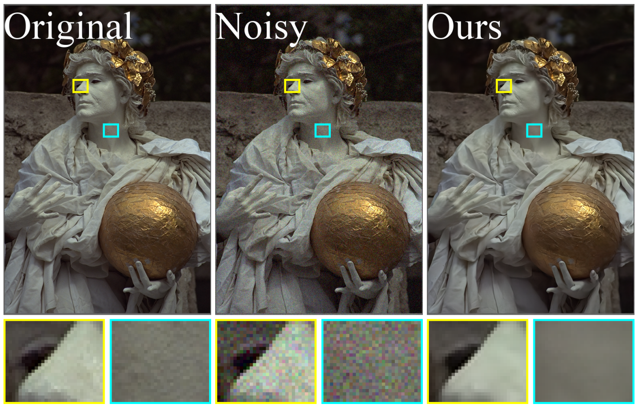

Figure 1. Results of our method applied under a heteroskedastic noise model with $\alpha = \beta = 0.05$

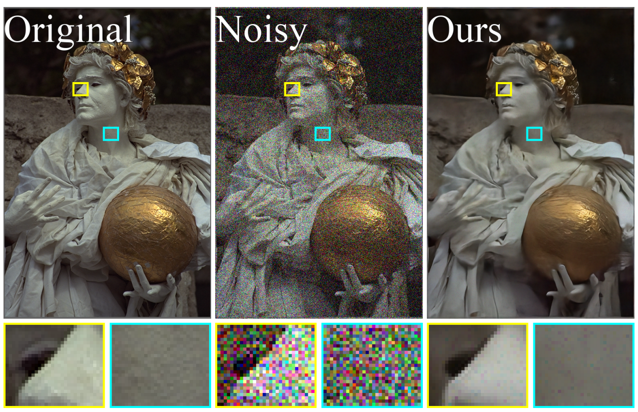

Figure 2. Results of our method applied under a heteroskedastic noise model with $\alpha = \beta = 0.2$.

A. Intensity-Dependent Heteroskedastic Noise

Real photographic noise is not typically purely additive. Instead, the magnitude of the noise at each pixel is correlated with the intensity of the clean image at that pixel. In this Appendix, we present an analysis of how this phenomenon affects our method. In particular, we consider a noise model where $Y = X + A \odot X + B$ where A and B have the same shape as X, $A \sim \mathcal{N}(0, \alpha^2)$, $B \sim \mathcal{N}(0, \beta^2)$ and '$\odot$' indicates pixel-wise multiplication. Note that the noise term is now $A \odot X + B$ and is dependent on X.

If we are able to draw a second noise sample from the same distribution, our previous analysis applies with minimal changes. Let our doubly-noisy image be $Z = X + A \odot X + B + C \odot X + D$, where C and D are drawn from the same distributions as A and B, respectively. Following the same reasoning as in Section 3.2, we have:

$$\mathbb{E}[Y|Z] = X + \frac{A+C}{2} \odot X + \frac{B+D}{2}$$

Performing the same manipulation as before, we take:

$$2\mathbb{E}[Y|Z] = 2X + (A+C) \odot X + (B+D)$$

Method	PSNR	SSIM
Noisy	25.31	0.552
Ours	33.51	0.907
Noise2Noise	34.98	0.924

Table 1. PSNR and SSIM comparison under a heteroskedastic noise model with $\alpha = \beta = 0.05$.

$$2\mathbb{E}[Y|Z] - Z = X$$

We thereby recover an estimate of the clean image using exactly the same procedure as in the case of purely additive noise. However, this analysis relies on our ability to draw a second noise sample from the same distribution as the first. This in turn requires us to have access to the clean image X, which is very likely unrealistic. Instead, we only have access to the noisy image, Y. We can attempt to approximate this second noise sample by instead using Y to compute the intensity-dependent component. This yields:

$$Z = X + A \odot X + B + C \odot (X + A \odot X + B) + D$$

$$= X + A \odot X + C \odot X + C \odot A \odot X + B + D + C \odot B$$

In our original analysis, we made use of the fact that $\mathbb{E}[N|Z] = \mathbb{E}[M|Z]$. In this setting, however, A and C are no longer exactly symmetric, nor are B and D, due to the presence of the $C \odot B$ term without a corresponding $A \odot D$ term. However, as long as the magnitude of the noise is relatively small, this extraneous term will be very small, as it is the pointwise product of two noise samples. Thus, we have $\mathbb{E}[A|Z] \approx \mathbb{E}[C|Z]$ and similarly for B and D, where the tightness of the approximation depends on the magnitude of the noise. We can therefore estimate $\mathbb{E}[A|Z] \approx \frac{\mathbb{E}[A|Z]+\mathbb{E}[C|Z]}{2}$ and $\mathbb{E}[C|Z] \approx \frac{\mathbb{E}[A|Z]+\mathbb{E}[C|Z]}{2}$, and similarly for B and D.

Using these approximations, we can compute (omitting expectations on the right-hand side for brevity):

$$\mathbb{E}[Y|Z] \approx X + \frac{A+C}{2} \odot X + \frac{B+D}{2} + \frac{(A+C)^2}{4} \odot X + \frac{A+C}{2} \odot \frac{B+D}{2} \quad (1)$$

Applying our correction step, we get:

$$2\mathbb{E}[Y|Z] - Z \approx X + \frac{(A+C)^2}{2} \cdot X + \frac{(A+C) \odot (B+D)}{2} \quad (2)$$

In other words, our reconstruction will still have remaining noise, with both an intensity dependent term and a purely additive term. However, both of these terms contain a pointwise product of noise samples, and thus have significantly lower magnitude than the original noise.

To quantitatively evaluate the performance of our method in this scenario, we consider a noise model with $\alpha = \beta = 0.05$. In other words, this model consists of a purely additive component with $\sigma = 0.05$ and an intensity-dependent component with $0 \leq \sigma \leq 0.05$.

Table 1 shows the PSNR and SSIM metrics for our method and Noise2Noise under this noise model. Note that unlike the experiments in the main text, we here train a Noise2Noise model from scratch on this specific noise distribution. Just as in the case of purely additive noise, our method is about 1.5dB worse than N2N. It may be possible to close this gap somewhat by deriving an analogue of the improvement described in Section 3.3.

Figure 1 shows an example of our method trained using this noise model. We note that the result is qualitatively similar to that acheived in the purely additive setting, in line with the metrics reported in Table 1.

By contrast, Figure 2 shows the results obtained when $\alpha = \beta = 0.2$. Here the noise magnitude is quite severe. As a consequence, the remaining noise term in Equation 2 is also large. When zoomed in, this remaining noise is clearly visible. Although the approximations made in deriving the correction step lead to subtle artifacts, our method still produces a significant improvement over the noisy image.