Consider the optimization problem:

$$\begin{aligned} &\min_{\boldsymbol{x},\boldsymbol{z}} \sum_{\ell=1}^{L} \frac{\mu_\ell}{2} \|\boldsymbol{z}_{\ell-1} - \boldsymbol{D}_\ell \boldsymbol{x}_\ell\|_2^2 + \lambda \|\boldsymbol{b}_\ell \cdot \boldsymbol{z}_\ell\|_1 \\ \text{subject to } &\boldsymbol{W}\boldsymbol{z}_0 - \boldsymbol{s} = 0 \\ &\sqrt{\mu_\ell}\boldsymbol{R}_\ell^{-1}\boldsymbol{z}_\ell - \sqrt{\mu_\ell}\boldsymbol{R}_\ell^{-1}\boldsymbol{x}_\ell = 0 \\ &\boldsymbol{z}_\ell \geq 0 \end{aligned} \tag{1}$$

This optimization problem has the augmented Lagrangian function:

$$\begin{aligned} \mathrm{L}_\rho(\boldsymbol{x},\boldsymbol{z},\boldsymbol{\gamma}) =& f(\boldsymbol{x},\boldsymbol{z}) + \frac{\rho}{2}\|\boldsymbol{W}\boldsymbol{z}_0 - \boldsymbol{s} + \frac{\gamma_0}{\rho}\|_2^2 + \frac{\rho}{2}\sum_{\ell=1}^{L}\|\sqrt{\mu_\ell}\boldsymbol{R}_\ell^{-1}(\boldsymbol{z}_\ell - \boldsymbol{x}_\ell) + \frac{\gamma_\ell}{\rho}\|_2^2 \\ &\text{subject to } \boldsymbol{z}_\ell \geq 0 \end{aligned} \tag{2}$$

1 Update Equation for $\boldsymbol{x}_\ell$

$$\nabla_{\boldsymbol{x}_\ell} f(\boldsymbol{x},\boldsymbol{z}) = \mu_\ell \boldsymbol{D}_\ell^H \boldsymbol{D}_\ell \boldsymbol{x}_\ell - \mu_\ell \boldsymbol{D}_\ell^H \boldsymbol{z}_{\ell-1} \tag{3}$$

$$\nabla_{\boldsymbol{x}_\ell} \frac{1}{2}\|\sqrt{\mu_\ell}\boldsymbol{R}_\ell^{-1}(\boldsymbol{z}_\ell - \boldsymbol{x}_\ell) + \frac{\gamma_\ell}{\rho}\|_2^2 = \mu_\ell \boldsymbol{R}_\ell^{-2}\boldsymbol{x} - \mu_\ell \boldsymbol{R}_\ell^{-2}\boldsymbol{z}_\ell - \frac{\sqrt{\mu_\ell}\boldsymbol{R}_\ell^{-1}\gamma_\ell}{\rho} \tag{4}$$

$$\nabla_{\boldsymbol{x}_\ell} \mathrm{L}_\rho(\boldsymbol{x},\boldsymbol{z},\boldsymbol{\gamma}) = \mu_\ell \boldsymbol{D}_\ell^H \boldsymbol{D}_\ell \boldsymbol{x}_\ell - \mu_\ell \boldsymbol{D}_\ell^H \boldsymbol{z}_{\ell-1} + \rho(\mu_\ell \boldsymbol{R}_\ell^{-2}\boldsymbol{x} - \mu_\ell \boldsymbol{R}_\ell^{-2}\boldsymbol{z}_\ell - \frac{\sqrt{\mu_\ell}\boldsymbol{R}_\ell^{-1}\boldsymbol{\gamma}_\ell}{\rho}) \tag{5}$$

For $\boldsymbol{x}$, $\boldsymbol{z}$, $\boldsymbol{\gamma}$, such that $\nabla_{\boldsymbol{x}_\ell} \mathrm{L}_\rho(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_L,\boldsymbol{z}_0,\ldots,\boldsymbol{z}_L,\boldsymbol{\gamma}_0,\ldots,\boldsymbol{\gamma}_L) = 0$:

$$\mu_\ell(\rho\boldsymbol{R}_\ell^{-2} + \boldsymbol{D}_\ell^H\boldsymbol{D}_\ell)\boldsymbol{x}_\ell = \mu_\ell\boldsymbol{D}_\ell^H\boldsymbol{z}_{\ell-1} + \rho\mu_\ell\boldsymbol{R}_\ell^{-2}\boldsymbol{z}_\ell + \sqrt{\mu_\ell}\boldsymbol{R}_\ell^{-1}\boldsymbol{\gamma}_\ell \tag{6}$$

$$(\rho\boldsymbol{R}_\ell^{-2} + \boldsymbol{D}_\ell^H\boldsymbol{D}_\ell)\boldsymbol{x}_\ell = \boldsymbol{D}_\ell^H\boldsymbol{z}_{\ell-1} + \rho\boldsymbol{R}_\ell^{-2}\boldsymbol{z}_\ell + \frac{\boldsymbol{R}_\ell^{-1}\boldsymbol{\gamma}_\ell}{\sqrt{\mu_\ell}} \tag{7}$$

$$\boldsymbol{x}_\ell = (\rho\boldsymbol{R}_\ell^{-2} + \boldsymbol{D}_\ell^H\boldsymbol{D}_\ell)^{-1}(\boldsymbol{D}_\ell^H\boldsymbol{z}_{\ell-1} + \rho\boldsymbol{R}_\ell^{-2}\boldsymbol{z}_\ell + \frac{\boldsymbol{R}_\ell^{-1}\boldsymbol{\gamma}_\ell}{\sqrt{\mu_\ell}}) \tag{8}$$

$$\boldsymbol{x}_\ell = \boldsymbol{R}_\ell(\rho\mathbf{I} + (\boldsymbol{D}_\ell\boldsymbol{R}_\ell)^H\boldsymbol{D}_\ell\boldsymbol{R}_\ell)^{-1}\boldsymbol{R}_\ell(\boldsymbol{D}_\ell^H\boldsymbol{z}_{\ell-1} + \rho\boldsymbol{R}_\ell^{-2}\boldsymbol{z}_\ell + \frac{\boldsymbol{R}_\ell^{-1}\boldsymbol{\gamma}_\ell}{\sqrt{\mu_\ell}}) \tag{9}$$

$$\boldsymbol{x}_\ell = \boldsymbol{R}_\ell(\rho\mathbf{I} + (\boldsymbol{D}_\ell\boldsymbol{R}_\ell)^H\boldsymbol{D}_\ell\boldsymbol{R}_\ell)^{-1}((\boldsymbol{D}_\ell\boldsymbol{R}_\ell)^H\boldsymbol{z}_{\ell-1} + \rho\boldsymbol{R}_\ell^{-1}\boldsymbol{z}_\ell + \frac{\boldsymbol{\gamma}_\ell}{\sqrt{\mu_\ell}}) \tag{10}$$

$$\boldsymbol{R}_\ell^{-1}\boldsymbol{x}_\ell = (\rho\mathbf{I} + (\boldsymbol{D}_\ell\boldsymbol{R}_\ell)^H\boldsymbol{D}_\ell\boldsymbol{R}_\ell)^{-1}((\boldsymbol{D}_\ell\boldsymbol{R}_\ell)^H\boldsymbol{z}_{\ell-1} + \rho(\boldsymbol{R}_\ell^{-1}\boldsymbol{z}_\ell + \frac{\gamma_\ell}{\rho\sqrt{\mu_\ell}})) \quad (11)$$

Note that here there is a dependance on an unscaled $\boldsymbol{z}_{\ell-1}$. If $\boldsymbol{R}_{\ell-1}^{-1}\boldsymbol{z}_{\ell-1}$ is used instead, it will have to be rescaled before the unnormalized dictionary is applied.

2 Update Equation for $\boldsymbol{z}_\ell$

Briefly ignore the positive constraint, I will return to it.

For $\boldsymbol{x}$, $\boldsymbol{z}$, $\boldsymbol{\gamma}$, such that $\nabla_{\boldsymbol{z}_\ell} \mathrm{L}_\rho(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_L,\boldsymbol{z}_0,\ldots,\boldsymbol{z}_L,\gamma_0,\ldots,\gamma_L) = 0$:

$$\nabla_{\boldsymbol{z}} \frac{\mu_{\ell+1}}{2}\|\boldsymbol{z}_\ell - \boldsymbol{D}_{\ell+1}\boldsymbol{x}_{\ell+1}\|_2^2 + \|\boldsymbol{b}_\ell \cdot \boldsymbol{z}_\ell\|_1 + \frac{\rho}{2}\|\sqrt{\mu_\ell}\boldsymbol{R}_\ell^{-1}(\boldsymbol{z}_\ell - \boldsymbol{x}_\ell) + \frac{\gamma_\ell}{\rho}\|_2^2 = 0 \quad (12)$$

Something important to note here is that each element of $\boldsymbol{z}_\ell$ can be treated independently, that is:

$$\nabla_{\boldsymbol{z}_\ell[i]} \frac{\mu_{\ell+1}}{2}(\boldsymbol{z}_\ell[i] - (\boldsymbol{D}_{\ell+1}\boldsymbol{x}_{\ell+1})[i])^2 + \boldsymbol{b}_\ell[i]|\boldsymbol{z}_\ell[i]| + \frac{\rho}{2}(\sqrt{\mu_\ell}\boldsymbol{R}_\ell^{-1}[i](\boldsymbol{z}_\ell[i] - \boldsymbol{x}_\ell[i]) + \frac{\gamma_\ell[i]}{\rho})^2 = 0 \quad (13)$$

$$\nabla_{\boldsymbol{z}_\ell[i]} \frac{\mu_{\ell+1}}{2}(\boldsymbol{z}_\ell[i] - (\boldsymbol{D}_{\ell+1}\boldsymbol{x}_{\ell+1})[i])^2 + \boldsymbol{b}_\ell[i]|\boldsymbol{z}_\ell[i]| + \frac{\rho\mu_\ell}{2\boldsymbol{R}_\ell^2[i]}(\boldsymbol{z}_\ell[i] - \boldsymbol{x}_\ell[i] + \frac{\boldsymbol{R}_\ell[i]\gamma_\ell[i]}{\rho\sqrt{\mu_\ell}})^2 = 0 \quad (14)$$

For the sake of brevity and convenience, I will now drop the indexing:

$$\nabla_{\boldsymbol{z}_\ell} \frac{\mu_{\ell+1}}{2}(\boldsymbol{z}_\ell^2 - 2(\boldsymbol{D}_{\ell+1}\boldsymbol{x}_{\ell+1})z_\ell) + \boldsymbol{b}_\ell|\boldsymbol{z}_\ell| + \frac{\rho\mu_\ell}{2\boldsymbol{R}_\ell^2}(\boldsymbol{z}_\ell^2 - 2\boldsymbol{x}_\ell\boldsymbol{z}_\ell + \frac{2\boldsymbol{R}_\ell\gamma_\ell\boldsymbol{z}_\ell}{\rho\sqrt{\mu_\ell}}) = 0 \quad (15)$$

$$\nabla_{\boldsymbol{z}_\ell} \frac{1}{2}(\mu_{\ell+1} + \rho\mu_\ell\boldsymbol{R}_\ell^{-2})\boldsymbol{z}_\ell^2 - \mu_{\ell+1}\boldsymbol{D}_{\ell+1}\boldsymbol{x}_{\ell+1}\boldsymbol{z}_\ell - \rho\mu_\ell\boldsymbol{R}_\ell^{-2}\boldsymbol{x}_\ell\boldsymbol{z}_\ell + \sqrt{\mu_\ell}\boldsymbol{R}_\ell^{-1}\gamma_\ell\boldsymbol{z}_\ell + \boldsymbol{b}_\ell|\boldsymbol{z}_\ell| = 0 \quad (16)$$

$$\nabla_{\boldsymbol{z}_\ell} \frac{1}{2}\boldsymbol{z}_\ell^2 - \frac{\mu_{\ell+1}\boldsymbol{D}_{\ell+1}\boldsymbol{x}_{\ell+1} + \rho\mu_\ell\boldsymbol{R}_\ell^{-2}\boldsymbol{x}_\ell - \sqrt{\mu_\ell}\boldsymbol{R}_\ell^{-1}\gamma_\ell}{\mu_{\ell+1} + \rho\mu_\ell\boldsymbol{R}_\ell^{-2}}\boldsymbol{z}_\ell + \frac{\boldsymbol{b}_\ell}{\mu_{\ell+1} + \rho\mu_\ell\boldsymbol{R}_\ell^{-2}}|\boldsymbol{z}_\ell| = 0 \quad (17)$$

$$\boldsymbol{z}_\ell = \mathrm{S}_{\frac{\boldsymbol{b}_\ell}{\mu_{\ell+1} + \rho\mu_\ell\boldsymbol{R}_\ell^{-2}}}\left(\frac{\mu_{\ell+1}(\boldsymbol{D}_{\ell+1}\boldsymbol{x}_{\ell+1}) + \rho\mu_\ell\boldsymbol{R}_\ell^{-2}(\boldsymbol{x}_\ell - \frac{\boldsymbol{R}_\ell\gamma_\ell}{\rho\sqrt{\mu_\ell}})}{\mu_{\ell+1} + \rho\mu_\ell\boldsymbol{R}_\ell^{-2}}\right) \quad (18)$$

$$\boldsymbol{z}_\ell = \frac{1}{\mu_{\ell+1} + \rho\mu_\ell \boldsymbol{R}_\ell^{-2}} \mathrm{S}_{\boldsymbol{b}_\ell}(\mu_{\ell+1}(\boldsymbol{D}_{\ell+1}\boldsymbol{R}_{\ell+1}\boldsymbol{R}_{\ell+1}^{-1}\boldsymbol{x}_{\ell+1}) + \rho\mu_\ell \boldsymbol{R}_\ell^{-1}(\boldsymbol{R}_\ell^{-1}\boldsymbol{x}_\ell - \frac{\gamma_\ell}{\rho\sqrt{\mu_\ell}})) \tag{19}$$

To add nonnegative constraint, swap the shrinkage operator for a rectified linear unit.

$$\boldsymbol{z}_\ell = \frac{1}{\mu_{\ell+1} + \rho\mu_\ell \boldsymbol{R}_\ell^{-2}} \mathrm{RELU}(\mu_{\ell+1}(\boldsymbol{D}_{\ell+1}\boldsymbol{R}_{\ell+1}\boldsymbol{R}_{\ell+1}^{-1}\boldsymbol{x}_{\ell+1}) + \rho\mu_\ell \boldsymbol{R}_\ell^{-1}(\boldsymbol{R}_\ell^{-1}\boldsymbol{x}_\ell - \frac{\gamma_\ell}{\rho\sqrt{\mu_\ell}}) - \boldsymbol{b}_\ell) \tag{20}$$

2.1 Relaxation

$$\nabla_{\boldsymbol{z}_\ell[i]} \frac{\mu_{\ell+1}}{2}(\boldsymbol{z}_\ell[i] - (\boldsymbol{D}_{\ell+1}\boldsymbol{x}_{\ell+1})[i])^2 + \boldsymbol{b}_\ell[i]|\boldsymbol{z}_\ell[i]| + \frac{\rho}{2}(\boldsymbol{A}_\ell[i]\boldsymbol{x}_\ell[i] + \boldsymbol{B}_\ell[i]\boldsymbol{z}_\ell[i] + \frac{\gamma_\ell[i]}{\rho})^2 = 0 \tag{21}$$

Once again, removing indexing for brevity:

$$\nabla_{\boldsymbol{z}_\ell} \frac{\mu_{\ell+1}}{2}(\boldsymbol{z}_\ell - (\boldsymbol{D}_{\ell+1}\boldsymbol{x}_{\ell+1}))^2 + \boldsymbol{b}_\ell|\boldsymbol{z}_\ell| + \frac{\rho}{2}(\boldsymbol{A}_\ell\boldsymbol{x}_\ell + \boldsymbol{B}_\ell\boldsymbol{z}_\ell + \frac{\gamma_\ell}{\rho})^2 = 0 \tag{22}$$

$$\nabla_{\boldsymbol{z}_\ell} \frac{\mu_{\ell+1}}{2}(\boldsymbol{z}_\ell^2 - 2(\boldsymbol{D}_{\ell+1}\boldsymbol{x}_{\ell+1})\boldsymbol{z}_\ell) + \boldsymbol{b}_\ell|\boldsymbol{z}_\ell| + \frac{\rho}{2}(2\boldsymbol{B}_\ell\boldsymbol{A}_\ell\boldsymbol{x}_\ell\boldsymbol{z}_\ell + \boldsymbol{B}_\ell^2\boldsymbol{z}_\ell^2 + \frac{2\boldsymbol{B}_\ell\gamma_\ell}{\rho}) = 0 \tag{23}$$

$$\nabla_{\boldsymbol{z}_\ell} \frac{1}{2}(\boldsymbol{z} - \frac{\mu_{\ell+1}\boldsymbol{D}_{\ell+1}\boldsymbol{x}_{\ell+1} - \rho(\boldsymbol{B}_\ell\boldsymbol{A}_\ell\boldsymbol{x}_\ell + \frac{\boldsymbol{B}_\ell\gamma_\ell}{\rho})}{\mu_{\ell+1} + \rho\boldsymbol{B}_\ell^2})^2 + \frac{\boldsymbol{b}_\ell}{\mu_{\ell+1} + \rho\boldsymbol{B}_\ell^2}|\boldsymbol{z}_\ell| = 0 \tag{24}$$

$$\nabla_{\boldsymbol{z}_\ell} \frac{1}{2}(\boldsymbol{z} - \frac{\mu_{\ell+1}\boldsymbol{D}_{\ell+1}\boldsymbol{x}_{\ell+1} - \rho(\sqrt{\mu_\ell}\boldsymbol{R}_\ell^{-1}\boldsymbol{A}_\ell\boldsymbol{x}_\ell + \frac{\sqrt{\mu_\ell}\boldsymbol{R}_\ell^{-1}\gamma_\ell}{\rho})}{\mu_{\ell+1} + \rho\mu_\ell\boldsymbol{R}_\ell^{-2}})^2 + \frac{\boldsymbol{b}_\ell}{\mu_{\ell+1} + \rho\mu_\ell\boldsymbol{R}_\ell^{-2}}|\boldsymbol{z}_\ell| = 0 \tag{25}$$

$$\nabla_{\boldsymbol{z}_\ell} \frac{1}{2}(\boldsymbol{z} - \frac{\mu_{\ell+1}\boldsymbol{D}_{\ell+1}\boldsymbol{x}_{\ell+1} - \rho\mu_\ell\boldsymbol{R}_\ell^{-1}(\frac{\boldsymbol{A}_\ell\boldsymbol{x}_\ell}{\sqrt{\mu_\ell}} + \frac{\gamma_\ell}{\rho\sqrt{\mu_\ell}})}{\mu_{\ell+1} + \rho\mu_\ell\boldsymbol{R}_\ell^{-2}})^2 + \frac{\boldsymbol{b}_\ell}{\mu_{\ell+1} + \rho\mu_\ell\boldsymbol{R}_\ell^{-2}}|\boldsymbol{z}_\ell| = 0 \tag{26}$$

$$\boldsymbol{z}_\ell = \mathrm{S}_{\frac{\boldsymbol{b}_\ell}{\mu_{\ell+1} + \rho\mu_\ell\boldsymbol{R}_\ell^{-2}}}(\frac{\mu_{\ell+1}\boldsymbol{D}_{\ell+1}\boldsymbol{x}_{\ell+1} - \rho\mu_\ell\boldsymbol{R}_\ell^{-1}(\frac{\boldsymbol{A}_\ell\boldsymbol{x}_\ell}{\sqrt{\mu_\ell}} + \frac{\gamma_\ell}{\rho\sqrt{\mu_\ell}})}{\mu_{\ell+1} + \rho\mu_\ell\boldsymbol{R}_\ell^{-2}}) \tag{27}$$

$$\boldsymbol{z}_\ell = \frac{1}{\mu_{\ell+1} + \rho\mu_\ell\boldsymbol{R}_\ell^{-2}} \mathrm{S}_{\boldsymbol{b}_\ell}(\mu_{\ell+1}\boldsymbol{D}_{\ell+1}\boldsymbol{x}_{\ell+1} - \rho\mu_\ell\boldsymbol{R}_\ell^{-1}(\frac{\boldsymbol{A}_\ell\boldsymbol{x}_\ell}{\sqrt{\mu_\ell}} + \frac{\gamma_\ell}{\rho\sqrt{\mu_\ell}})) \tag{28}$$

For over-relaxation or under-relaxation, replace $\boldsymbol{A}_\ell \boldsymbol{x}_\ell^{(k+1)}$ with the expression $\alpha_k \boldsymbol{A}_\ell \boldsymbol{x}_\ell^{(k+1)} + (1-\alpha_k)(\boldsymbol{B}_\ell \boldsymbol{z}_\ell^{(k)} + \boldsymbol{C})$, where $\alpha_k \in (0,2)$ is the relaxation factor.

$$\boldsymbol{z}_\ell^{(k+1)} = \frac{1}{\mu_{\ell+1} + \rho\mu_\ell \boldsymbol{R}_\ell^{-2}} \operatorname{S}_{\boldsymbol{b}_\ell}(\mu_{\ell+1}\boldsymbol{D}_{\ell+1}\boldsymbol{x}_{\ell+1}^{(k+1)} - \rho\mu_\ell \boldsymbol{R}_\ell^{-1}((1-\alpha_k)\boldsymbol{R}_\ell^{-1}\boldsymbol{z}_\ell^{(k)} - \alpha_k \boldsymbol{R}_\ell^{-1}\boldsymbol{x}_\ell^{(k+1)} + \frac{\boldsymbol{\gamma}_\ell^{(k)}}{\rho\sqrt{\mu_\ell}})) \tag{29}$$

To add the nonnegativity constraint, use a rectified linear unit instead of a shrinkage operator:

$$\boldsymbol{z}_\ell = \frac{1}{\mu_{\ell+1} + \rho\mu_\ell \boldsymbol{R}_\ell^{-2}} \operatorname{RELU}(\mu_{\ell+1}\boldsymbol{D}_{\ell+1}\boldsymbol{x}_{\ell+1} - \rho\mu_\ell \boldsymbol{R}_\ell^{-1}(\frac{\boldsymbol{A}_\ell \boldsymbol{x}_\ell}{\sqrt{\mu_\ell}} + \frac{\boldsymbol{\gamma}_\ell}{\rho\sqrt{\mu_\ell}}) - \boldsymbol{b}_\ell) \tag{30}$$

$$\boldsymbol{z}_\ell^{(k+1)} = \frac{1}{\mu_{\ell+1} + \rho\mu_\ell \boldsymbol{R}_\ell^{-2}} \operatorname{RELU}(\mu_{\ell+1}\boldsymbol{D}_{\ell+1}\boldsymbol{x}_{\ell+1}^{(k+1)} - \rho\mu_\ell \boldsymbol{R}_\ell^{-1}((1-\alpha_k)\boldsymbol{R}_\ell^{-1}\boldsymbol{z}_\ell^{(k)} - \alpha_k \boldsymbol{R}_\ell^{-1}\boldsymbol{x}_\ell^{(k+1)} + \frac{\boldsymbol{\gamma}_\ell^{(k)}}{\rho\sqrt{\mu_\ell}}) - \boldsymbol{b}_\ell) \tag{31}$$

2.2 Update for $\boldsymbol{z}_L$

$$\nabla_{\boldsymbol{z}_L} \frac{\rho}{2}\|\boldsymbol{A}_L\boldsymbol{x}_L + \boldsymbol{B}_L\boldsymbol{z}_L + \frac{\boldsymbol{\gamma}_L}{\rho}\|_2^2 + \|\boldsymbol{b}_L \cdot \boldsymbol{z}_L\|_1 = 0 \tag{32}$$

$$\nabla_{\boldsymbol{z}_L} \frac{\rho}{2}\|\boldsymbol{A}_L\boldsymbol{x}_L + \sqrt{\mu_L}\boldsymbol{R}_L^{-1}\boldsymbol{z}_L + \frac{\boldsymbol{\gamma}_L}{\rho}\|_2^2 + \|\boldsymbol{b}_L \cdot \boldsymbol{z}_L\|_1 = 0 \tag{33}$$

$$\nabla_{\boldsymbol{z}_L[i]} \frac{\rho}{2}((\boldsymbol{A}_L\boldsymbol{x}_L)[i] + \sqrt{\mu_L}\boldsymbol{R}_L^{-1}[i]\boldsymbol{z}_L[i] + \frac{\boldsymbol{\gamma}_L[i]}{\rho})^2 + \boldsymbol{b}_L[i]|\boldsymbol{z}_L[i]| = 0 \tag{34}$$

$$\nabla_{\boldsymbol{z}_L[i]} \frac{\rho\mu_L\boldsymbol{R}_L^{-2}}{2}(\boldsymbol{z}_L[i] + \frac{(\boldsymbol{A}_L\boldsymbol{x}_L)[i] + \frac{\boldsymbol{\gamma}_L[i]}{\rho}}{\sqrt{\mu_L}\boldsymbol{R}_L^{-1}})^2 + \boldsymbol{b}_L[i]|\boldsymbol{z}_L[i]| = 0 \tag{35}$$

$$\nabla_{\boldsymbol{z}_L[i]} \frac{1}{2}(\boldsymbol{z}_L[i] + \frac{(\boldsymbol{A}_L\boldsymbol{x}_L)[i] + \frac{\boldsymbol{\gamma}_L[i]}{\rho}}{\sqrt{\mu_L}\boldsymbol{R}_L^{-1}})^2 + \frac{\boldsymbol{b}_L[i]|}{\rho\mu_L\boldsymbol{R}_L^{-2}}|\boldsymbol{z}_L[i]| = 0 \tag{36}$$

$$\boldsymbol{z}_L = \operatorname{S}_{\frac{\boldsymbol{R}_L^2\boldsymbol{b}_L}{\rho\mu_L}}(-\frac{\boldsymbol{R}_L\boldsymbol{A}_L\boldsymbol{x}_L}{\sqrt{\mu_L}} - \frac{\boldsymbol{R}_L\boldsymbol{\gamma}_L[i]}{\rho\sqrt{\mu_L}}) \tag{37}$$

$$\boldsymbol{z}_L = \boldsymbol{R}_\ell \operatorname{S}_{\frac{\boldsymbol{R}_L\boldsymbol{b}_L}{\rho\mu_L}}(-\frac{\boldsymbol{A}_L\boldsymbol{x}_L}{\sqrt{\mu_L}} - \frac{\boldsymbol{\gamma}_L}{\rho\sqrt{\mu_L}}) \tag{38}$$

So, using a relaxation parameter, I have:

$$\boldsymbol{z}_L^{(k+1)} = \boldsymbol{R}_\ell \operatorname{S}_{\frac{\boldsymbol{R}_L\boldsymbol{b}_L}{\rho\mu_L}}(\alpha_k\boldsymbol{R}_\ell^{-1}\boldsymbol{x}_\ell^{(k+1)} - (1-\alpha_k)\boldsymbol{R}_\ell^{-1}\boldsymbol{z}_\ell^{(k)} - \frac{\boldsymbol{\gamma}_L^{(k)}}{\rho\sqrt{\mu_L}}) \tag{39}$$

Or, if not using an overrelaxation parameter:

$$\boldsymbol{z}_L^{(k+1)} = \boldsymbol{R}_\ell \, \mathrm{S}_{\frac{\boldsymbol{R}_L \boldsymbol{b}_L}{\rho \mu_L}} (\boldsymbol{R}_L^{-1} \boldsymbol{x}_L^{(k+1)} - \frac{\boldsymbol{\gamma}_L^{(k)}}{\rho \sqrt{\mu_L}}) \tag{40}$$

To constrain $\boldsymbol{z}_L$ to be nonnegative replace the shrinkage operator with a rectified linear unit.

$$\boldsymbol{z}_L^{(k+1)} = \boldsymbol{R}_\ell \, \mathrm{RELU}(-\frac{\boldsymbol{A}_L \boldsymbol{x}_L^{(k+1)}}{\sqrt{\mu_L}} - \frac{\boldsymbol{\gamma}_L^{(k)}}{\rho \sqrt{\mu_L}} - \frac{\boldsymbol{R}_L \boldsymbol{b}_L}{\rho \mu_L}) \tag{41}$$

$$\boldsymbol{z}_L^{(k+1)} = \boldsymbol{R}_\ell \, \mathrm{RELU}(\alpha_k \boldsymbol{R}_\ell^{-1} \boldsymbol{x}_\ell^{(k+1)} - (1 - \alpha_k) \boldsymbol{R}_\ell^{-1} \boldsymbol{z}_\ell^{(k)} - \frac{\boldsymbol{\gamma}_L}{\rho \sqrt{\mu_L}} - \frac{\boldsymbol{R}_L \boldsymbol{b}_L}{\rho \mu_L}) \tag{42}$$

Or, if not using overrelaxation:

$$\boldsymbol{z}_L^{(k+1)} = \boldsymbol{R}_\ell \, \mathrm{RELU}(\boldsymbol{R}_L^{-1} \boldsymbol{x}_L^{(k+1)} - \frac{\boldsymbol{\gamma}_L^{(k)}}{\rho \sqrt{\mu_L}} - \frac{\boldsymbol{R}_L \boldsymbol{b}_L}{\rho \mu_L}) \tag{43}$$

2.3 Update for $\boldsymbol{z}_0$

$$\nabla_{\boldsymbol{z}_0} \frac{\mu_1}{2} \|\boldsymbol{z}_0 - \boldsymbol{D}_1 \boldsymbol{x}_1\|_2^2 + \frac{\rho}{2} \|\boldsymbol{A}_0 \boldsymbol{x}_0 + \boldsymbol{B} \boldsymbol{z}_0 + \boldsymbol{C}_0 + \frac{\boldsymbol{\gamma}_0}{\rho}\|_2^2 = 0 \tag{44}$$

where $\boldsymbol{A}_0 \boldsymbol{x}_0$ is zero unless overrelaxation is used.

$$\nabla_{\boldsymbol{z}_0} \frac{\mu_1}{2} \|\boldsymbol{z}_0 - \boldsymbol{D}_1 \boldsymbol{x}_1\|_2^2 + \frac{\rho}{2} \|\boldsymbol{A}_0 \boldsymbol{x}_0 + \boldsymbol{W} \boldsymbol{z}_0 - \boldsymbol{s} + \frac{\boldsymbol{\gamma}_0}{\rho}\|_2^2 = 0 \tag{45}$$

$$\mu_1 (\boldsymbol{z}_0 - \boldsymbol{D}_1 \boldsymbol{x}_1) + \rho (\boldsymbol{W}^T \boldsymbol{W} \boldsymbol{z}_0 + \boldsymbol{W}^T (\boldsymbol{A}_0 \boldsymbol{x}_0 - \boldsymbol{s} + \frac{\boldsymbol{\gamma}_0}{\rho})) = 0 \tag{46}$$

$$(\mu_1 \mathbf{I} + \rho \boldsymbol{W}^T \boldsymbol{W}) \boldsymbol{z}_0 = \mu_1 \boldsymbol{D}_1 \boldsymbol{x}_1 + \rho \boldsymbol{W}^T (\boldsymbol{s} - \boldsymbol{A}_0 \boldsymbol{x}_0 - \frac{\boldsymbol{\gamma}_0}{\rho}) \tag{47}$$

$$\boldsymbol{z}_0 = (\mu_1 \mathbf{I} + \rho \boldsymbol{W}^T \boldsymbol{W})^{-1} (\mu_1 \boldsymbol{D}_1 \boldsymbol{x}_1 + \rho \boldsymbol{W}^T (\boldsymbol{s} - \boldsymbol{A}_0 \boldsymbol{x}_0 - \frac{\boldsymbol{\gamma}_0}{\rho})) \tag{48}$$

$$\boldsymbol{z}_0^{(k+1)} = (\mu_1 \mathbf{I} + \rho \boldsymbol{W}^T \boldsymbol{W})^{-1} (\mu_1 \boldsymbol{D}_1 \boldsymbol{x}_1^{(k+1)} + \rho \boldsymbol{W}^T (\boldsymbol{s} - (1 - \alpha_k)(\boldsymbol{W} \boldsymbol{z}_0^{(k)} - \boldsymbol{s}) - \frac{\boldsymbol{\gamma}_0^{(k)}}{\rho})) \tag{49}$$

$$\boldsymbol{z}_0^{(k+1)} = (\mu_1 \mathbf{I} + \rho \boldsymbol{W}^T \boldsymbol{W})^{-1} (\mu_1 \boldsymbol{D}_1 \boldsymbol{x}_1^{(k+1)} + \rho \boldsymbol{W}^T ((2 - \alpha_k) \boldsymbol{s} - (1 - \alpha_k) \boldsymbol{W} \boldsymbol{z}_0^{(k)} - \frac{\boldsymbol{\gamma}_0^{(k)}}{\rho})) \tag{50}$$

$$\mu_1 \mathbf{I} + \rho \boldsymbol{W}^T \boldsymbol{W} = \mu_1 (\mathbf{I} - \boldsymbol{W}^T \boldsymbol{W} + \boldsymbol{W}^T \boldsymbol{W}) + \rho \boldsymbol{W}^T \boldsymbol{W} \tag{51}$$

$$\mu_1 \mathbf{I} + \rho \boldsymbol{W}^T \boldsymbol{W} = \mu_1 (\mathbf{I} - \boldsymbol{W}^T \boldsymbol{W}) + (\rho + \mu_1) \boldsymbol{W}^T \boldsymbol{W} \tag{52}$$

$$(\mu_1 \mathbf{I} + \rho \boldsymbol{W}^T \boldsymbol{W})^{-1} = \frac{1}{\mu_1} (\mathbf{I} - \boldsymbol{W}^T \boldsymbol{W}) + \frac{1}{\rho + \mu_1} \boldsymbol{W}^T \boldsymbol{W} \tag{53}$$

$$\boldsymbol{z}_0^{(k+1)} = (\frac{1}{\mu_1} (\mathbf{I} - \boldsymbol{W}^T \boldsymbol{W}) + \frac{1}{\rho + \mu_1} \boldsymbol{W}^T \boldsymbol{W}) (\mu_1 \boldsymbol{D}_1 \boldsymbol{x}_1^{(k+1)} + \rho \boldsymbol{W}^T ((2 - \alpha_k) \boldsymbol{s} - (1 - \alpha_k) \boldsymbol{W} \boldsymbol{z}_0^{(k)} - \frac{\boldsymbol{\gamma}_0^{(k)}}{\rho})) \tag{54}$$

$$\boldsymbol{z}_0^{(k+1)} = ((\mathbf{I} - \boldsymbol{W}^T \boldsymbol{W}) + \frac{\mu_1}{\rho + \mu_1} \boldsymbol{W}^T \boldsymbol{W}) \boldsymbol{D}_1 \boldsymbol{x}_1^{(k+1)} + \frac{\rho \boldsymbol{W}^T}{\rho + \mu_1} ((2 - \alpha_k) \boldsymbol{s} - (1 - \alpha_k) \boldsymbol{W} \boldsymbol{z}_0^{(k)} - \frac{\boldsymbol{\gamma}_0^{(k)}}{\rho} \tag{55}$$

3 Update for $\boldsymbol{\gamma}_\ell$

$$\boldsymbol{\gamma}_\ell^{(k+1)} = \boldsymbol{\gamma}_\ell^{(k)} + \rho (\boldsymbol{A}_\ell \boldsymbol{x}_\ell^{(k+1)} + \boldsymbol{B}_\ell \boldsymbol{z}_\ell^{(k+1)}) \tag{56}$$

$$\boldsymbol{\gamma}_\ell^{(k+1)} = \boldsymbol{\gamma}_\ell^{(k)} + \rho ((1 - \alpha_k) \sqrt{\mu_\ell} \boldsymbol{R}_\ell^{-1} \boldsymbol{z}_\ell^{(k)} - \alpha_k \sqrt{\mu_\ell} \boldsymbol{R}^{-1} \boldsymbol{x}_\ell^{(k+1)} + \sqrt{\mu_\ell} \boldsymbol{R}_\ell^{-1} \boldsymbol{z}_\ell^{(k+1)}) \tag{57}$$

$$\frac{\boldsymbol{\gamma}_\ell^{(k+1)}}{\rho \sqrt{\mu_\ell}} = \frac{\boldsymbol{\gamma}_\ell^{(k)}}{\rho \sqrt{\mu_\ell}} + (1 - \alpha_k) \boldsymbol{R}_\ell^{-1} \boldsymbol{z}_\ell^{(k)} - \alpha_k \boldsymbol{R}^{-1} \boldsymbol{x}_\ell^{(k+1)} + \boldsymbol{R}_\ell^{-1} \boldsymbol{z}_\ell^{(k+1)} \tag{58}$$

Or, if $\alpha_k = 1$:

$$\boldsymbol{\gamma}_\ell^{(k+1)} = \boldsymbol{\gamma}_\ell^{(k)} + \rho (\sqrt{\mu_\ell} \boldsymbol{R}_\ell^{-1} \boldsymbol{z}_\ell^{(k+1)} - \sqrt{\mu_\ell} \boldsymbol{R}_\ell^{-1} \boldsymbol{x}_\ell^{(k+1)}) \tag{59}$$

$$\frac{\boldsymbol{\gamma}_\ell^{(k+1)}}{\rho \sqrt{\mu_\ell}} = \frac{\boldsymbol{\gamma}_\ell^{(k)}}{\rho \sqrt{\mu_\ell}} + \boldsymbol{R}_\ell^{-1} \boldsymbol{z}_\ell^{(k+1)} - \boldsymbol{R}_\ell^{-1} \boldsymbol{x}_\ell^{(k+1)} \tag{60}$$

4 Update for γ_0

$$\boldsymbol{\gamma}_0^{(k+1)} = \boldsymbol{\gamma}_0^{(k)} + \rho (\boldsymbol{A}_0 \boldsymbol{x}_0^{(k+1)} + \boldsymbol{B}_0 \boldsymbol{z}_0^{(k+1)} + \boldsymbol{C}_0) \tag{61}$$

$$\boldsymbol{\gamma}_0^{(k+1)} = \boldsymbol{\gamma}_0^{(k)} + \rho ((1 - \alpha_k) (\boldsymbol{W} \boldsymbol{z}_0^{(k)} - \boldsymbol{s}) + \boldsymbol{W} \boldsymbol{z}_0^{(k+1)} - \boldsymbol{s}) \tag{62}$$

$$\boldsymbol{\gamma}_0^{(k+1)} = \boldsymbol{\gamma}_0^{(k)} + \rho ((1 - \alpha_k) \boldsymbol{W} \boldsymbol{z}_0^{(k)} + \boldsymbol{W} \boldsymbol{z}_0^{(k+1)} - (2 - \alpha_k) \boldsymbol{s}) \tag{63}$$

$$\frac{\boldsymbol{\gamma}_0^{(k+1)}}{\rho} = \frac{\boldsymbol{\gamma}_0^{(k)}}{\rho} + (1 - \alpha_k)\boldsymbol{W}\boldsymbol{z}_0^{(k)} + \boldsymbol{W}\boldsymbol{z}_0^{(k+1)} - (2 - \alpha_k)\boldsymbol{s} \tag{64}$$

Or, if $\alpha_k = 1$:

$$\boldsymbol{\gamma}_0^{(k+1)} = \boldsymbol{\gamma}_0^{(k)} + \rho(\boldsymbol{W}\boldsymbol{z}_0^{(k+1)} - \boldsymbol{s}) \tag{65}$$

$$\frac{\boldsymbol{\gamma}_0^{(k+1)}}{\rho} = \frac{\boldsymbol{\gamma}_0^{(k)}}{\rho} + \boldsymbol{W}\boldsymbol{z}_0^{(k+1)} - \boldsymbol{s} \tag{66}$$