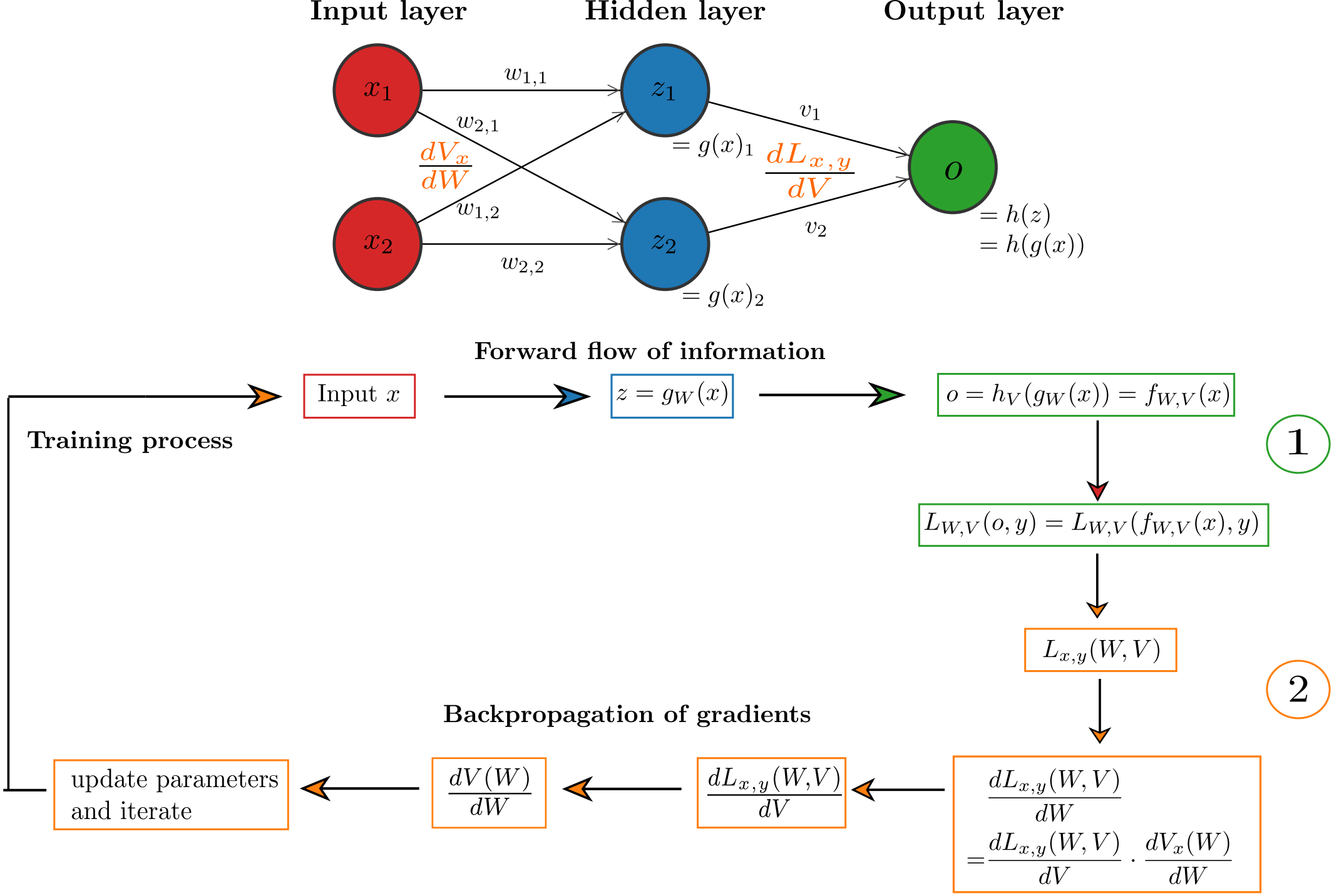

1 FNN as function of input x for fixed parameters W, V

$f_{W,V} : \mathbb{R}^2 \to \mathbb{R}$,

$x \mapsto f_{W,V}(x) = (g(x) \mapsto h(g(x)))$

$= x \mapsto \psi.(V(\phi.(Wx)))$

Loss as function of output o and label y

$L_{W,V} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$,

$(o, y) \mapsto L_{W,V}(o, y) = L_{W,V}(f_{W,V}(x), y) = \mid f_{W,V}(x) - y \mid^2$

2 FNN as function of parameters W, V for fixed input x

$f_x : \mathbb{R}^{2\times 2} \times \mathbb{R}^2 \to \mathbb{R}$,

$(W, V) \mapsto f_x(W, V) = \psi.(V(\phi.(Wx)))$

Loss as function of parameters W, V

$L_{x,y} : \mathbb{R}^{2\times 2} \times \mathbb{R}^2 \to \mathbb{R}$,

$(W, V) \mapsto L_{x,y}(W, V) = \mid f_x(W, V) - y \mid^2$

differentiable w.r.t W, V.