Algorithm 1 Low-Precision Training for SGHMC.

given: L layers DNN $\{f_1 \dots, f_L\}$. Weight, gradient, activation, and error quantizers Q_W, Q_G, Q_A, Q_E. Variance-corrected quantization Q^{vc}, and quantization gap of weights Δ. Data batch sequence $\{(\theta_k, h_k)\}_{k=1}^K$, where the θ_k is the input, and h_k is the target. The loss function $\mathcal{L}(\mathbf{a}, \mathbf{h})$ measures the loss between the prediction $\mathbf{a}$ and target $\mathbf{h}$. And $\mathbf{x}_k^{fp}$ denotes the full-precision buffer of the weight. Let $\text{Var}_{\mathbf{v}}^{hmc} = u(1-e^{-2\gamma\eta})$ and $\text{Var}_{\mathbf{x}}^{hmc} = u\gamma^{-2}(2\gamma\eta + 4e^{-\gamma\eta} - e^{-2\gamma\eta} - 3)$ and $S_{\mathbf{v}} = 1$. {Initialize the scaling parameter}

for $k = 1 : K$ **do**

1. Forward Propagation:

$a_k^{(0)} = \theta_k$

$a_k^{(l)} = Q_A(f_l(a_k^{(l-1)}, \mathbf{x}_k^l)), \forall l \in [1, L]$

2. Backward Propagation:

$e^{(L)} = \nabla_{a_k^{(L)}} \mathcal{L}(a_k^{(L)}, h_k)$

$e^{(l-1)} = Q_E\left(\frac{\partial f_l(a_k^{(l)})}{\partial a_k^{(l-1)}} e_k^{(l)}\right), \forall l \in [1, L]$

$g_k^{(l)} = Q_G\left(\frac{\partial f_l}{\partial \theta_k^{(l)}} e_k^{(l)}\right), \forall l \in [1, L]$

3. SGHMC Update:

full-precision gradient accumulators:

$\mathbf{v}_{k+1}^{(l)} \leftarrow \mathbf{v}_k^{(l)} - u\gamma^{-1}(1-e^{-\gamma\eta})g_k^{(l)} + \xi_k^{\mathbf{v}}, \forall l \in [1, L],$

$\mathbf{x}_{k+1}^{(l),fp} \leftarrow \mathbf{x}_k^{(l),fp} + \gamma^{-1}(1-e^{-\gamma\eta})\mathbf{v}_k^{(l)} + u\gamma^{-2}(\gamma\eta + e^{-\gamma\eta} - 1)g_k^{(l)} + \xi_k^{\mathbf{x}}, \quad \mathbf{x}_{k+1}^{(l)} \leftarrow Q_W\left(\mathbf{x}_{k+1}^{(l),fp}\right), \forall l \in [1, L]$

low-precision gradient accumulators:

$\mathbf{v}_k^{(l)} = \mathbf{v}_k^{(l)} * S_{\mathbf{v}}^{(l)}, \forall l \in [1, L]$ {Restore the velocity before update}

$\mu(\mathbf{v}_{k+1}^{(l)}) \leftarrow \mathbf{v}_k^{(l)} e^{-\gamma\eta} - u\gamma^{-1}(1-e^{-\gamma\eta})g_k^{(l)}, \forall l \in [1, L]$

$S_{\mathbf{v}}^{(l)} = \frac{\left\|\mu(\mathbf{v}_{k+1}^{(l)})\right\|_\infty}{\bar{U}}, \forall l \in [1, L]$ {Update the Scaling}

$\mathbf{v}_{k+1}^{(l)} \leftarrow Q_W(\left(\mu(\mathbf{v}_{k+1}^{(l)}) + \xi_k^{\mathbf{v}}\right)/S_{\mathbf{v}}^{(l)}), \forall l \in [1, L]$

$\mathbf{x}_{k+1}^{(l)} \leftarrow Q_W\left(\mathbf{x}_k^{(l)} + \gamma^{-1}(1-e^{-\gamma\eta})\mathbf{v}_k^{(l)} + u\gamma^{-2}(\gamma\eta + e^{-\gamma\eta} - 1)g_k^{(l)} + \xi_k^{\mathbf{x}}\right), \forall l \in [1, L]$

Variance-corrected low-precision gradient accumulators:

$\mathbf{v}_k^{(l)} = \mathbf{v}_k^{(l)} * S_v^{(l)}, \forall l \in [1, L]$ {Restore the velocity before update}

$\mu(\mathbf{v}_{k+1}^{(l)}) = \mathbf{v}_k^{(l)} e^{-\gamma\eta} - u\gamma^{-1}(1-e^{-\gamma\eta})g_k^{(l)}, \forall l \in [1, L]$

$\mu(\mathbf{x}_{k+1}^{(l)}) = \mathbf{x}_k^{(l)} + \gamma^{-1}(1-e^{-\gamma\eta})\mathbf{v}_k^{(l)} + u\gamma^{-2}(\gamma\eta + e^{-\gamma\eta} - 1)g_k^{(l)}, \forall l \in [1, L]$

$S_{\mathbf{v}}^{(l)} = \frac{\left\|\mu(\mathbf{v}_{k+1}^{(l)})\right\|_\infty}{\bar{U}}, \forall l \in [1, L]$ {Update the Scaling}

$\mathbf{v}_{k+1}^{(l)} \leftarrow Q^{vc}\left(\mu(\mathbf{v}_{k+1}^{(l)})/S_{\mathbf{v}}^{(l)}, Var_{\mathbf{v}}^{hmc}/(S_{\mathbf{v}}^{(l)})^2, \Delta\right), \forall l \in [1, L]$

$\mathbf{x}_{k+1}^{(l)} \leftarrow Q^{vc}\left(\mu(\mathbf{x}_{k+1}^{(l)}), Var_{\mathbf{x}}^{hmc}, \Delta\right), \forall l \in [1, L]$

end for

output: samples $\{(\mathbf{v}_k^{(l)}, \mathbf{x}_k^{(l)})\}$