Denote

$$G^t = I^t - \mathbf{W}_O^t I_O - \mathbf{W}_O^t A \circ \mathbf{W}_B^t \hat{I}_B - \mathbf{W}_O^t \hat{A} \circ \mathbf{W}_B^t I_B + \mathbf{W}_O^t \hat{A} \circ \mathbf{W}_B^t \hat{I}_B \tag{1}$$

Then derivatives of Eq.15 in the supplementary paper can be calculated as:

$$\begin{aligned}\frac{\partial E}{\partial I_O} =& w_1 \sum_t \left(-\mathbf{W}_O^{tT} G^t \right) \\ &+ \lambda_2 w_2 (D_x^T D_x + D_y^T D_y) I_O \\ &+ \lambda_3 \sum_x ||\nabla \hat{I}_B(x)||^2 (D_x^T D_x + D_y^T D_y) I_O \\ &+ \lambda_p(min(0, I_O) + max(0, I_O - 1)) = 0\end{aligned} \tag{2}$$

$$\begin{aligned}\frac{\partial E}{\partial I_B} =& w_1 \sum_t \left(-\mathbf{W}_O^t \hat{A} \circ \mathbf{W}_B^{tT} G^t \right) \\ &+ \lambda_2 w_3 (D_x^T D_x + D_y^T D_y) I_B \\ &+ \lambda_3 || \sum_x ||\nabla \hat{I}_O(x)||^2 (D_x^T D_x + D_y^T D_y) I_B \\ &+ \lambda_p(min(0, I_B) + max(0, I_B - 1)) = 0\end{aligned} \tag{3}$$

$$\begin{aligned}\frac{\partial E}{\partial A} =& w_1 \sum_t \left(-\mathbf{W}_B^t \hat{I}_B \circ \mathbf{W}_O^{tT} G^t \right) \\ &+ \lambda_1 (D_x^T D_x + D_y^T D_y) A \\ &+ \lambda_p(min(0, A) + max(0, A - 1)) = 0\end{aligned} \tag{4}$$

Arrange the three equations into the terms including I_O, I_B, A, and constants separately.

$$\begin{aligned}\frac{\partial E}{\partial I_O} =& \Big(\sum_t w_1 \mathbf{W}_O^{tT} \mathbf{W}_O^t + \lambda_2 w_2 (D_x^T D_x + D_y^T D_y) + \lambda_3 \sum_x ||\nabla \hat{I}_B(x)||^2 (D_x^T D_x + D_y^T D_y) + \lambda_p w_{po} \Big) I_O \\ &+ \sum_t w_1 \mathbf{W}_O^{tT} (\mathbf{W}_O^t \hat{A} \circ \mathbf{W}_B^t I_B) \\ &+ \sum_t w_1 \mathbf{W}_O^{tT} (\mathbf{W}_B^t \hat{I}_B \circ \mathbf{W}_O^t A) \\ &- \sum_t w_1 \mathbf{W}_O^{tT} (I^t + \mathbf{W}_O^t \hat{A} \circ \mathbf{W}_B^t \hat{I}_B) + \lambda_p w_{poc} = 0.\end{aligned} \tag{5}$$

w_{po} equals to 1 when I_O is bigger than 1 or less than 0, otherwise w_{po} equals to 0. w_{poc} equals to -1 when I_O is bigger than 1 otherwise w_{poc} equals to 0.

Similarly,

$$\begin{aligned}
\frac{\partial E}{\partial I_B} =& (w_1 \sum_t \mathbf{W}_O^t \hat{A} \circ \mathbf{W}_B^{tT} \mathbf{W}_O^t) I_O \\
&+ \Big(\sum_t w_1 \mathbf{W}_O^t \hat{A} \circ \mathbf{W}_O^t \hat{A} \circ \mathbf{W}_B^{tT} \mathbf{W}_B^t + \lambda_2 w_3 (D_x^T D_x + D_y^T D_y) \\
&+ \lambda_3 \sum_x ||\nabla \hat{I}_O(x)||^2 (D_x^T D_x + D_y^T D_y) + \lambda_p w_{pb} \Big) I_B \\
&+ (\sum_t w_1 \mathbf{W}_O^t \hat{A} \circ \mathbf{W}_B^t \hat{I}_B \circ \mathbf{W}_B^{tT} \mathbf{W}_O^t) A \\
&- \sum_t w_1 \mathbf{W}_O^t \hat{A} \circ \mathbf{W}_B^{tT} (I^t + \mathbf{W}_O^t \hat{A} \circ \mathbf{W}_B^t \hat{I}_B) + \lambda_p w_{pbc} = 0.
\end{aligned} \tag{6}$$

w_{pb} equals to 1 when I_B is bigger than 1 or less than 0, otherwise w_{pb} equals to 0. w_{pbc} equals to -1 when I_B is bigger than 1 otherwise w_{pbc} equals to 0.

$$\begin{aligned}
\frac{\partial E}{\partial A} =& (w_1 \sum_t \mathbf{W}_B^t \hat{I}_B \circ \mathbf{W}_O^{tT} \mathbf{W}_O^t) I_O \\
&+ \Big(w_1 \sum_t \mathbf{W}_B^t \hat{I}_B \circ \mathbf{W}_O^t \hat{A} \circ \mathbf{W}_O^{tT} \mathbf{W}_B^t \Big) I_B \\
&+ \Big(\sum_t w_1 \mathbf{W}_B^t \hat{I}_B \circ \mathbf{W}_B^t \hat{I}_B \circ \mathbf{W}_O^{tT} \mathbf{W}_O^t + \lambda_1 (D_x^T D_x + D_y^T D_y) + \lambda_p w_{pa} \Big) A \\
&- \sum_t w_1 \mathbf{W}_B^t \hat{I}_B \circ \mathbf{W}_O^{tT} (I^t + \mathbf{W}_O^t \hat{A} \circ \mathbf{W}_B^t \hat{I}_B) + \lambda_p w_{pac} = 0.
\end{aligned} \tag{7}$$

w_{pa} equals to 1 when A is bigger than 1 or less than 0, otherwise w_{pa} equals to 0. w_{pac} equals to -1 when A is bigger than 1 otherwise w_{pac} equals to 0.