

trainable step sizes

$$\Gamma^+[\mathbf{v}_k; \zeta_{\mathbf{v}}] \equiv \mathbf{v}_k \odot \exp\left(\frac{\varepsilon_{\mathbf{v}}^k}{2} s_{\mathbf{v}}^k(\zeta_{\mathbf{v}_k})\right) - \frac{\varepsilon_{\mathbf{v}}^k}{2} \left[\partial_x S(x_k) \odot \exp\left(\varepsilon_{\mathbf{v}}^k q_{\mathbf{v}}^k(\zeta_{\mathbf{v}_k})\right) \right] + t_{\mathbf{v}}^k(\zeta_{\mathbf{v}_k})$$

$$\Lambda^+[\bar{\mathbf{x}}_k; \zeta_{\bar{\mathbf{x}}_k}] \equiv \bar{\mathbf{x}}_k \odot \exp\left(\varepsilon_{\bar{\mathbf{x}}}^k s_{\bar{\mathbf{x}}}^k(\zeta_{\bar{\mathbf{x}}_k})\right) + \varepsilon_{\bar{\mathbf{x}}}^k \left[v'_k \odot \exp\left(\varepsilon_{\mathbf{x}}^k q_{\mathbf{x}}^k(\zeta_{\bar{\mathbf{x}}_k})\right) \right] + t_{\mathbf{x}}^k(\zeta_{\bar{\mathbf{x}}_k})$$

The diagram illustrates the structure of two functions, Γ^+ and Λ^+ , which are composed of trainable step sizes and various scaling and translation components.

Top Equation (Γ^+):

- Left Term:** $\mathbf{v}_k \odot \exp\left(\frac{\varepsilon_{\mathbf{v}}^k}{2} s_{\mathbf{v}}^k(\zeta_{\mathbf{v}_k})\right)$. This term is associated with **v scaling**.
- Middle Term:** $-\frac{\varepsilon_{\mathbf{v}}^k}{2} \left[\partial_x S(x_k) \odot \exp\left(\varepsilon_{\mathbf{v}}^k q_{\mathbf{v}}^k(\zeta_{\mathbf{v}_k})\right) \right]$. This term is associated with **force scaling**.
- Right Term:** $+ t_{\mathbf{v}}^k(\zeta_{\mathbf{v}_k})$. This term is associated with **translation**.

Bottom Equation (Λ^+):

- Left Term:** $\bar{\mathbf{x}}_k \odot \exp\left(\varepsilon_{\bar{\mathbf{x}}}^k s_{\bar{\mathbf{x}}}^k(\zeta_{\bar{\mathbf{x}}_k})\right)$. This term is associated with **x scaling**.
- Middle Term:** $+ \varepsilon_{\bar{\mathbf{x}}}^k \left[v'_k \odot \exp\left(\varepsilon_{\mathbf{x}}^k q_{\mathbf{x}}^k(\zeta_{\bar{\mathbf{x}}_k})\right) \right]$. This term is associated with **v scaling**.
- Right Term:** $+ t_{\mathbf{x}}^k(\zeta_{\bar{\mathbf{x}}_k})$. This term is associated with **translation**.

Arrows indicate the flow of trainable step sizes ($\varepsilon_{\mathbf{v}}^k$ and $\varepsilon_{\bar{\mathbf{x}}}^k$) from the top equation to the bottom equation.