Let $T : \mathbb{R}^n \times \Theta \rightarrow \mathbb{R}$, $(\boldsymbol{\epsilon}, \psi) \mapsto T(\boldsymbol{\epsilon}, \psi)$ be a differentiable bijection s.t. $T(\boldsymbol{\epsilon}, \psi) = \mathbf{z}$. Furthermore, let $p_\psi(\boldsymbol{\epsilon})$ define the pdf of $\boldsymbol{\epsilon}$ where its distribution depends weakly on ψ.

(I) Then the gradient of the reparameterized ELBO is:

$$\nabla_\psi ELBO(q_\psi(\mathbf{z})) = \mathbb{E}_{p_\psi(\boldsymbol{\epsilon})}[\nabla_{\mathbf{z}} f(\mathbf{z})|_{\mathbf{z}=T(\boldsymbol{\epsilon},\psi)} \nabla_\psi T(\boldsymbol{\epsilon}, \psi)] + \nabla_\psi \mathbb{H}(q_\psi(\mathbf{z}))$$

(II) Then the BBVI gradient is:

$$\nabla_\psi ELBO(q_\psi(\mathbf{z})) = \mathbb{E}_{q_\psi(\mathbf{z})}[f(\mathbf{z}) \nabla_\psi log(q_\psi(\mathbf{z}))] + \nabla_\psi \mathbb{H}(q_\psi(\mathbf{z}))$$