A. Convergence Proof

Definition 4. *Let Φ be the set of all functions $\phi : \mathbb{E} \to \mathbb{R}_+$ which are lower semi-continuous function satisfying the following properties:*

$$\begin{array}{cl} (i) & \phi(\mathbf{0}) = 0, \\ (ii) & \phi(\mathbf{x}) = \phi(-\mathbf{x})(\textit{symmetry}), \\ (iii) & \phi(\mathbf{x}+\mathbf{y}) \le \phi(\mathbf{x}) + \phi(\mathbf{y})(\textit{subadditivity}). \end{array}$$

Here $\mathbb{E}$ is a finite dimensional Euclidean space.

We can verify that the function of matrix Z involved in the definition of the k-BDMS, *i.e.*, $\mathrm{rank}(L_W)$ with $W = (|Z| + |Z^T|)/2$, falls in the above set Φ.

Definition 5 (SRIP(k, α))**.** *We say the SRIP(k, α) holds for an affine operator $\mathcal{A}$ if there exist $\nu_k, \mu_k > 0$ satisfying $\mu_k/\nu_k < \alpha$ such that*

$$\nu_k \|\mathbf{x}\| \le \|\mathcal{A}(\mathbf{x})\| \le \mu_k \|\mathbf{x}\|, \forall \mathbf{x} \in \mathcal{C}_k,$$

where $\mathcal{C}_k := \{\mathbf{x} : \phi(\mathbf{x}) \le k\}$ is a nonconvex constraint set parameterized by k.

We have the following convergence guarantee for applying the gradient projection algorithm (Algorithm 1) to optimize the function f_1 in Eqn. (1).

Theorem 2 (Convergence of Alg. 1 for BD-SSC)**.** *Consider the Gradient Projection (GP) method with a constant stepsize $\eta_t = \eta \in [\mu_k^2, 2\nu_k^2)$ and suppose that SRIP$(k, \sqrt{2})$ is satisfied. Then*

$$f_1(Z_{t+1}) - f_1(Z^*) \le \left(\rho - \frac{1}{2}\right)(f_1(Z_t) - f_1(Z^*)), \forall t \ge 0$$

with $\rho = \eta/2\nu_k^2$. As a consequence,

$$f_1(Z_{t+1}) - f_1(Z^*) \le \left(\rho - \frac{1}{2}\right)^t (f_1(Z_0) - f_1(Z^*)), \forall t \ge 0$$

and $f_1(Z_t) \to f_1(Z^*)$ as $t \to \infty$.

Proof. Let

$$q_t(Z, Z_t) := f_1(Z_t) + \langle Z - Z_t, \partial f_1(Z_t)\rangle + \frac{\eta_t}{2}\|Z - Z_t\|_F^2.$$

Then the GP method can be equivalently rewritten as

$$Z_{t+1} \in \arg\min \{q_t(Z, Z_t) : Z \in \mathcal{K}\},$$

and hence, for the global optimum $Z^* \in \mathcal{K}$ it holds that

$$q_t(Z_{t+1}, Z_t) \le q_t(Z^*, Z_t). \tag{6}$$

Now, since $f_1(Z) = \frac{\lambda}{2}\|XZ - X\|_F^2 + \|Z\|_1$, it follows that

$$\begin{aligned} & f_1(Z_{t+1}) && (7) \\ = \;& f_1(Z_t) + \langle Z_{t+1} - Z_t, \partial f_1(Z_t)\rangle + \frac{1}{2}\|X(Z_{t+1} - Z_t)\|_F^2 \\ \overset{\text{SRIP}}{\le} \;& f_1(Z_t) + \langle Z_{t+1} - Z_t, \partial f_1(Z_t)\rangle + \frac{\eta_t}{2}\|Z_{t+1} - Z_t\|_F^2, \end{aligned}$$

where the last inequality follows from the fact that $Z_t - Z_{t+1} \in \mathcal{C}_k$ (by the subadditivity and symmetry of the function $\phi \in \Phi$) and from the fact that the definition of the stepsize implies that $\|X(Z_{t+1} - Z_t)\|_F \le \sqrt{\eta_t}\|Z_{t+1} - Z_t\|$. Therefore, we have shown that $f_1(Z_{t+1}) \le q_t(Z_{t+1}, Z_t)$ so that

$$f_1(Z_{t+1}) = q_t(Z_{t+1}, Z_t) \overset{(6)}{\le} q_t(Z^*, Z_t).$$

On the other hand,

$$\begin{aligned} & q_k(Z^*, Z_t) \\ & = f_1(Z_t) + \langle Z^* - Z_t, \partial f_1(Z_t)\rangle + \frac{\eta_t}{2}\|Z^* - Z_t\|_F^2 \\ & \overset{\text{SRIP}}{\le} f_1(Z_t) + \langle Z^* - Z_t, \partial f_1(Z_t)\rangle + \frac{\eta_t}{2\nu_k^2}\|X(Z^* - Z_t)\|_F^2 \\ & \overset{(7)}{=} f_1(Z^*) + \left(\frac{\eta_t}{2\nu_k^2} - \frac{1}{2}\right)\|XZ^* - XZ_t\|_F^2 \\ & \le f_1(Z^*) + \left(\frac{\eta_t}{2\nu_k^2} - \frac{1}{2}\right)(f_1(Z_t) - f_1(Z^*)). \end{aligned}$$

Therefore, we have,

$$f_1(Z_{t+1}) - f_1(Z^*) \le \left(\frac{\eta_t}{2\nu_k^2} - \frac{1}{2}\right)(f_1(Z_t) - f_1(Z^*)).$$

□

Similarly, we have the convergence guarantee for applying the gradient projection algorithm (Algorithm 1) to optimize the function f_* in Eqn. (2).

Theorem 3 (Convergence of Alg. 1 for BD-LRR)**.** *Also consider the Gradient Projection (GP) method with a constant stepsize $\eta_t = \eta \in [\mu_k^2, 2\nu_k^2)$ and suppose that SRIP$(k, \sqrt{2})$ is satisfied. Then*

$$f_*(Z_{t+1}) - f_*(Z^*) \le \left(\rho - \frac{1}{2}\right)(f_*(Z_t) - f_*(Z^*)), \forall t \ge 0$$

with $\rho = \eta/2\nu_k^2$. As a consequence,

$$f_*(Z_{t+1}) - f_*(Z^*) \le \left(\rho - \frac{1}{2}\right)^t (f_*(Z_0) - f_*(Z^*)), \forall t \ge 0$$

and $f_*(Z_t) \to f_*(Z^*)$ as $t \to \infty$.

Proof. The proof exactly follows the procedure of proving Theorem 2. □