From the classical EM derivation at iteration t, we have, with $\boldsymbol{X}$ the observed random variables and $\boldsymbol{Z}$ the hidden random variables, for all sets of parameters $\boldsymbol{\theta}$ and for a fixed set of parameters $\boldsymbol{\theta}^{(t)}$:

$$\underbrace{\log p_{\boldsymbol{\theta}}(\boldsymbol{x})}_{l(\boldsymbol{\theta})} = \underbrace{\mathbb{E}_{p_{\boldsymbol{\theta}^{(t)}}(\boldsymbol{z}|\boldsymbol{x})}[\log p_{\boldsymbol{\theta}}(\boldsymbol{z},\boldsymbol{x})]}_{Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)})} \underbrace{-\mathbb{E}_{p_{\boldsymbol{\theta}^{(t)}}(\boldsymbol{z}|\boldsymbol{x})}[\log p_{\boldsymbol{\theta}}(\boldsymbol{z}|\boldsymbol{x})]}_{H(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)})},$$

where l is the model log-likelihood, Q is the classical EM quantity and H is the entropy. Hence, $\forall \boldsymbol{\theta}$:

$$\begin{aligned}
& l(\boldsymbol{\theta}) = Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)}) + H(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)}) \\
\implies \; & l(\boldsymbol{\theta}) \geq Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)}) + H(\boldsymbol{\theta}^{(t)}|\boldsymbol{\theta}^{(t)}) \\
& \qquad \text{because } \forall \boldsymbol{\theta}, H(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)}) \geq H(\boldsymbol{\theta}^{(t)}|\boldsymbol{\theta}^{(t)}) \\
\implies \; & l(\boldsymbol{\theta}) - Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)}) \geq H(\boldsymbol{\theta}^{(t)}|\boldsymbol{\theta}^{(t)}).
\end{aligned}$$

Then we have that the minimum of $l(\boldsymbol{\theta}) - Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)})$ is $H(\boldsymbol{\theta}^{(t)}|\boldsymbol{\theta}^{(t)})$ but we know that $l(\boldsymbol{\theta}^{(t)}) - Q(\boldsymbol{\theta}^{(t)}|\boldsymbol{\theta}^{(t)}) = H(\boldsymbol{\theta}^{(t)}|\boldsymbol{\theta}^{(t)})$. Thus $\boldsymbol{\theta} = \boldsymbol{\theta}^{(t)}$ is the place of the minimum, and by differentiating with respect to $\boldsymbol{\theta}$:

$$\begin{aligned}
& \nabla_{\boldsymbol{\theta}} \big\{ l(\boldsymbol{\theta}) - Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)}) \big\} |_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(t)}} = 0 \\
\implies \; & \nabla_{\boldsymbol{\theta}} \big\{ l(\boldsymbol{\theta}) \big\} |_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(t)}} = \nabla_{\boldsymbol{\theta}} \big\{ Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)}) \big\} |_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(t)}}.
\end{aligned}$$

Thus a gradient ascent over the log-likelihood is equivalent to a gradient ascent over Q.