## A Proof of Lemma 2.2 and Main Theorem 1.1

**Proof** [Lemma 2.2] Clearly, $\widehat{\boldsymbol{L}}_{\rm root}(\mu), \widehat{\boldsymbol{S}}_{\rm root}(\mu)$ satisfies the constraint for the StablePCP$_c$ with $\delta = \delta(\mu)$. And by optimality of $\widehat{\boldsymbol{L}}_{\rm root}(\mu), \widehat{\boldsymbol{S}}_{\rm root}(\mu)$ for the $\sqrt{\rm PCP}$ problem, we get $\forall \boldsymbol{L}, \boldsymbol{S}\ s.t. \|\boldsymbol{D} - \boldsymbol{L} - \boldsymbol{S}\|_F \le \delta(\mu)$,

$$\|\boldsymbol{L}\|_* + \lambda\|\boldsymbol{S}\|_1 \ge \|\widehat{\boldsymbol{L}}_{\rm root}(\mu)\|_* + \lambda\|\widehat{\boldsymbol{S}}_{\rm root}(\mu)\|_1. \tag{A.1}$$

This shows the optimality of $\widehat{\boldsymbol{L}}_{\rm root}(\mu), \widehat{\boldsymbol{S}}_{\rm root}(\mu)$ for the StablePCP$_c$ problem with $\delta = \delta(\mu)$. ■

**Lemma A.1 (Adapted from Theorem 2 in [1])** *Under the same conditions as Theorem 2.1, with probability at least $1 - c' n_1^{-10}$, for any $\boldsymbol{Z}_0$ with $\|\boldsymbol{Z}_0\|_F \ge \delta$, the solution $\widehat{\boldsymbol{X}} = (\widehat{\boldsymbol{L}}, \widehat{\boldsymbol{S}})$ to the StablePCP$_c$ problem 1.5 with $\lambda = 1/\sqrt{n_1}$ satisfies*

$$\|\widehat{\boldsymbol{X}} - \boldsymbol{X}_0\|_F \le \sqrt{720 n_1 n_2 + 2} \cdot \|\boldsymbol{Z}_0\|_F.$$

The proof of this lemma is essentially the same as the proof of Theorem 2.1 in [1]. The main differences are: we need to replace every occurrence of $\delta$ with $\|\boldsymbol{Z}_0\|_F$, and now since $(\widehat{\boldsymbol{L}_0}, \widehat{\boldsymbol{S}_0})$ is not a feasible solution, we need to change the bound of the objective to $\|\widehat{\boldsymbol{L}}\|_* + \lambda\|\widehat{\boldsymbol{S}}\|_1 \le \|\widehat{\boldsymbol{L}} + \widehat{\boldsymbol{Z}_0}\|_* + \lambda\|\widehat{\boldsymbol{S}}\|_1 \le \|\widehat{\boldsymbol{L}}\|_* + \sqrt{n_2}\|\boldsymbol{Z}_0\|_F + \lambda\|\widehat{\boldsymbol{S}}\|_1$.

**Proof** [Theorem 1.1] Define $\widehat{\boldsymbol{\delta}_L} = \widehat{\boldsymbol{L}} - \boldsymbol{L}_0$ and $\widehat{\boldsymbol{\delta}_S} = \widehat{\boldsymbol{S}} - \boldsymbol{S}_0$. By the optimality of $\widehat{\boldsymbol{L}}, \widehat{\boldsymbol{S}}$ and triangle inequality,

$$\begin{aligned}
&(\|\boldsymbol{L}_0\|_* + \lambda\|\boldsymbol{S}_0\|_1) - (\|\widehat{\boldsymbol{L}}\|_* + \lambda\|\widehat{\boldsymbol{S}}\|_1) \\
&\ge \mu(\|\boldsymbol{D} - \widehat{\boldsymbol{L}} - \widehat{\boldsymbol{S}}\|_F - \|\boldsymbol{Z}_0\|_F) \\
&\ge \mu(\|\widehat{\boldsymbol{\delta}_L} + \widehat{\boldsymbol{\delta}_S}\|_F - 2\|\boldsymbol{Z}_0\|_F).
\end{aligned} \tag{A.2}$$

Treating the dual certificate in Lemma 2.3 as an approximate subgradient for the norm $\|\cdot\|_*$ and $\|\cdot\|_1$,

$$\begin{aligned}
&(\|\widehat{\boldsymbol{L}}\|_* - \|\boldsymbol{L}_0\|_*) + \lambda(\|\widehat{\boldsymbol{S}}\|_1 - \|\boldsymbol{S}_0\|_1) \\
&\ge \langle \widehat{\boldsymbol{\delta}_L}, \boldsymbol{U}\boldsymbol{V}^* + \boldsymbol{W}\rangle + \langle \widehat{\boldsymbol{\delta}_S}, \lambda(\operatorname{sign}(\boldsymbol{S}_0) + \boldsymbol{F})\rangle \\
&= \underbrace{\left\langle \widehat{\boldsymbol{\delta}_L} + \widehat{\boldsymbol{\delta}_S}, \boldsymbol{U}\boldsymbol{V}^* + \boldsymbol{W} - \frac{\lambda P_\Omega \boldsymbol{H}}{2} \right\rangle}_{\Theta_1} + \underbrace{\lambda \left\langle \widehat{\boldsymbol{\delta}_L} - \widehat{\boldsymbol{\delta}_S}, \frac{P_\Omega \boldsymbol{H}}{2} \right\rangle}_{\Theta_2}.
\end{aligned} \tag{A.3}$$

Next, we bound $\Theta_1, \Theta_2$ in (A.3).

$$\begin{aligned}
\Theta_1 &\ge -\|\widehat{\boldsymbol{\delta}_L} + \widehat{\boldsymbol{\delta}_S}\|_F \cdot \|\boldsymbol{U}\boldsymbol{V}^* + \boldsymbol{W} - \lambda P_\Omega \boldsymbol{H}/2\|_F \\
&\ge -\|\widehat{\boldsymbol{\delta}_L} + \widehat{\boldsymbol{\delta}_S}\|_F \cdot (\sqrt{r + n_2/4} + \lambda\|P_\Omega \boldsymbol{H}\|_F/2) \\
&\ge -\|\widehat{\boldsymbol{\delta}_L} + \widehat{\boldsymbol{\delta}_S}\|_F \cdot (\sqrt{7/10} + 1/520)\sqrt{n_2/2} \\
&\ge -0.85\mu\|\widehat{\boldsymbol{\delta}_L} + \widehat{\boldsymbol{\delta}_S}\|_F.
\end{aligned}$$

where the second inequality follows from triangle inequality and properties of $\boldsymbol{W}$ from Lemma 2.3. The third inequality follows from the condition (1.3) $r \le n_2/10$. For $\Theta_2$, by the dual construction,

$$\Theta_2 \ge -(\lambda\|P_\Omega \boldsymbol{H}\|_F/2) \cdot (\|\widehat{\boldsymbol{\delta}_L}\|_F + \|\widehat{\boldsymbol{\delta}_S}\|_F) \ge -(\|\widehat{\boldsymbol{\delta}_L}\|_F + \|\widehat{\boldsymbol{\delta}_S}\|_F)/(520\sqrt{2n_1}).$$

Combining (A.2) and (A.3) with the bounds on $\Theta_1$ and $\Theta_2$,

$$\frac{1}{520\sqrt{n_1 n_2}}(\|\widehat{\boldsymbol{\delta}_L}\|_F + \|\widehat{\boldsymbol{\delta}_S}\|_F) + 2\|\boldsymbol{Z}_0\|_F \ge 0.15\|\widehat{\boldsymbol{\delta}_L} + \widehat{\boldsymbol{\delta}_S}\|_F. \tag{A.4}$$

From Lemma 2.2, $(\widehat{\boldsymbol{L}}, \widehat{\boldsymbol{S}})$ is also the solution to the StablePCP$_c$ problem parameterized by $\delta = \|\widehat{\boldsymbol{\delta}_L} + \widehat{\boldsymbol{\delta}_S} - \boldsymbol{Z}_0\|_F$. If $\delta \le \|\boldsymbol{Z}_0\|_F$, Lemma A.1 gives the bound that

$$\|(\widehat{\boldsymbol{\delta}_L}, \widehat{\boldsymbol{\delta}_S})\|_F \le 27\sqrt{n_1 n_2}\|\boldsymbol{Z}_0\|_F.$$

If $\delta \geq \|\boldsymbol{Z}_0\|_F$, from Theorem 2.1, together with the trivial inequality that $\sqrt{320n_1n_2+4} \leq 26\sqrt{n_1n_2/2}$,

$$\|(\widehat{\boldsymbol{\delta}_L}, \widehat{\boldsymbol{\delta}_S})\|_F \leq 26\sqrt{n_1n_2/2} \cdot \|\widehat{\boldsymbol{\delta}_L} + \widehat{\boldsymbol{\delta}_S} - \boldsymbol{Z}_0\|_F. \tag{A.5}$$

Combining (A.4) and (A.5), we get

$$\|\widehat{\boldsymbol{\delta}_L}\|_F + \|\widehat{\boldsymbol{\delta}_S}\|_F \leq 26\sqrt{n_1n_2}(\|\widehat{\boldsymbol{\delta}_L} + \widehat{\boldsymbol{\delta}_S}\|_F + \|\boldsymbol{Z}_0\|_F) \leq \frac{1}{3}(\|\widehat{\boldsymbol{\delta}_L}\|_F + \|\widehat{\boldsymbol{\delta}_S}\|_F) + 373\sqrt{n_1n_2}\|\boldsymbol{Z}_0\|_F.$$

which proves the claim that

$$\|\widehat{\boldsymbol{X}} - \boldsymbol{X}_0\|_F \leq \|\widehat{\boldsymbol{\delta}_L}\|_F + \|\widehat{\boldsymbol{\delta}_S}\|_F \leq 560\sqrt{n_1n_2}\|\boldsymbol{Z}_0\|_F.$$

■

## B Stopping Criteria in Algorithm 1

The function `helper()` containing the stopping criteria and updates for $\rho$, adapted from [2], is presented in Algorithm 1.

**Algorithm 1** Function `helper()`: update $\rho$ and check convergence

**Input:** $\boldsymbol{D}, \boldsymbol{L}_1, \boldsymbol{L}_2, \boldsymbol{L}_2', \boldsymbol{S}_1, \boldsymbol{S}_2, \boldsymbol{S}_2', \boldsymbol{Z}, \boldsymbol{Y}_1, \boldsymbol{Y}_2, \boldsymbol{Y}_3 \in \mathbb{R}^{n_1 \times n_2}, \rho, \epsilon_{\text{abs}}, \epsilon_{\text{rel}}$.
**Output:** $\rho_+$, ifConverge.

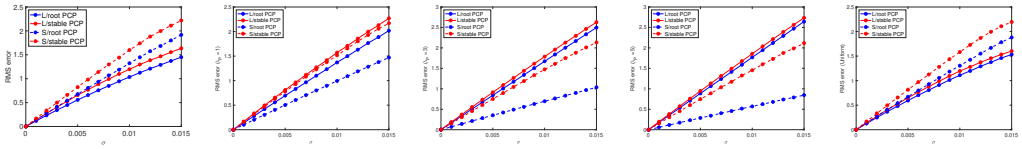
*# Calculate residuals*
$r_{\text{primal}} \leftarrow \|(\boldsymbol{L}_1 - \boldsymbol{L}_2, \boldsymbol{S}_1 - \boldsymbol{S}_2, \boldsymbol{Z} + \boldsymbol{L}_2 + \boldsymbol{S}_2 - \boldsymbol{D})\|_F$
$r_{\text{dual}} \leftarrow \rho \cdot \|(\boldsymbol{L}_2 - \boldsymbol{L}_2', \boldsymbol{S}_2 - \boldsymbol{S}_2', \boldsymbol{L}_2 + \boldsymbol{S}_2 - \boldsymbol{L}_2' - \boldsymbol{S}_2')\|_F$
*# Calculate thresholds*
$\theta_{\text{primal}} \leftarrow \epsilon_{\text{rel}} \max\{\|(\boldsymbol{L}_1, \boldsymbol{S}_1, \boldsymbol{Z})\|_F, \|(\boldsymbol{L}_2, \boldsymbol{S}_2, \boldsymbol{L}_2 + \boldsymbol{S}_2)\|_F, \|\boldsymbol{D}\|_F\} + \epsilon_{\text{abs}}\sqrt{3n_1n_2}$
$\theta_{\text{dual}} \leftarrow \epsilon_{\text{rel}}\|(\boldsymbol{Y}_1, \boldsymbol{Y}_2, \boldsymbol{Y}_3)\|_F + \epsilon_{\text{abs}}\sqrt{3n_1n_2}$
*# Update $\rho$*
$\rho_+ \leftarrow \rho$
**if** $r_{\text{primal}} > 10 \cdot r_{\text{dual}}$ **then**
  $\rho_+ \leftarrow 2\rho$
**else if** $r_{\text{dual}} > 10 \cdot r_{\text{primal}}$ **then**
  $\rho_+ \leftarrow \rho/2$
**end if**
*# Check convergence*
ifConverge $\leftarrow$ False
**if** $r_{\text{primal}} < \theta_{\text{primal}}$ and $r_{\text{dual}} < \theta_{\text{dual}}$ **then**
  ifConverge $\leftarrow$ True
**end if**
**return** $\rho_+$, ifConverge

## C Experiments and Settings

### C.1 Experiments with Different Distributions of Noise for Section 4.1

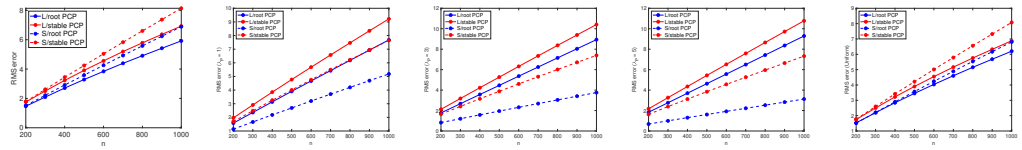
We test our $\sqrt{\text{PCP}}$ and StablePCP$_u$on simulation experiments with different noise distributions. All the setup and parameters are the same as in Section 4.1 except that now, instead of adding Gaussian noise which follows $\mathcal{N}(0, \sigma^2)$, we add (scaled) Poisson noise, $l \cdot Poisson(\lambda_P)$ where we choose $\lambda_P \in \{1, 3, 5\}$[1] with scale $l = \frac{\sigma}{\sqrt{\lambda_P + \lambda_P^2}}$, and Uniform noise $Uniform(-\sqrt{3}\sigma, \sqrt{3}\sigma)$. We choose $(\lambda_P, l)$ and the range of the Uniform distribution in this way such that $E[(\boldsymbol{Z}_0)_{ij}^2] = \sigma^2$. Results are presented in Figures 1 and 2.

[1]Since a Poisson variable with parameter $\lambda_P$ equals 0 with probability $e^{-\lambda_P}$, our choices of $\lambda_P$ give $\boldsymbol{Z}_0$'s which are approximately $36.79\%$, $4.98\%$, and $0.67\%$ sparse.



(a) Gaussian (b) Poisson $\lambda_P = 1$ (c) Poisson $\lambda_P = 3$ (d) Poisson $\lambda_P = 5$ (e) Uniform

**Figure 1:** StablePCP$_u$ vs $\sqrt{\text{PCP}}$ : effect of varying $\sigma$ for different noise distributions



(a) Gaussian (b) Poisson $\lambda_P = 1$ (c) Poisson $\lambda_P = 3$ (d) Poisson $\lambda_P = 5$ (e) Uniform

**Figure 2:** StablePCP$_u$ vs $\sqrt{\text{PCP}}$ : effect of varying $n$ for different noise distributions

### C.2 Additional Results for Section 4.2

We run the experiments on a laptop with 2.3 GHz Dual-Core Intel Core i5, and we set the maximal iteration number of our ADMM to be 5000. For experiments that don't converge in 5000 steps, the number of iterations is represented as $5000+$.

For the hall dataset, we present the relative error for $\boldsymbol{L}, \boldsymbol{S}$, the running time, and the number of iteration in Table 1 and Table 2.

**Table 1:** $\sqrt{\text{PCP}}$ : Hall dataset

| $\sigma$ | $\frac{\|\widehat{\boldsymbol{L}}-\boldsymbol{L}_0\|_F}{\|\boldsymbol{L}_0\|_F}$ | $\frac{\|\widehat{\boldsymbol{S}}-\boldsymbol{S}_0\|_F}{\|\boldsymbol{S}_0\|_F}$ | Time ($\times 10^3$ s) | Iter |
|---|---|---|---|---|
| 0 | 0.0019 | 0.0266 | 0.6894 | 1688 |
| 30 | 0.0445 | 0.7403 | 0.5381 | 1293 |
| 60 | 0.0737 | 1.3525 | 0.9470 | 2425 |
| 90 | 0.0968 | 1.9288 | 1.4040 | 3258 |
| 120 | 0.1200 | 2.5067 | 1.6230 | 4113 |

**Table 2:** StablePCP$_u$: Hall dataset

| $\sigma$ | $\frac{\|\widehat{\boldsymbol{L}}-\boldsymbol{L}_0\|_F}{\|\boldsymbol{L}_0\|_F}$ | $\frac{\|\widehat{\boldsymbol{S}}-\boldsymbol{S}_0\|_F}{\|\boldsymbol{S}_0\|_F}$ | Time ($\times 10^3$ s) | Iter |
|---|---|---|---|---|
| 0 | 0.0019 | 0.0266 | 1.7340 | 4423 |
| 30 | 0.0443 | 0.7495 | 1.8936 | 4918 |
| 60 | 0.0740 | 1.3494 | 1.3006 | 3362 |
| 90 | 0.0974 | 1.8922 | 1.8050 | 4563 |
| 120 | 0.1214 | 2.4349 | 1.9886 | 5000+ |

In Figure 3 and 4, we present more results for frame 1, 20 for varying $\sigma$ for this hall dataset.

We also apply our algorithms to the dataset lights. We present the relative error for $\boldsymbol{L}, \boldsymbol{S}$, the running time, and the number of iteration in Table 3 and Table 4.

**Table 3:** $\sqrt{\text{PCP}}$ : Lights dataset

| $\sigma$ | $\frac{\|\widehat{\boldsymbol{L}}-\boldsymbol{L}_0\|_F}{\|\boldsymbol{L}_0\|_F}$ | $\frac{\|\widehat{\boldsymbol{S}}-\boldsymbol{S}_0\|_F}{\|\boldsymbol{S}_0\|_F}$ | Time ($\times 10^3$ s) | Iter |
|---|---|---|---|---|
| 0 | 0.0013 | 0.1052 | 0.7100 | 1605 |
| 30 | 0.0520 | 2.9707 | 0.5120 | 1155 |
| 60 | 0.0938 | 5.8908 | 1.7101 | 3880 |
| 90 | 0.1323 | 8.7761 | 2.1983 | 5000+ |
| 120 | 0.1689 | 11.5848 | 2.5344 | 5000+ |

**Table 4:** StablePCP$_u$: Lights dataset

| $\sigma$ | $\frac{\|\widehat{\boldsymbol{L}}-\boldsymbol{L}_0\|_F}{\|\boldsymbol{L}_0\|_F}$ | $\frac{\|\widehat{\boldsymbol{S}}-\boldsymbol{S}_0\|_F}{\|\boldsymbol{S}_0\|_F}$ | Time ($\times 10^3$ s) | Iter |
|---|---|---|---|---|
| 0 | 0.0013 | 0.1052 | 2.2069 | 5000+ |
| 30 | 0.0527 | 2.7920 | 1.4677 | 3377 |
| 60 | 0.0951 | 5.5085 | 1.2136 | 2804 |
| 90 | 0.1342 | 8.2023 | 1.6699 | 3886 |
| 120 | 0.1710 | 10.8930 | 2.0856 | 4902 |

In Figure 5 and 6, we present more results for frame 1, 20 for varying $\sigma$ for this lights dataset.

In Figure 7(a), we present the RMS error. Again, we see that the error is linear in the noise level $\sigma$.

### C.3 Additional Experiments Using Face Dataset

In addition to the video dataset in the previous section, we also test $\sqrt{\text{PCP}}$ and StablePCP$_u$ on datasets of face images. It has been pointed out in [3] that under distant illumination,

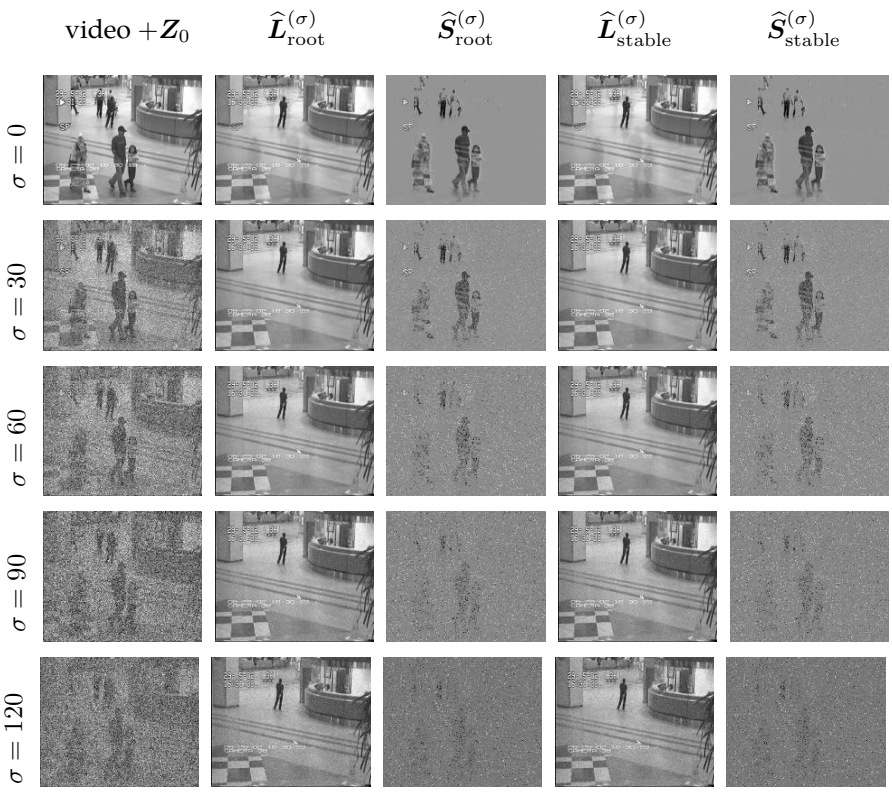



**Figure 3:** hall: recovered $\widehat{\boldsymbol{L}}$, $\widehat{\boldsymbol{S}}$ for frame 1

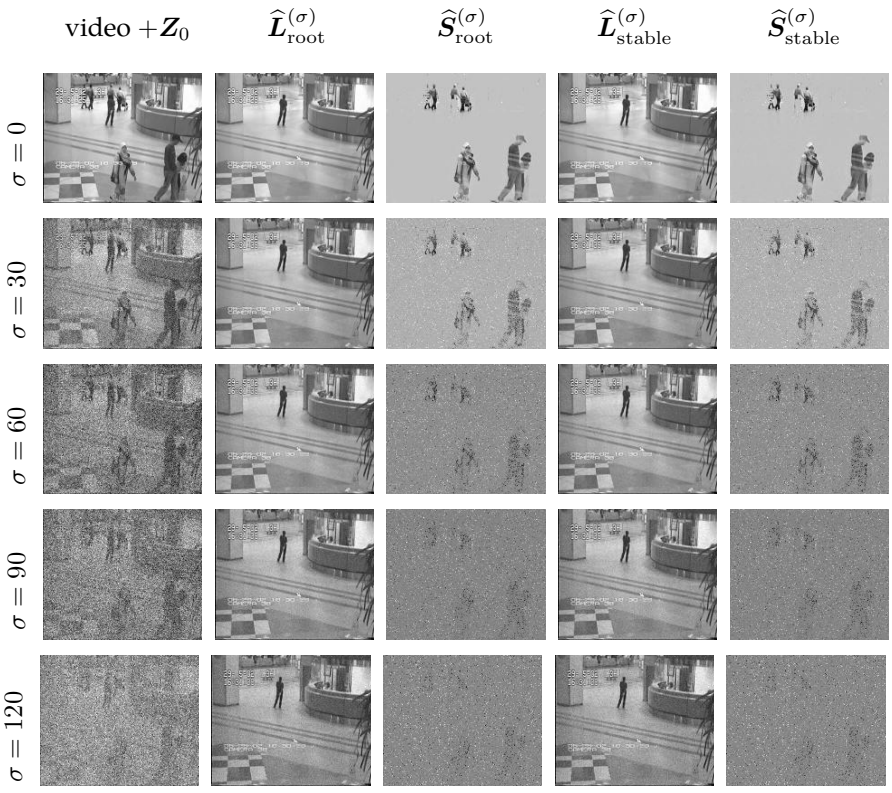



**Figure 4:** hall: recovered $\widehat{\boldsymbol{L}}$, $\widehat{\boldsymbol{S}}$ for frame 20

images of a convex Lambertian object lie near a low dimensional linear subspace called

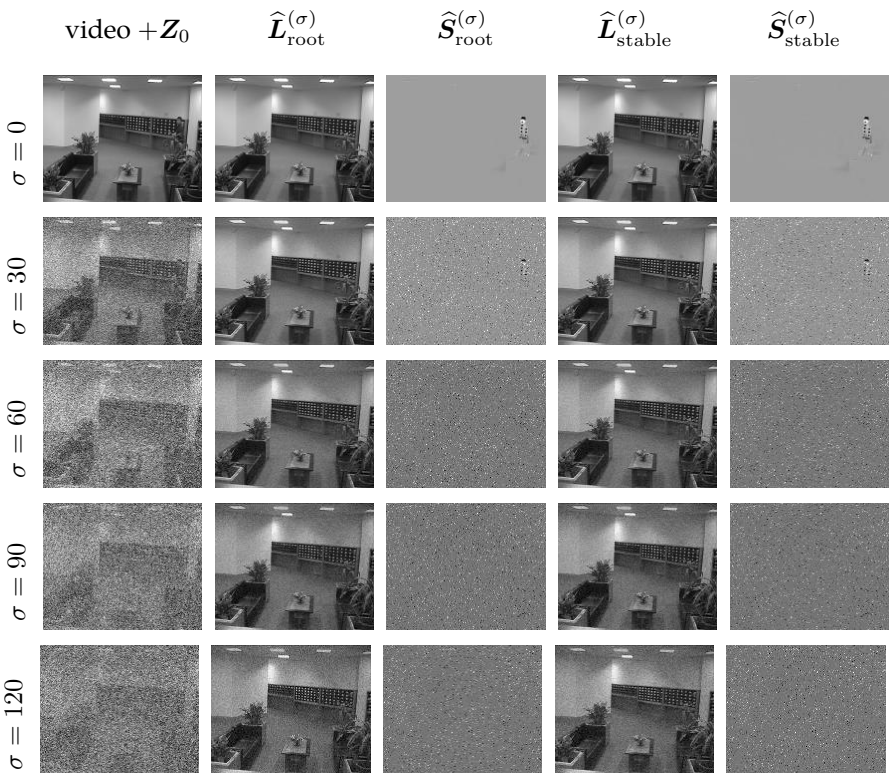



**Figure 5:** lights: recovered $\widehat{\boldsymbol{L}}, \widehat{\boldsymbol{S}}$ for frame 1

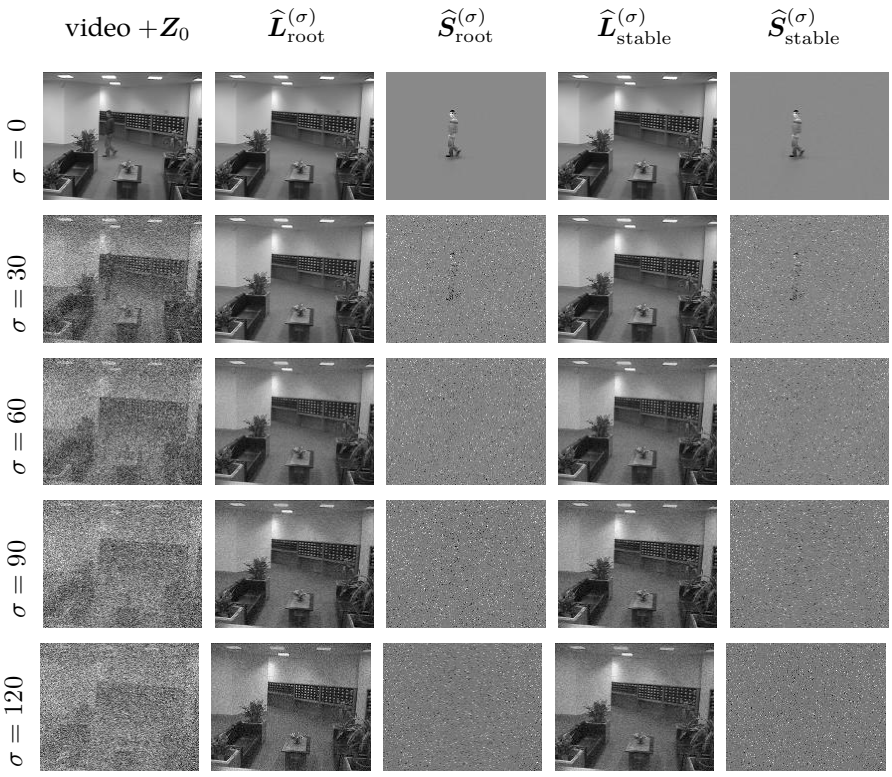
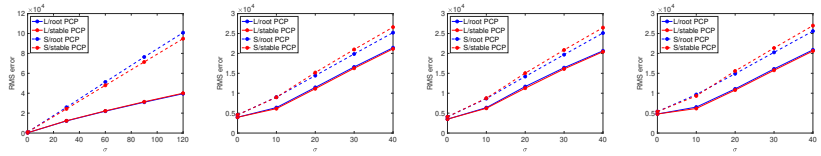



**Figure 6:** lights: recovered $\widehat{\boldsymbol{L}}, \widehat{\boldsymbol{S}}$ for frame 20

the *harmonic plane*. However, real images of faces are often corrupted by shadows and specularities, which can have large magnitudes but are sparse in the spatial domain. This




(a) "lights"+noise (b) yaleB01 + noise (c) yaleB02 + noise (d) yaleB03 + noise

**Figure 7:** StablePCP$_u$ vs $\sqrt{\text{PCP}}$ : real datasets

fits well into our low-rank and sparse model [2], and our goal is to remove these shadows and specularities from the noisy and corrupted observation.

To be precise, the dataset we use is from Yale B face database [4]. For each face, there are 65 images of resolution $192 \times 168$ under various illuminations, so we have $n_1 = 192 \times 168 = 32256$ and $n_2 = 65$. Similar to the experiments on the video dataset, we assume that there is no noise in these images, and add $\boldsymbol{Z}_0$ with $\sigma \in \{0, 10, 20, 30, 40\}$.

In Figures 8, 9, 10, 11, 12, and 13, we present the recovered low rank and sparse matrices for frame 1 and frame 20 using $\sqrt{\text{PCP}}$ and StablePCP$_u$.

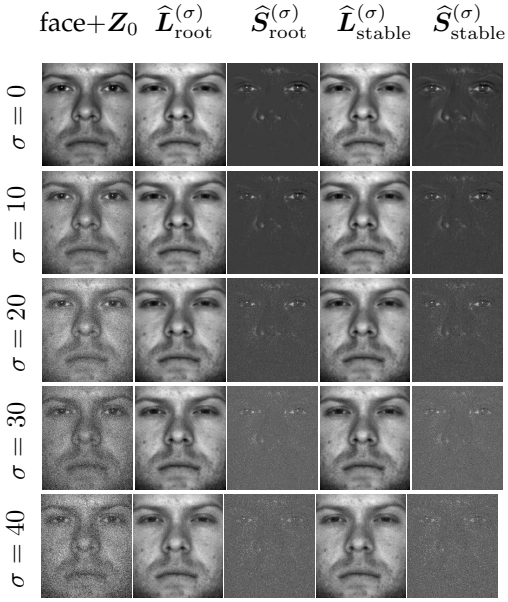



**Figure 8:** yaleB01: recovered $\widehat{\boldsymbol{L}}, \widehat{\boldsymbol{S}}$ for frame 1

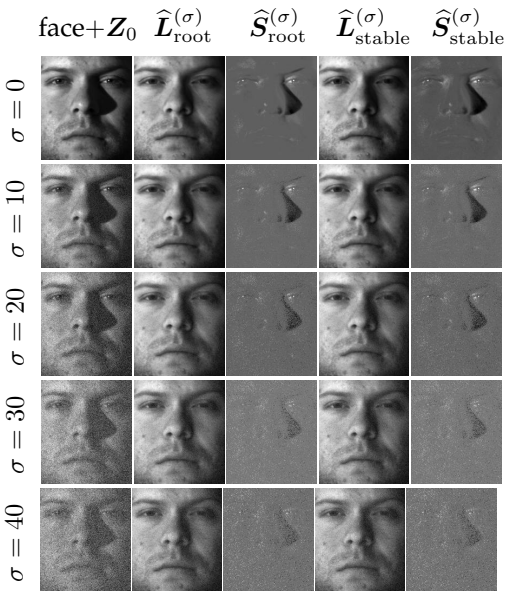



**Figure 9:** yaleB01: recovered $\widehat{\boldsymbol{L}}, \widehat{\boldsymbol{S}}$ for frame 20

In Table 5, 7, and 9, we show the relative error, running time and iteration for $\sqrt{\text{PCP}}$ on yaleB01, yaleB02, and yaleB03 datasets. In Table 6, 8, and 10, we show the relative error, running time and iteration for StablePCP$_u$on yaleB01, yaleB02, and yaleB03 datasets.

**Table 5:** $\sqrt{\text{PCP}}$ : yaleB01

| $\sigma$ | $\frac{\|\widehat{\boldsymbol{L}}-\boldsymbol{L}_0\|_F}{\|\boldsymbol{L}_0\|_F}$ | $\frac{\|\widehat{\boldsymbol{S}}-\boldsymbol{S}_0\|_F}{\|\boldsymbol{S}_0\|_F}$ | Time (s) | Iter |
|---|---|---|---|---|
| 0 | 0.0298 | 0.1934 | 255.6449 | 2054 |
| 10 | 0.0481 | 0.3823 | 171.0386 | 1366 |
| 20 | 0.0863 | 0.6117 | 184.6412 | 1480 |
| 30 | 0.1250 | 0.8421 | 210.4474 | 1693 |
| 40 | 0.1613 | 1.0692 | 493.5861 | 3953 |

**Table 6:** StablePCP$_u$: yaleB01

| $\sigma$ | $\frac{\|\widehat{\boldsymbol{L}}-\boldsymbol{L}_0\|_F}{\|\boldsymbol{L}_0\|_F}$ | $\frac{\|\widehat{\boldsymbol{S}}-\boldsymbol{S}_0\|_F}{\|\boldsymbol{S}_0\|_F}$ | Time (s) | Iter |
|---|---|---|---|---|
| 0 | 0.0298 | 0.1934 | 602.0781 | 5000+ |
| 10 | 0.0461 | 0.3808 | 599.3534 | 5000+ |
| 20 | 0.0838 | 0.6441 | 594.2583 | 5000+ |
| 30 | 0.1225 | 0.8898 | 596.9568 | 5000+ |
| 40 | 0.1591 | 1.1276 | 602.5028 | 5000+ |

[2] Although faces are not convex Lambertian objects and the harmonic plane may not apply here, previous experiments in [3] have shown the effectiveness of PCP in this task.

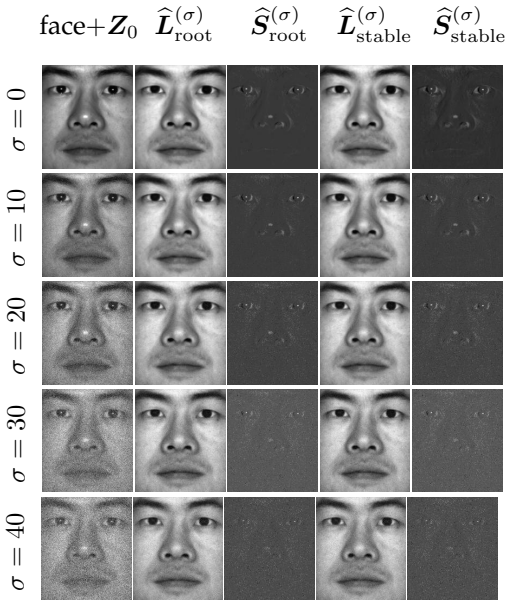



**Figure 10:** yaleB02: recovered $\widehat{L}, \widehat{S}$ for frame 1

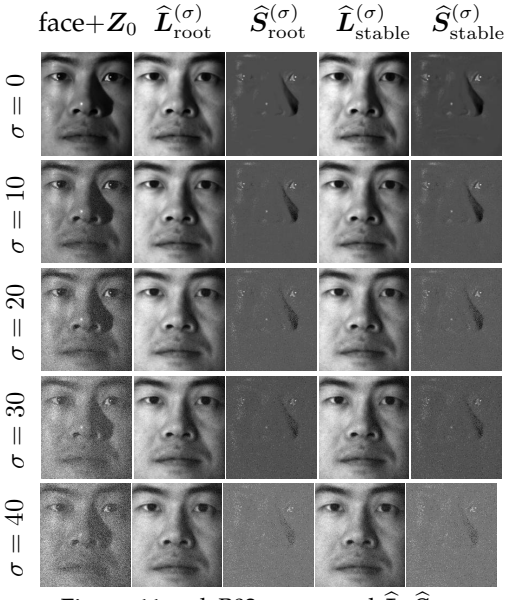



**Figure 11:** yaleB02: recovered $\widehat{L}, \widehat{S}$ for frame 20

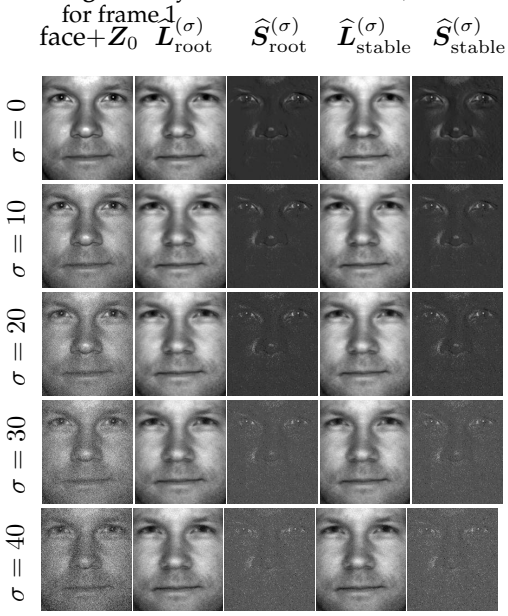



**Figure 12:** yaleB03: recovered $\widehat{L}, \widehat{S}$ for frame 1

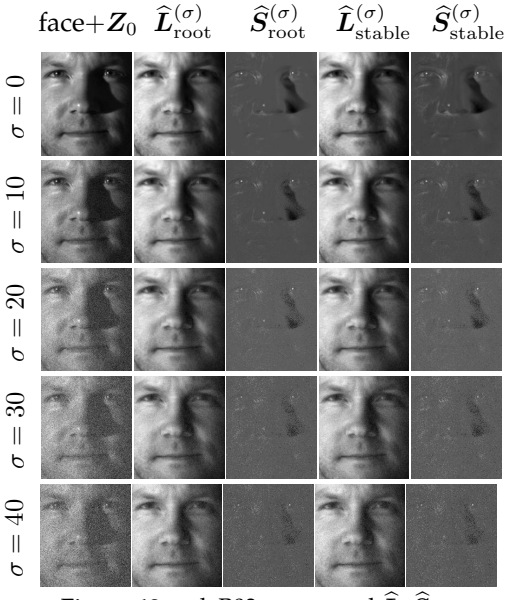



**Figure 13:** yaleB03: recovered $\widehat{L}, \widehat{S}$ for frame 20

**Table 7:** $\sqrt{\text{PCP}}$ : yaleB02

| $\sigma$ | $\frac{\|\widehat{L}-L_0\|_F}{\|L_0\|_F}$ | $\frac{\|\widehat{S}-S_0\|_F}{\|S_0\|_F}$ | Time (s) | Iter |
|---|---|---|---|---|
| 0 | 0.0261 | 0.1720 | 272.5150 | 2190 |
| 10 | 0.0477 | 0.3657 | 167.7344 | 1352 |
| 20 | 0.0873 | 0.5993 | 174.0215 | 1409 |
| 30 | 0.1228 | 0.8311 | 407.7892 | 3289 |
| 40 | 0.1545 | 1.0599 | 244.3114 | 1970 |

**Table 8:** StablePCP$_u$: yaleB02

| $\sigma$ | $\frac{\|\widehat{L}-L_0\|_F}{\|L_0\|_F}$ | $\frac{\|\widehat{S}-S_0\|_F}{\|S_0\|_F}$ | Time (s) | Iter |
|---|---|---|---|---|
| 0 | 0.0261 | 0.1720 | 595.5254 | 5000+ |
| 10 | 0.0465 | 0.3697 | 595.0926 | 5000+ |
| 20 | 0.0843 | 0.6348 | 594.2518 | 5000+ |
| 30 | 0.1204 | 0.8808 | 599.8056 | 5000+ |
| 40 | 0.1528 | 1.1174 | 593.7441 | 5000+ |

### C.4 Additional Results for Section 4.3

We provide frame 30, 60, and 90 for the three videos in Figures 14, 15, and 16.

**Table 9:** $\sqrt{\text{PCP}}$ : yaleB03

| $\sigma$ | $\frac{\|\widehat{\boldsymbol{L}}-\boldsymbol{L}_0\|_F}{\|\boldsymbol{L}_0\|_F}$ | $\frac{\|\widehat{\boldsymbol{S}}-\boldsymbol{S}_0\|_F}{\|\boldsymbol{S}_0\|_F}$ | Time (s) | Iter |
|---|---|---|---|---|
| 0 | 0.0336 | 0.2128 | 272.8555 | 2194 |
| 10 | 0.0458 | 0.3823 | 175.4205 | 1417 |
| 20 | 0.0780 | 0.5887 | 261.3707 | 1519 |
| 30 | 0.1133 | 0.8008 | 221.5380 | 1751 |
| 40 | 0.1469 | 1.0114 | 512.7160 | 4051 |

**Table 10:** StablePCP$_u$: yaleB03

| $\sigma$ | $\frac{\|\widehat{\boldsymbol{L}}-\boldsymbol{L}_0\|_F}{\|\boldsymbol{L}_0\|_F}$ | $\frac{\|\widehat{\boldsymbol{S}}-\boldsymbol{S}_0\|_F}{\|\boldsymbol{S}_0\|_F}$ | Time (s) | Iter |
|---|---|---|---|---|
| 0 | 0.0336 | 0.2128 | 594.5691 | 5000+ |
| 10 | 0.0434 | 0.3687 | 594.4220 | 5000+ |
| 20 | 0.0760 | 0.6162 | 596.9680 | 5000+ |
| 30 | 0.1109 | 0.8440 | 839.4338 | 5000+ |
| 40 | 0.1449 | 1.0662 | 601.5731 | 5000+ |

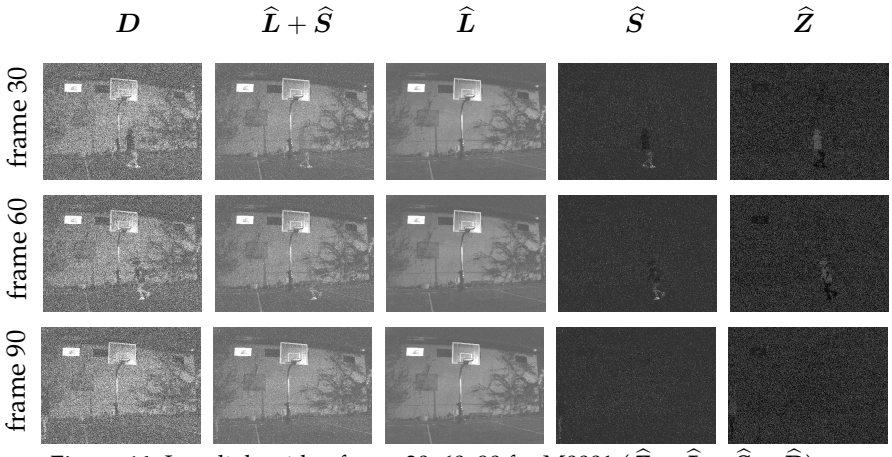



**Figure 14:** Low light video frame 30, 60, 90 for M0001 ($\widehat{\boldsymbol{Z}} = \widehat{\boldsymbol{L}} + \widehat{\boldsymbol{S}} - \widehat{\boldsymbol{D}}$).

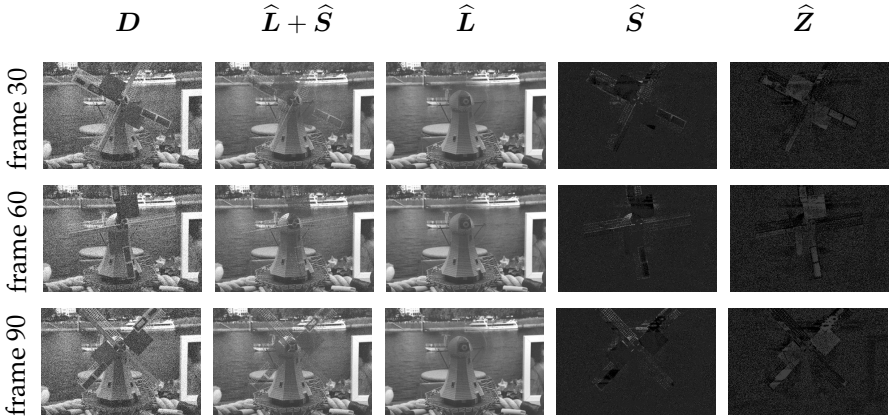



**Figure 15:** Low light video frame 30, 60, 90 for M0004 ($\widehat{\boldsymbol{Z}} = \widehat{\boldsymbol{L}} + \widehat{\boldsymbol{S}} - \widehat{\boldsymbol{D}}$).

### C.5 Additional Results for Section 4.4

Results for frame 150 and 200 are presented in Figure 17.

### C.6 Settings and Additional Results for Section 4.5

Table 11 below lists the settings for this set of experiments.

**Table 11:** Parameters for Simulation: $\rho_S = 0.1$, $|(\boldsymbol{S}_0)_{(i,j)\in\Omega}| = 0.05$

| $n_1$ | $n_2$ | $\rho_L$ | $\sigma$ |
|---|---|---|---|
| $\{200, 300, \ldots, 1000\}$ | $n_2 = n_1$ | 0.1 | 0.01 |
| $\{300, \ldots, 1000\}$ | 300 | 0.1 | 0.01 |
| 300 | 300 | $\{0.05, 0.1, \ldots, 0.5\}$ | 0.01 |
| 300 | 300 | 0.1 | $\{0.005, 0.01, \ldots, 0.05\}$ |

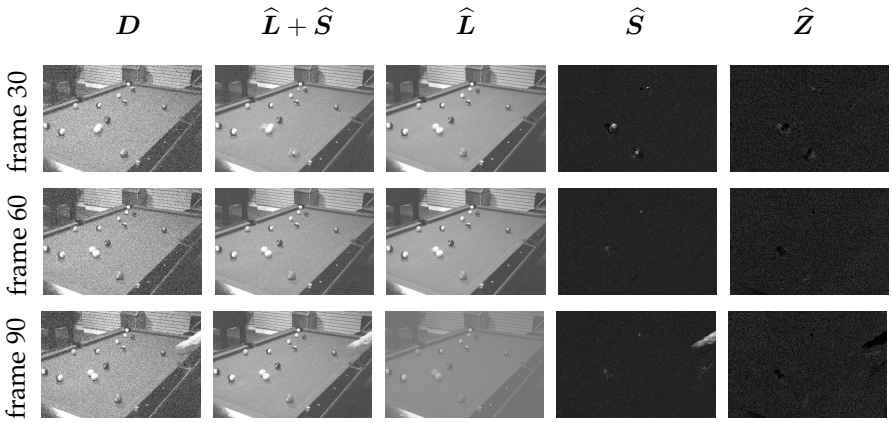
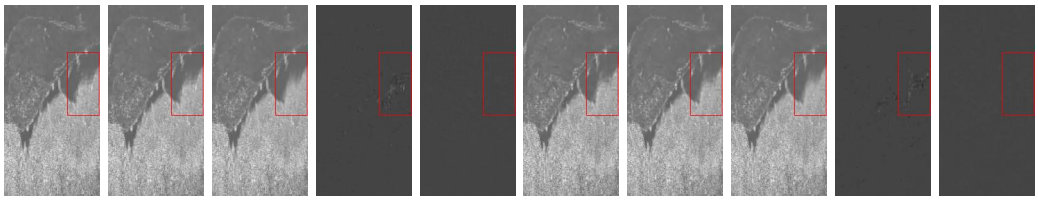


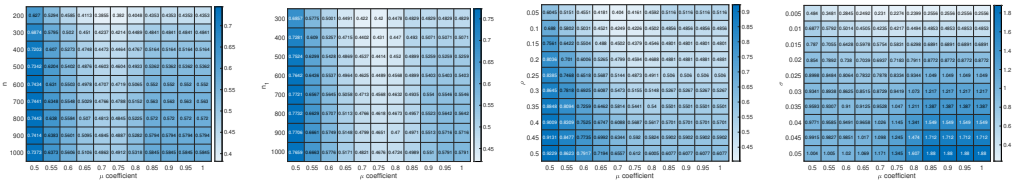

**Figure 16:** Low light video frame 30, 60, 90 for M0009 ($\widehat{\boldsymbol{Z}} = \widehat{\boldsymbol{L}} + \widehat{\boldsymbol{S}} - \widehat{\boldsymbol{D}}$).



(a) $\boldsymbol{D}$ (b) $\widehat{\boldsymbol{L}}+\widehat{\boldsymbol{S}}$ (c) $\widehat{\boldsymbol{L}}$ (d) $\widehat{\boldsymbol{S}}$ (e) $\widehat{\boldsymbol{Z}}$ (f) $\boldsymbol{D}$ (g) $\widehat{\boldsymbol{L}}+\widehat{\boldsymbol{S}}$ (h) $\widehat{\boldsymbol{L}}$ (i) $\widehat{\boldsymbol{S}}$ (j) $\widehat{\boldsymbol{Z}}$

**Figure 17:** OCT, a-e): frame 150, f-j): frame 200 ($\widehat{\boldsymbol{Z}} = \widehat{\boldsymbol{L}} + \widehat{\boldsymbol{S}} - \widehat{\boldsymbol{D}}$).

We also provide heatmaps for the relative recovery error, i.e. the 10-average of $\eta_{\text{rel}}(\mu) := \|(\widehat{\boldsymbol{L}}(\mu), \widehat{\boldsymbol{S}}(\mu)) - (\boldsymbol{L}_0, \boldsymbol{S}_0)\|_F / \|(\boldsymbol{L}_0, \boldsymbol{S}_0)\|_F$, in Figure 18.



(a) vary $n = n_1 = n_2$ (b) vary $n_1$ (c) vary $\rho_L$ (d) vary $\sigma$

**Figure 18:** $\eta_{\text{rel}}(\mu)$ under different varying parameters

## References


[1] Zihan Zhou, Xiaodong Li, John Wright, Emmanuel Candes, and Yi Ma. Stable principal component pursuit. In *2010 IEEE international symposium on information theory*, pages 1518–1522. IEEE, 2010.

[2] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.*, 3(1):1–122, January 2011.

[3] Emmanuel J Candès, Xiaodong Li, Yi Ma, and John Wright. Robust principal component analysis? *Journal of the ACM (JACM)*, 58(3):1–37, 2011.

[4] A.S. Georghiades, P.N. Belhumeur, and D.J. Kriegman. From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Trans. Pattern Anal. Mach. Intelligence*, 23(6):643–660, 2001.