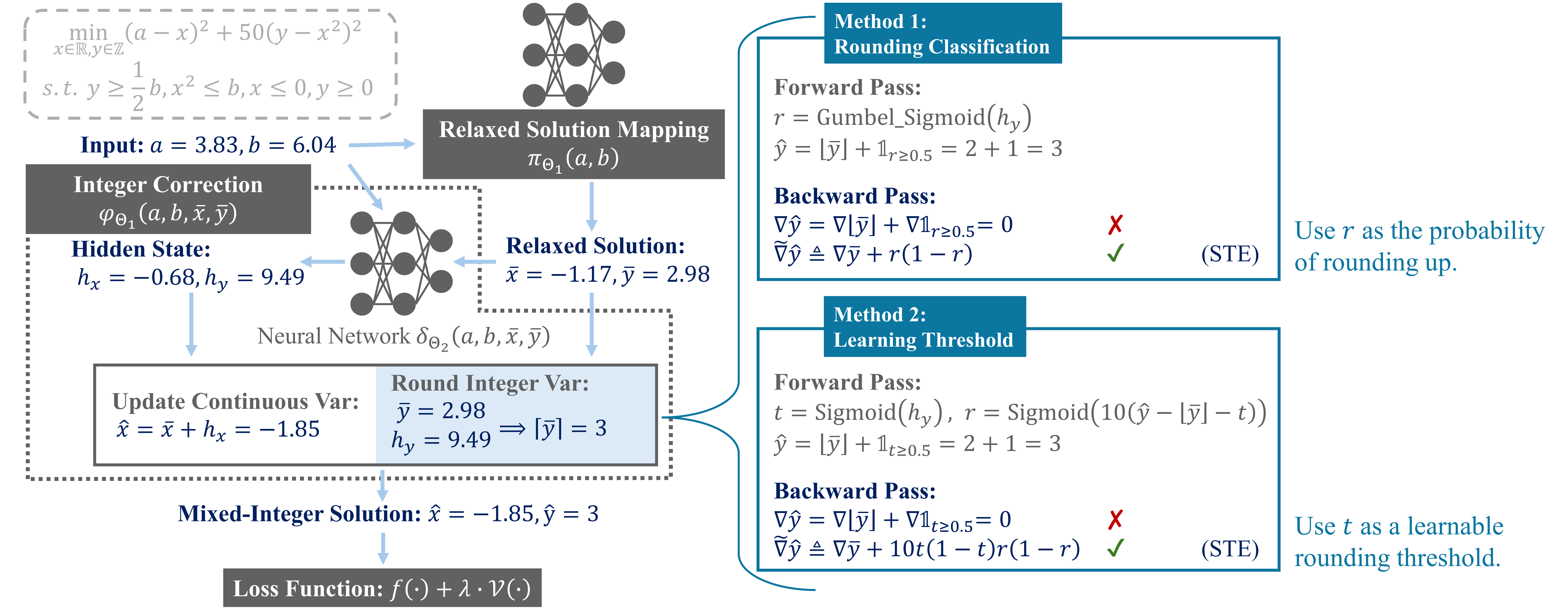

$\min_{x\in\mathbb{R},y\in\mathbb{Z}} (a-x)^2 + 50(y-x^2)^2$
$s.t.\ y \geq \frac{1}{2}b, x^2 \leq b, x \leq 0, y \geq 0$
Input: $a = 3.83, b = 6.04$
Relaxed Solution Mapping
$\pi_{\Theta_1}(a,b)$
Integer Correction
$\varphi_{\Theta_1}(a,b,\bar{x},\bar{y})$
Hidden State:
$h_x = -0.68, h_y = 9.49$
Relaxed Solution:
$\bar{x} = -1.17, \bar{y} = 2.98$
Neural Network $\delta_{\Theta_2}(a,b,\bar{x},\bar{y})$
Update Continuous Var:
$\hat{x} = \bar{x} + h_x = -1.85$
Round Integer Var:
$\bar{y} = 2.98$
$h_y = 9.49 \Rightarrow \lceil \bar{y} \rfloor = 3$
Mixed-Integer Solution: $\hat{x} = -1.85, \hat{y} = 3$
Loss Function: $f(\cdot) + \lambda \cdot \mathcal{V}(\cdot)$
Method 1:
Rounding Classification
Forward Pass:
$r = \text{Gumbel\_Sigmoid}(h_y)$
$\hat{y} = \lfloor \bar{y} \rfloor + \mathbb{1}_{r\geq 0.5} = 2 + 1 = 3$
Backward Pass:
$\nabla\hat{y} = \nabla\lfloor \bar{y} \rfloor + \nabla\mathbb{1}_{r\geq 0.5} = 0$ ✗
$\tilde{\nabla}\hat{y} \triangleq \nabla\bar{y} + r(1-r)$ ✓ (STE)
Use r as the probability of rounding up.
Method 2:
Learning Threshold
Forward Pass:
$t = \text{Sigmoid}(h_y),\ r = \text{Sigmoid}(10(\hat{y} - \lfloor \bar{y} \rfloor - t))$
$\hat{y} = \lfloor \bar{y} \rfloor + \mathbb{1}_{t\geq 0.5} = 2 + 1 = 3$
Backward Pass:
$\nabla\hat{y} = \nabla\lfloor \bar{y} \rfloor + \nabla\mathbb{1}_{t\geq 0.5} = 0$ ✗
$\tilde{\nabla}\hat{y} \triangleq \nabla\bar{y} + 10t(1-t)r(1-r)$ ✓ (STE)
Use t as a learnable rounding threshold.