Appendix A: Derivation for DCFNet

We derive the backward formulas for optimizing the CNN parameters θ for DCFNet. We treat the elements in $g_{\boldsymbol{\theta}}(\mathbf{z})$ as variables which are relevant with $\varphi^l(\mathbf{z})$ and $\varphi^l(\mathbf{x})$ and simplify $\partial g_{\boldsymbol{\theta}}(\mathbf{z})$ as ∂g. It is assumed that the channels are independent. By using the chain rule to return the errors to the two branches of the network individually, we can obtain:

$$\begin{aligned}\frac{\partial L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} &= \sum_{l=1}^{D} \frac{\partial L(\boldsymbol{\theta})}{\partial \varphi^l(\mathbf{z})} \frac{\partial \varphi^l(\mathbf{z})}{\partial \boldsymbol{\theta}} + \sum_{l=1}^{D} \frac{\partial L(\boldsymbol{\theta})}{\partial \varphi^l(\mathbf{x})} \frac{\partial \varphi^l(\mathbf{x})}{\partial \boldsymbol{\theta}} + 2\gamma\boldsymbol{\theta} \\ &= \sum_{l=1}^{D} \mathrm{F}^{-1}\left(\frac{\partial L(\boldsymbol{\theta})}{\partial \hat{\varphi}^l(\mathbf{z})^*}\right) \frac{\partial \varphi^l(\mathbf{z})}{\partial \boldsymbol{\theta}} + \sum_{l=1}^{D} \mathrm{F}^{-1}\left(\frac{\partial L(\boldsymbol{\theta})}{\partial \hat{\varphi}^l(\mathbf{x})^*}\right) \frac{\partial \varphi^l(\mathbf{x})}{\partial \boldsymbol{\theta}} + 2\gamma\boldsymbol{\theta}.\end{aligned} \tag{a}$$

According to (a), it is necessary to compute $\partial L / \partial(\varphi^l(\mathbf{z}))$ and $\partial L / \partial(\varphi^l(\mathbf{x}))$. We start with $\partial L / \partial g$. According to (4) and (5) in the main paper,

$$\frac{\partial L}{\partial g} = 2\left(\sum_{l=1}^{D} \mathbf{w}^l \star \varphi^l(\mathbf{x}) - \mathbf{y}\right). \tag{b}$$

The chain rule is a little complicated since the intermediate variables are complex-valued variables. The definitions of the discrete Fourier transform and inverse discrete Fourier transform are used to derive their gradients [29]. Let N be the number of components in vector g. It is apparent that

$$\hat{g}_f = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} g_n e^{-j\frac{2\pi}{N}nf} = \underset{n\to f}{F}(g_n), \tag{c}$$

$$g_n = \frac{1}{\sqrt{N}} \sum_{f=0}^{N-1} \hat{g}_f e^{-j\frac{2\pi}{N}nf} = \underset{f\to n}{F^{-1}}(\hat{g}_f), \tag{d}$$

with g_n as the discrete time signal, $\hat{g}_f$ as the transformed signal in the frequency domain, and n and f having the range $\{0,...,N\}$. Partial derivatives of $\hat{g}_f$ with respect to g_n and g_n^* are:

$$\begin{cases} \dfrac{\partial \hat{g}_f}{\partial g_n} = \dfrac{1}{\sqrt{N}} e^{-j\frac{2\pi}{N}nf}, \\ \dfrac{\partial \hat{g}_f}{\partial g_n^*} = 0, \end{cases} \tag{e}$$

Partial derivatives of g_n with respect to $\hat{g}_f$ and $\hat{g}_f^*$ are:

$$\begin{cases} \dfrac{\partial g_n}{\partial \hat{g}_f} = \dfrac{1}{\sqrt{N}} e^{j\frac{2\pi}{N}nf}, \\ \dfrac{\partial g_n}{\partial \hat{g}_f^*} = 0. \end{cases} \tag{f}$$

Then, we can obtain the following inferences:

$$\frac{\partial L}{\partial g_n^*} = \sum_{f=0}^{N-1} \left(\left(\frac{\partial L}{\partial \hat{g}_f^*} \right)^* 0 + \frac{1}{\sqrt{N}} \frac{\partial L}{\partial \hat{g}_f^*} \left(e^{-j\frac{2\pi}{N}nf} \right)^* \right) = \frac{1}{\sqrt{N}} \sum_{f=0}^{N-1} \frac{\partial L}{\partial \hat{g}_f^*} e^{j\frac{2\pi}{N}nf} = F^{-1} \left(\frac{\partial L}{\partial \hat{g}_f^*} \right), \tag{g}$$

$$\frac{\partial L}{\partial \hat{g}_f^*} = \sum_{n=0}^{N-1} \left(\left(\frac{\partial L}{\partial \hat{g}_n^*} \right)^* 0 + \frac{1}{\sqrt{N}} \frac{\partial L}{\partial \hat{g}_n^*} \left(e^{j\frac{2\pi}{N}nf} \right)^* \right) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \frac{\partial L}{\partial \hat{g}_n^*} e^{-j\frac{2\pi}{N}nf} = F \left(\frac{\partial L}{\partial \hat{g}_f^*} \right). \tag{h}$$

Since g is the correlation response which is a real-valued vector, it holds that $g = g^*$. The discrete Fourier transform and inverse discrete Fourier transform of the derivatives with respect to g have the following relations:

$$\begin{cases} \hat{g} = \mathrm{F}(g), \\ \dfrac{\partial L}{\partial \hat{g}^*} = \mathrm{F}\left(\dfrac{\partial L}{\partial g^*} \right) = \mathrm{F}\left(\dfrac{\partial L}{\partial g} \right), \\ \dfrac{\partial L}{\partial g} = \dfrac{\partial L}{\partial g^*} = \mathrm{F}^{-1}\left(\dfrac{\partial L}{\partial \hat{g}^*} \right). \end{cases} \tag{i}$$

Since the operations in the forward pass process only contain the element-based Hadamard product and division, we calculate the derivative in (i) per pixel (u, v) in the frequency domain:

$$\frac{\partial L}{\partial \hat{g}_{uv}^*} = \left(\mathrm{F}\left(\frac{\partial L}{\partial g} \right) \right)_{uv}. \tag{j}$$

For the back-propagation of the tracking branch, the partial differential $\partial L / \partial(\varphi^l(\mathbf{z}))$ for $\mathbf{z}$ is required to compute. According to (4) in the main paper,

$$\frac{\partial \hat{g}_{uv}^*(\mathbf{z})}{\partial (\hat{\varphi}_{uv}^l(\mathbf{z}))^*} = \hat{w}_{uv}^l. \tag{k}$$

Then,

$$\frac{\partial L}{\partial (\hat{\varphi}_{uv}^l(\mathbf{z}))^*} = \frac{\partial L}{\partial \hat{g}_{uv}^*} \frac{\partial \hat{g}_{uv}^*(\mathbf{z})}{\partial (\hat{\varphi}_{uv}^l(\mathbf{z}))^*} = \frac{\partial L}{\partial \hat{g}_{uv}^*} \hat{w}_{uv}^l, \tag{l}$$

where $\partial L / \partial \hat{g}_{uv}^*$ is computed using (b) and (j). It holds that

$$\frac{\partial L}{\partial \varphi^l(\mathbf{z})} = \mathcal{F}^{-1}\left(\frac{\partial L}{\partial (\hat{\varphi}^l(\mathbf{z}))^*}\right). \tag{m}$$

For the back-propagation of the branch of learning the correlation filter, the partial differential $\partial L / \partial(\varphi^l(\mathbf{x}))$ for $\mathbf{x}$ is required to compute. We compute $\partial L / \partial \hat{\varphi}^l_{uv}(\mathbf{x})$ and $\partial L / \partial(\hat{\varphi}^l_{uv}(\mathbf{x}))^*$ independently, and then combine them to compute $\partial L / \partial(\varphi^l(\mathbf{x}))$. It is apparent that

$$\begin{cases} \dfrac{\partial L}{\partial \hat{\varphi}^l_{uv}(\mathbf{x})} = \dfrac{\partial L}{\partial \hat{g}^*_{uv}} \dfrac{\partial \hat{g}^*_{uv}}{\partial \hat{\varphi}^l_{uv}(\mathbf{x})}, \\ \dfrac{\partial \hat{g}^*_{uv}}{\partial \hat{\varphi}^l_{uv}(\mathbf{x})} = \dfrac{\partial \hat{g}^*_{uv}}{\partial \hat{\mathbf{w}}^l_{uv}} \dfrac{\partial \hat{\mathbf{w}}^l_{uv}}{\partial \hat{\varphi}^l_{uv}(\mathbf{x})} = (\hat{\varphi}^l_{uv}(\mathbf{z}))^* \dfrac{\partial \hat{\mathbf{w}}^l_{uv}}{\partial \hat{\varphi}^l_{uv}(\mathbf{x})}, \end{cases} \tag{n}$$

where according to (4) in the main paper,

$$\frac{\partial \hat{g}^*_{uv}}{\partial \hat{\mathbf{w}}^l_{uv}} = (\hat{\varphi}^l_{uv}(\mathbf{z}))^*. \tag{o}$$

Let μ be the denominator of (3) in the main paper:

$$\mu = \sum_{k=1}^{D} \hat{\varphi}^k_{uv}(\mathbf{x})(\hat{\varphi}^k_{uv}(\mathbf{x}))^* + \lambda. \tag{p}$$

According to (3) in the main paper,

$$\frac{\partial \hat{\mathbf{w}}^l_{uv}}{\partial \hat{\varphi}^l_{uv}(\mathbf{x})} = \frac{\hat{\mathbf{y}}^*_{uv}\mu - \hat{\mathbf{y}}^*_{uv}\hat{\varphi}^l_{uv}(\mathbf{x})(\hat{\varphi}^l_{uv}(\mathbf{x}))^*}{\mu^2}. \tag{q}$$

Substitution of (q) into (n) yields

$$\frac{\partial \hat{g}^*_{uv}}{\partial \hat{\varphi}^l_{uv}(\mathbf{x})} = \frac{\hat{\mathbf{y}}^*_{uv}\mu(\hat{\varphi}^l_{uv}(\mathbf{z}))^* - \hat{\mathbf{y}}^*_{uv}\hat{\varphi}^l_{uv}(\mathbf{x})(\hat{\varphi}^l_{uv}(\mathbf{x}))^*(\hat{\varphi}^l_{uv}(\mathbf{z}))^*}{\mu^2}. \tag{r}$$

According to (3) in the main paper,

$$\mu = \frac{\hat{\varphi}^l_{uv}(\mathbf{x})\hat{\mathbf{y}}^*_{uv}}{\hat{\mathbf{w}}^l_{uv}}, \tag{s}$$

and $\mu\hat{\mathbf{w}}^l_{uv} = \hat{\varphi}^l_{uv}(\mathbf{x})\hat{\mathbf{y}}^*_{uv}$. Substitution of (s) into one μ in the denominator of (r) yields

$$\frac{\partial \hat{g}^*_{uv}}{\partial \hat{\varphi}^l_{uv}(\mathbf{x})} = \frac{\hat{\mathbf{y}}^*_{uv}\mu(\hat{\varphi}^l_{uv}(\mathbf{z}))^*\hat{\mathbf{w}}^l_{uv} - \hat{\mathbf{y}}^*_{uv}\hat{\varphi}^l_{uv}(\mathbf{x})(\hat{\varphi}^l_{uv}(\mathbf{x}))^*(\hat{\varphi}^l_{uv}(\mathbf{z}))^*\hat{\mathbf{w}}^l_{uv}}{\mu\hat{\varphi}^l_{uv}(\mathbf{x})\hat{\mathbf{y}}^*_{uv}}. \tag{t}$$

Replacing $\mu\hat{\mathbf{w}}^l_{uv}$ in the numerator of (t) with $\hat{\varphi}^l_{uv}(\mathbf{x})\hat{\mathbf{y}}^*_{uv}$ yields

$$\begin{aligned}\frac{\partial \hat{g}_{uv}^*}{\partial \hat{\varphi}_{uv}^l(\mathbf{x})} &= \frac{\hat{\mathbf{y}}_{uv}^*(\hat{\varphi}_{uv}^l(\mathbf{z}))^*\hat{\varphi}_{uv}^l(\mathbf{x})\hat{\mathbf{y}}_{uv}^* - \hat{\varphi}_{uv}^l(\mathbf{x})\hat{\mathbf{y}}_{uv}^*(\hat{\varphi}_{uv}^l(\mathbf{x}))^*(\hat{\varphi}_{uv}^l(\mathbf{z}))^*\hat{\mathbf{w}}_{uv}^l}{\mu\hat{\varphi}_{uv}^l(\mathbf{x})\hat{\mathbf{y}}_{uv}^*} \\ &= \frac{\hat{\mathbf{y}}_{uv}^*(\hat{\varphi}_{uv}^l(\mathbf{z}))^* - (\hat{\varphi}_{uv}^l(\mathbf{x}))^*(\hat{\varphi}_{uv}^l(\mathbf{z}))^*\hat{\mathbf{w}}_{uv}^l}{\mu},\end{aligned} \tag{u}$$

where $\hat{\varphi}_{uv}^l(\mathbf{x})\hat{\mathbf{y}}_{uv}^*$ in the denominator and the numerator is canceled. Substitution of (p) into (u) yields

$$\frac{\partial \hat{g}_{uv}^*}{\partial \hat{\varphi}_{uv}^l(\mathbf{x})} = \frac{\hat{\mathbf{y}}_{uv}^*(\hat{\varphi}_{uv}^l(\mathbf{z}))^* - (\hat{\varphi}_{uv}^l(\mathbf{x}))^*(\hat{\varphi}_{uv}^l(\mathbf{z}))^*\hat{\mathbf{w}}_{uv}^l}{\sum_{k=1}^{D}\hat{\varphi}_{uv}^k(\mathbf{x})(\hat{\varphi}_{uv}^k(\mathbf{x}))^* + \lambda}. \tag{v}$$

Substitution of (v) and (j) into the upper part of (n) yields the solution for the partial differential $\partial L / \partial \hat{\varphi}_{uv}^l(\mathbf{x})$. We derive $\partial L / \partial(\hat{\varphi}_{uv}^l(\mathbf{x}))^*$ below. It is apparent that

$$\begin{cases} \dfrac{\partial L}{\partial(\hat{\varphi}_{uv}^l(\mathbf{x}))^*} = \dfrac{\partial L}{\partial \hat{g}_{uv}^*}\dfrac{\partial \hat{g}_{uv}^*}{\partial(\hat{\varphi}_{uv}^l(\mathbf{x}))^*}, \\ \dfrac{\partial \hat{g}_{uv}^*}{\partial(\hat{\varphi}_{uv}^l(\mathbf{x}))^*} = \dfrac{\partial \hat{g}_{uv}^*}{\partial \hat{\mathbf{w}}_{uv}^l}\dfrac{\partial \hat{\mathbf{w}}_{uv}^l}{\partial(\hat{\varphi}_{uv}^l(\mathbf{x}))^*}. \end{cases} \tag{w}$$

According to (3) in the main paper,

$$\frac{\partial \hat{\mathbf{w}}_{uv}^l}{\partial(\hat{\varphi}_{uv}^l(\mathbf{x}))^*} = -\frac{\hat{\mathbf{y}}^*\hat{\varphi}_{uv}^l(\mathbf{x})\hat{\varphi}_{uv}^l(\mathbf{x})}{\mu^2}. \tag{x}$$

Substitution of (o) and (x) into the lower part of (w) yields

$$\frac{\partial \hat{g}_{uv}^*}{\partial(\hat{\varphi}_{uv}^l(\mathbf{x}))^*} = -\frac{\hat{\mathbf{y}}_{uv}^*\hat{\varphi}_{uv}^l(\mathbf{x})\hat{\varphi}_{uv}^l(\mathbf{x})(\hat{\varphi}_{uv}^l(\mathbf{z}))^*}{\mu^2}. \tag{y}$$

Substitution of (s) into one μ in the denominator of (y) yields

$$\frac{\partial \hat{g}_{uv}^*}{\partial(\hat{\varphi}_{uv}^l(\mathbf{x}))^*} = -\frac{\hat{\mathbf{y}}_{uv}^*\hat{\varphi}_{uv}^l(\mathbf{x})\hat{\varphi}_{uv}^l(\mathbf{x})(\hat{\varphi}_{uv}^l(\mathbf{z}))^*\hat{\mathbf{w}}_{uv}^l}{\mu\hat{\mathbf{y}}_{uv}^*\hat{\varphi}_{uv}^l(\mathbf{x})} = -\frac{\hat{\varphi}_{uv}^l(\mathbf{x})(\hat{\varphi}_{uv}^l(\mathbf{z}))^*\hat{\mathbf{w}}_{uv}^l}{\mu}, \tag{z}$$

where $\hat{\mathbf{y}}_{uv}^*\hat{\varphi}_{uv}^l(\mathbf{x})$ in the denominator and the numerator is canceled. Substitution of (p) into (z) yields

$$\frac{\partial \hat{g}_{uv}^*}{\partial(\hat{\varphi}_{uv}^l(\mathbf{x}))^*} = \frac{-\hat{\varphi}_{uv}^l(\mathbf{x})(\hat{\varphi}_{uv}^l(\mathbf{z}))^*\hat{\mathbf{w}}_{uv}^l}{\sum_{k=1}^{D}\hat{\varphi}_{uv}^k(\mathbf{x})(\hat{\varphi}_{uv}^k(\mathbf{x}))^* + \lambda}. \tag{A}$$

Substitution of (j) and (A) into the upper part of (w) yields the solution for the partial differential $\partial L / \partial(\hat{\varphi}_{uv}^l(\mathbf{x}))^*$. Combination of $\partial L / \partial \hat{\varphi}_{uv}^l(\mathbf{x})$ and $\partial L / \partial(\hat{\varphi}_{uv}^l(\mathbf{x}))^*$ yields

$$\frac{\partial L}{\partial \varphi^l(\mathbf{x})} = \mathrm{F}^{-1}\left(\frac{\partial L}{\partial(\hat{\varphi}^l(\mathbf{x}))^*} + \left(\frac{\partial L}{\partial \hat{\varphi}^l(\mathbf{x})}\right)^*\right). \tag{B}$$

Appendix B: Derivation for Convolutional-Deconvolutional DCFNet

The parameters $\boldsymbol{\theta}$ of the CNN in the convolutional-deconvolutional DCFNet are optimized by minimizing (32) in the main paper. The derivatives of L_{low} in (32) in the main paper are obtained:

$$\frac{\partial L_{low}}{\partial \hat{g}^*(\mathbf{z}_s)} = 2(\hat{g}(\mathbf{z}_s) - \hat{\mathbf{y}}_s). \tag{C}$$

The back-propagation of the tracking branch for search image **z** is represented by:

$$\frac{\partial L_{low}}{\partial \varphi^l(\mathbf{z}_s)} = \mathrm{F}^{-1}\left(\frac{\partial L_{low}}{\partial \hat{g}^*(\mathbf{z}_s)} \odot \hat{\mathbf{w}}^{l*}\right). \tag{D}$$

For the back-propagation of the learning branch for the target image **x**, according to (31) in the main paper, the gradient of back-propagation of the filter w is computed by:

$$\frac{\partial L_{low}}{\partial \hat{\mathbf{w}}^l} = \sum_{s=1}^{S} \frac{\partial L_{low}}{\partial \hat{g}^* \varphi^l(\mathbf{z}_s)} \odot (\hat{\varphi}^l(\mathbf{z}_s))^*. \tag{E}$$

Let μ be the denominator of $\hat{\mathbf{w}}$ in (30) in the main paper:

$$\mu = \sum_{s=1}^{S} (\hat{\varphi}^l(\mathbf{x}_s^0))^* \odot \hat{\varphi}^l(\mathbf{x}_s^0) + \lambda_1 + \lambda_2 \sum_{i=1}^{k} (\hat{\varphi}^l(\mathbf{x}^i))^* \odot \hat{\varphi}^l(\mathbf{x}^i). \tag{F}$$

We compute the gradients for the object image patch's features $\varphi(\mathbf{x}_s^0)$ and the context image patch's features $\varphi(\mathbf{x}^i)$ as follows: According to (30) in the main paper, the following partial differentials hold:

$$\frac{\partial \hat{\mathbf{w}}^l}{\partial (\hat{\varphi}^l(\mathbf{x}_s^0))^*} = \frac{\mu \hat{\mathbf{y}}_s - \hat{\varphi}^l(\mathbf{x}_s^0) \odot (\hat{\varphi}^l(\mathbf{x}_s^0))^* \odot \hat{\mathbf{y}}_s}{\mu^2} = \frac{\hat{\mathbf{y}}_s - \hat{\varphi}^l(\mathbf{x}_s^0) \odot \hat{\mathbf{w}}^l}{\mu}, \tag{G}$$

$$\frac{\partial \hat{\mathbf{w}}^l}{\partial \hat{\varphi}(\mathbf{x}_s^0)} = \frac{-\hat{\varphi}^l(\mathbf{x}_s^0)^* \odot \hat{\varphi}^l(\mathbf{x}_s^0)^* \odot \hat{\mathbf{y}}_s}{\mu^2} = \frac{-\hat{\varphi}^l(\mathbf{x}_s^0)^* \odot \hat{\mathbf{w}}^l}{\mu}, \tag{H}$$

$$\frac{\partial \hat{\mathbf{w}}^l}{\partial (\hat{\varphi}^l(\mathbf{x}^i))^*} = \frac{\sum_{s=1}^{S} -(\hat{\varphi}^l(\mathbf{x}^i))^* \odot (\hat{\varphi}^l(\mathbf{x}_s^0))^* \odot \hat{\mathbf{y}}_s}{\mu^2} = \frac{\sum_{s=1}^{S} -(\hat{\varphi}^l(\mathbf{x}_s^0))^* \odot \hat{\mathbf{w}}^l}{\mu}, \tag{I}$$

$$\frac{\partial \hat{\mathbf{w}}^l}{\partial \hat{\varphi}^l(\mathbf{x}^i)} = \frac{\sum_{s=1}^{S} -(\hat{\varphi}^l(\mathbf{x}^i))^* \odot (\hat{\varphi}^l(\mathbf{x}_s^0))^* \odot \hat{\mathbf{y}}_s}{\mu^2} = \frac{\sum_{s=1}^{S} -(\hat{\varphi}^l(\mathbf{x}_s^0))^* \odot \hat{\mathbf{w}}^l}{\mu}. \tag{J}$$

Let $\mathrm{Re}(\cdot)$ be the real part of a complex-valued matrix. The back-propagation of the learning branch is carried out by:

$$
\begin{aligned}
\frac{\partial L_{low}}{\partial \varphi^{l}(\mathbf{x}_s^0)} &= \mathcal{F}^{-1}\left(\frac{\partial L_{low}}{\partial \hat{\mathbf{w}}^{l}}\frac{\partial \hat{\mathbf{w}}}{\partial (\hat{\varphi}^{l}(\mathbf{x}_s^0))^{*}} + \left(\frac{\partial L_{low}}{\partial \hat{\mathbf{w}}^{l}}\frac{\partial \hat{\mathbf{w}}^{l}}{\partial \hat{\varphi}(\mathbf{x}_s^0)}\right)^{*}\right) \\
&= \mathcal{F}^{-1}\left(\frac{\partial L_{low}}{\partial \hat{\mathbf{w}}^{l}} \odot \frac{\hat{\mathbf{y}}_s^{*} - 2\mathrm{Re}((\hat{\varphi}^{l}(\mathbf{x}_s^0))^{*} \odot \hat{\mathbf{w}}^{l})\mathrm{e}}{\mu}\right),
\end{aligned}
\tag{K}
$$

$$
\begin{aligned}
\frac{\partial L_{low}}{\partial \hat{\varphi}^{l}(\mathbf{x}^{i})} &= \mathcal{F}^{-1}\left(\frac{\partial L_{low}}{\partial \hat{\mathbf{w}}^{l}}\frac{\partial \hat{\mathbf{w}}^{l}}{\partial (\hat{\varphi}^{l}(\mathbf{x}^{i}))^{*}} + \left(\frac{\partial L_{low}}{\partial \hat{\mathbf{w}}^{l}}\frac{\partial \hat{\mathbf{w}}^{l}}{\partial \hat{\varphi}(\mathbf{x}^{i})}\right)^{*}\right) \\
&= \mathcal{F}^{-1}\left(\frac{\partial L_{low}}{\partial \hat{\mathbf{w}}^{l}} \odot \frac{-2\mathrm{Re}((\hat{\varphi}^{l}(\mathbf{x}^{i}))^{*} \odot \hat{\mathbf{w}}^{l})\mathrm{e}}{\mu}\right).
\end{aligned}
\tag{L}
$$