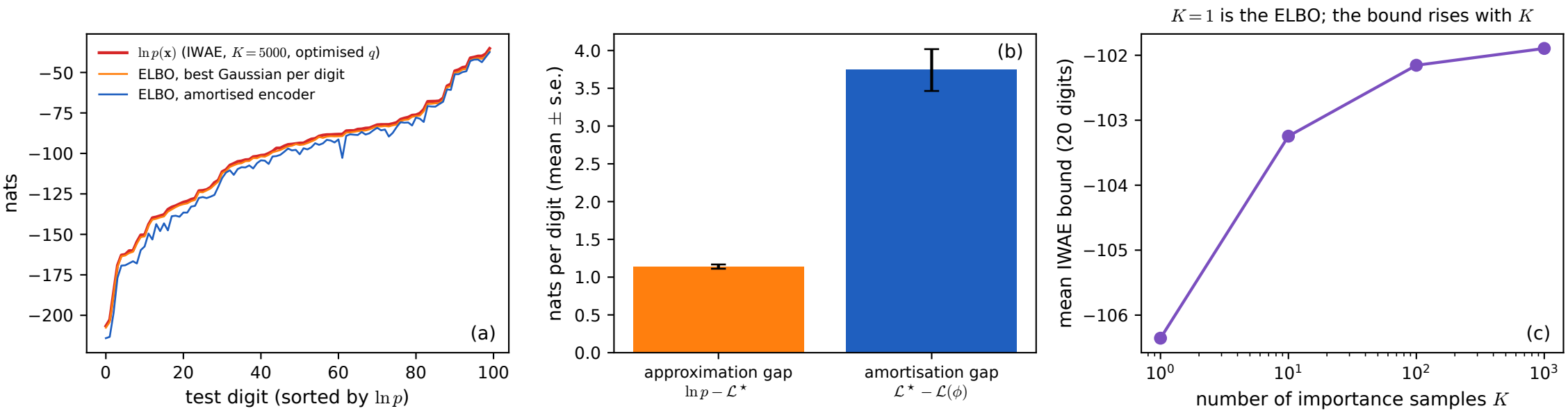

$\ln p(\mathbf{x})$ (IWAE, $K=5000$, optimised $q$)
ELBO, best Gaussian per digit
ELBO, amortised encoder
nats
test digit (sorted by $\ln p$)
(a)
nats per digit (mean $\pm$ s.e.)
approximation gap
$\ln p - \mathcal{L}^\star$
amortisation gap
$\mathcal{L}^\star - \mathcal{L}(\phi)$
(b)
$K=1$ is the ELBO; the bound rises with $K$
mean IWAE bound (20 digits)
number of importance samples $K$
(c)