Building on PCRandCS (via DW),

- I'm thinking it will be better to do this under the model in the paper. This means $U_{p-d} = 0 \Rightarrow R_d = 0$, but no harm in keeping it around for now.
- Eq.3 $\rightarrow$ Eq.4, $\tilde{U}_d$ became U_d, so $V\Lambda$ needs to be $V_d\Lambda_d$ in Eq.4 (fixed in Eq.5).
- Eq.4 $\rightarrow$ Eq.5, factoring out $U_d^\top Y$, the second half got missed. So starting from Eq.6, we should have

$$
\begin{aligned}
&= \left\| U_d(F)\Lambda_d(F)^{-1}U_d(F)^\top V_d\Lambda_d - V_d\Lambda_d^{-1} \right\| M_d + R_d \text{ (dropping } R_d \text{ now)} && (1)\\
&\leq \left\| U_d(F)\Lambda_d(F)^{-1}U_d(F)^\top V_d\Lambda_d - U_d(F)\Lambda_d(F)^{-1}\Lambda_d \right\| M_d && (2)\\
&\quad + \left\| U_d(F)\Lambda_d(F)^{-1}\Lambda_d - V_d\Lambda_d^{-1} \right\| M_d && (3)\\
&\leq \left\| U_d(F)\Lambda_d(F)^{-1} \right\| \left\| U_d(F)^\top V_d - I \right\| \left\| \Lambda_d \right\| M_d && (4)\\
&\quad + \left\| U_d(F)\Lambda_d(F)^{-1/2}\Lambda_d(F)^{-1/2}\Lambda_d - V_d\Lambda_d^{-1} \right\| M_d && (5)\\
&\leq \left\| U_d(F)\Lambda_d(F)^{-1} \right\| \left\| U_d(F)^\top V_d - I \right\| \left\| \Lambda_d \right\| M_d && (6)\\
&\quad + \left\| U_d(F)\Lambda_d(F)^{-1/2} \right\| \left\| \Lambda_d(F)^{-1/2}\Lambda_d - I \right\| M_d + \left\| U_d(F)\Lambda_d(F)^{-1/2} - V_d\Lambda_d^{-1} \right\| M_d && (7)
\end{aligned}
$$

- Is there a relationship between $\|\Lambda_d\|$ and $\|\Lambda_d(F)\|$? This would be nice.
- M_d seems like it will be a pain: $\Theta(n)$.
- My thinking (up to now) had been to mimic Paul, Bair, et. al:
  1. Show that $\|\sin(\mathcal{E}, \mathcal{F})\|$ is small where $\mathcal{E}$ is the span of V_d and $\mathcal{F}$ is the span of $U_d(F)$.
  2. Show that $\|\Lambda(F)_d - \Lambda_d\|$ is small.
  3. See whether this gives anything about $\hat{\beta}_d$.
- For the first step, this would amount to examining a function of $V_dV_d^\top - U_d(F)U_d(F)^\top$. I was thinking with Lemma 4.2 or Corollary 4.1 in Lei and Vu's sparse PCA paper. Although, this again is just a different way of measuring the approximation accuracy of $U_d(F)$.
- My thoughts on the target journal here is JCGS. To that end, I think we need some or all of the following:
  1. Minor theoretical contributions along the lines above. Get as far as we can before it gets painful, likely under strong assumptions.
  2. Do the Nystrom version as well. (Already done in simulations, it's a bit worse, though not terrible)
  3. Implement GLMs.