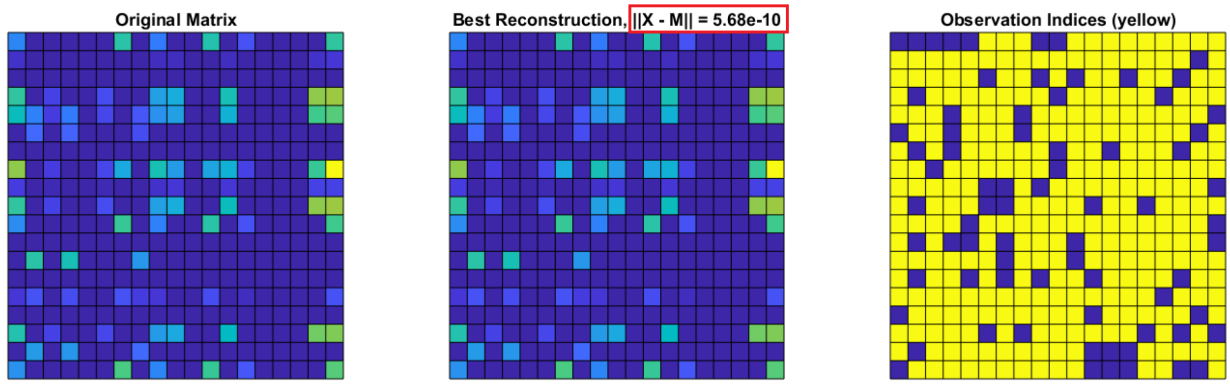
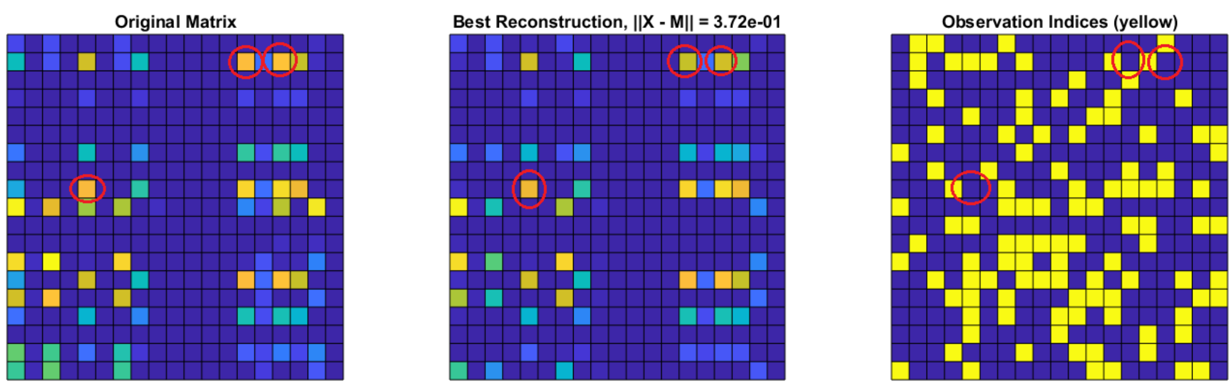
If we observe enough of the entries (80% in this example), we get perfect reconstruction for sparse matrices with relatively high probability (perfect reconstruction roughly half the time) using nuclear norm minimization

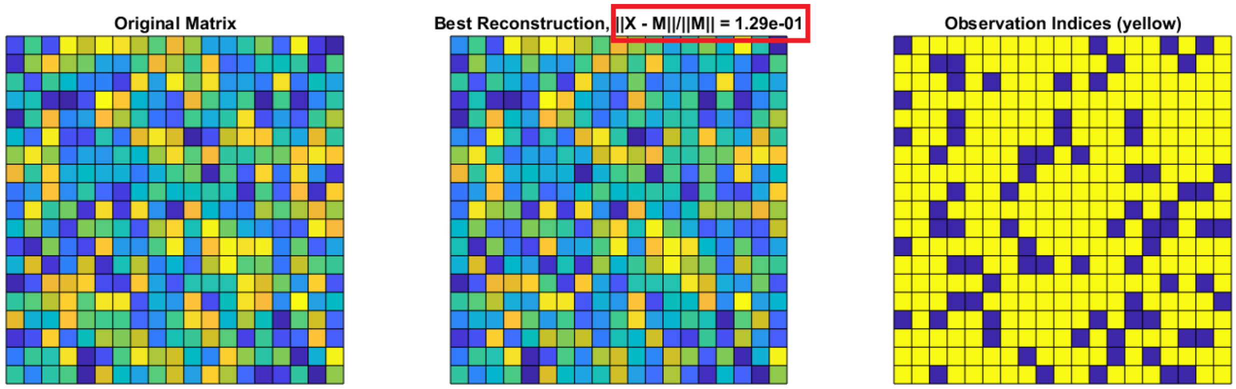
$$\min \|X\|_* \text{ subject to } P_\Omega(X) = P_\Omega(M).$$



When we sample much less (30% in this example), we no longer get perfect reconstruction, but we do preserve some of the structure of the matrix. Note the entries circled in red: these are large entries in the original matrix and, though they are unobserved, the reconstruction matches them fairly well. [Of course, the reconstruction also misses some large unobserved entries.



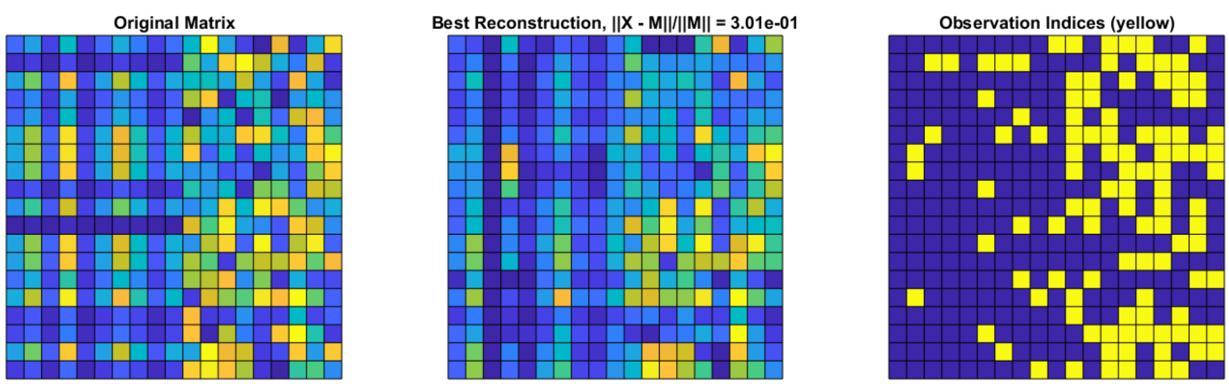
Moving to dense matrices, we cannot expect perfect reconstruction even when observing 80% of entries.



We want to consider matrice with certain columns correlated. As a first attempt, we let the first ten columns be "perfectly correlated" (i.e., they are all constant multiples of a single vector so that their correlation matrix is a 10×10 matrix where every entry is 1). In this case, we can selectively sample the entries. Here we sample 30% of the entire matrix (120 entries), but we take only 20% (24 entries) of these samples from the correlated columns and the remaining 80% (96 entries) from the uncorrelated columns. Note, we are still using the nuclear norm minimization

$$\min \|X\|_* \text{ subject to } P_\Omega(X) = P_\Omega(M).$$

This selective sampling does worse than if we just sample evenly in the same scenario. This should be expected since we haven't actually used the correlation yet.



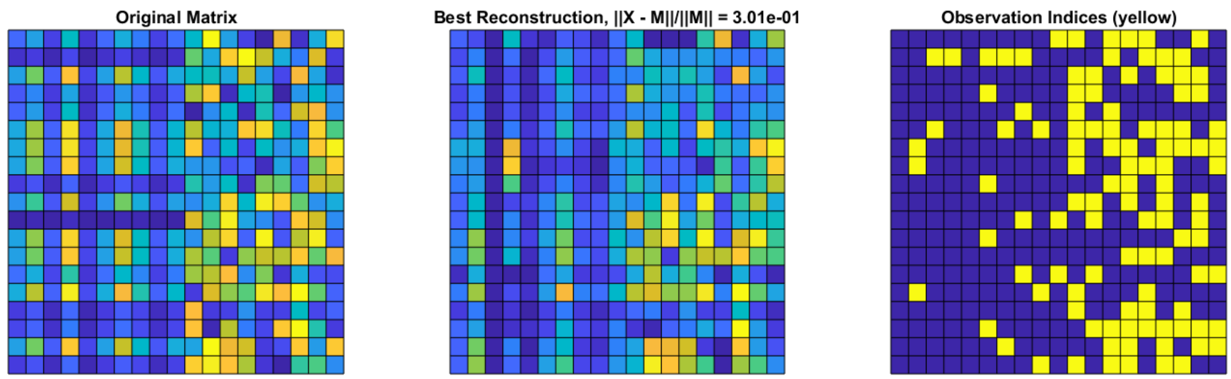
We tried to think of ways to incorporate the correlation into the problem. A first thought might be to change the functional to incorporate the correlation matrix itself, perhaps:

$$\min \|X\|_* + \gamma \|\mathrm{Corr}(X_\tau) + \mathrm{Corr}(M_\tau)\|_F^2 \text{ subject to } P_\Omega(X) = P_\Omega(M)$$

where τ represents the collection of columns that are known to be correlated. However, the map $X \mapsto \mathrm{Corr}(X)$ is non-linear and so that map $X \mapsto \|\mathrm{Corr}(X_\tau) - \mathrm{Corr}(M_\tau)\|_F^2$ is non-convex. If we wanted to do this, we would need some more sophisticated non-convex optimization routines. Instead, it may be better to use some portion of the samples to produce an estimation $\tilde{M}_\tau$ to M_τ and then insert a fidelity term in order to match X_τ to $\tilde{M}_\tau$ and combine this with nuclear norm minimization:

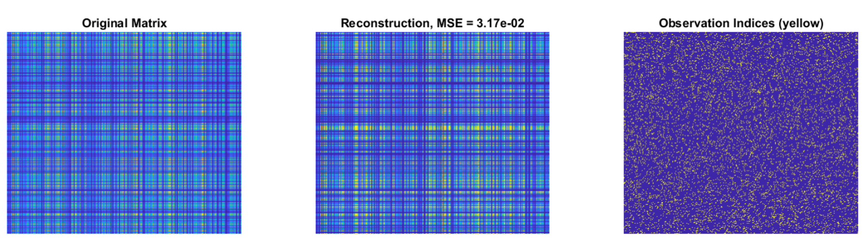
$$\min \|X\|_* + \gamma\|X_\tau + \tilde{M}_\tau\|_F^2 \ \text{ subject to } \ P_\Omega(X) = P_\Omega(M).$$

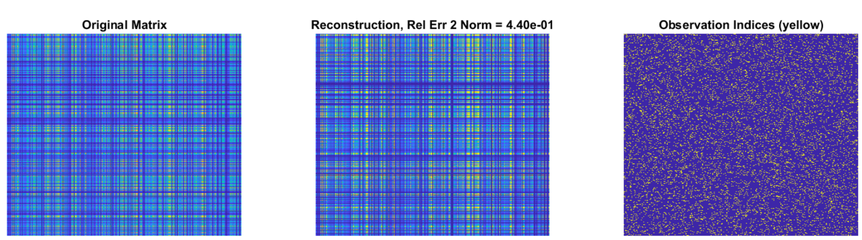
This idea shows some promise. Indeed assuming we can produce $\tilde{M}_\tau$ with some amount of accuracy, we can sample at a much lower rate from the correlated columns since they are accounted for by the fidelity term. Indeed, we did a small parameter sweep and it suggested that taking $\gamma\tilde{1}$ and taking all the samples from the uncorrelated columns is ideal (here we assume we have used 40 samples to construct $\tilde{M}_\tau$ and then take the remaining 80 samples from the uncorrelated portion).



Next we are focusing on how to accurately reconstruct M_τ with relatively few samples, given that we know that the columns are correlated (or an easier problems: given that we know exactly how the columns are correlated; i.e., we know $\mathrm{Corr}(M_\tau)$ exactly).

We've found that singular value threshholding (SVT) works pretty well for large, low rank matrices so we may be able to use SVT for $\tilde{M}_\tau$ if we assume that M_τ is low-rank. In the following picture, we are using SVT to approximately reconstruct at rank 1, 300×300 matrix. We are still not using any information about the correlation (How to incorporate this into SVT?) but even so the reconstruction is very good when we are only observing 10% of the entries.





Otherwise, we may be able to use multivariate regression to identify the correlation. PLEASE WRITE A SMALL AMOUNT ABOUT THE MV REGRESSION