

**Table A.1.** Common notations used throughout the manuscript.

Context	Symbol	Description	Example Value
Dataset	$n$	Number of data points (e.g., images)	50,000 for CIFAR-10
	$d$	Number of model outputs (e.g., classes, labels)	10 for CIFAR-10
	$m$	Size of NTK matrix is $m \times m$ where $m = nd$	500,000 for CIFAR-10
MEMDET	$n_b$	Number of row/column blocks (i.e., $n_b^2$ blocks in total)	e.g., 16
	$b$	Block size; typically $b \approx m/n_b$ , for $b \times b$ submatrices	e.g., 31,250
	$c$	Available memory capacity (in bytes)	
	$\beta$	Precision (in bytes) of a floating-point number	
FLODANCE	$n_s$	Number of sampled data points from $n$	e.g., 2000
	$m_s$	Size of sampled NTK matrix $m_s = n_s d$	e.g., 20,000
	$n_0$	Start of fitting interval $[n_0, n_s]$	e.g., 100
	$q$	Truncation order of Laurent series	e.g., 3
	$c_0, \nu_0, \dots, \nu_q$	Regression coefficients	
SLQ	$l$	Lanczos iterations (Krylov subspace size)	e.g., 100
	$s$	Number of Monte Carlo samples	e.g., 100
Variables	$\mathbf{K}_n$	NTK matrix with $n$ data points (matrix of size $m$ )	
	$\mathbf{K}_{n_s}$	Sampled NTK matrix with $n_s$ data points (matrix of size $m_s$ )	
	$\ell_n$	Log-determinant of NTK with $n$ data points (matrix of size $m$ )	
	$\hat{\ell}_n$	Estimated log-determinant	
	$L_n$	Normalized log-determinant $L_n = n^{-1} \ell_n$	
	$\theta$	Vectorization of neural network parameters	
	$p$	Dimension of $\theta$	
Functions	<b>pdet</b>	Pseudo-determinant (product of nonzero eigenvalues)	
	<b>logabsdet</b>	Natural logarithm of the absolute value of determinant	