Table 1: Theoretical results of the achieved 2-Wasserstein distance and the required gradient complexity for both log-concave (*italic*) and non-log-concave (**bold**) target distributions, where ϵ is any sufficiently small constant, Δ is the quantization error, and μ^* and λ^* denote the contraction rate of underdamped and overdamped Langevin dynamics respectively. Under non-log-concave target distributions, low-precision SGHMC achieves a better upper bound within shorter iterations compared with low-precision SGLD.

	Condition	Gradient Complexity	Achieved 2-Wasserstein
Full-precision gradient accumulators			
SGLD/SGHMC (Theorem 4)	*Strongly log-concave*	$\tilde{\mathcal{O}}\left(\log\left(\epsilon^{-1}\right)\epsilon^{-2}\right)$	$\tilde{\mathcal{O}}\left(\epsilon+\Delta\right)$
SGLD (Theorem 7)	**Non-log-concave**	$\tilde{\mathcal{O}}\left(\epsilon^{-4}\lambda^{*-1}\log^5\left(\epsilon^{-1}\right)\right)$	$\tilde{\mathcal{O}}\left(\epsilon+\log\left(\epsilon^{-1}\right)\sqrt{\Delta}\right)$
SGHMC (Theorem 1)	**Non-log-concave**	$\tilde{\mathcal{O}}\left(\epsilon^{-2}\mu^{*-2}\log^2\left(\epsilon^{-1}\right)\right)$	$\tilde{\mathcal{O}}\left(\epsilon+\sqrt{\log\left(\epsilon^{-1}\right)\Delta}\right)$
Low-precision gradient accumulators			
SGLD/SGHMC (Theorem 5)	*Strongly log-concave*	$\tilde{\mathcal{O}}\left(\log\left(\epsilon^{-1}\right)\epsilon^{-2}\right)$	$\tilde{\mathcal{O}}\left(\epsilon+\epsilon^{-1}\Delta\right)$
VC SGLD/VC SGHMC (Theorem 6)	*Strongly log-concave*	$\tilde{\mathcal{O}}\left(\log\left(\epsilon^{-1}\right)\epsilon^{-2}\right)$	$\tilde{\mathcal{O}}\left(\epsilon+\sqrt{\Delta}\right)$
SGLD (Theorem 8)	**Non-log-concave**	$\tilde{\mathcal{O}}\left(\epsilon^{-4}\lambda^{*-1}\log^5\left(\epsilon^{-1}\right)\right)$	$\tilde{\mathcal{O}}\left(\epsilon+\log^5\left(\epsilon^{-1}\right)\epsilon^{-4}\sqrt{\Delta}\right)$
VC SGLD (Theorem 9)	**Non-log-concave**	$\tilde{\mathcal{O}}\left(\epsilon^{-4}\lambda^{*-1}\log^3\left(\epsilon^{-1}\right)\right)$	$\tilde{\mathcal{O}}\left(\epsilon+\log^3\left(\epsilon^{-1}\right)\epsilon^{-2}\sqrt{\Delta}\right)$
SGHMC (Theorem 2)	**Non-log-concave**	$\tilde{\mathcal{O}}\left(\epsilon^{-2}\mu^{*-2}\log^2\left(\epsilon^{-1}\right)\right)$	$\tilde{\mathcal{O}}\left(\epsilon+\log^{3/2}\left(\epsilon^{-1}\right)\epsilon^{-2}\sqrt{\Delta}\right)$
VC SGHMC (Theorem 3)	**Non-log-concave**	$\tilde{\mathcal{O}}\left(\epsilon^{-2}\mu^{*-2}\log^2\left(\epsilon^{-1}\right)\right)$	$\tilde{\mathcal{O}}\left(\epsilon+\log\left(\epsilon^{-1}\right)\epsilon^{-1}\sqrt{\Delta}\right)$