig.4). Also shown is a reference Gaussian PDF matching the mean (17.17 nm) and standard deviation (1.15 nm) of the modeled CD data.
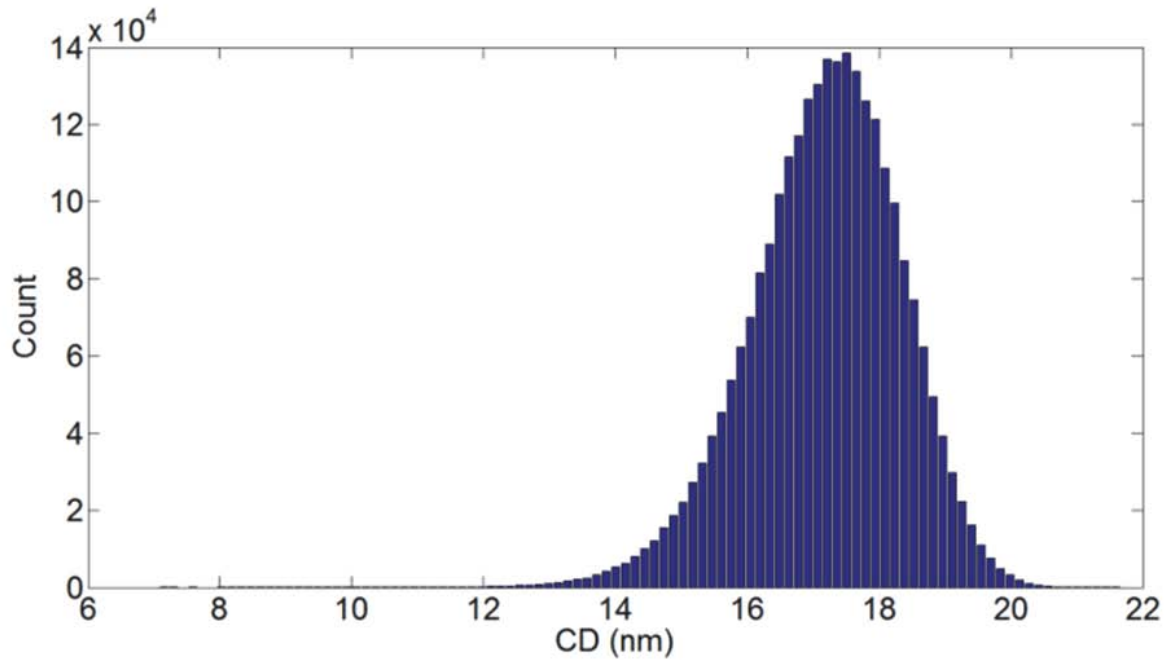

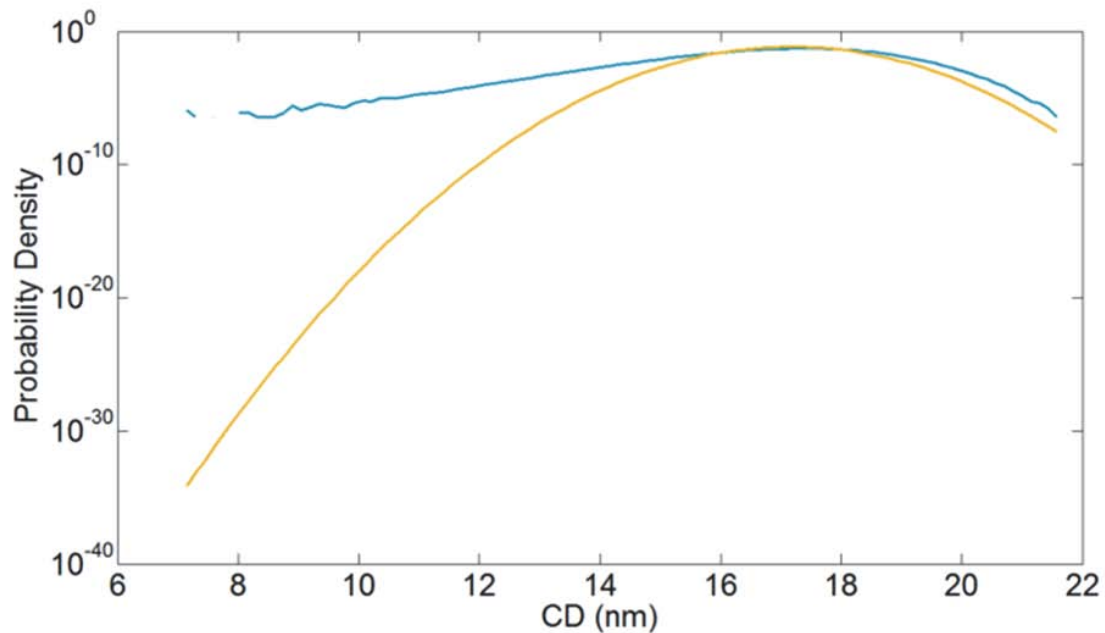Fig. 3. CD probability distribution function (PDF) for 2.6M contacts in resist as described with reference to Fig. 2.
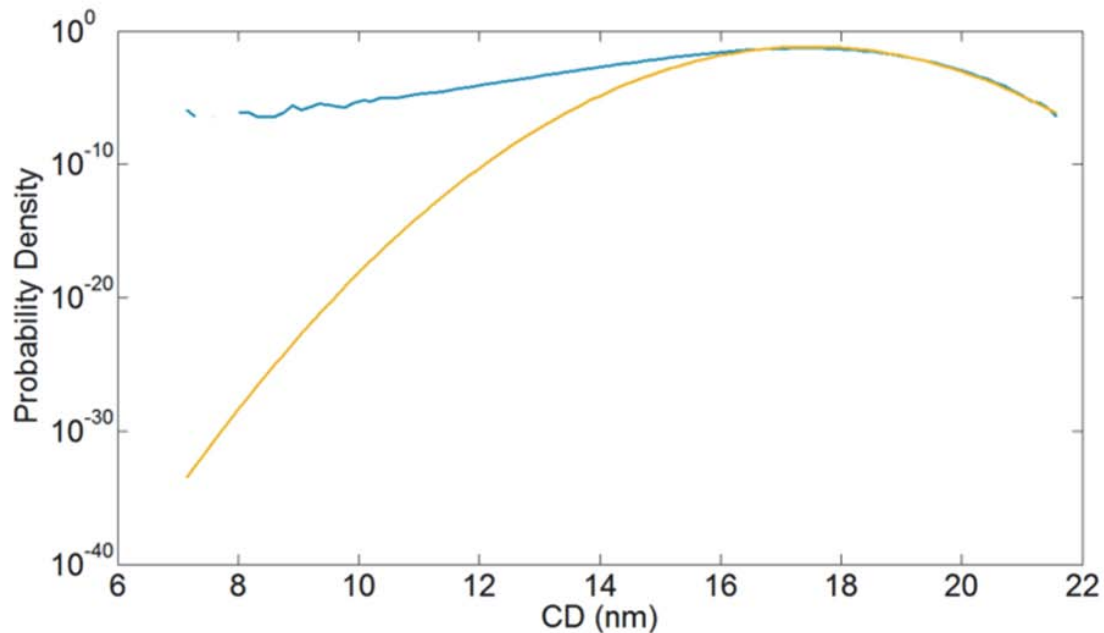

Fig. 4. PDF from Fig. 3 (blue/top) with a logarithmic y axis along with Gaussian reference (orange/bottom)matching the mean and standard deviation of the modeled CD data.

If we instead determine the Gaussian reference based on the best fit to the raw data to the right of the peak of the distribution, we get the results in Fig. 5. Here the mean is 17.50 nm and the standard deviation is 1.20 nm. From this perspective the modeled CD is well represented by a Gaussian distribution on the large CD side while it deviates considerably on the small CD side. At the small CD extreme the modeled CD data has probabilities more than 20 orders of magnitude larger than a strictly Gaussian distribution would predict.

Fig. 5. PDF from Fig. 3 (blue/top) with a logarithmic y axis along with Gaussian reference (orange/bottom) matching the mean and standard deviation of the modeled CD data to the right of peak in the raw PDF.

An alternative visualization tool for such data is the cumulative distribution function (CDF) quantile plot. In such a representation (Fig. 6), Gaussian data will appear as a straight line. The y axis in this plot should be interpreted as Gaussian standard deviations. For example we see that CD events at 8 nm and smaller occur at rates of greater than 4☐ whereas a true Gaussian would have such events occur at a rate of only 9☐ Thus, whereas the Gaussian distribution meets the criteria described in Table 1, the actual modeled data is very far from meeting the requirement.
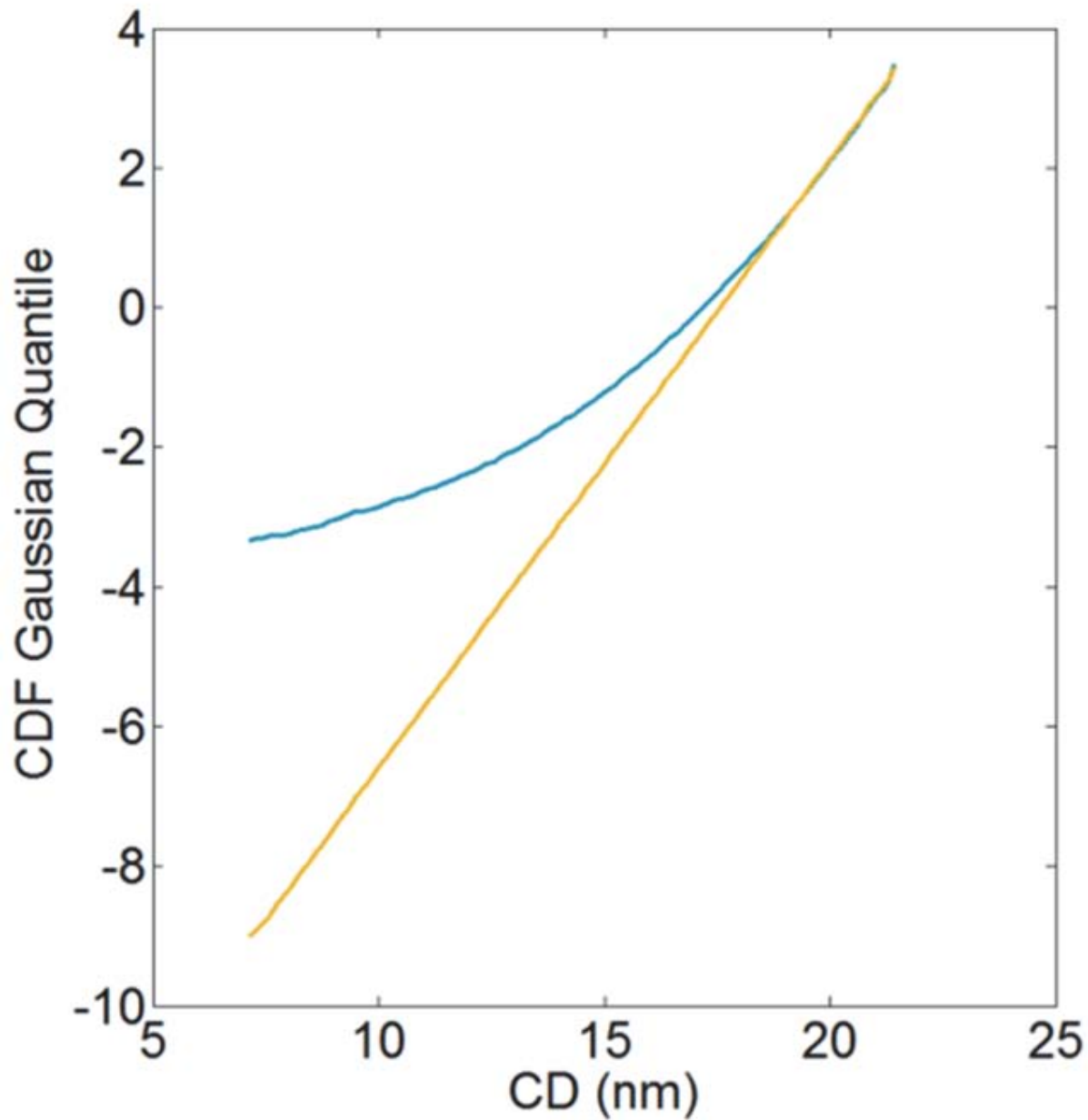
Fig. 6. CDF from Fig. 3 (blue/top) on a Gaussian quantile y axis along with Gaussian reference (orange/bottom).

Comparing the results in Fig. 6 to experimental patterning results presented in Ref. 1, we see quite similar behavior with a good match to Gaussian on the large CD side and substantial deviation on the small CD side. Although Ref. 1 does not quantify the specific CDs in question for the experimental data, it does show the CDF bottoming out around 5☐as compared to the 4☐seen in the modeling data presented here. Assuming the experimental data in Ref. 1 to be larger than the 16-nm contacts modeled here, this better 5☐performance is to be expected.

5. Are poisson statistics to blame?
Given that the true physical statistics are expected to be Poisson instead of Gaussian and

that the MPPM uses Poisson, we next consider the question as to whether or not these Poisson statistics are the cause of the deviation observed above. To this end, we directly compare Poisson and Gaussian statistics using the PDF and CDF plots described above. Figure 7 shows the Poisson PDF and CDF compared to Gaussian for mean values of 10, 100, and 1000 respectively. We do see that at small mean values (less than 100) the Poisson does indeed deviate noticeably from the Gaussian reference, however, in the opposite direction as seen with the modeled CD data. That is to say that Poisson statistics alone would cause extremes on the small side to be even less likely than for a Gaussian distribution.

6. CD versus dose

The results in Section 5 demonstrate that Poisson statistics are not the cause of the skewed CD distributions in Section 4. Given that the input noise is not the cause of the deviation, we next consider the error propagation functions. Because all the stochastic terms described in Section 2 can be seen as either directly affecting the local dose or resist sensitivity and that sensitivity variations can arguably be mapped to dose variations, it is reasonable to assume that the deterministic CD versus dose function is a good representation of the stochastic error propagation function. This function can be determined by running the MPPM in deterministic mode instead of stochastic mode.
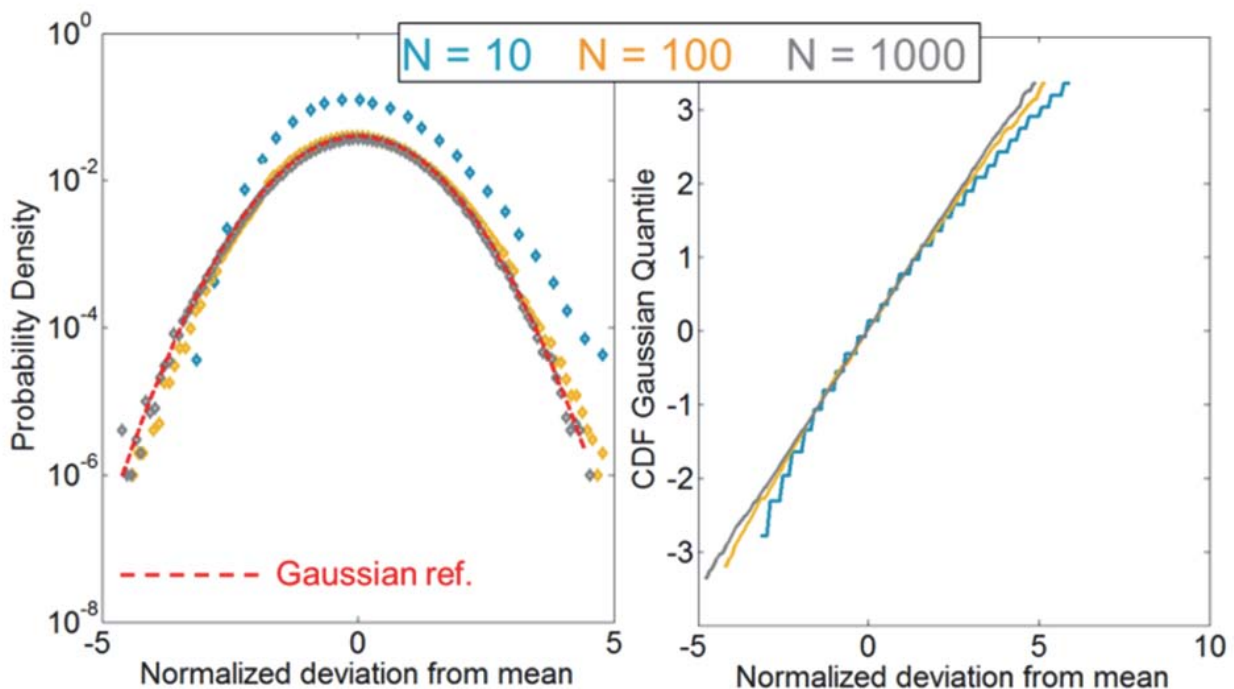

Fig. 7. Direct comparison of Poisson and Gaussian PDF and CDF for various Poisson mean values (N). Smaller N values show larger deviation from Gaussian but in opposite direction compared to the modeled CD data.
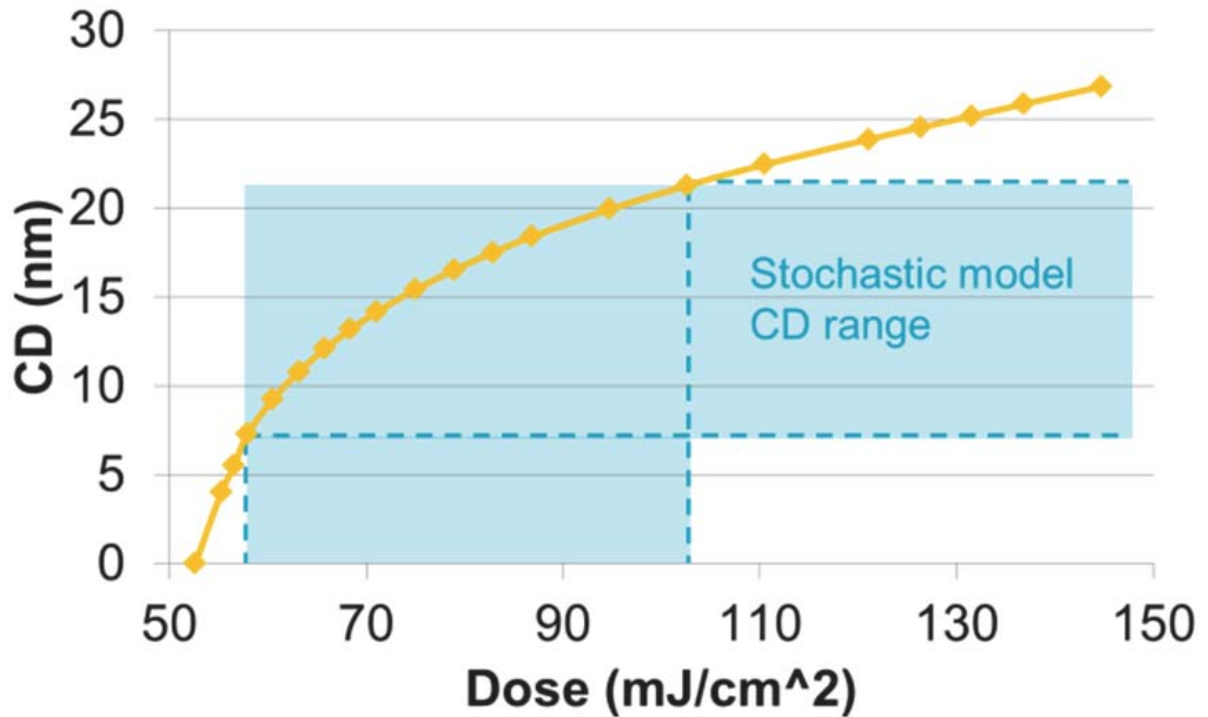
Fig. 8. Modeled deterministic CD as a function of dose.

Next assuming that we can effectively represent the stochastics in the system simply as Gaussian dose noise, we can use the function in Fig. 8 to transform the dose noise to a CD probability distribution function. Figure 9 shows the direct comparison of the PDF computed in this way compared to the stochastic modeled PDF. Very good agreement is seen between the two when the mean dose is set to 82 mJ/cm$_2$ with a standard deviation of 4.7 mJ/cm$_2$ indicating that the Gaussian dose noise assumption is indeed valid. The Gaussian dose noise is determined from the measured large CD Gaussian CD statistics and simply mapped to dose based on the linear portion of the CD versus dose plot.
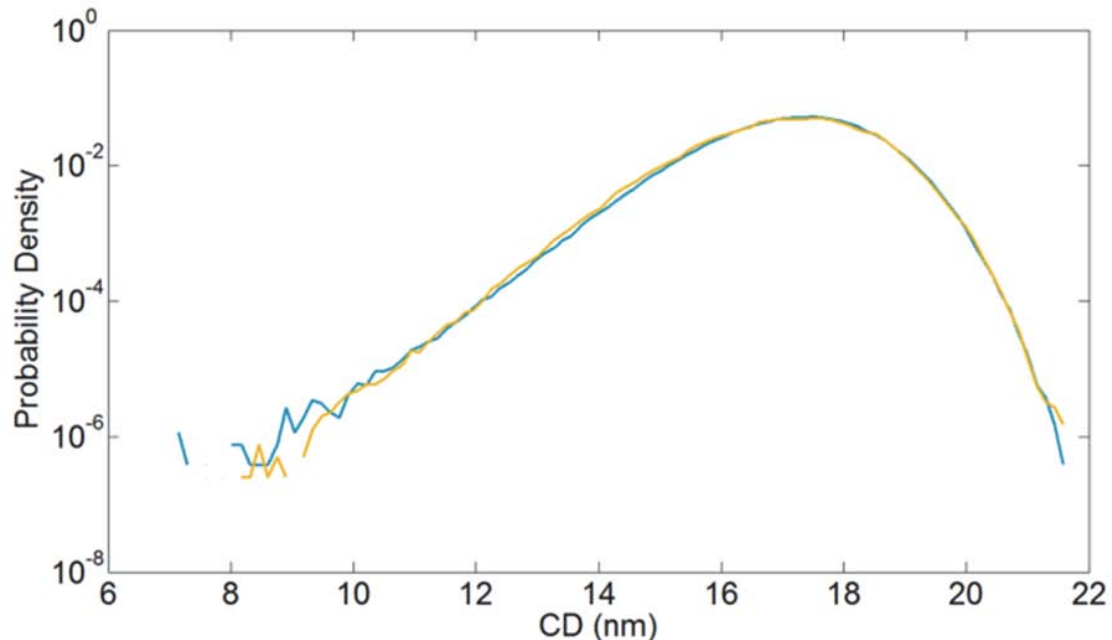
Fig. 9. Mapped Gaussian PDF (orange) compared to modeled stochastic PDF (blue).

7. Predicting error rates

Being able to represent the statistics of interest as Gaussian greatly facilitates the determination of failure rates. For example, considering the data in Fig. 8 and noting the separation between the vanishing-CD dose and the mean-CD (17.5 nm) dose, relative to the effective dose noise standard deviation, we can quickly evaluate how the process compares to the requirements called out in Table 1. In this case (Fig. 10) we find a margin of only 6.4☐compared to the required 7☐or 8☐☐☐

An obvious result of this perspective is that it is beneficial to bias the printed CD to larger values by increasing the nominal dose. Achieving a margin of 8☐simply requires the nominal dose to be increased to 90 mJ/cm$_2$ (or by 10%) yielding a mean CD of approximately 18 nm instead of 16 nm. Achieving the same performance based on reduction of the dose noise instead would have required the dose noise to be reduced by 25%. If we simply assume the dose noise to go as the square root of the mean dose, a 25% reduction in noise would require a 56% increase in dose. This arguably represents the "shot noise limit" [8-19], yet w e f ind t hat b y b iasing t he C D b y 2 n m w e significantly outperform this "limit". Figure 11 shows the corresponding PDFs for the original case in Fig. 9 which already includes a CD bias of about 1 nm and the case with the CD bias set to 2 nm (mean dose set to 90 mJ/cm$_2$).

Another result of this perspective is that it is beneficial to improve the CD versus dose performance on the small CD end (i.e. the resolution limit of the process). To this end it is reasonable to assume that decreased resist blur would be beneficial. Reducing the acid diffusion from 12 nm to 4 nm yields the results in Fig. 12.
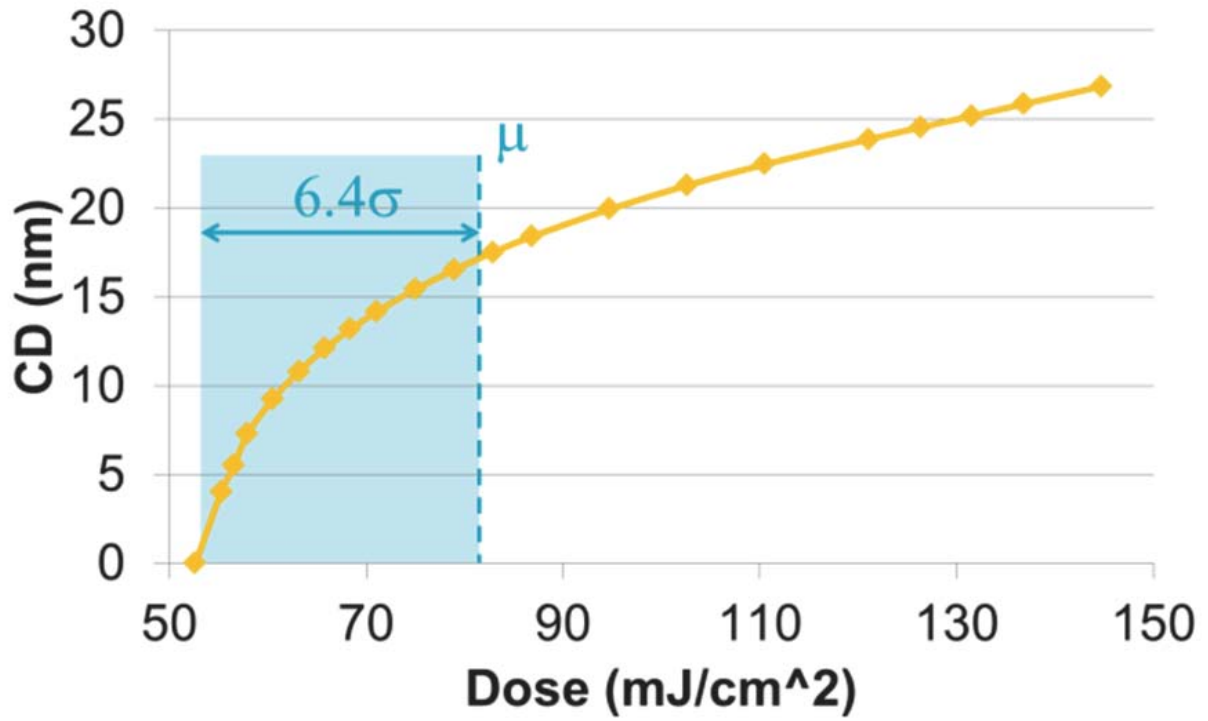
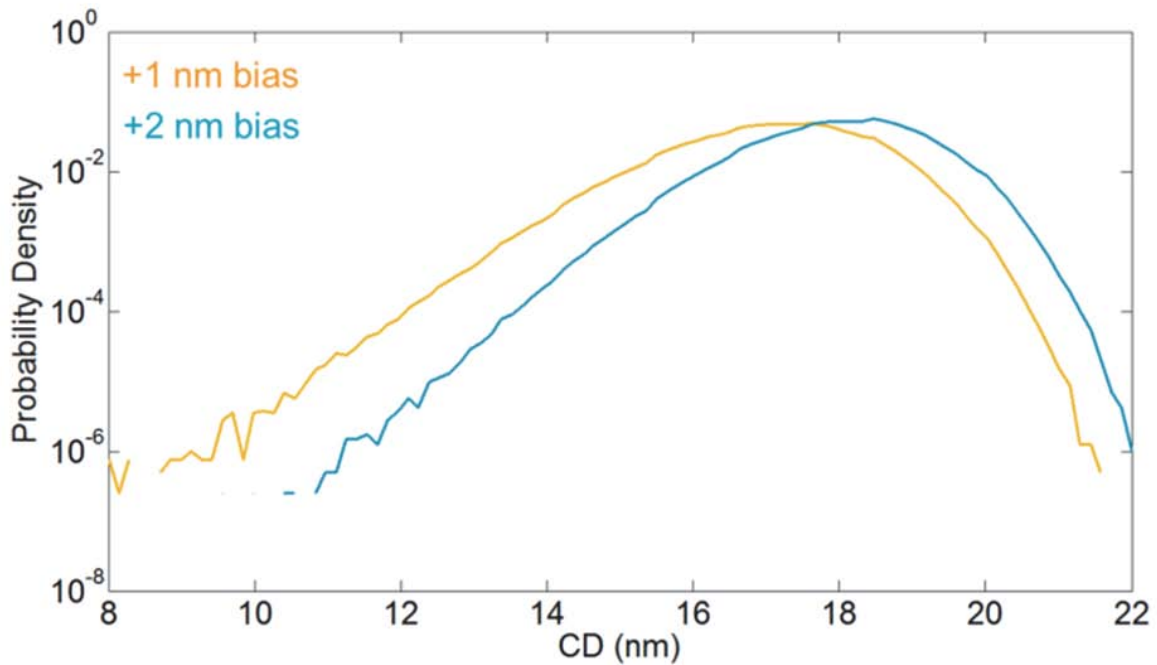Fig. 10. Using Gaussian dose noise and CD versus dose function to evaluate failure rate.

Fig. 11. Comparison of CD probability distribution functions for cases of CD bias of 1 nm (orange) and 2 nm (blue).
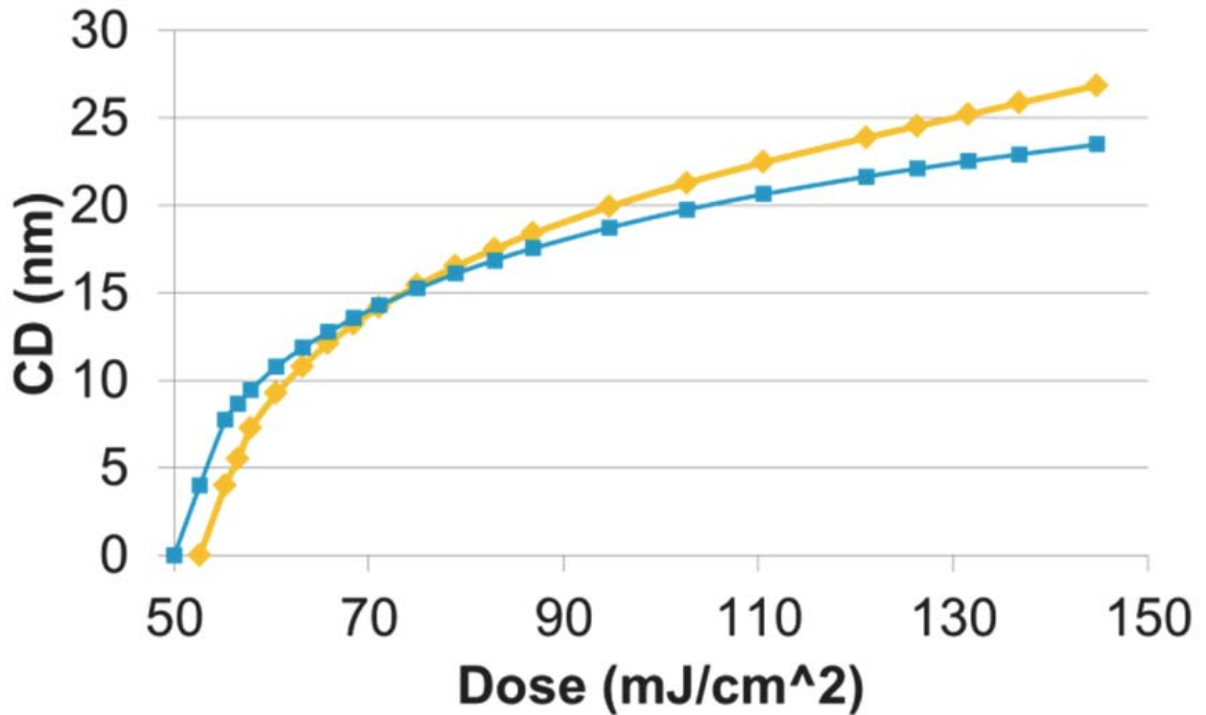
Fig. 12. CD versus dose for acid diffusion range of 12 nm (orange) and 4 nm (blue).

Before directly using this new error propagation function we must also determine the new effective dose noise. Based on the fact that we decreased the resist blur one should expect the effective dose noise to increase in accordance to the well documented tradeoff between resolution and noise [9-11]. Repeating the stochastic modeling and analysis processes described above for this new case yields the PDF results in Fig. 13. The plot shows three cases: (orange line) acid diffusion range of 12 nm and CD bias of 1 nm. (blue line) acid diffusion range of 12 nm and CD bias of 2 nm. (gray lines) acid diffusion range of 4 nm and CD bias of 2 nm. For the third case, both the raw stochastic results (dashed) as well as the Gaussian dose noise mapped results are shown.
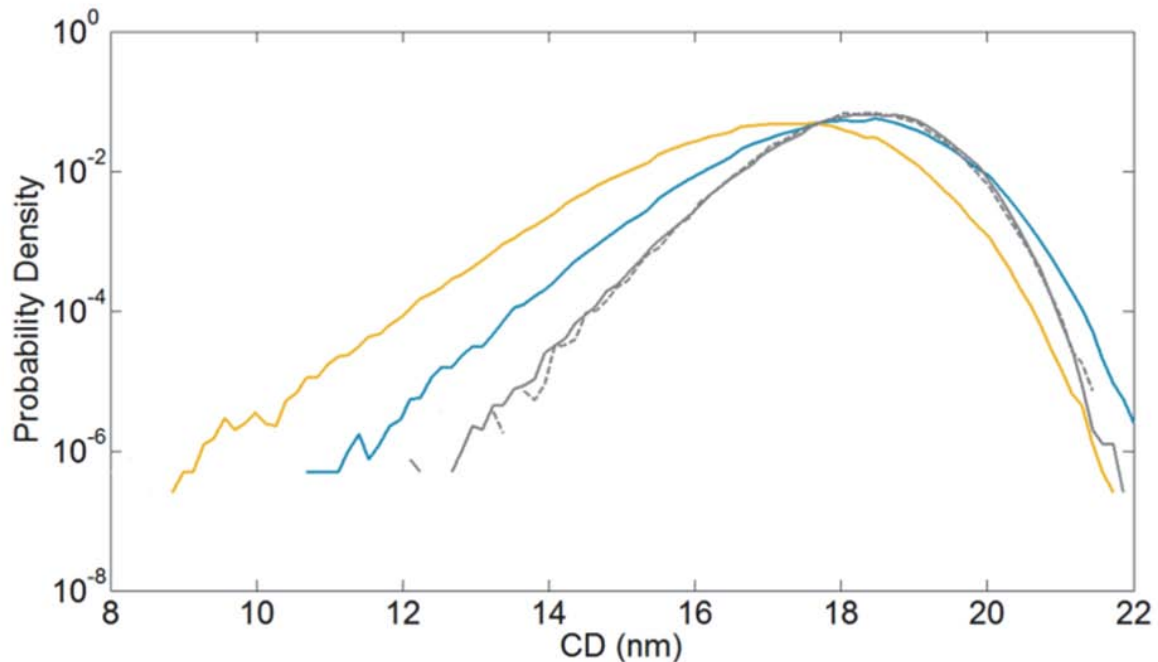
Fig. 13. PDF for three different cases: (orange) acid diffusion range of 12 nm and CD bias of 1 nm. (blue) acid diffusion range of 12 nm and CD bias of 2 nm. (gray) acid diffusion range of 4 nm and CD bias of 2 nm. For the third case, both raw stochastic results (dashed) and Gaussian dose noise mapped results are shown.

For the reduced blur case, the effective dose noise increases to 5.8 mJ/cm$_2$, yet the PDF shows a decreased failure rate without sacrifice in dose. Thus we have arguably again outperformed the "shot noise tradeoff limit". Further extending this analysis to a direct activation resist with an electron blur of 3 nm and chemical blur of 2 nm yields the results in Fig. 14. In this case, both the noise and the error propagation function improve relative to the 4-nm blur chemically amplified blur case considered in Fig. 13. The effective dose noise is found to be 4.7 mJ/cm$_2$ with the reduction coming from the elimination of quencher which is the sparsest chemical element in the chemically amplified resist model.

Finally, we consider the case of improving the error propagation function by increasing the numerical aperture (NA) of the lithography tool while keeping the target patterned CD fixed. The computed CD versus dose plots comparing NAs of 0.33 and 0.55 are shown in Fig. 15. The resulting probability densities are plotted in Fig. 16. For both NAs, we have assumed the same 4-nm blur chemically amplified resist model as described with respect to Fig. 12. We note that for this comparison we have also set the patterned CD bias to zero. The impressive impact of increased NA on reduced variability is clearly evident.
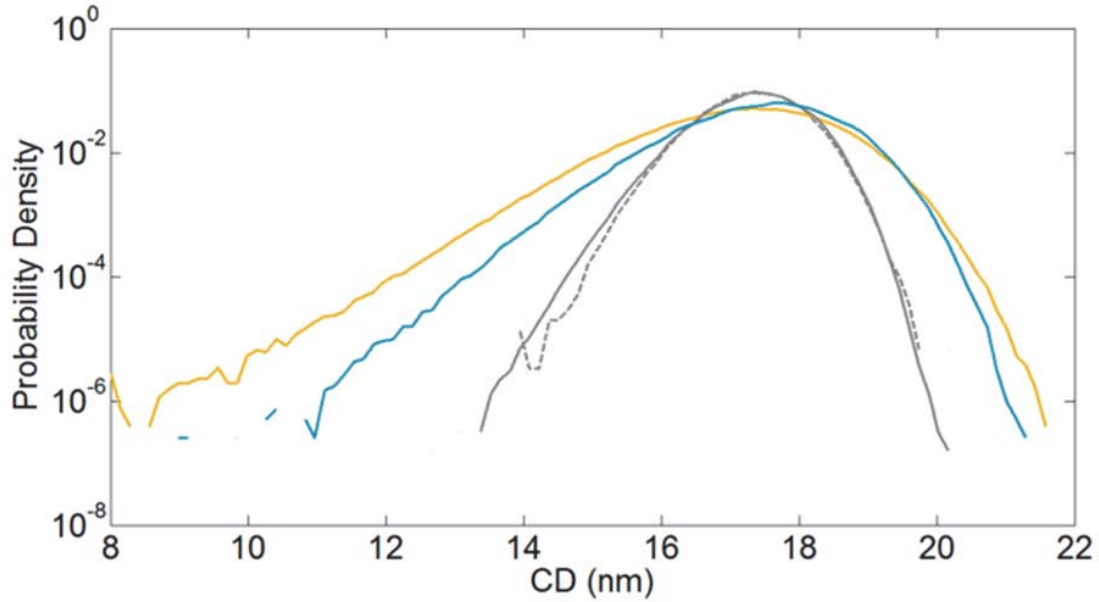
Fig. 14. PDF for three different cases: (orange) acid diffusion range of 12 nm and CD bias of 1 nm. (blue) acid diffusion range of 12 nm and CD bias of 2 nm. (gray) direct activation resist with electron blur of 3 nm, chemical blur of 4 nm and CD bias of 1 nm. For the third case, both raw stochastic results (dashed) as well as Gaussian dose noise mapped results are shown.
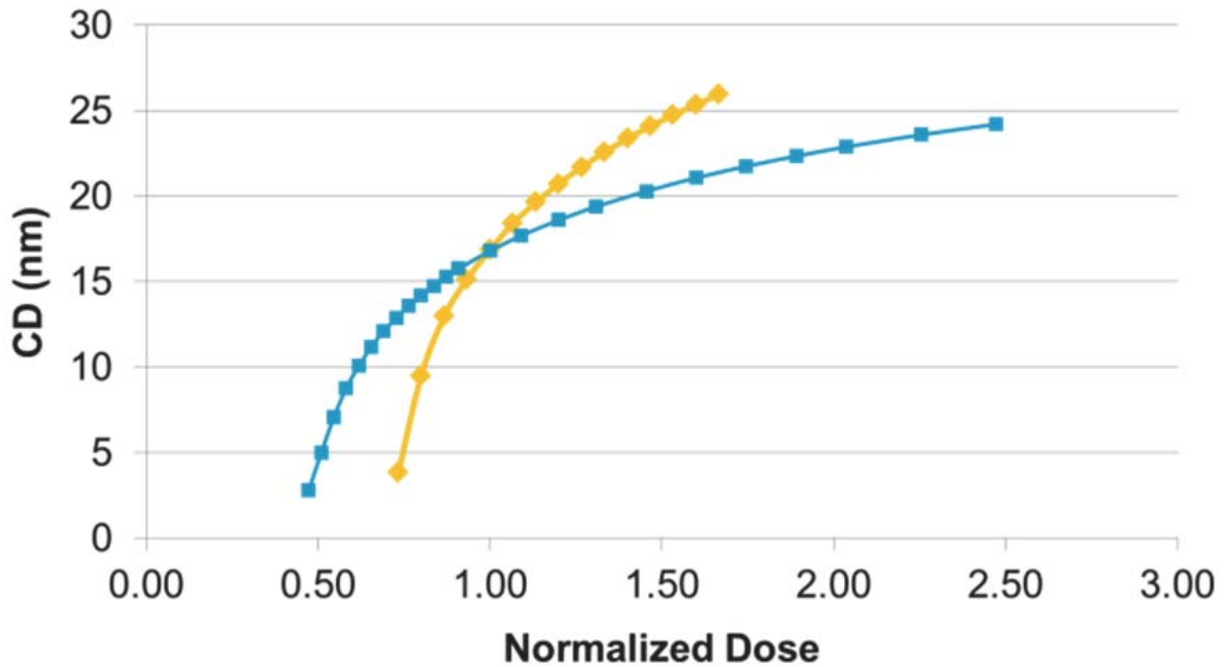

Fig. 15. CD versus dose for NA of 0.33 (orange) and 0.55 (blue). Both cases assume 4-nm blur chemically amplified resist from Fig. 12.
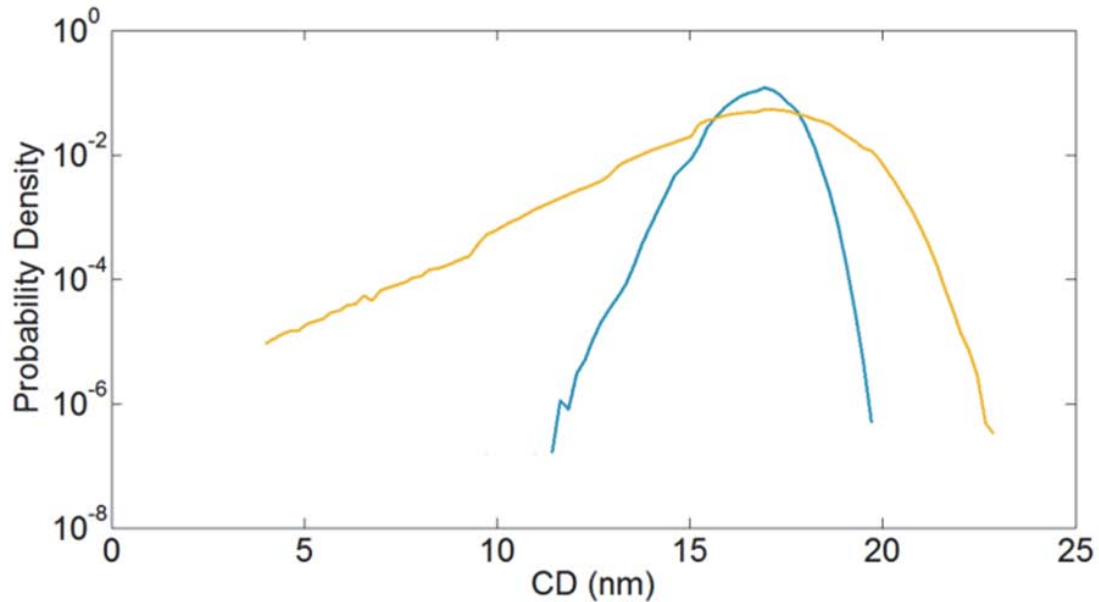
Fig. 16. PDF of contact CD for NA of 0.33 (orange) and 0.55 (blue). Both cases assume 4-nm blur chemically amplified resist from Fig. 12.

8. Conclusion

A method has been described allowing Gaussian statistics to be used to predict contact failure rates even at levels of 8□and beyond by way of a non-linear error transfer function derived from the deterministic CD versus dose performance of the system. The method has been used to demonstrate the benefits of patterned CD bias, resist resolution improvement, and increased NA. The method also clearly demonstrates methodologies for improving on the often-referenced shot noise tradeoff limit.