Answer:

Simplest way to think about the setup is that we have $a(\theta) = l_{\mathsf{unsup}}(\theta)$, $b(\theta) = l_{\mathsf{sup}}(\theta)$ and $\tilde{a}(Q, \theta) = \mathrm{ELBO}(Q, \theta)$. We have from previous theory by Jensen's inequality that for any Q, we have $a(\theta) \leq \tilde{a}(Q, \theta)$, but that by argmaxing over Q we have $Q_i(z_i) = p(z_i | x_i; \theta)$ s.t. $a(\theta) = \tilde{a}(Q^*, \theta)$.

Thus at the E-step we set Q to max out this lower bound, $a(\theta^{(t)}) = \tilde{a}(Q^{(t)}, \theta)$, and then at the M-step we argmax over θ the sum of $\tilde{a}(Q, \theta) + b(\theta)$ to get $\theta^{(t+1)}$. Hence at the E-step we raise $\tilde{a}$ such that $a = \tilde{a}$ and so $a + b = \tilde{a} + b$. And then on the M-step we overall raise $\tilde{a} + b$. This means that after the M-step, our new $a(\theta^{(t+1)}) + b(\theta^{(t+1)}) \geq a(\theta^{(t)}) + b(\theta^{(t)})$.