Single-iteration cost of a conventional GD-based optimizer									
Specifications		**FLOPs**				**Ratios**			
Implementation	N	*net*	*loss-trace*	*GD-step*	*total*	$C_{net} = \frac{net}{K}$	$\frac{loss\text{-}trace}{KC_{net}}$	$\frac{GD\text{-}step}{loss\text{-}trace}$	$\frac{total}{KC_{net}}$
UltraPINNs	2	1.72e+08	1.72e+08	3.54e+08	5.25e+08	8.58e+06	1.00e+00	2.06e+00	3.06e+00
	4	3.43e+08	3.43e+08	6.91e+08	1.03e+09	8.58e+06	1.00e+00	2.01e+00	3.01e+00
	8	6.87e+08	6.87e+08	1.37e+09	2.05e+09	8.59e+06	1.00e+00	1.99e+00	2.99e+00
	16	1.38e+09	1.38e+09	2.72e+09	4.10e+09	8.61e+06	1.00e+00	1.97e+00	2.98e+00
	32	2.77e+09	2.77e+09	5.45e+09	8.22e+09	8.64e+06	1.00e+00	1.97e+00	2.97e+00
	64	5.57e+09	5.59e+09	1.10e+10	1.66e+10	8.70e+06	1.00e+00	1.96e+00	2.97e+00
	128	1.13e+10	1.14e+10	2.23e+10	3.37e+10	8.84e+06	1.01e+00	1.96e+00	2.98e+00
	256	2.33e+10	2.36e+10	4.59e+10	6.95e+10	9.11e+06	1.01e+00	1.94e+00	2.98e+00
	512	4.94e+10	5.06e+10	9.72e+10	1.48e+11	9.64e+06	1.02e+00	1.92e+00	2.99e+00
VPINNs (forward AD)	2	1.72e+08	5.09e+08	1.05e+09	1.56e+09	8.58e+06	2.97e+00	2.06e+00	9.08e+00
	4	3.43e+08	1.02e+09	2.06e+09	3.08e+09	8.58e+06	2.97e+00	2.02e+00	8.97e+00
	8	6.87e+08	2.04e+09	4.09e+09	6.13e+09	8.59e+06	2.97e+00	2.00e+00	8.92e+00
	16	1.38e+09	4.09e+09	8.15e+09	1.22e+10	8.61e+06	2.97e+00	1.99e+00	8.89e+00
	32	2.77e+09	8.22e+09	1.63e+10	2.46e+10	8.64e+06	2.97e+00	1.99e+00	8.88e+00
	64	5.57e+09	1.66e+10	3.29e+10	4.95e+10	8.70e+06	2.97e+00	1.99e+00	8.88e+00
	128	1.13e+10	3.37e+10	6.68e+10	1.00e+11	8.84e+06	2.98e+00	1.98e+00	8.88e+00
	256	2.33e+10	6.95e+10	1.38e+11	2.07e+11	9.11e+06	2.98e+00	1.98e+00	8.89e+00
	512	4.94e+10	1.48e+11	2.92e+11	4.39e+11	9.64e+06	2.99e+00	1.97e+00	8.90e+00
VPINNs (backward AD)	2	1.72e+08	3.41e+08	7.04e+08	1.04e+09	8.58e+06	1.99e+00	2.07e+00	6.09e+00
	4	3.43e+08	6.81e+08	1.38e+09	2.06e+09	8.58e+06	1.99e+00	2.02e+00	6.00e+00
	8	6.87e+08	1.36e+09	2.73e+09	4.09e+09	8.59e+06	1.99e+00	2.00e+00	5.96e+00
	16	1.38e+09	2.74e+09	5.44e+09	8.18e+09	8.61e+06	1.99e+00	1.99e+00	5.94e+00
	32	2.77e+09	5.49e+09	1.09e+10	1.64e+10	8.64e+06	1.99e+00	1.98e+00	5.93e+00
	64	5.57e+09	1.11e+10	2.19e+10	3.30e+10	8.70e+06	1.99e+00	1.98e+00	5.93e+00
	128	1.13e+10	2.25e+10	4.46e+10	6.71e+10	8.84e+06	1.99e+00	1.98e+00	5.93e+00
	256	2.33e+10	4.66e+10	9.18e+10	1.38e+11	9.11e+06	2.00e+00	1.97e+00	5.94e+00
	512	4.94e+10	9.92e+10	1.94e+11	2.94e+11	9.64e+06	2.01e+00	1.96e+00	5.95e+00

Table 1: FLOP measurements and ratios when performing a single iteration of a conventional GD-based optimizer. The columns of the table are organized as follows: *net* is the reference cost of evaluating $u^{\boldsymbol{\alpha},\boldsymbol{\omega}}$ over the batch of integration points; *loss-trace* is the cost of evaluating the loss function within the `tf.GradientTape` environment; *GD-step* is the cost of extracting the values of the gradient of the loss with respect to all trainable parameters ($\boldsymbol{\alpha}$ and $\boldsymbol{\omega}$) and applying them on a GD step; and *total* = *loss-trace* + *GD-step*.

Single-iteration cost of our hybrid LS/GD optimizer															
Specifications		**FLOPs**							**Ratios**						
Implementation	N	*net*	*vect-net*	*mat-build*	*LS-solve*	*loss-trace*	*GD-step*	*total*	$C_{net} = \frac{net}{K}$	$\frac{vect\text{-}net}{KC_{net}}$	$\frac{mat\text{-}build}{KMN}$	$\frac{LS\text{-}solve}{MN^2}$	$\frac{loss\text{-}trace}{KC_{net}}$	$\frac{GD\text{-}step}{loss\text{-}trace}$	$\frac{total}{KC_{net}}$
UltraPINNs	2	1.72e+08	1.72e+08	1.02e+04	7.67e+03	1.72e+08	3.54e+08	6.97e+08	8.58e+06	1.00e+00	2.54e+01	1.92e+02	1.00e+00	2.06e+00	4.06e+00
	4	3.43e+08	3.43e+08	4.03e+04	8.50e+03	3.43e+08	6.91e+08	1.38e+09	8.58e+06	1.00e+00	1.26e+01	2.66e+01	1.00e+00	2.01e+00	4.01e+00
	8	6.87e+08	6.87e+08	1.87e+05	1.41e+04	6.87e+08	1.37e+09	2.74e+09	8.59e+06	1.00e+00	7.30e+00	5.49e+00	1.00e+00	1.99e+00	3.99e+00
	16	1.38e+09	1.38e+09	9.61e+05	5.50e+04	1.38e+09	2.72e+09	5.48e+09	8.61e+06	1.00e+00	4.69e+00	2.69e+00	1.00e+00	1.97e+00	3.98e+00
	32	2.77e+09	2.77e+09	5.64e+06	3.70e+05	2.77e+09	5.45e+09	1.10e+10	8.64e+06	1.00e+00	3.44e+00	2.26e+00	1.00e+00	1.97e+00	3.97e+00
	64	5.57e+09	5.57e+09	3.54e+07	2.89e+06	5.59e+09	1.10e+10	2.22e+10	8.70e+06	1.00e+00	2.70e+00	2.21e+00	1.00e+00	1.96e+00	3.98e+00
	128	1.13e+10	1.13e+10	2.47e+08	2.27e+07	1.14e+10	2.23e+10	4.53e+10	8.84e+06	1.00e+00	2.36e+00	2.17e+00	1.01e+00	1.96e+00	4.00e+00
	256	2.33e+10	2.33e+10	1.83e+09	1.75e+08	2.36e+10	4.59e+10	9.50e+10	9.11e+06	1.00e+00	2.18e+00	2.09e+00	1.01e+00	1.94e+00	4.07e+00
	512	4.94e+10	4.94e+10	1.41e+10	1.40e+09	5.06e+10	9.72e+10	2.13e+11	9.64e+06	1.00e+00	2.09e+00	2.08e+00	1.02e+00	1.92e+00	4.32e+00
VPINNs (forward AD)	2	1.72e+08	5.09e+08	1.03e+04	7.70e+03	5.09e+08	1.05e+09	2.07e+09	8.58e+06	2.97e+00	2.57e+01	1.92e+02	2.97e+00	2.06e+00	1.21e+01
	4	3.43e+08	1.02e+09	4.08e+04	8.53e+03	1.02e+09	2.06e+09	4.10e+09	8.58e+06	2.97e+00	1.28e+01	2.67e+01	2.97e+00	2.02e+00	1.19e+01
	8	6.87e+08	2.04e+09	1.88e+05	1.41e+04	2.04e+09	4.09e+09	8.17e+09	8.59e+06	2.97e+00	7.34e+00	5.51e+00	2.97e+00	2.00e+00	1.19e+01
	16	1.38e+09	4.09e+09	9.63e+05	5.50e+04	4.09e+09	8.15e+09	1.63e+10	8.61e+06	2.97e+00	4.70e+00	2.69e+00	2.97e+00	1.99e+00	1.19e+01
	32	2.77e+09	8.21e+09	5.64e+06	3.70e+05	8.22e+09	1.63e+10	3.28e+10	8.64e+06	2.97e+00	3.44e+00	2.26e+00	2.97e+00	1.99e+00	1.18e+01
	64	5.57e+09	1.65e+10	3.54e+07	2.89e+06	1.66e+10	3.29e+10	6.60e+10	8.70e+06	2.97e+00	2.70e+00	2.21e+00	2.97e+00	1.99e+00	1.19e+01
	128	1.13e+10	3.36e+10	2.47e+08	2.27e+07	3.37e+10	6.68e+10	1.34e+11	8.84e+06	2.97e+00	2.36e+00	2.17e+00	2.98e+00	1.98e+00	1.19e+01
	256	2.33e+10	6.92e+10	1.83e+09	1.75e+08	6.95e+10	1.38e+11	2.79e+11	9.11e+06	2.97e+00	2.18e+00	2.09e+00	2.98e+00	1.98e+00	1.19e+01
	512	4.94e+10	1.47e+11	1.41e+10	1.40e+09	1.48e+11	2.92e+11	6.02e+11	9.64e+06	2.97e+00	2.09e+00	2.08e+00	2.99e+00	1.97e+00	1.22e+01
VPINNs (backward AD)	2	1.72e+08	5.09e+08	1.03e+04	5.94e+03	3.40e+08	7.04e+08	1.55e+09	8.58e+06	2.97e+00	2.57e+01	1.48e+02	1.99e+00	2.07e+00	9.06e+00
	4	3.43e+08	1.70e+09	4.08e+04	6.77e+03	6.81e+08	1.38e+09	3.76e+09	8.58e+06	4.94e+00	1.28e+01	2.11e+01	1.99e+00	2.02e+00	1.09e+01
	8	6.87e+08	6.10e+09	1.88e+05	1.23e+04	1.36e+09	2.73e+09	1.02e+10	8.59e+06	8.88e+00	7.34e+00	4.82e+00	1.99e+00	2.00e+00	1.48e+01
	16	1.38e+09	2.31e+10	9.63e+05	5.33e+04	2.74e+09	5.44e+09	3.13e+10	8.61e+06	1.67e+01	4.70e+00	2.60e+00	1.99e+00	1.99e+00	2.27e+01
	32	2.77e+09	8.99e+10	5.64e+06	3.68e+05	5.49e+09	1.09e+10	1.06e+11	8.64e+06	3.25e+01	3.44e+00	2.25e+00	1.99e+00	1.98e+00	3.84e+01
	64	5.57e+09	3.57e+11	3.54e+07	2.89e+06	1.11e+10	2.19e+10	3.90e+11	8.70e+06	6.40e+01	2.70e+00	2.21e+00	1.99e+00	1.98e+00	7.00e+01
	128	1.13e+10	1.44e+12	2.47e+08	2.27e+07	2.25e+10	4.46e+10	1.50e+12	8.84e+06	1.27e+02	2.36e+00	2.17e+00	1.99e+00	1.98e+00	1.33e+02
	256	2.33e+10	5.90e+12	1.83e+09	1.75e+08	4.66e+10	9.18e+10	6.04e+12	9.11e+06	2.53e+02	2.18e+00	2.09e+00	2.00e+00	1.97e+00	2.59e+02
	512	4.94e+10	2.49e+13	1.41e+10	1.40e+09	9.92e+10	1.94e+11	2.53e+13	9.64e+06	5.05e+02	2.09e+00	2.08e+00	2.01e+00	1.96e+00	5.12e+02

Table 2: FLOP measurements and ratios when performing a single iteration of our hybrid LS/GD optimizer. The columns of the table are organized as follows: *net* is the reference cost of evaluating $u^{\boldsymbol{\alpha},\boldsymbol{\omega}}$ over the batch of integration points; *vect-net* is the cost of evaluating the vector-valued neural network ($\mathbf{u}^{\boldsymbol{\alpha}}$ in UltraPINNs and $\nabla\mathbf{u}^{\boldsymbol{\alpha}}$ in VPINNs) over the batch of integration points; *mat-build* is the cost of constructing matrix $\mathbf{B}^{\boldsymbol{\alpha}}$ given the values of *vect-net*; *LS-solve* is the cost of computing $\boldsymbol{\omega}^{\boldsymbol{\alpha}}$ via the LS solver; *loss-trace* is the cost of evaluating the loss function within the `tf.GradientTape` environment; *GD-step* is the cost of extracting the gradient of the loss with respect to $\boldsymbol{\alpha}$ and applying a GD step; and *total* = *vect-net* + *mat-build* + *LS-solve* + *loss-trace* + *GD-step*.