



Md. Rafiul Amin<sup>1</sup>, Rose T. Faghih<sup>1</sup>,

**1** Department of Electrical and Computer Engineering, University of Houston, Houston, Texas, USA

**S1 Appendix. Expectation Maximization.** We can re-write the maximum likelihood estimation as the following marginal likelihood function of  $p(Y, X; \theta)$ :

$$\max_{\theta} \log p(Y; \theta) = \max_{\theta} \log \int_X p(Y, X; \theta) dX. \quad (1)$$

Obtaining the marginal likelihood by the integration operation in Equation 1 is difficult, especially with edge computation on wearable devices or smart-phones. This problem is usually modified as follows:

$$\begin{aligned} \max_{\theta} \log \int_X p(Y, X; \theta) dX &= \max_{\theta} \log \left( \int_X \frac{p(Y, X; \theta)}{q(X)} q(X) dX \right) \\ &\geq \max_{\theta} \int_X q(X) \log \left( \frac{p(Y, X; \theta)}{q(X)} \right) dX \quad [\text{Jensen's inequality}] \\ &= \underbrace{\max_{\theta} \int_X q(X) \log (p(Y, X; \theta)) dX}_{\text{function of } \theta} - \underbrace{\int_X q(X) \log (q(X)) dX}_{\text{constant}}, \end{aligned}$$

where  $q(X)$  is any probability density function. Therefore, the original problem is defined as the following expectation maximization (EM) approach,

$$\max_{\theta} \log p(Y; \theta) = \max_{\theta} \mathbb{E}_{X \sim q(X)} \{\log p(Y, X; \theta)\}. \quad (2)$$

As it is expressed in Equation 2, the unknowns can be estimated by iteratively maximizing the expectation of the joint log-likelihood  $\log p(Y, X; \theta)$ .