3.2 Objetivo variacional

Entrenamos conjuntamente el codificador $q_\phi\left(z \mid x\right)$ y el decodificador $p_\theta\left(x \mid z\right)$ maximizando la cota inferior de evidencia sobre la log-verosimilitud marginal de cada observación x :

$$\mathcal{L}(\theta, \phi; x) = \mathbb{E}_{q_\phi(z|x)}\Big[\log p_\theta(x \mid z)\Big] - D_{\mathrm{KL}}\Big(q_\phi(z \mid x) \,\|\, p(z)\Big). \tag{4}$$

El primer término recompensa una reconstrucción fiel; el segundo mantiene la posterior aproximada cerca de la prior $p\left(z\right) = \mathcal{N}\left(0, I\right)$. La tabla 2 reporta el error para cuatro tamaños latentes.

Tabla 2: Error cuadrático medio de reconstrucción en el conjunto de validación, en milésimas.

Tamaño latente	Línea base	Nuestro método	Cambio
16	9.84	8.71	−1.13
32	7.62	6.05	−1.57
64	6.48	4.93	−1.55
128	6.31	4.88	−1.43

Aumentar d más allá de 64 produce rendimientos decrecientes, lo cual es consistente con el análisis de capacidad de la sección 5. Por lo tanto, fijamos $d = 64$ en todos los experimentos restantes.