
Article:
Rockett, P. (2026) Solving ordinary differential equations with genetic programming with hard initial/boundary value constraints. Genetic Programming and Evolvable Machines, 27 (1). 11. ISSN: 1389-2576

https://doi.org/10.1007/s10710-026-09536-x

Supplementary Material for "Solving Ordinary Differential Equations with Genetic Programming with Hard Initial/Boundary Value Constraints"

Peter Rockett
School of Electrical and Electronic Engineering
University of Sheffield
Mappin Street
Sheffield S1 3JD
UK

24th September 2025

Abstract

This document contains the supplementary material for the paper "Solving Ordinary Differential Equations with Genetic Programming with Hard Initial/Boundary Value Constraints" submitted to Genetic Programming and Evolvable Machines. The full results of statistical testing are included for reference.

1 Introduction

This document records the complete set of results from solving the benchmark set of ten ordinary differential equations (ODEs) (Seaton et al., 2010). and provides commentary on individual results.

2 Research Questions

1. **RQ1:** The fundamental method employed here (and in related GP and deep neural network research) is collocation. How does the number of collocation points affect quality of the evolved result? In particular, is there any evidence that the GP model overfits—namely, that evolved solution achieves very small residual errors at two adjacent collocation points, but that the fitted model exhibits spurious, oscillatory behavior between the collocation points? One would expect that a larger number of collocation points (= more data) would produce less overfitting than fewer collocation points.

2. **RQ2:** We have proposed a novel method for handling constraints that has been designed specifically to obviate the 'yes/no' outcome of the death penalty approach to enforcing constraints. Is there any evidence that the new constraint penalization approach produces superior results to the death penalty? Further, one contribution of this paper is to propose a new methodology for imposing the boundary/initial values (BIVs) of the DE as *hard* constraints in the parameter optimization, as opposed to incorporating BIVs as soft constraints in the loss function—the so-called *penalty method*. Is there any evidence that this novel hard constraints approach produces either superior solution quality in terms of better fit to the underlying solutions, and/or tighter matching of the BIVs compared to ether the death penalty or the soft-constraint penalty method?

3. **RQ3:** Is the restricted function set used in this work able to synthesize accurate solutions without the need to embed to transcendental functions?

3 Statistical significance

Results which suggest statistical significance are highlighted in pink in the following tables. Statistical significance is assessed on the basis of the posterior probability being outside the range $[0.0909, 0.909]$ *and* the confidence interval (CI) not enclosing zero. The former condition is equivalent to a posterior odds ratio of at least 10:1 meaning that the data are at least ten times more likely to be explained by one model compared to the alternative. The only exception to the above is when the CI is smaller than the region of probable equivalence (ROPE), here taken as $[-10^{-6}, +10^{-6}]$, in which case, we regard the effect size as being of no practical importance.

4 RQ1: Influence of the numbers of collocation points

The layout of the (rather elaborate) tables of results in this section follows the pattern described below. Taking Table S1 as an exemplar, p1 is the probability that model with 10 collocation points has a superior test error to the model with 20 points. The entry p2 is the probability that the 10-point model has a smaller initial value violation than the 20-point model. The corresponding figures for the 10-versus-50 point comparisons are p3 and p4, respectively; Likewise, the 10-versus-100 point comparisons are given by p5 and p6.

As mentioned above, the confidence intervals (CIs) are also needed to fully gauge statistical significance. Thus l1 is the 5% CI of the differences between the test errors of the 10-point and 20-point models. The median test error difference is m1, while the 95% CI is indicated by h1. The corresponding CIs for the 10-versus-20 point comparisons of initial value violations is given by the tuple (l2, m2, h2). In a similar vein, the test error CIs for the 10-versus-50 point comparison is given by (l3, m3, h3), and the corresponding CIs for the 10-versus-50 initial value violations given by (l4, m4 h4).

The second group of rows compares 20 collocation points against 50 and 100 points, and the last group of rows compares 50 versus 100 points. Obviously, the results of comparing 20 against 50 points are essentially identical to comparing 50 with 20 points—the probability values are complements and the CI values simply change sign. These redundant values are omitted.

Finally, for the 2^{nd}-order equations there are two initial/boundary conditions: the initial function value and the initial value of the derivative. The each comparison in the tables for ODE-7 to 10 thus includes additional column for this second initial derivative value.

Table S1: Explanatory table for the comparisons of numbers of collocation points. for a first-order ODE. See text for further details.

		No. of collocation points					
		20		**50**		**100**	
No. of collocation points	**10**	p1 l1 m1 h1	p2 l2 m2 h2	p3 l3 m3 h3	p4 l4 m4 h4	p5 l5 m5 h5	p6 l6 m6 h6
	20			p7 l7 m7 h7	p8 l8 m8 h8	p9 l9 m9 h9	p10 l10 m10 h10
	50					p11 l11 m11 h11	p12 l12 m12 h12

4.1 Penalized loss function

Results for ODE-4, ODE-7 and ODE-8 *without* the threshold constraint have been omitted because significant numbers of the 30 repeated runs failed to converge. Throughout, we interpret 'convergence' as meaning not producing the trivial solution of $y = 0$. Table S2 shows the numbers of converged runs for each number of collocation points for the problematic ODEs. While adding the threshold constraint mostly rectifies the trivial solution problem, it is by no means foolproof. See the main paper for a discussion.

Table S2: Numbers of converged runs from the 30 repetitions ODE-4, ODE-7 and ODE-8 with and without the additional threshold constraint for varying numbers of collocation points for the penalty method.

	# Collocation pts.			
ODE	**10**	**20**	**50**	**100**
ODE-4 without	10	19	25	23
ODE-4 with	30	30	30	30
ODE-7 without	0	0	0	0
ODE-7 with	0	1	1	6
ODE-8 without	0	0	4	9
ODE-8 with	30	30	30	30

The table of results for ODE-7 *with* threshold has been included (as Table S9) although it is clear from Table S2 that these data should be treated with caution since few of the runs have converged. This accounts for the large numbers of zero values in Table S9 as many pairs of identical, trivial solutions are being compared.

In terms of judging statistical significance, the very first entry in Table S4 is noteworthy. While the probability value of 0.0133 (< 0.0909) implies there is a significant difference between the test errors of the 10 and 20 point variants, the 5%-95% confidence interval of $[-3.74 \times 10^{-6}, 3.73 \times 10^{-5}]$ includes the no-difference value of zero meaning there is no evidence of an actual difference between the results. Similarly in Table S7, the comparison of the IV violations for the 10-versus-20 points has a probability value of 0.99462 but a CI of $[-7.11 \times 10^{-9}, -6.23 \times 10^{-10}]$. This CI thus falls entirely within the region-of-probable-equivalence (ROPE) of $[-10^{-6}, +10^{-6}]$ and although there superficially appears to be strong evidence of a difference based on the probability value, we can judge the *effect size* to be insignificant. Other tables contain similar examples.

For the remaining ODEs, we can observe that there is no compelling evidence of any difference between the numbers of collocation points, at least for the penalty method. There is a single, isolated instance of statistical significance in ODE-2 where 50 collocation points appear better than 10, but this result does not form part of any consistent pattern and so is most likely a false positive.

Table S3: Comparisons of numbers of collocation points for ODE-1 for penalty-method. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points					
		20		**50**		**100**	
No. of collocation points	**10**	0.12444	0.5283	0.06668	0.42774	0.14218	0.30804
		-1.16e-9	-6.01e-5	-5.07e-9	-8.78e-5	-2.80e-9	-3.76e-5
		2.69e-10	4.95e-6	3.10e-9	7.28e-6	2.02e-9	3.02e-6
		1.06e-7	4.12e-5	1.42e-7	1.26e-4	1.28e-7	7.95e-5
	20			0.23094	0.26956	0.27196	0.20512
				-1.53e-8	-1.86e-5	-5.27e-9	-7.26e-5
				3.35e-9	5.63e-6	-1.04e-19	4.43e-6
				1.33e-7	2.27e-4	1.06e-7	1.72e-4
	50					0.34066	0.28528
						-3.12e-9	-4.13e-5
						3.69e-14	8.87e-7
						1.88e-8	8.42e-5

Table S4: Comparisons of numbers of collocation points for ODE-2 for penalty-method. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points					
		20		**50**		**100**	
No. of collocation points	**10**	0.0133	0.28346	0.0305	0.13232	0.08768	0.52724
		-3.74e-6	-2.76e-3	9.31e-7	-1.44e-3	-1.04e-5	-4.97e-3
		9.54e-6	-2.36e-4	1.42e-5	2.96e-4	1.31e-5	-1.28e-3
		3.73e-5	1.81e-3	4.83e-5	3.30e-3	3.99e-5	1.92e-3
	20			0.4552	0.59932	0.9143	0.97596
				-8.66e-6	-2.59e-3	-2.02e-5	-4.16e-3
				8.67e-7	1.49e-4	-7.04e-6	-6.74e-4
				6.92e-6	1.05e-3	2.69e-6	3.75e-5
	50					0.80748	0.9782
						-1.94e-5	-3.21e-3
						-4.09e-6	-9.51e-4
						4.25e-6	3.80e-4

Table S5: Comparisons of numbers of collocation points for ODE-3 for penalty-method. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points					
		20		**50**		**100**	
No. of collocation points	**10**	0.0536	0.41946	0.15596	0.46402	0.208	0.67638
		-2.61e-17	-6.97e-14	-1.03e-16	-5.04e-14	-1.02e-16	-2.88e-13
		2.71e-17	0.00e+0	6.30e-17	0.00e+0	6.51e-18	0.00e+0
		1.43e-15	1.18e-13	8.20e-16	3.18e-14	1.48e-15	3.18e-14
	20			0.7393	0.44354	0.61614	0.72678
				-6.56e-16	-1.41e-13	-1.78e-16	-6.28e-13
				-2.53e-17	1.27e-15	7.67e-18	0.00e+0
				1.27e-16	9.79e-14	4.46e-17	2.49e-13
	50					0.46178	0.8699
						-1.29e-16	-6.08e-13
						-2.05e-18	0.00e+0
						1.27e-16	0.00e+0

Table S6: Comparisons of numbers of collocation points for ODE-4_threshold for penalty-method. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points					
		20		**50**		**100**	
No. of collocation points	**10**	0.02378	0.36156	0.00026	0.76242	0.00028	0.22706
		-8.00e-16	-8.71e-11	2.61e-15	-3.73e-10	3.19e-16	-9.27e-11
		6.20e-15	8.22e-12	5.36e-14	-5.00e-11	3.69e-14	5.58e-11
		1.11e-13	1.66e-10	1.14e-13	1.76e-10	1.16e-13	2.15e-10
	20			0.53584	0.46286	0.45256	0.41102
				-9.03e-15	-5.13e-11	-1.32e-14	-7.72e-11
				-9.90e-17	8.28e-12	1.12e-15	-1.69e-11
				3.82e-15	7.02e-11	7.02e-15	7.61e-11
	50					0.32806	0.22498
						-6.11e-15	-2.58e-11
						1.28e-15	1.97e-11
						7.78e-15	1.73e-10

Table S7: Comparisons of numbers of collocation points for ODE-5 for penalty-method. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points					
		20		**50**		**100**	
No. of collocation points	**10**	0.0	0.99462	2e-5	0.995	0.0	0.9947
		1.85e-12	-7.11e-9	2.77e-12	-6.59e-9	2.18e-12	-4.83e-9
		7.30e-12	-3.35e-9	7.41e-12	-1.29e-9	6.99e-12	-1.90e-9
		1.29e-11	-6.23e-10	1.09e-11	-2.42e-10	1.18e-11	-5.37e-10
	20			0.0991	0.4286	0.55724	0.58892
				-8.03e-13	-3.76e-9	-7.17e-13	-3.95e-9
				4.87e-13	-6.14e-10	-8.66e-14	-2.84e-10
				2.17e-12	4.30e-9	6.45e-13	3.36e-9
	50					0.94088	0.5117
						-1.43e-12	-1.73e-9
						-4.17e-13	2.38e-10
						1.99e-13	4.81e-9

Table S8: Comparisons of numbers of collocation points for ODE-6 for penalty-method. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points					
		20		**50**		**100**	
No. of collocation points	**10**	0.3789	0.79714	0.9622	0.89922	0.56544	0.78008
		-1.81e-30	-5.55e-17	-3.24e-25	-1.39e-16	-1.20e-29	-5.55e-17
		1.28e-32	0.00e+0	-2.95e-30	0.00e+0	-2.66e-33	0.00e+0
		4.45e-29	0.00e+0	1.95e-32	0.00e+0	2.12e-31	0.00e+0
	20			0.98924	0.87216	0.70672	0.78534
				-5.38e-25	-1.39e-16	-1.40e-28	-1.11e-16
				-2.93e-30	0.00e+0	-3.38e-31	0.00e+0
				-6.24e-33	0.00e+0	6.19e-31	2.78e-17
	50					0.09938	0.33648
						-4.28e-32	-5.55e-17
						3.84e-31	0.00e+0
						2.08e-26	1.11e-16

Table S9: Comparisons of numbers of collocation points for ODE-7_threshold for penalty-method. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points								
			20			**50**			**100**	
No. of collocation points	**10**	0.0	1.0	0.27736	0.76566	1.0	0.85878	0.0	0.86912	0.81312
		0.00e+0	0.00e+0	0.00e+0	0.00e+0	-4.44e-15	0.00e+0	0.00e+0	-1.14e-13	0.00e+0
		0.00e+0	0.00e+0	0.00e+0	0.00e+0	0.00e+0	0.00e+0	0.00e+0	0.00e+0	0.00e+0
		0.00e+0	1.78e-15	0.00e+0	0.00e+0	8.88e-16	0.00e+0	0.00e+0	8.88e-16	0.00e+0
	20				0.8796	0.05542	0.8257	0.0	0.01008	0.9685
					0.00e+0	-9.33e-16	0.00e+0	0.00e+0	-1.14e-13	0.00e+0
					0.00e+0	0.00e+0	0.00e+0	0.00e+0	0.00e+0	0.00e+0
					0.00e+0	0.00e+0	0.00e+0	0.00e+0	0.00e+0	0.00e+0
	50							0.0017	0.48326	0.624
								0.00e+0	-1.14e-13	0.00e+0
								0.00e+0	0.00e+0	0.00e+0
								0.00e+0	1.78e-15	0.00e+0

Table S10: Comparisons of numbers of collocation points for ODE-8_threshold for penalty-method. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points								
		20			**50**			**100**		
No. of collocation points	**10**	0.0 1.94e-8 3.13e-7 1.25e-6	0.64916 -2.71e-8 -3.96e-9 1.55e-15	0.94114 -3.77e-8 2.75e-15 8.83e-9	0.0 4.13e-8 3.07e-7 1.77e-6	0.89794 -3.69e-8 -9.25e-9 -2.45e-9	0.98046 -6.76e-8 -1.03e-8 9.16e-10	0.0 4.26e-8 3.14e-7 1.28e-6	0.79026 -9.14e-8 -2.87e-8 -2.00e-9	0.98012 -9.82e-8 -7.63e-10 3.33e-9
	20				0.03548 -7.82e-11 9.94e-10 3.30e-9	0.51038 -4.37e-8 -5.18e-9 1.53e-8	0.61732 -5.72e-8 -3.88e-9 6.27e-8	0.01596 1.49e-10 1.79e-9 5.02e-9	0.513 -4.21e-8 6.46e-11 3.74e-8	0.64872 -3.71e-8 1.66e-9 6.03e-8
	50							0.1375 -2.42e-11 1.34e-10 3.72e-10	0.6496 -6.68e-8 -1.85e-8 1.23e-8	0.95326 -5.19e-8 -1.17e-9 3.80e-8

Table S11: Comparisons of numbers of collocation points for ODE-9 for penalty-method. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points								
		20			**50**			**100**		
No. of collocation points	**10**	0.0	0.86742	0.40514	0.0	0.88566	0.37082	0.0	0.98238	0.86964
		1.53e-13	-5.00e-11	-2.51e-11	1.70e-13	-1.76e-11	-1.26e-10	1.78e-13	-6.06e-10	-1.14e-9
		1.39e-12	-2.81e-12	0.00e+0	1.40e-12	1.47e-12	0.00e+0	1.40e-12	-8.06e-11	-2.22e-12
		7.88e-12	2.58e-10	0.00e+0	8.64e-12	2.43e-10	0.00e+0	7.99e-12	1.72e-11	0.00e+0
	20				0.10818	0.46986	0.70974	0.00636	0.95226	0.92264
					-4.07e-15	-5.28e-11	-3.15e-11	9.36e-18	-5.14e-10	-1.10e-9
					1.86e-15	-5.45e-12	0.00e+0	9.54e-15	-1.62e-10	-1.85e-11
					6.15e-14	3.58e-11	7.93e-12	6.90e-14	-1.44e-11	0.00e+0
	50							0.23134	0.96198	0.81064
								-3.31e-15	-5.78e-10	-1.14e-9
								1.77e-15	-8.63e-11	-2.22e-12
								1.03e-14	2.24e-12	8.75e-15

Table S12: Comparisons of numbers of collocation points for ODE-10 for penalty-method. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points								
		20			**50**			**100**		
No. of collocation points	**10**	0.0	0.22054	0.5965	0.0	0.71508	0.81708	0.0	0.47212	0.52516
		7.40e-11	-4.39e-9	-2.08e-9	1.14e-10	-4.74e-9	-7.00e-9	1.57e-10	-5.98e-9	-3.98e-9
		3.30e-10	-5.27e-10	3.00e-9	3.73e-10	-1.24e-9	-1.76e-9	3.68e-10	6.60e-10	1.94e-9
		7.80e-10	4.28e-9	9.24e-9	7.77e-10	3.69e-9	8.49e-9	7.82e-10	4.77e-9	5.05e-9
	20				0.00628	0.93816	0.88906	8e-5	0.72392	0.38844
					-1.29e-12	-4.97e-9	-9.93e-9	7.37e-13	-2.91e-9	-9.73e-9
					8.31e-12	-1.13e-9	-9.64e-10	1.67e-11	-3.33e-10	-3.96e-10
					4.36e-11	-9.03e-11	1.00e-9	7.05e-11	5.29e-9	1.86e-9
	50							0.0019	0.50246	0.34598
								3.22e-14	-5.36e-9	-1.24e-8
								2.36e-12	5.97e-10	-5.03e-10
								1.04e-11	6.76e-9	6.35e-9

4.2 Death penalty

The numbers of successful runs for the problematic set of ODEs with the death penalty technique are summarized in Table S13. Results for ODE-7 and ODE-8 with hard constraints imposed via the death penalty but *without* the additional threshold constraint have been omitted because all 30 runs converged to trivial solutions. The majority of solutions for ODE-4 without threshold are trivial, hence these results are omitted. All of the solutions for ODE-4 with threshold, on the other hand, are non-trivial. See the main paper for a discussion.

Table S13: Numbers of non-trivial solutions from the 30 repetitions ODE-4, ODE-7 and ODE-8 with and without the additional threshold constraint for varying numbers of collocation points for the death penalty.

	# Collocation pts.			
ODE	**10**	**20**	**50**	**100**
ODE-4 without	0	6	5	8
ODE-4 with	30	30	30	30
ODE-7 without	0	0	0	0
ODE-7 with	30	30	30	30
ODE-8 without	0	0	0	0
ODE-8 with	24	26	26	26

As is the case for the penalty method in Section 4.1, there is little evidence favoring one number of collocation points over another for the death penalty. The only statistically significant results are for:

- ODE-4 with threshold although again, as before, this is an isolated result and so possibly a false positive.
- ODE-7 where 10 points appears too few although there is no difference between larger numbers of collocation points.
- ODE-8 where small numbers of collocation points (10 or 20) appear to be superior to larger numbers although the pattern is inconsistent so the results are probably distorted by the numbers of trivial solutions—see Table S13.
- ODE-10 where, as with ODE-7, 10 points appears too few although there are no differences for larger numbers of collocation points.

Table S14: Comparisons of numbers of collocation points for ODE-1 death-penalty for hard-constraints. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points					
		20		**50**		**100**	
No. of collocation points	**10**	0.25034 -1.44e-29 8.99e-31 2.51e-18	0.39262 -7.11e-15 0.00e+0 9.59e-14	0.61866 -8.76e-23 -1.34e-31 6.99e-26	0.80838 -2.12e-12 -1.78e-15 1.78e-15	0.46278 -1.23e-22 4.23e-31 1.40e-25	0.71442 -2.52e-12 -1.78e-15 7.11e-15
	20			0.66038 -3.02e-21 -3.94e-31 2.36e-18	0.79008 -1.25e-12 -1.78e-15 1.78e-15	0.14032 -3.46e-31 4.74e-31 3.49e-17	0.68802 -7.11e-15 0.00e+0 5.33e-15
	50					0.71278 -9.42e-18 3.74e-31 1.34e-28	0.78746 -4.81e-11 0.00e+0 3.55e-15

Table S15: Comparisons of numbers of collocation points for ODE-2 death-penalty for hard-constraints. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points					
		20		**50**		**100**	
No. of collocation points	**10**	0.591 -3.87e-9 -3.63e-10 2.15e-9	0.2256 -1.73e-10 7.67e-11 2.33e-9	0.73044 -4.05e-9 -3.28e-10 1.31e-9	0.73924 -2.72e-9 -4.17e-11 2.70e-10	0.87366 -2.59e-8 -2.45e-10 4.81e-10	0.93872 -2.08e-8 -1.36e-9 6.01e-11
	20			0.2899 -1.21e-9 4.26e-10 5.74e-9	0.67274 -9.60e-10 1.50e-12 1.25e-10	0.68862 -1.65e-8 4.20e-10 4.31e-9	0.99514 -1.66e-8 -1.32e-9 -1.07e-10
	50					0.88954 -2.84e-8 -3.02e-10 2.42e-9	0.9447 -2.28e-8 -3.41e-9 1.18e-11

Table S16: Comparisons of numbers of collocation points for ODE-3 death-penalty for hard-constraints. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points					
		20		**50**		**100**	
No. of collocation points	**10**	8e-5	0.53808	0.0	0.50016	0.0	0.84716
		2.72e-18	-1.90e-25	6.93e-18	-7.36e-27	1.11e-17	-1.63e-17
		5.04e-17	0.00e+0	1.19e-16	0.00e+0	1.28e-16	0.00e+0
		6.72e-16	2.14e-29	1.08e-15	2.19e-23	1.08e-15	0.00e+0
	20			0.05364	0.7348	0.1544	0.82928
				1.93e-20	-9.32e-25	-2.99e-21	-1.63e-17
				1.43e-18	0.00e+0	1.16e-18	0.00e+0
				1.07e-17	8.76e-27	4.88e-18	4.67e-27
	50					0.6888	0.53486
						-2.95e-18	-2.58e-18
						-8.12e-20	0.00e+0
						8.05e-19	1.91e-26

Table S17: Comparisons of numbers of collocation points for ODE-4 death-penalty + threshold for hard-constraints. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points					
		20		**50**		**100**	
No. of collocation points	**10**	0.90824	0.94978	0.28914	0.90526	0.03252	0.26632
		-9.10e-2	-2.94e-11	8.36e-3	-3.70e-11	1.19e-2	-7.77e-16
		6.94e-3	-1.01e-12	2.31e-2	-2.37e-13	2.78e-2	9.33e-15
		1.77e-2	5.07e-14	6.09e-2	2.97e-13	6.10e-2	6.36e-13
	20			0.59516	0.33922	0.04708	0.1019
				-9.16e-3	-2.46e-12	-3.28e-3	0.00e+0
				-5.94e-3	3.30e-14	9.33e-3	4.03e-14
				1.25e+0	1.69e-11	6.10e-1	2.02e-12
	50					0.29406	0.0826
						-1.28e-2	0.00e+0
						2.16e-3	7.22e-15
						1.79e-2	4.64e-11

Table S18: Comparisons of numbers of collocation points for ODE-5 death-penalty for hard-constraints. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points					
		20		**50**		**100**	
No. of collocation points	**10**	0.0	0.92792	0.0	0.9974	0.0	0.98482
		3.30e-12	-1.02e-11	3.25e-12	-2.46e-11	3.39e-12	-1.64e-11
		1.05e-11	-2.32e-12	9.89e-12	-7.86e-12	1.06e-11	-4.22e-12
		2.80e-11	4.13e-13	2.82e-11	-2.10e-12	2.82e-11	-4.63e-13
	20			0.01558	0.8013	0.00204	0.22946
				6.37e-16	-1.50e-11	3.29e-14	-3.97e-12
				8.06e-14	-4.06e-12	1.00e-13	6.76e-13
				2.20e-13	1.69e-12	2.14e-13	1.59e-11
	50					0.0979	0.0801
						-6.77e-15	-3.10e-12
						1.50e-14	5.52e-12
						6.36e-14	2.26e-11

Table S19: Comparisons of numbers of collocation points for ODE-6 death-penalty for hard-constraints. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points					
		20		**50**		**100**	
No. of collocation points	**10**	0.38494	0.20568	0.13768	0.78024	0.24258	0.0155
		-6.01e-32	0.00e+0	-4.03e-32	-5.55e-17	-2.06e-32	0.00e+0
		3.18e-33	0.00e+0	1.94e-32	0.00e+0	1.43e-33	0.00e+0
		3.65e-31	5.55e-17	1.93e-31	0.00e+0	4.33e-32	5.55e-17
	20			0.36226	0.89444	0.60892	0.03988
				-2.92e-32	-5.55e-17	-4.60e-32	0.00e+0
				-7.52e-33	0.00e+0	-8.29e-33	0.00e+0
				6.71e-32	0.00e+0	8.54e-32	2.78e-17
	50					0.84006	0.0001
						-2.56e-31	0.00e+0
						-1.86e-32	5.55e-17
						9.36e-33	1.11e-16

Table S20: Comparisons of numbers of collocation points for ODE-7 death-penalty + threshold for hard-constraints. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points								
		20			**50**			**100**		
No. of collocation points	**10**	0.0072	0.95482	0.09076	0.0	0.42076	0.7033	8e-5	0.37356	0.3852
		1.82e-2	-1.46e-10	-8.53e-14	4.22e-2	-1.26e-7	-1.25e-20	2.80e-2	-1.04e-9	-1.44e-19
		2.53e-1	1.05e-10	-1.55e-18	2.81e-1	-1.69e-14	0.00e+0	1.60e-1	1.51e-14	0.00e+0
		3.45e-1	7.97e-8	1.98e-21	3.13e-1	6.56e-8	1.25e-17	3.40e-1	3.88e-8	1.57e-17
	20				0.5698	0.06556	0.97996	0.82098	0.08294	0.72286
					-1.63e-2	-1.26e-7	-2.01e-22	-2.17e-2	-3.36e-8	-6.64e-24
					-1.82e-3	-1.66e-9	4.72e-19	-1.49e-2	-1.77e-11	1.07e-18
					1.08e-2	9.77e-15	3.66e-12	5.14e-3	1.14e-8	4.26e-14
	50							0.92044	0.16104	0.19378
								-2.05e-2	-4.54e-10	0.00e+0
								-1.12e-2	7.67e-10	0.00e+0
								4.71e-3	7.07e-8	2.44e-15

Table S21: Comparisons of numbers of collocation points for ODE-8 death-penalty + threshold for hard-constraints. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points								
		20			**50**			**100**		
No. of collocation points	**10**	0.91208	0.55958	0.40734	0.91284	0.4871	0.85078	0.88532	0.90374	0.66058
		-1.95e-2	0.00e+0	0.00e+0	-1.26e-1	-7.47e-8	0.00e+0	-7.00e-2	0.00e+0	-4.03e-22
		0.00e+0	0.00e+0	0.00e+0	0.00e+0	0.00e+0	0.00e+0	0.00e+0	0.00e+0	0.00e+0
		0.00e+0	0.00e+0	0.00e+0	0.00e+0	0.00e+0	0.00e+0	0.00e+0	0.00e+0	0.00e+0
	20				0.99888	0.14462	0.48776	0.95762	0.88512	0.6041
					-2.74e-1	0.00e+0	0.00e+0	-3.80e-2	0.00e+0	-3.84e-22
					0.00e+0	0.00e+0	0.00e+0	0.00e+0	0.00e+0	0.00e+0
					0.00e+0	0.00e+0	1.50e-27	0.00e+0	0.00e+0	0.00e+0
	50							0.2648	0.94372	0.3303
								0.00e+0	0.00e+0	-1.36e-20
								0.00e+0	0.00e+0	0.00e+0
								0.00e+0	0.00e+0	0.00e+0

Table S22: Comparisons of numbers of collocation points for ODE-9 death-penalty for hard-constraints. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points								
		20			**50**			**100**		
No. of collocation points	**10**	0.0 2.17e-13 9.70e-13 1.79e-7	0.06262 -2.37e-14 -1.94e-15 -4.44e-16	1.0 2.08e-32 1.08e-19 4.58e-17	0.0 2.44e-13 1.02e-12 1.79e-7	0.43232 -4.46e-14 -2.66e-15 -4.44e-16	1.0 -1.60e-22 1.31e-25 2.95e-17	0.0 2.50e-13 1.03e-12 3.57e-7	0.82476 -3.28e-15 -9.99e-16 -5.55e-17	0.99954 -7.27e-17 -1.44e-24 5.55e-17
	20				0.10272 -5.15e-15 1.44e-15 2.33e-14	0.95212 -8.27e-15 -2.78e-16 2.22e-16	0.74878 -8.58e-18 -2.41e-24 0.00e+0	0.00016 2.72e-15 7.25e-15 1.61e-14	0.96616 -6.66e-16 2.78e-16 2.39e-15	0.37474 -1.37e-16 -7.15e-25 0.00e+0
	50							0.00114 5.48e-16 3.17e-15 1.32e-14	0.75326 -2.16e-15 6.66e-16 1.82e-14	0.21042 -2.78e-17 -5.69e-27 5.70e-25

Table S23: Comparisons of numbers of collocation points for ODE-10 death-penalty for hard-constraints. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points								
		20			**50**			**100**		
No. of collocation points	**10**	0.0 6.10e-11 6.86e-10 1.28e-6	0.0317 0.00e+0 0.00e+0 4.44e-16	0.03132 0.00e+0 0.00e+0 4.44e-16	0.00014 3.64e-11 2.77e-10 1.28e-6	0.018 0.00e+0 0.00e+0 6.66e-16	0.01802 0.00e+0 0.00e+0 6.66e-16	0.00024 6.14e-11 6.92e-10 1.28e-6	0.08978 0.00e+0 0.00e+0 2.22e-16	0.08894 0.00e+0 0.00e+0 2.22e-16
	20				0.03306 2.39e-14 1.15e-12 4.16e-12	0.4984 0.00e+0 0.00e+0 0.00e+0	0.49998 0.00e+0 0.00e+0 0.00e+0	0.02548 9.50e-14 1.62e-12 3.98e-12	0.94348 0.00e+0 0.00e+0 0.00e+0	0.94428 0.00e+0 0.00e+0 0.00e+0
	50							0.62324 -3.86e-13 -6.50e-14 3.40e-13	0.81944 0.00e+0 0.00e+0 0.00e+0	0.81952 0.00e+0 0.00e+0 0.00e+0

4.3 Hard constraints

Table S24 shows the numbers of non-trivial solutions for the ODEs solved by the hard constraint method where this is an issue. The additional threshold suppresses trivial solutions for ODE-4, ODE-7 and ODE-8 completely successfully. See the main paper for a discussion.

Table S24: Numbers of non-trivial solutions from the 30 repetitions ODE-4, ODE-7 and ODE-8 with and without the additional threshold constraint for varying numbers of collocation points for hard constraints.

	# Collocation pts.			
ODE	**10**	**20**	**50**	**100**
ODE-4 without	8	11	22	21
ODE-4 with	30	30	30	30
ODE-7 without	0	0	0	0
ODE-7 with	30	30	30	30
ODE-8 without	0	0	0	0
ODE-8 with	30	30	30	30

Comparison tables for ODE-7 and ODE-8 with hard constraints but without the additional threshold constraint have been omitted because all 30 runs converged to trivial solutions—see Table S24.

- For all the first-order ODEs, there is no evidence of any statistically-significant differences in any of the measures.
- For ODE-7 with threshold, the test error indicates that 10 points is too few; there is also evidence that 100 points has superior to 20 points although 50 and 100 points appear equivalent.
- For ODE-8 with threshold, again 10 collocation points appears too few.

Table S25: Comparisons of numbers of collocation points for ODE-1 for hard-constraints. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points					
		20		**50**		**100**	
No. of collocation points	**10**	0.88074 -1.53e-14 -1.05e-29 1.61e-29	0.98636 -1.46e-12 -3.55e-15 0.00e+0	0.88654 -4.21e-17 -1.58e-19 1.57e-30	0.99946 -5.18e-11 -1.28e-12 0.00e+0	0.94766 -3.37e-16 -7.22e-21 1.79e-30	0.99972 -1.27e-10 -1.31e-13 -1.78e-15
	20			0.64228 -4.32e-14 -1.58e-19 2.23e-20	0.8296 -1.64e-11 -7.11e-15 3.89e-13	0.68266 -5.06e-18 -8.87e-23 4.11e-23	0.9841 -7.42e-11 -6.04e-14 0.00e+0
	50					0.58844 -1.54e-15 1.86e-26 3.57e-17	0.32864 -3.57e-12 1.78e-15 4.30e-11

Table S26: Comparisons of numbers of collocation points for ODE-2 for hard-constraints. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points					
		20		**50**		**100**	
No. of collocation points	**10**	0.39664 -6.03e-9 -3.85e-10 6.14e-9	0.49882 -1.55e-9 -2.43e-11 6.06e-10	0.8504 -2.31e-8 -3.76e-9 2.04e-9	0.75554 -4.72e-9 -5.28e-10 8.69e-10	0.10516 -6.99e-10 1.31e-9 2.20e-8	0.40842 -4.15e-10 -1.95e-14 5.65e-10
	20			0.98746 -3.52e-8 -1.67e-8 2.04e-9	0.88034 -2.32e-9 -2.91e-10 9.83e-12	0.36612 -3.06e-9 -9.05e-11 1.11e-8	0.64974 -3.26e-9 -9.05e-12 2.15e-9
	50					0.09002 -1.66e-9 2.46e-9 3.17e-8	0.22596 -2.58e-9 4.40e-11 5.61e-9

Table S27: Comparisons of numbers of collocation points for ODE-3 for hard-constraints. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points					
		20		**50**		**100**	
No. of collocation points	**10**	0.43696 -2.65e-16 -1.06e-17 2.51e-16	0.26394 -1.93e-27 0.00e+0 7.58e-25	0.06776 -8.81e-18 2.50e-17 3.64e-16	0.88096 -1.23e-25 0.00e+0 0.00e+0	0.12806 -3.75e-17 4.45e-18 6.11e-16	0.55036 -1.16e-26 0.00e+0 7.05e-27
	20			0.15246 -2.65e-17 6.04e-17 2.93e-16	0.47086 -2.22e-27 0.00e+0 3.42e-26	0.07368 -3.59e-17 1.62e-16 5.11e-16	0.51686 -1.04e-26 0.00e+0 4.92e-27
	50					0.76006 -2.10e-16 -2.53e-17 3.23e-17	0.47444 -1.14e-26 0.00e+0 6.51e-18

Table S28: Comparisons of numbers of collocation points for ODE-4 threshold for hard-constraints. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points					
		20		**50**		**100**	
No. of collocation points	**10**	0.01918	0.17328	0.0	0.55636	0.00064	0.37412
		-7.42e-15	-2.22e-16	1.40e-14	-6.66e-16	7.85e-15	-6.11e-16
		2.13e-14	0.00e+0	4.23e-14	0.00e+0	3.70e-14	0.00e+0
		1.45e-13	1.44e-15	1.02e-13	9.99e-16	1.07e-13	1.50e-15
	20			0.12816	0.98778	0.04126	0.85652
				-2.26e-15	-1.17e-15	-8.35e-16	-9.44e-16
				6.35e-15	-2.22e-16	2.73e-15	-1.11e-16
				1.69e-14	0.00e+0	2.48e-14	1.11e-16
	50					0.53822	0.4128
						-4.50e-15	-5.00e-16
						-4.45e-16	0.00e+0
						5.53e-15	3.39e-16

Table S29: Comparisons of numbers of collocation points for ODE-5 for hard-constraints. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points					
		20		**50**		**100**	
No. of collocation points	**10**	0.0	0.94192	0.0	0.98676	0.0	0.7531
		2.59e-12	-8.02e-12	2.04e-12	-1.08e-11	2.66e-12	-2.50e-12
		3.90e-12	-3.21e-13	5.10e-12	-1.90e-12	6.44e-12	-5.06e-14
		1.04e-11	2.49e-13	1.17e-11	-3.73e-14	1.16e-11	4.62e-13
	20			0.26364	0.4349	0.28102	0.41126
				-3.41e-13	-6.30e-12	-9.13e-13	-8.41e-13
				3.76e-13	5.85e-13	3.53e-13	6.75e-14
				5.88e-13	6.82e-12	1.08e-12	3.65e-12
	50					0.31614	0.35686
						-7.09e-13	-1.86e-12
						4.45e-13	-5.06e-14
						8.97e-13	6.17e-12

Table S30: Comparisons of numbers of collocation points for ODE-6 for hard-constraints. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points					
		20		**50**		**100**	
No. of collocation points	**10**	0.18938	0.22546	0.45758	0.88052	0.13652	0.14178
		-1.09e-32	0.00e+0	-3.71e-32	-8.33e-17	-4.35e-33	0.00e+0
		2.95e-33	0.00e+0	-1.09e-33	0.00e+0	4.05e-33	0.00e+0
		1.01e-31	5.55e-17	4.09e-32	0.00e+0	3.19e-32	5.55e-17
	20			0.69492	0.91676	0.53078	0.51242
				-2.73e-32	-5.55e-17	-2.12e-32	0.00e+0
				-1.45e-33	0.00e+0	5.45e-34	0.00e+0
				1.13e-32	0.00e+0	1.98e-32	0.00e+0
	50					0.47766	0.0572
						-2.45e-32	0.00e+0
						7.81e-33	0.00e+0
						2.16e-32	1.11e-16

Table S31: Comparisons of numbers of collocation points for ODE-7 threshold for hard-constraints. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points								
			20			**50**			**100**	
No. of collocation points	**10**	0.0003	0.91812	0.71614	0.00018	0.80978	0.96708	0.0	0.69438	0.97478
		1.01e-3	-1.78e-14	-1.82e-20	5.78e-4	-3.97e-11	-1.21e-22	1.74e-3	-9.92e-11	-4.39e-23
		4.53e-2	0.00e+0	0.00e+0	6.79e-3	0.00e+0	0.00e+0	4.53e-2	0.00e+0	0.00e+0
		3.66e-1	0.00e+0	0.00e+0	2.76e-1	0.00e+0	0.00e+0	3.74e-1	0.00e+0	0.00e+0
	20				0.3048	0.48954	0.7018	0.00626	0.40024	0.9593
					-4.31e-6	-2.09e-11	-1.21e-22	1.28e-8	-9.92e-11	-4.39e-23
					4.27e-7	0.00e+0	0.00e+0	1.05e-6	-3.55e-15	0.00e+0
					2.37e-5	8.88e-15	1.82e-20	8.39e-5	0.00e+0	1.71e-18
	50							0.09924	0.38162	0.612
								-3.01e-7	-5.68e-14	0.00e+0
								3.74e-8	0.00e+0	0.00e+0
								9.64e-6	7.11e-15	6.22e-35

Table S32: Comparisons of numbers of collocation points for ODE-8 threshold for hard-constraints. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points								
		20			**50**			**100**		
No. of collocation points	**10**	2e-5 2.24e-7 1.06e-6 6.33e-5	0.91046 -7.53e-13 -2.76e-14 4.44e-16	0.98604 -1.75e-15 -7.10e-24 5.83e-31	0.00022 2.18e-7 1.78e-6 6.33e-5	0.5539 -2.20e-13 -1.79e-14 -4.44e-16	0.9668 -9.41e-26 6.04e-36 1.03e-29	2e-5 2.24e-7 1.78e-6 6.33e-5	0.91344 -2.02e-12 -2.40e-14 0.00e+0	0.99588 -3.61e-16 -4.26e-24 5.82e-28
	20				0.0186 -1.65e-13 3.88e-10 8.32e-10	0.30482 -1.82e-13 1.07e-14 6.11e-13	0.24948 -9.75e-26 1.99e-25 8.30e-22	0.11436 -3.21e-11 4.94e-11 5.52e-10	0.1398 -8.24e-13 2.22e-16 1.03e-13	0.56018 -2.21e-26 8.55e-24 1.75e-15
	50							0.44632 -3.02e-11 8.52e-12 4.02e-11	0.39408 -7.84e-13 -1.11e-16 7.11e-14	0.71086 -6.42e-24 0.00e+0 7.20e-26

Table S33: Comparisons of numbers of collocation points for ODE-9 for hard-constraints. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points								
		20			**50**			**100**		
No. of collocation points	**10**	0.0	0.0657	0.99998	0.0	0.43158	0.99996	0.0	0.82448	0.99946
		2.30e-13	-2.37e-14	2.08e-32	2.44e-13	-4.46e-14	-3.19e-22	2.48e-13	-3.00e-15	-7.27e-17
		9.70e-13	-1.94e-15	1.08e-19	1.02e-12	-2.66e-15	1.31e-25	1.03e-12	-9.99e-16	-1.44e-24
		1.79e-7	-4.44e-16	4.58e-17	3.57e-7	-4.44e-16	2.95e-17	3.57e-7	-5.55e-17	5.55e-17
	20				0.10034	0.95352	0.75228	8e-5	0.9658	0.37152
					-5.15e-15	-9.44e-15	-8.58e-18	2.72e-15	-6.66e-16	-1.37e-16
					1.44e-15	-2.78e-16	-2.41e-24	7.25e-15	2.78e-16	-7.15e-25
					2.33e-14	2.22e-16	0.00e+0	1.66e-14	2.39e-15	0.00e+0
	50							0.00112	0.75576	0.21224
								5.48e-16	-2.16e-15	-2.78e-17
								3.17e-15	6.66e-16	-5.69e-27
								1.32e-14	1.82e-14	4.30e-25

Table S34: Comparisons of numbers of collocation points for ODE-10 for hard-constraints. Each cell contains the posterior probability that the row-wise entry is superior to the column-wise entry, and the confidence interval for the same comparison. See text for further details.

		No. of collocation points								
		20			**50**			**100**		
No. of collocation points	**10**	8e-5	0.6307	0.33314	0.0	0.20776	0.48072	0.0	0.8775	0.83162
		1.38e-10	-8.88e-16	-2.22e-16	1.43e-10	-1.33e-15	0.00e+0	1.52e-10	-2.19e-13	-1.62e-14
		2.82e-9	0.00e+0	0.00e+0	2.76e-9	0.00e+0	0.00e+0	2.82e-9	-4.44e-16	0.00e+0
		1.23e-8	2.66e-15	0.00e+0	1.23e-8	1.33e-15	0.00e+0	1.23e-8	4.44e-16	0.00e+0
	20				0.26272	0.05414	0.3064	6e-5	0.71622	0.78194
					-3.40e-12	-1.33e-15	0.00e+0	5.62e-13	-6.39e-14	-1.62e-14
					-1.58e-13	4.44e-16	0.00e+0	3.95e-12	-4.44e-16	0.00e+0
					8.08e-12	1.87e-14	6.66e-16	1.01e-11	8.88e-16	2.22e-16
	50							0.00736	0.85566	0.9169
								-1.75e-14	-1.98e-13	-2.22e-16
								6.22e-13	-8.88e-16	0.00e+0
								6.08e-12	0.00e+0	0.00e+0

5 RQ2: Does the evolutionary method affect the performance?

Here we consider the second research question of whether, for a given ODE and a given number of collocation points, there is any difference between the three methods considered in this paper: hard-constraints (**hc**) with the novel ranking method, hard constraints enforced with a death penalty (**dp**) mechanism, and the penalized loss function (**pm**).

We omit comparisons for ODE-4, ODE-7 and ODE-8 *without* threshold since the results indicate very frequent convergence to trivial solutions for these configurations.

The following tables summarize the statistical tests and the confidence intervals for each pairwise comparison. Here we make comparisons between pairs of methods for a given number of collocation points, and report both posterior probabilities of differences and the confidence intervals (CIs). To take Table S35 for 10 collocation points as an example, the probability that the hard constraint (**hc**) method has a superior test error than the death penalty (**dp**) method is 0.26628 with a median difference of 1.11×10^{-30}, and a confidence interval of $[-1.67 \times 10^{-23}, 2.69 \times 10^{-22}]$. The probability for comparing the initial value violations for this **hc**-**dp** pairing is 0.90512, and the corresponding confidence interval $[-1.51 \times 19^{-13}.1.78 \times 10^{-15}]$. The tabular entries for the remaining pairwise method comparisons follow accordingly.

We can make the following observations for each of the ODEs:

- For ODE-1, for each of the numbers of collocation points, there is no evidence for any differences between the hard constraint (**hc**) and death penalty (**dp**) approaches whereas there is strong evidence of statistical superiority of both **hc** and **dp** over the penalty method (**pm**) for initial value (IV) attainment, but not test error.
- For ODE-2, again there is no difference between **hc** and **dp**, but both these methods are superior to **pm** for test error and and IV violation.
- For ODE-3, there is no evidence of any difference between any of the compared methods.
- For ODE-4 with threshold, **hc** is consistently superior to **dp** for both the test error, while **pm** is superior to **dp** in test error. For ODE-4 with threshold, the death penalty (**dp**) thus appears to perform especially poorly.
- For ODE-5 and ODE-6, there is no evidence of any difference between any of the compared methods.
- For ODE-7 with threshold, **hc** is superior to both **dp** and **pm** on test error apart from for 10 collocation points where the evidence of difference between **hc** and **dp** is very modest. For this ODE, the penalty method appears to generate noticeably poorer results than either **hc** or **dp**.
- For ODE-8 with threshold, **hc** is consistently superior to **dp** on test error while being superior on initial value violation for 10 and 50 points; the statistical test for IV violation for 20 and 100 points is on the margins of superiority. **pm**, on the other hand, is consistently superior to **dp** on test error. Again, the **dp** solutions appears noticeably the worst.
- For ODE-9 and ODE-10, there is no difference between methods. An additional observation for ODE-10 is that many of the CIs are zero, which probably arises because the solution to this ODE is a monotonically increasing function that passes through the origin. It is probably straightforward for GP to evolve a tree that does not add a constant term into the root node thus achieving a 'perfect' match to the IV.

Table S35: Comparison of the performance of the three methods for ODE-1 and 10 collocation points. See text for further details.

Method			Method			
			Death penalty		Penalty method	
Hard constr.			0.26628	0.90512	1.0	1.0
			-1.67e-23	-1.51e-13	-1.77e-7	-1.78e-4
			1.11e-30	-1.78e-15	-4.88e-9	-5.37e-5
			2.69e-22	1.78e-15	-5.81e-10	-2.39e-5
Death penalty					1.0	1.0
					-1.81e-7	-1.78e-4
					-4.88e-9	-5.37e-5
					-5.82e-10	-2.39e-5

Table S36: Comparison of the performance of the three methods for ODE-1 and 20 collocation points. See text for further details.

Method			Method			
			Death penalty		Penalty method	
Hard constr.			0.00944	0.42082	1.0	1.0
			1.10e-30	-8.88e-15	-1.65e-7	-3.35e-4
			2.23e-20	3.55e-15	-5.52e-8	-8.52e-5
			2.07e-12	8.88e-15	-6.47e-10	-7.64e-6
Death penalty					1.0	1.0
					-1.65e-7	-3.35e-4
					-5.57e-8	-8.52e-5
					-6.46e-10	-7.64e-6

Table S37: Comparison of the performance of the three methods for ODE-1 and 50 collocation points. See text for further details.

Method			Method			
			Death penalty		Penalty method	
Hard constr.			0.07018	0.10416	1.0	1.0
			-2.49e-31	-1.78e-15	-2.41e-8	-1.17e-4
			4.35e-19	1.00e-12	-4.71e-9	-3.45e-5
			4.21e-15	5.77e-11	-5.97e-11	-1.21e-5
Death penalty					1.0	1.0
					-2.26e-8	-1.26e-4
					-4.75e-9	-3.45e-5
					-5.97e-11	-1.28e-5

Table S38: Comparison of the performance of the three methods for ODE-1 and 100 collocation points. See text for further details.

Method			Method			
			Death penalty		Penalty method	
Hard constr.			0.13288	0.33138	1.0	1.0
			-3.21e-25	-5.33e-15	-3.43e-8	-2.09e-4
			8.87e-23	5.33e-15	-4.57e-9	-2.51e-5
			5.06e-18	2.29e-11	-9.28e-13	-1.09e-6
Death penalty					1.0	1.0
					-4.28e-8	-2.04e-4
					-4.57e-9	-2.51e-5
					-3.32e-14	-1.13e-6

Table S39: Comparison of the performance of the three methods for ODE-2 and 10 collocation points. See text for further details.

Method			Method			
			Death penalty		Penalty method	
Hard	constr.		0.25872	0.76732	0.99998	1.0
			-4.12e-9	-1.65e-9	-5.89e-5	-4.66e-3
			2.93e-11	-1.18e-10	-4.08e-5	-1.70e-3
			6.42e-9	4.98e-10	-1.10e-5	-6.32e-4
Death	penalty				1.0	1.0
					-5.90e-5	-3.93e-3
					-4.09e-5	-1.70e-3
					-1.22e-5	-6.33e-4

Table S40: Comparison of the performance of the three methods for ODE-2 and 20 collocation points. See text for further details.

Method			Method			
			Death penalty		Penalty method	
Hard	constr.		0.20556	0.49862	1.0	1.0
			-2.69e-9	-1.49e-10	-1.07e-5	-3.34e-3
			5.03e-10	-5.14e-12	-7.15e-6	-1.42e-3
			1.10e-8	5.19e-10	-3.81e-6	-9.24e-4
Death	penalty				1.0	1.0
					-1.06e-5	-3.34e-3
					-7.15e-6	-1.42e-3
					-3.72e-6	-9.86e-4

Table S41: Comparison of the performance of the three methods for ODE-2 and 50 collocation points. See text for further details.

Method		Method			
		Death penalty		Penalty method	
Hard constr.		0.00906	0.62504	1.0	1.0
		-1.25e-9	-4.87e-9	-2.14e-5	-3.39e-3
		1.01e-8	-4.31e-12	-8.79e-6	-1.93e-3
		3.69e-8	8.82e-10	-2.30e-6	-9.05e-4
Death penalty				1.0	1.0
				-2.14e-5	-3.38e-3
				-9.80e-6	-1.93e-3
				-2.32e-6	-9.05e-4

Table S42: Comparison of the performance of the three methods for ODE-2 and 100 collocation points. See text for further details.

Method		Method			
		Death penalty		Penalty method	
Hard constr.		0.77628	0.9957	1.0	1.0
		-2.25e-8	-2.42e-8	-3.11e-5	-7.50e-3
		-2.18e-11	-5.71e-9	-2.03e-5	-4.54e-3
		1.56e-9	-1.01e-10	-7.27e-6	-1.50e-3
Death penalty				1.0	1.0
				-3.20e-5	-7.50e-3
				-2.03e-5	-4.54e-3
				-6.55e-6	-1.50e-3

Table S43: Comparison of the performance of the three methods for ODE-3 and 10 collocation points. See text for further details.

Method			Method			
			Death penalty		Penalty method	
Hard constr.			0.31266	0.6693	0.65804	0.99186
			-4.77e-17	-9.68e-27	-1.07e-15	-3.72e-13
			3.26e-18	0.00e+0	-6.50e-17	0.00e+0
			1.25e-16	6.73e-27	1.54e-16	0.00e+0
Death penalty					0.4522	0.99056
					-1.31e-16	-3.72e-13
					5.66e-17	0.00e+0
					5.23e-16	0.00e+0

Table S44: Comparison of the performance of the three methods for ODE-3 and 20 collocation points. See text for further details.

Method			Method			
			Death penalty		Penalty method	
Hard constr.			8e-5	0.23514	0.31052	1.0
			4.27e-17	0.00e+0	-9.32e-17	-3.35e-13
			2.41e-16	0.00e+0	2.81e-17	-6.97e-14
			4.77e-16	4.51e-27	3.47e-16	0.00e+0
Death penalty					0.9982	1.0
					-3.03e-16	-3.35e-13
					-3.50e-17	-6.97e-14
					-7.53e-19	0.00e+0

Table S45: Comparison of the performance of the three methods for ODE-3 and 50 collocation points. See text for further details.

			Method			
			Death penalty		Penalty method	
Method	Hard constr.		0.00222	0.50364	0.87612	0.99752
			-7.32e-20	-9.40e-25	-4.89e-16	-4.96e-13
			5.40e-17	0.00e+0	-1.62e-17	-1.08e-15
			1.47e-16	3.25e-18	3.94e-17	0.00e+0
	Death penalty				0.99994	0.99576
					-8.14e-16	-4.96e-13
					-8.60e-17	-1.08e-15
					-4.04e-18	0.00e+0

Table S46: Comparison of the performance of the three methods for ODE-3 and 100 collocation points. See text for further details.

			Method			
			Death penalty		Penalty method	
Method	Hard constr.		0.0	0.7999	0.73752	0.99392
			4.67e-17	-4.93e-18	-2.92e-16	-7.41e-13
			8.22e-17	0.00e+0	-1.42e-17	0.00e+0
			3.28e-16	1.82e-29	1.23e-16	0.00e+0
	Death penalty				0.99994	0.99368
					-6.90e-16	-7.41e-13
					-1.11e-16	0.00e+0
					-5.99e-18	0.00e+0

Table S47: Comparison of the performance of the three methods for ODE-4_threshold and 10 collocation points. See text for further details.

			Method			
			Death penalty		Penalty method	
Method	Hard constr.		1.0	0.99996	0.5328	1.0
			-1.18e+00	-1.07e-12	-6.40e-14	-3.26e-10
			-7.73e-02	-2.35e-13	4.65e-15	-1.25e-10
			-5.86e-02	-1.57e-14	6.34e-14	-7.22e-11
	Death penalty				0.0	0.99182
					5.86e-02	-2.62e-10
					7.73e-02	-1.07e-10
					1.18e+00	-5.12e-12

Table S48: Comparison of the performance of the three methods for ODE-4_threshold and 20 collocation points. See text for further details.

			Method			
			Death penalty		Penalty method	
Method	Hard constr.		1.0	1.0	0.1152	1.0
			-1.30e+00	-3.22e-11	-4.18e-16	-2.05e-10
			-5.09e-02	-3.18e-12	3.44e-15	-9.60e-11
			-4.07e-02	-2.46e-13	1.51e-14	-2.87e-11
	Death penalty				0.0	0.88794
					4.07e-02	-1.73e-10
					5.09e-02	-3.63e-11
					1.30e+00	-9.39e-13

Table S49: Comparison of the performance of the three methods for ODE-4_threshold and 50 collocation points. See text for further details.

Method		Method			
		Death penalty		Penalty method	
Hard constr.		1.0	0.999	0.39986	1.0
		-4.94e-02	-4.64e-11	-6.41e-15	-4.27e-10
		-4.89e-02	-4.69e-13	3.20e-16	-2.14e-10
		-3.89e-02	0.00e+00	9.11e-15	-5.59e-11
Death penalty				0.0	0.9892
				3.89e-02	-2.54e-10
				4.89e-02	-7.74e-11
				4.94e-02	-1.98e-11

Table S50: Comparison of the performance of the three methods for ODE-4_threshold and 100 collocation points. See text for further details.

Method		Method			
		Death penalty		Penalty method	
Hard constr.		1.0	0.8872	0.54834	1.0
		-4.71e-02	-1.57e-13	-7.57e-15	-2.57e-10
		-4.47e-02	0.00e+00	-6.46e-17	-1.06e-10
		-3.13e-02	3.33e-16	4.53e-15	-2.74e-11
Death penalty				0.0	0.97668
				3.13e-02	-1.49e-10
				4.47e-02	-7.16e-11
				4.71e-02	-1.33e-11

Table S51: Comparison of the performance of the three methods for ODE-5 and 10 collocation points. See text for further details.

Method			Method			
			Death penalty		Penalty method	
Hard constr.			0.9336	0.6302	0.72158	1.0
			-2.02e-11	-5.10e-13	-8.59e-12	-1.35e-09
			-4.04e-12	-5.06e-14	-1.08e-12	-5.47e-10
			4.62e-12	1.79e-13	4.21e-12	-3.17e-10
Death penalty					0.028	0.99948
					-5.02e-12	-9.80e-10
					5.58e-12	-5.03e-10
					2.08e-11	-1.99e-10

Table S52: Comparison of the performance of the three methods for ODE-5 and 20 collocation points. See text for further details.

Method			Method			
			Death penalty		Penalty method	
Hard constr.			0.0	0.85968	0.96552	1.0
			1.89e-13	-8.64e-12	-1.57e-12	-1.16e-08
			9.29e-13	-2.08e-13	-6.97e-13	-4.60e-09
			1.75e-12	4.30e-13	-8.15e-15	-2.84e-09
Death penalty					1.0	0.99996
					-3.20e-12	-9.28e-09
					-1.42e-12	-4.58e-09
					-8.78e-13	-2.84e-09

Table S53: Comparison of the performance of the three methods for ODE-5 and 50 collocation points. See text for further details.

Method		Method			
		Death penalty		Penalty method	
Hard constr.		0.0	0.9684	0.79312	1.0
		4.71e-13	-2.25e-11	-1.22e-12	-8.25e-09
		6.99e-13	-7.19e-12	-1.24e-13	-3.57e-09
		1.16e-12	-3.08e-13	3.73e-13	-2.10e-09
Death penalty				1.0	1.0
				-2.07e-12	-8.60e-09
				-1.39e-12	-3.99e-09
				-7.72e-13	-2.27e-09

Table S54: Comparison of the performance of the three methods for ODE-5 and 100 collocation points. See text for further details.

Method		Method			
		Death penalty		Penalty method	
Hard constr.		0.0	0.94928	0.97914	1.0
		2.98e-13	-1.37e-11	-2.27e-12	-7.18e-09
		6.20e-13	-4.94e-12	-9.89e-13	-2.93e-09
		1.54e-12	-2.81e-13	1.43e-13	-1.48e-09
Death penalty				1.0	0.99994
				-2.94e-12	-7.21e-09
				-2.02e-12	-3.01e-09
				-1.38e-12	-1.54e-09

Table S55: Comparison of the performance of the three methods for ODE-6 and 10 collocation points. See text for further details.

Method			Method			
			Death penalty		Penalty method	
Hard	constr.		0.56454	0.51546	0.97146	0.81542
			-9.17e-32	-2.78e-17	-4.32e-28	-2.78e-17
			2.88e-33	0.00e+00	-1.65e-31	0.00e+00
			3.04e-32	2.78e-17	-1.10e-32	0.00e+00
Death	penalty				0.96476	0.95284
					-4.32e-28	-1.11e-16
					-1.96e-31	0.00e+00
					9.15e-34	0.00e+00

Table S56: Comparison of the performance of the three methods for ODE-6 and 20 collocation points. See text for further details.

Method			Method			
			Death penalty		Penalty method	
Hard	constr.		0.4838	0.46118	0.8744	0.89478
			-5.52e-32	0.00e+00	-1.55e-30	-8.33e-17
			3.98e-34	0.00e+00	-9.78e-33	-2.78e-17
			6.03e-32	2.78e-17	7.15e-33	0.00e+00
Death	penalty				0.89344	0.9559
					-5.93e-30	-1.11e-16
					-2.16e-32	0.00e+00
					9.59e-33	0.00e+00

Table S57: Comparison of the performance of the three methods for ODE-6 and 50 collocation points. See text for further details.

Method			Method			
			Death penalty		Penalty method	
Hard constr.			0.27224	0.32964	0.9984	0.90808
			-3.15e-32	0.00e+00	-4.86e-24	-1.94e-16
			1.87e-33	0.00e+00	-2.98e-30	0.00e+00
			2.56e-32	5.55e-17	2.53e-34	8.33e-17
Death penalty					1.0	0.79614
					-7.45e-24	-2.22e-16
					-2.36e-29	0.00e+00
					-6.55e-31	8.33e-17

Table S58: Comparison of the performance of the three methods for ODE-6 and 100 collocation points. See text for further details.

Method			Method			
			Death penalty		Penalty method	
Hard constr.			0.74204	0.04716	0.92498	0.9934
			-2.99e-32	0.00e+00	-1.11e-27	-1.68e-16
			-1.16e-32	0.00e+00	-7.46e-31	0.00e+00
			2.55e-33	5.55e-17	6.87e-33	0.00e+00
Death penalty					0.99448	0.99988
					-4.63e-27	-1.67e-16
					-6.77e-30	0.00e+00
					6.77e-33	0.00e+00

Table S59: Comparison of the performance of the three methods for ODE-7_threshold and 10 collocation points. See text for further details.

Method		Method					
		Death penalty			Penalty method		
Hard constr.		0.86952	0.99838	0.99996	0.99984	0.0	0.2998
		-2.01e-01	-7.64e-08	-1.57e-17	-4.76e-01	-1.78e-15	0.00e+00
		-3.64e-02	-5.74e-10	-5.28e-22	-4.32e-01	0.00e+00	0.00e+00
		2.66e-02	-1.51e-14	0.00e+00	-1.03e-01	3.55e-15	9.71e-31
Death penalty					0.97422	0.0	0.0
					-3.92e-01	2.86e-13	5.29e-315
					-1.42e-01	3.10e-09	4.84e-21
					-5.49e-02	7.97e-08	1.23e-16

Table S60: Comparison of the performance of the three methods for ODE-7_threshold and 20 collocation points. See text for further details.

Method		Method					
		Death penalty			Penalty method		
Hard constr.		0.99966	0.9989	0.99534	0.99998	0.06948	0.12332
		-5.87e-02	-2.95e-08	-3.66e-12	-4.77e-01	0.00e+00	0.00e+00
		-5.13e-02	-6.93e-10	-1.79e-16	-4.77e-01	0.00e+00	0.00e+00
		-4.02e-02	0.00e+00	0.00e+00	-4.77e-01	1.60e-14	1.82e-20
Death penalty					0.99686	0.00024	0.00038
					-4.35e-01	2.75e-14	5.03e-28
					-4.26e-01	1.86e-09	1.81e-16
					-4.12e-01	2.95e-08	3.66e-12

Table S61: Comparison of the performance of the three methods for ODE-7_threshold and 50 collocation points. See text for further details.

Method		Method					
		Death penalty			Penalty method		
Hard constr.		0.99944	0.98364	0.97042	0.9998	0.00212	0.18422
		-5.98e-02	-1.29e-07	-2.44e-15	-4.77e-01	0.00e+00	0.00e+00
		-4.82e-02	-8.62e-10	-1.03e-19	-4.77e-01	0.00e+00	0.00e+00
		-3.77e-02	-3.55e-15	0.00e+00	-4.77e-01	2.09e-11	6.04e-21
Death penalty					1.0	0.0	0.00126
					-4.37e-01	2.84e-14	0.00e+00
					-4.29e-01	7.14e-09	1.09e-19
					-4.14e-01	1.26e-07	2.44e-15

Table S62: Comparison of the performance of the three methods for ODE-7_threshold and 100 collocation points. See text for further details.

Method		Method					
		Death penalty			Penalty method		
Hard constr.		1.0	0.81918	0.67806	1.0	0.0034	0.57228
		-6.59e-02	-5.69e-09	-2.87e-19	-4.77e-01	-1.14e-13	0.00e+00
		-5.77e-02	0.00e+00	0.00e+00	-4.77e-01	0.00e+00	0.00e+00
		-5.01e-02	1.42e-14	0.00e+00	-4.77e-01	7.11e-15	2.32e-21
Death penalty					0.99984	0.0001	0.16786
					-4.24e-01	0.00e+00	0.00e+00
					-4.15e-01	8.88e-15	4.66e-22
					-3.90e-01	5.74e-09	9.39e-16

Table S63: Comparison of the performance of the three methods for ODE-8_threshold and 10 collocation points. See text for further details.

Method		Method					
		Death penalty			Penalty method		
Hard constr.		0.99994	0.95756	0.99998	0.05786	0.99976	0.99992
		-1.94e+02	-1.91e+00	-1.38e-09	-2.10e-07	-1.28e-08	-2.35e-08
		-1.76e+02	-9.84e-02	0.00e+00	8.63e-08	-3.13e-09	-2.04e-09
		-1.60e+02	-1.90e-02	6.04e-36	6.33e-05	-1.98e-10	-2.10e-15
Death penalty					0.0	0.46646	0.0
					1.63e+02	2.54e-02	-1.71e-09
					1.84e+02	1.46e-01	-1.11e-15
					1.95e+02	2.00e+00	5.15e-18

Table S64: Comparison of the performance of the three methods for ODE-8_threshold and 20 collocation points. See text for further details.

Method		Method					
		Death penalty			Penalty method		
Hard constr.		0.9999	0.8984	0.99998	0.95716	0.99984	0.99972
		-1.95e+02	-2.00e+00	-3.35e-06	-4.96e-09	-6.23e-08	-1.10e-07
		-1.75e+02	-1.87e-01	0.00e+00	-1.08e-09	-1.11e-08	-1.39e-08
		-1.57e+02	-9.36e-03	4.72e-23	-1.86e-10	-5.98e-09	-3.08e-09
Death penalty					0.0	0.22932	0.0
					1.58e+02	1.59e-02	-1.24e-08
					1.75e+02	1.87e-01	0.00e+00
					1.95e+02	2.00e+00	3.25e-06

Table S65: Comparison of the performance of the three methods for ODE-8_threshold and 50 collocation points. See text for further details.

Method		Method					
		Death penalty			Penalty method		
Hard constr.		0.99994	0.94094	1.0	0.99094	0.99994	0.99974
		-1.95e+02	-2.00e+00	-4.92e-10	-8.69e-10	-5.66e-08	-1.56e-07
		-1.70e+02	-1.94e-01	-3.73e-34	-2.19e-10	-2.40e-08	-6.40e-08
		-1.58e+02	-7.12e-03	2.20e-28	-2.02e-11	-5.18e-09	-1.15e-08
Death penalty					0.0	0.58146	0.0
					1.59e+02	8.69e-03	-6.40e-08
					1.75e+02	1.94e-01	-4.79e-09
					1.95e+02	2.00e+00	1.36e-20

Table S66: Comparison of the performance of the three methods for ODE-8_threshold and 100 collocation points. See text for further details.

Method		Method					
		Death penalty			Penalty method		
Hard constr.		0.99996	0.84042	0.99996	0.78732	0.99976	0.99976
		-1.95e+02	-2.00e+00	-5.12e-08	-8.25e-11	-9.16e-08	-1.27e-07
		-1.74e+02	-1.75e-01	0.00e+00	-4.75e-11	-2.91e-08	-1.32e-08
		-1.61e+02	-1.23e-03	1.41e-20	-5.46e-13	-8.90e-09	-2.02e-09
Death penalty					0.0	0.2983	0.0
					1.61e+02	3.77e-03	-1.69e-08
					1.74e+02	1.75e-01	-2.05e-09
					1.95e+02	2.00e+00	9.95e-06

Table S67: Comparison of the performance of the three methods for ODE-9 and 10 collocation points. See text for further details.

		Method					
		Death penalty			Penalty method		
Method	Hard constr.	0.33974	0.0537	0.66238	0.1723	0.94372	1.0
		-1.57e-11	-2.22e-16	-1.28e-32	-7.91e-13	-5.44e-10	-5.05e-12
		2.67e-14	0.00e+00	1.21e-17	4.81e-14	-5.64e-11	5.82e-34
		7.01e-13	2.22e-16	7.83e-17	7.62e-10	-1.84e-11	1.05e-17
	Death penalty				0.05634	0.95458	1.0
					1.96e-14	-5.44e-10	-5.05e-12
					2.03e-11	-5.64e-11	0.00e+00
					5.65e-10	-1.84e-11	6.84e-30

Table S68: Comparison of the performance of the three methods for ODE-9 and 20 collocation points. See text for further details.

		Method					
		Death penalty			Penalty method		
Method	Hard constr.	0.00512	0.39364	0.2934	0.6986	0.99986	1.0
		1.10e-15	-2.05e-15	-1.28e-24	-3.56e-14	-2.02e-10	-4.99e-11
		1.19e-14	1.67e-16	2.25e-29	3.45e-16	-6.02e-11	-1.42e-12
		3.59e-14	2.30e-14	5.45e-18	8.40e-15	-3.76e-11	0.00e+00
	Death penalty				0.91204	0.99268	1.0
					-8.20e-14	-2.02e-10	-4.99e-11
					-4.48e-16	-4.45e-11	-1.43e-12
					2.97e-15	-2.84e-11	2.86e-28

Table S69: Comparison of the performance of the three methods for ODE-9 and 50 collocation points. See text for further details.

			Method					
			Death penalty			Penalty method		
Method	Hard constr.		0.0	0.07848	0.61662	0.35152	0.99984	1.0
			1.51e-15	-6.72e-15	0.00e+00	-5.03e-15	-2.30e-10	-1.27e-10
			6.46e-15	-4.44e-16	3.39e-25	1.17e-15	-7.89e-11	-6.30e-11
			3.13e-14	2.55e-15	1.94e-22	1.79e-14	-2.03e-11	2.68e-28
	Death penalty					1.0	0.99998	0.99994
						-1.92e-14	-2.30e-10	-1.27e-10
						-5.79e-15	-8.05e-11	-6.30e-11
						-4.96e-15	-2.03e-11	0.00e+00

Table S70: Comparison of the performance of the three methods for ODE-9 and 100 collocation points. See text for further details.

			Method					
			Death penalty			Penalty method		
Method	Hard constr.		0.0012	0.09258	0.49754	0.98556	0.99998	1.0
			1.75e-16	-5.55e-16	0.00e+00	-9.85e-15	-8.45e-10	-1.33e-09
			7.23e-16	-5.55e-17	2.03e-24	-1.38e-15	-3.33e-10	-5.17e-11
			3.09e-15	9.44e-16	5.96e-17	6.93e-17	-1.00e-10	0.00e+00
	Death penalty					0.9999	1.0	1.0
						-1.55e-14	-8.46e-10	-1.27e-09
						-4.10e-15	-3.33e-10	-5.83e-11
						-4.31e-16	-8.91e-11	0.00e+00

Table S71: Comparison of the performance of the three methods for ODE-10 and 10 collocation points. See text for further details.

Method		Method					
		Death penalty			Penalty method		
Hard constr.		0.55948	0.57756	0.08848	0.1123	1.0	0.99998
		-3.49e-08	0.00e+00	0.00e+00	-5.47e-10	-1.04e-08	-1.94e-08
		4.25e-13	0.00e+00	0.00e+00	7.50e-11	-3.50e-09	-8.51e-09
		6.24e-09	2.66e-15	0.00e+00	1.10e-08	-1.54e-09	-4.73e-09
Death penalty					0.09798	1.0	1.0
					-2.59e-10	-1.07e-08	-1.94e-08
					3.21e-11	-4.09e-09	-8.51e-09
					1.26e-06	-2.05e-09	-4.73e-09

Table S72: Comparison of the performance of the three methods for ODE-10 and 20 collocation points. See text for further details.

Method		Method					
		Death penalty			Penalty method		
Hard constr.		0.05262	0.002	0.0004	0.94854	1.0	1.0
		-1.12e-12	0.00e+00	0.00e+00	-4.32e-11	-9.83e-09	-1.07e-08
		9.08e-13	8.88e-16	0.00e+00	-3.45e-12	-3.29e-09	-4.68e-09
		1.07e-11	1.91e-14	1.33e-15	5.85e-13	-1.28e-09	-1.33e-09
Death penalty					0.98994	1.0	1.0
					-7.03e-11	-9.83e-09	-1.07e-08
					-9.27e-12	-3.29e-09	-4.68e-09
					5.94e-15	-1.28e-09	-1.33e-09

Table S73: Comparison of the performance of the three methods for ODE-10 and 50 collocation points. See text for further details.

Method		Method					
		Death penalty			Penalty method		
Hard constr.		0.24638	0.13382	0.0036	0.8281	1.0	1.0
		-6.07e-13	0.00e+00	0.00e+00	-6.86e-12	-1.23e-08	-2.54e-08
		2.93e-13	0.00e+00	0.00e+00	-2.39e-12	-5.73e-09	-6.26e-09
		3.29e-12	1.69e-14	0.00e+00	2.77e-13	-2.78e-09	-3.56e-09
Death penalty					0.97082	1.0	1.0
					-7.24e-12	-1.24e-08	-2.54e-08
					-4.63e-12	-5.71e-09	-6.99e-09
					-6.79e-13	-2.78e-09	-3.77e-09

Table S74: Comparison of the performance of the three methods for ODE-10 and 100 collocation points. See text for further details.

Method		Method					
		Death penalty			Penalty method		
Hard constr.		0.28446	0.15784	0.0144	0.87432	1.0	1.0
		-2.76e-13	0.00e+00	0.00e+00	-2.57e-12	-9.45e-09	-1.73e-08
		1.83e-13	8.88e-16	0.00e+00	-2.93e-13	-5.99e-09	-6.68e-09
		1.60e-12	1.09e-13	4.00e-15	3.72e-13	-1.57e-09	-1.79e-09
Death penalty					0.87018	1.0	1.0
					-1.95e-12	-9.45e-09	-1.72e-08
					-3.03e-13	-5.99e-09	-6.68e-09
					6.89e-14	-1.58e-09	-1.79e-09

6 Overall summary of statistical results

Reviewing the results as a whole, we can draw a number of conclusions. In terms of statistical significance, there are many cases where the Bayesian Wilcoxon test implies a strong significant difference, but the confidence interval for the statistic encompasses the null outcome of zero, and thus suggests no difference. There are also many cases where the Wilcoxon test implies a difference, but the CI falls inside the region of probable equivalence (ROPE) so the effect size is judged of no practical importance. Simply reporting the outcome of the statistical test would have resulted in far more 'statistically significant' results, many of which would be (highly) questionable.

In terms of RQ1 and whether there are differences between the numbers of collocation points and performance. The answer to this research question appears to be that, generally, there is no statistical difference between 10, 20, 50 or 100 points although results a few individual ODEs indicate that 10 or sometimes 20 points are insufficient. ODE-7, ODE-8 and to a lesser degree ODE-10, seem to stand out in this regard.

Regarding RQ2 and whether any of the three methods stands out as exhibiting superior performance.

Insofar as test error performance is concerned, most of the pairwise comparisons yield no evidence of statistical difference. This is surprising, and we conjecture that this is a consequence of the parameter tuning. Where there is evidence of difference, sometimes **pm** is better than **dp**; sometimes the reverse is true. Although **hc** is not consistently superior to either or both of the other two methods, where there is a statistical difference involving **hc**, it always favors the hard-constraint approach. This suggest that **hc** should be the preferred approach since it is either superior to its competitors, or at worst, performs equally well.

A corollary is that while the death penalty approach frequently performs well—maybe far better than the criticisms of the technique would suggest—occasionally it does produce poor results.

Apart from ODE-1 and ODE-2, a noteworthy feature of these results is that the ability of all methods to satisfy the initial conditions is remarkably consistent. One might assume that the penalty method (**pm**) would struggle to meet the initial conditions because these constraints are 'soft'. We speculate that this is a consequence of constant tuning coupled with the fact that a function that accurately satisfies the IVs (at the edge of the domain) also accurately satisfies the ODE in the rest of the domain: for an ODE there is no conflict/trade-off between satisfying the initial conditions an solving the ODE. There is evidence that **pm** is deficient for ODE-1 and ODE-2, but what properties of these two ODEs lead to this is unclear.

7 The Best Results Obtained

Although the statistical analyses in the preceding sections address the questions of whether one method is better/worse than another 'on average', in practice, a numerical analyst would adopt the best solution for practical use. Therefore, in this section, the best results obtained for each method and for each number of collocation points are presented. Three points need to be borne in mind, however, when interpreting these results:

1. The results here are the 'extreme' points in the distributions of approximations. They are thus not typical of a given method/number of collocation points, but they do give some indication of the (upper bound on the) accuracy that can be obtained with a particular approach.

2. In general, the solution for a given ODE will not be known so presenting the result that is closest to the known, analytical solution is deceptive; in practice, the value of the technique will be for ODEs for which a closed-form solution is *not* known. We therefore present the results for the GP run which gave the smallest training loss. This is 'unusual' in the sphere of machine learning where a more elaborate model selection routine is typically used. Here, however, we are not learning a function from (noisy) data, but rather searching for the solution to an equation.

3. Finally, we present plots of *relative error* since this is the most useful practical measure, and effectively gauges the number of accurate decimal digits in the GP approximation.

Since the results in this section are 'extreme' results, there is unsurprisingly no consistent pattern of behavior. For example, ranking by peak magnitude of the relative error does not appear to provide any useful insight. Nonetheless, we can make the following observations:

- For two of the benchmark ODEs (ODE-1 and ODE-6), the GP framework used here appears able to produce the analytical solution to within double-precision floating-point rounding error—see Figures S1-S3. As previously noted Rockett (2025), this is interesting since it implies that the GP function set used in this work is able to synthesize general functions. The solution to ODE-1 is $y = x + \frac{2}{x}$ and the solution to ODE-6 $y = \frac{1}{1+x}$, neither of which can be directly generated by the present function set as it does not include the division operator. These observations provide at least a partial evidence to answer to RQ3.

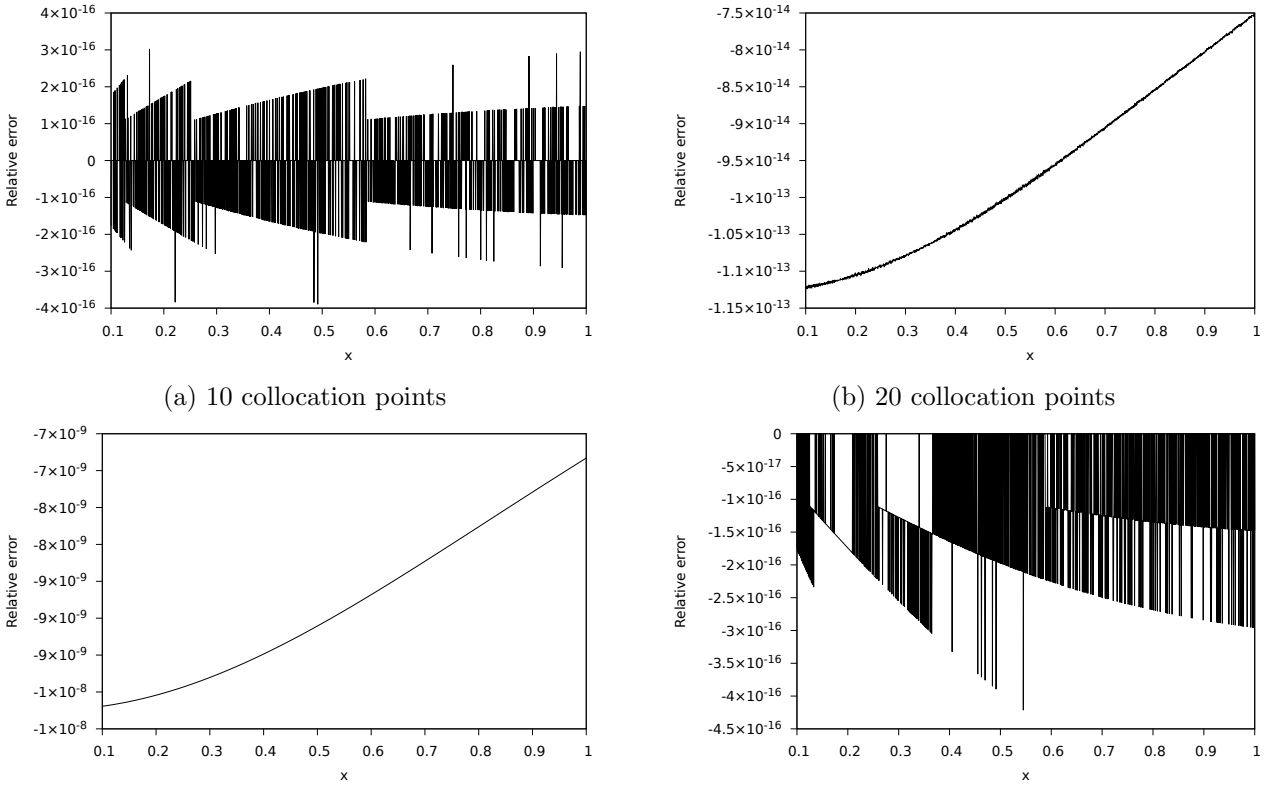
- Generally, the relative error plots mostly follow the well-known oscillatory form typically observed in function approximation.

- There is no evidence in the best-of-run plots to support the view that more collocation points leads to a more accurate (best) approximation. In many cases, fewer collocation points produces a smaller peak relative error magnitude.

- The peak relative error magnitudes are generally very small, with some values down to the scale of 10^{-11}, and generally below the 10^{-6} level we have taken as indicating effective 'equivalence'. See the main paper for comparisons with other techniques, such as deep neural networks.

- The three ODEs (ODE-4, ODE-7 and ODE-8) that proved problematic in terms of trivial solutions were still able to generate acceptable best solutions with the addition of the threshold constraint. The relative errors were on the scale of 10^{-8} for ODE-4 and ODE-8. ODE-7, however, exhibits 'spiking' behavior in the relative error plots although this, as we have pointed out, is largely a consequence of the true function attaining very small values and, since it is the denominator in the calculation of relative error, can lead to a large value of relative error. More generally, the absolute errors are on the scale of 10^{-4} to 10^{-5}.

7.1 ODE-1



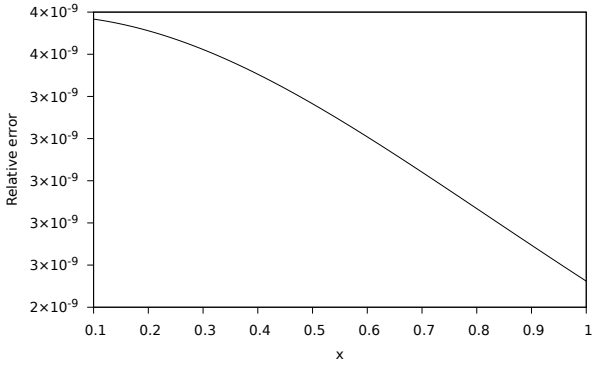
(a) 10 collocation points

(b) 20 collocation points

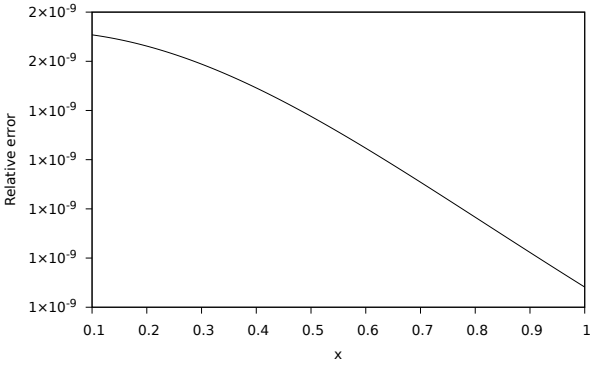
(c) 50 collocation points

(d) 100 collocation points

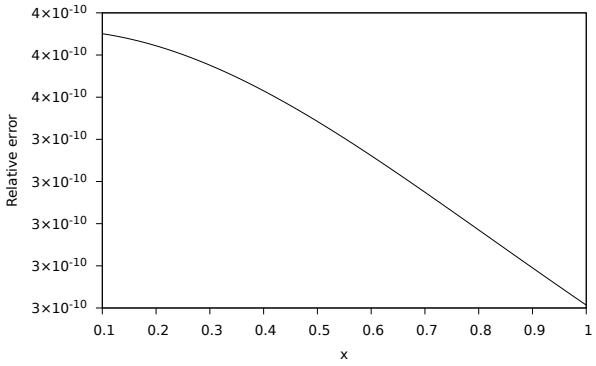
Figure S1: Best results for ODE-1 (hc).



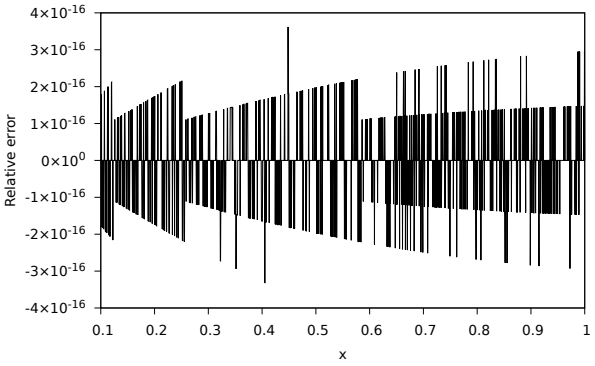
(a) 10 collocation points



(b) 20 collocation points

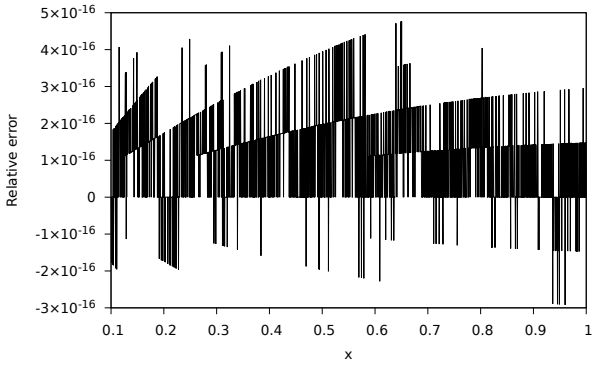


(c) 50 collocation points

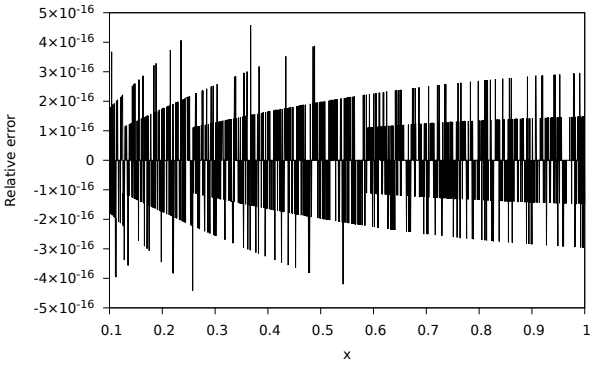


(d) 100 collocation points

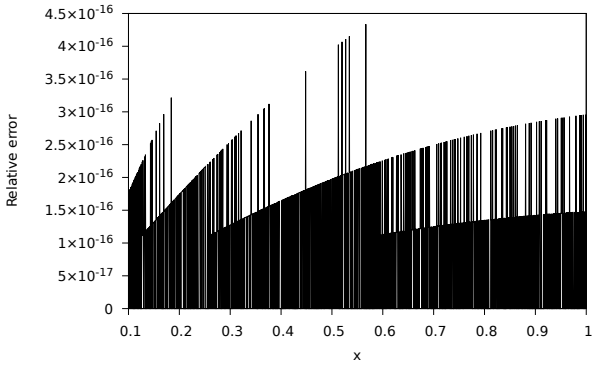
Figure S2: Best results for ODE-1 with death penalty (dp).



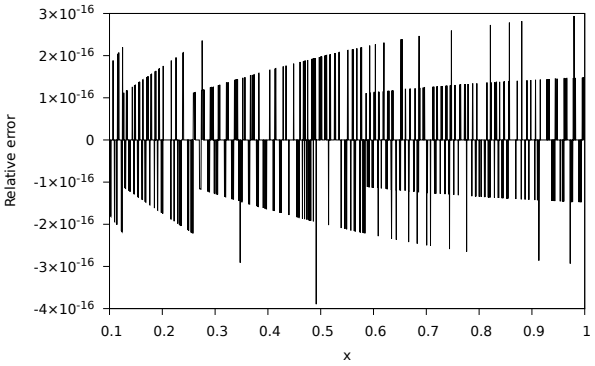
(a) 10 collocation points



(b) 20 collocation points



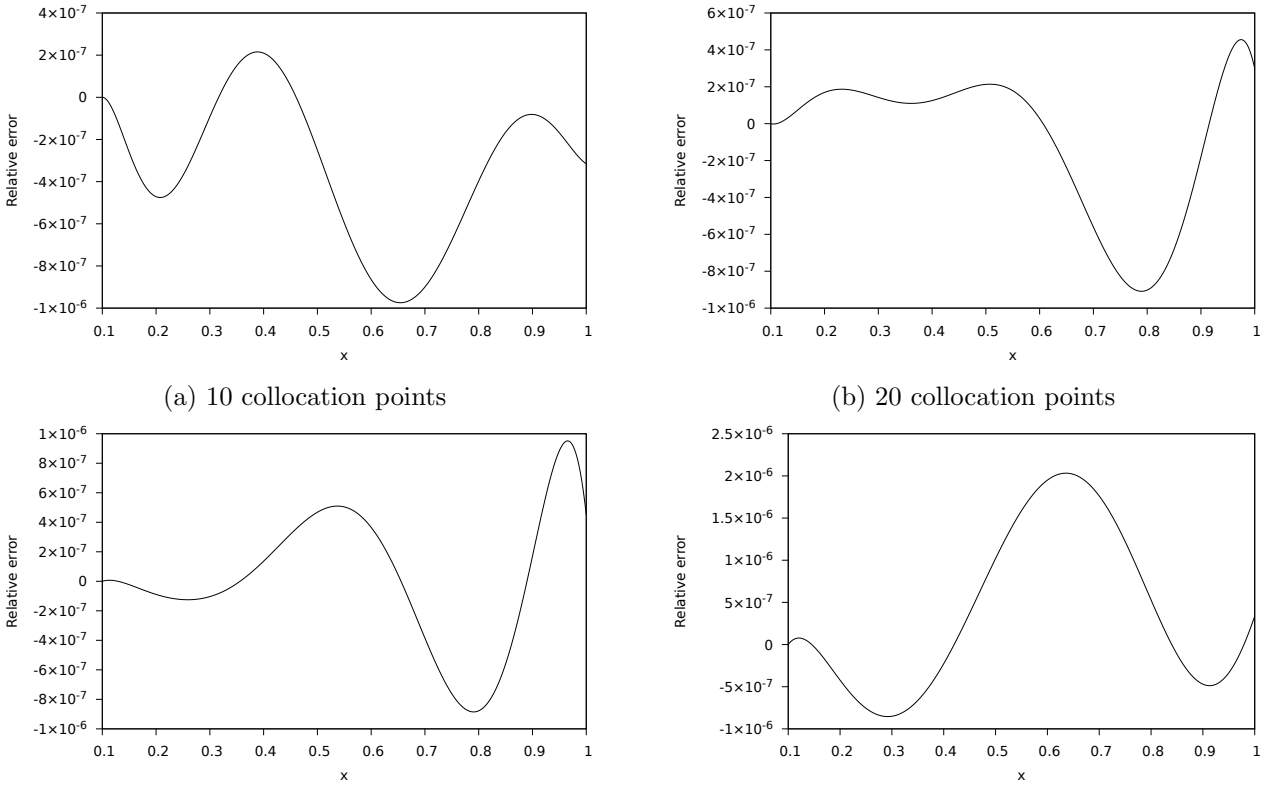
(c) 50 collocation points



(d) 100 collocation points

Figure S3: Best results for ODE-1 with the penalty method (pm).

7.2 ODE-2



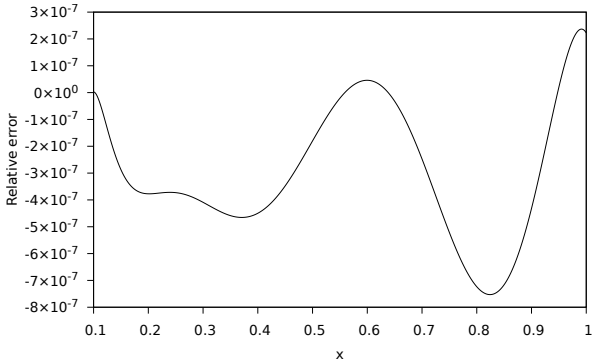
(a) 10 collocation points

(b) 20 collocation points

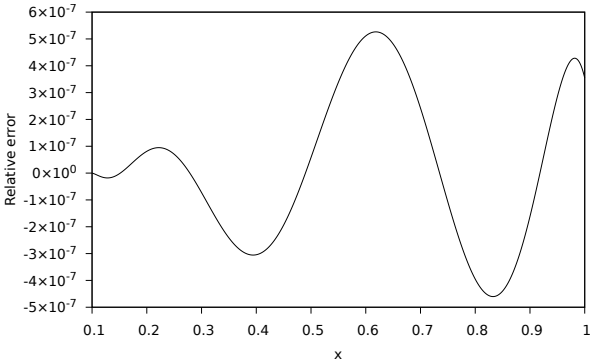
(c) 50 collocation points

(d) 100 collocation points

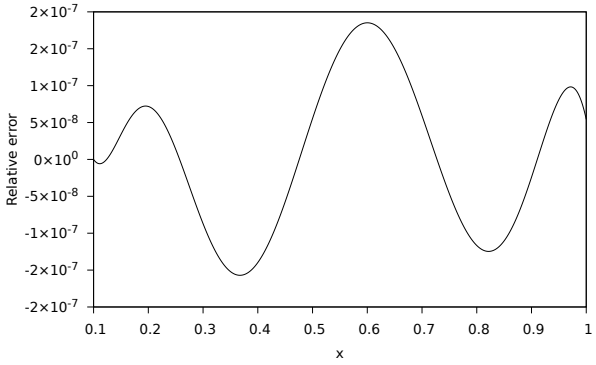
Figure S4: Best results for ODE-2 (hc).



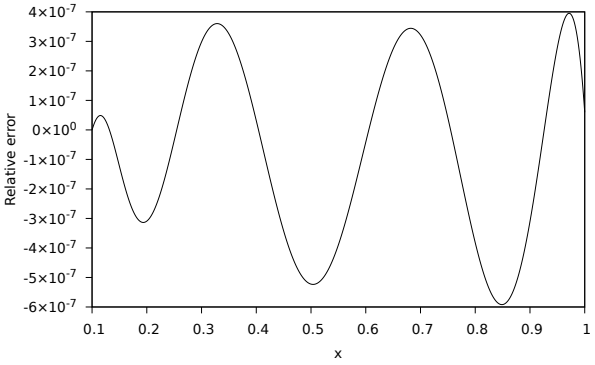
(a) 10 collocation points



(b) 20 collocation points

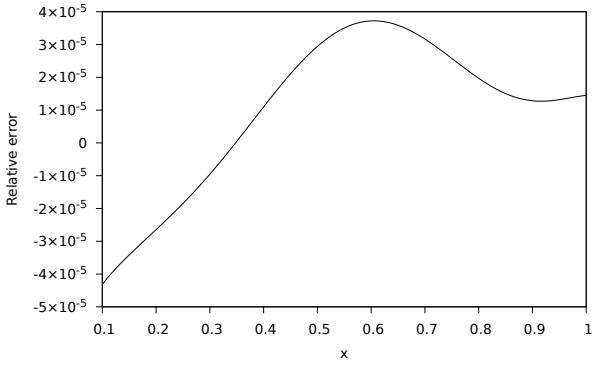


(c) 50 collocation points

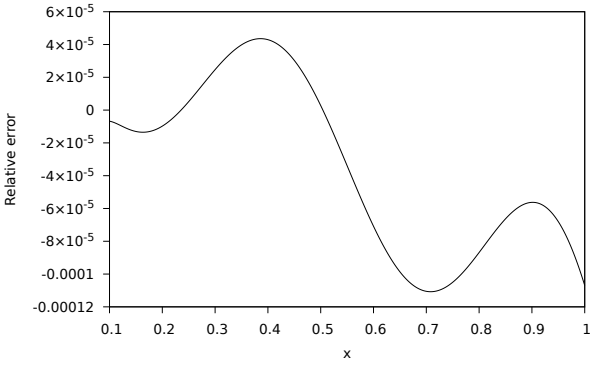


(d) 100 collocation points

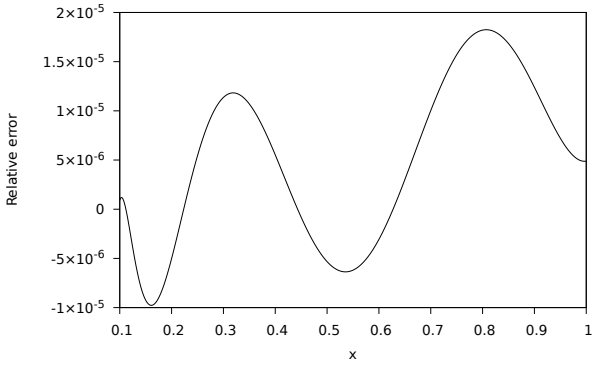
Figure S5: Best results for ODE-2 with death penalty (dp).



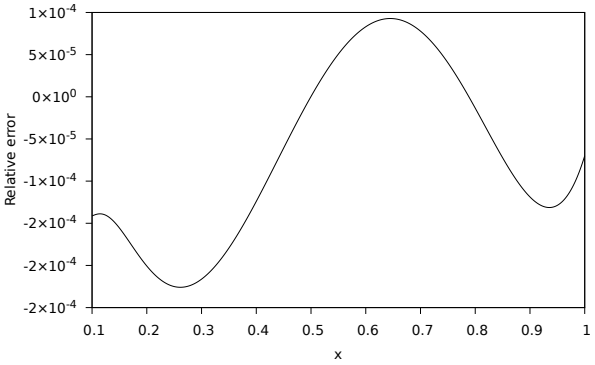
(a) 10 collocation points



(b) 20 collocation points



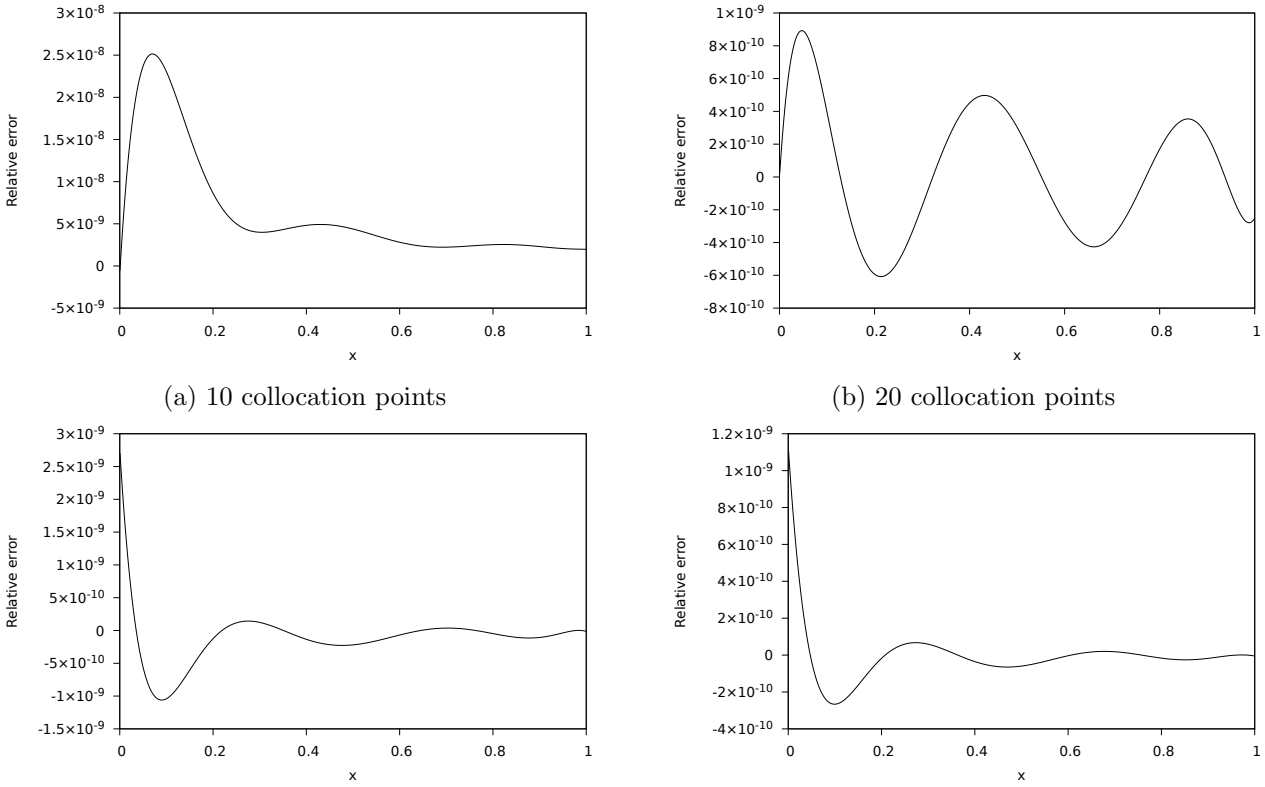
(c) 50 collocation points



(d) 100 collocation points

Figure S6: Best results for ODE-2 with the penalty method (pm).

7.3 ODE-3



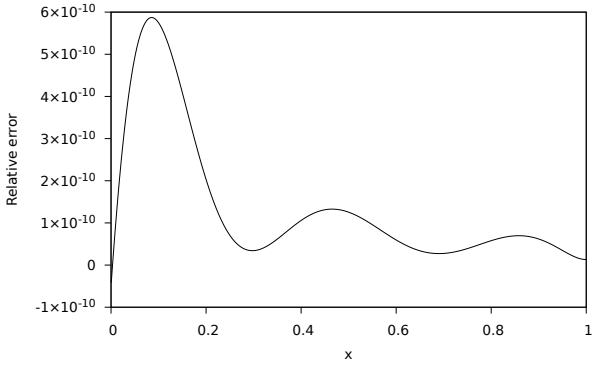
(a) 10 collocation points

(b) 20 collocation points

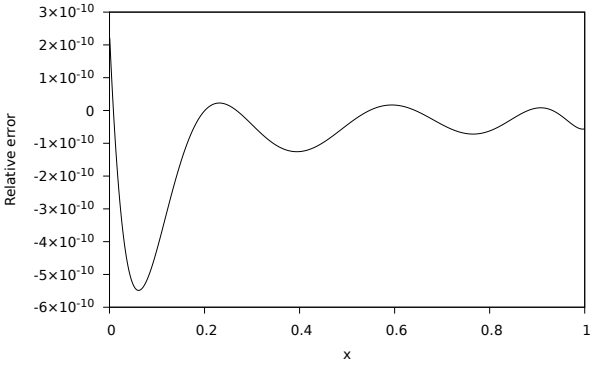
(c) 50 collocation points

(d) 100 collocation points

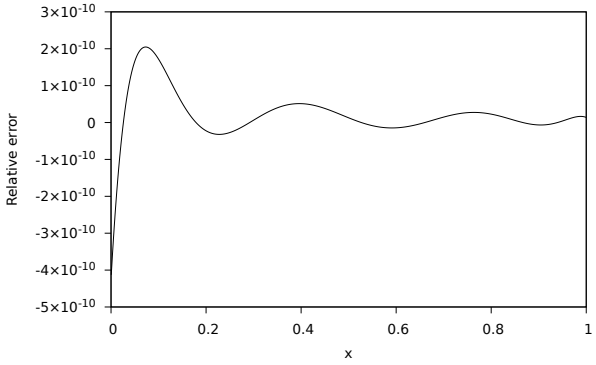
Figure S7: Best results for ODE-3 (hc).



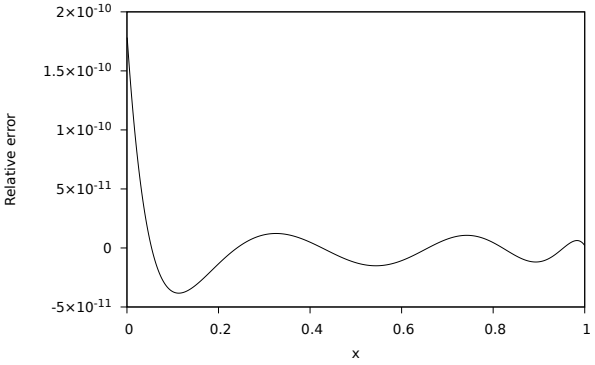
(a) 10 collocation points



(b) 20 collocation points

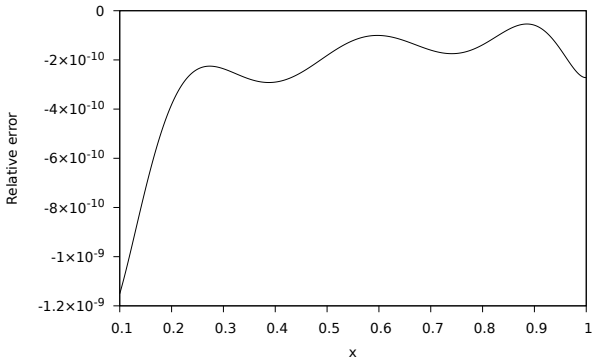


(c) 50 collocation points

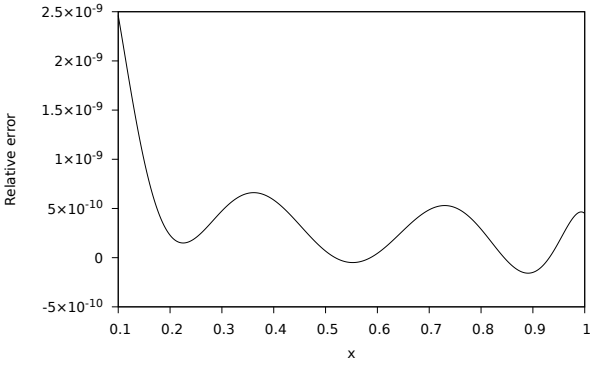


(d) 100 collocation points

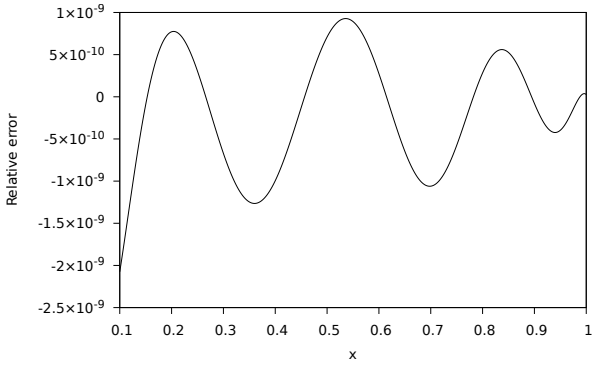
Figure S8: Best results for ODE-3 with death penalty (dp).



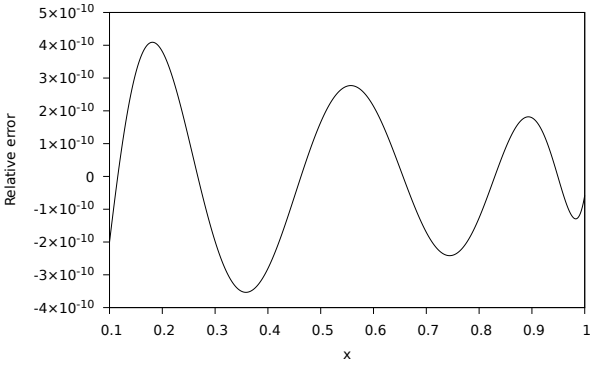
(a) 10 collocation points



(b) 20 collocation points



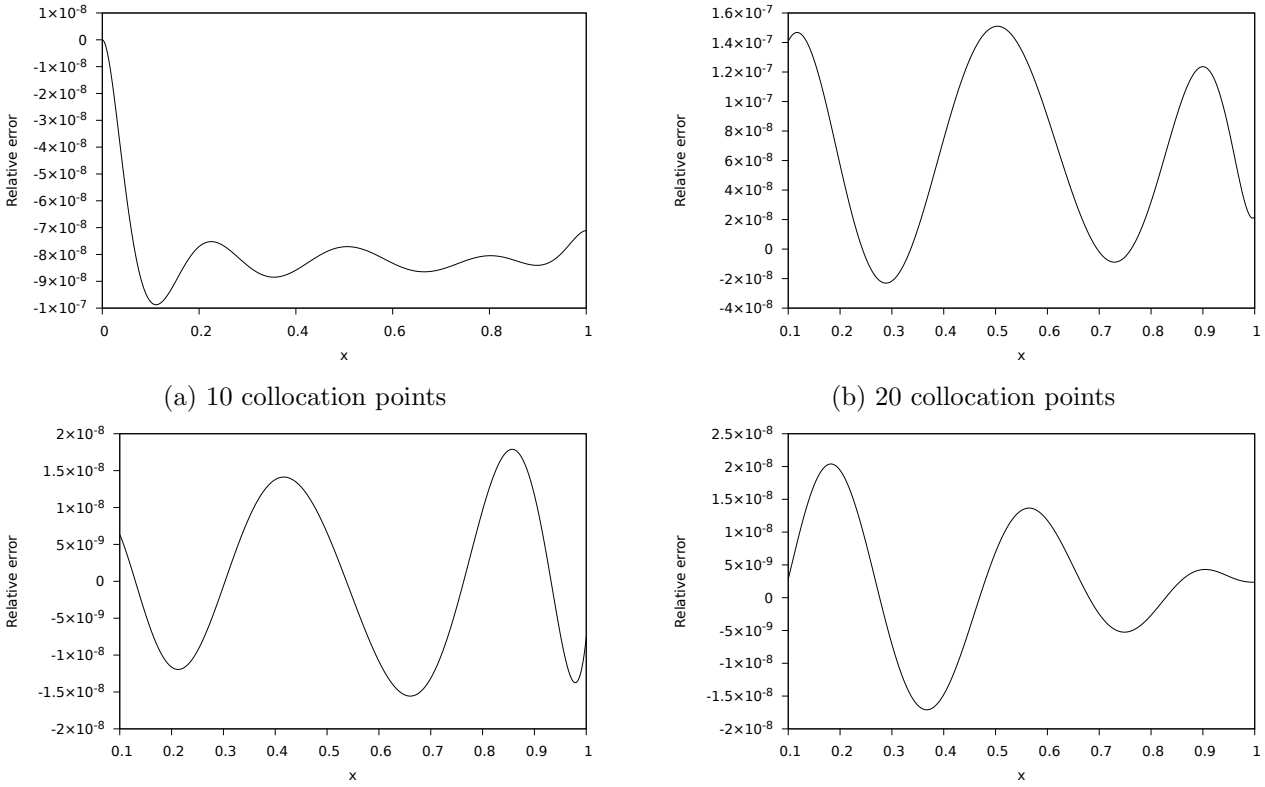
(c) 50 collocation points



(d) 100 collocation points

Figure S9: Best results for ODE-3 with the penalty method (pm).

7.4 ODE-4



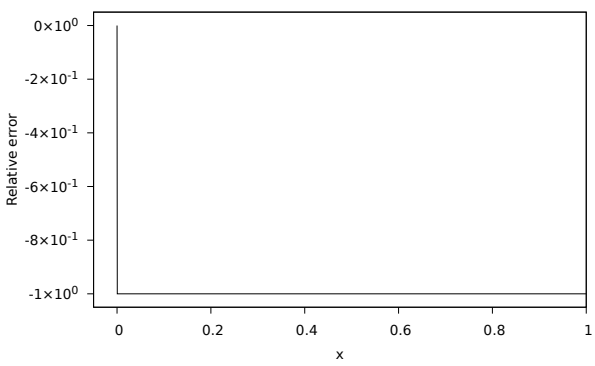
(a) 10 collocation points

(b) 20 collocation points

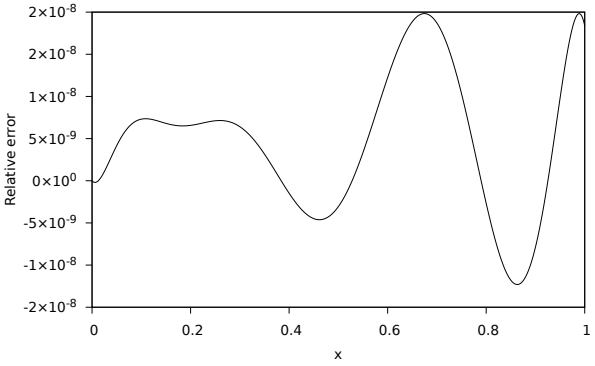
(c) 50 collocation points

(d) 100 collocation points

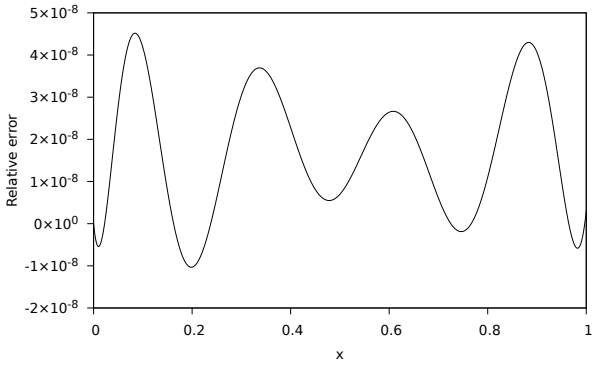
Figure S10: Best results for ODE-4 (hc).



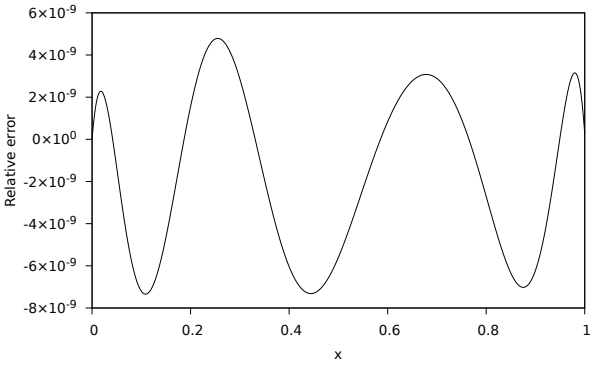
(a) 10 collocation points



(b) 20 collocation points

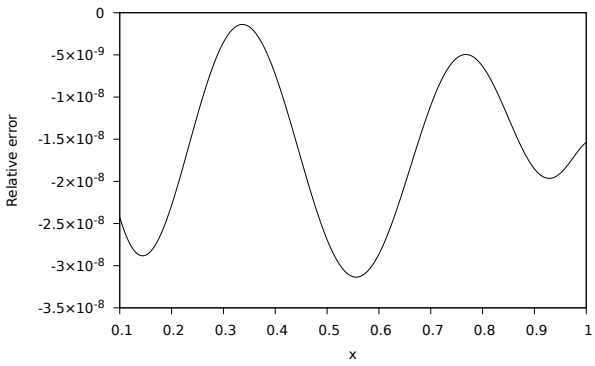


(c) 50 collocation points

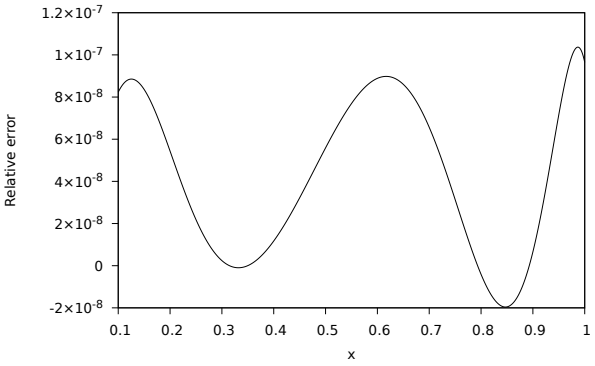


(d) 100 collocation points

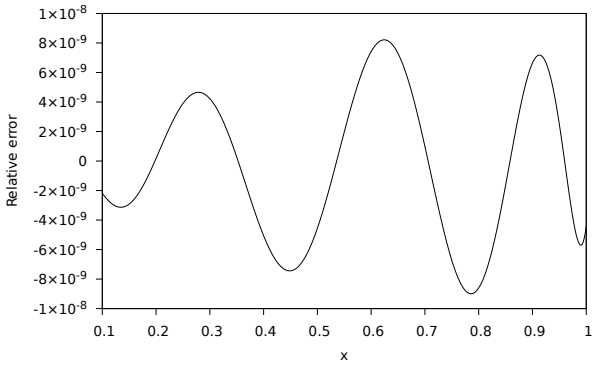
Figure S11: Best results for ODE-4 with death penalty (dp).



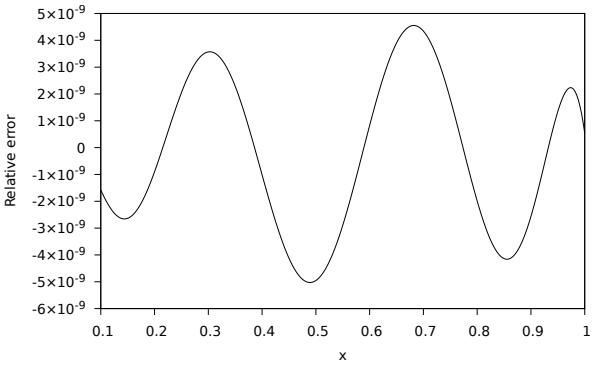
(a) 10 collocation points



(b) 20 collocation points



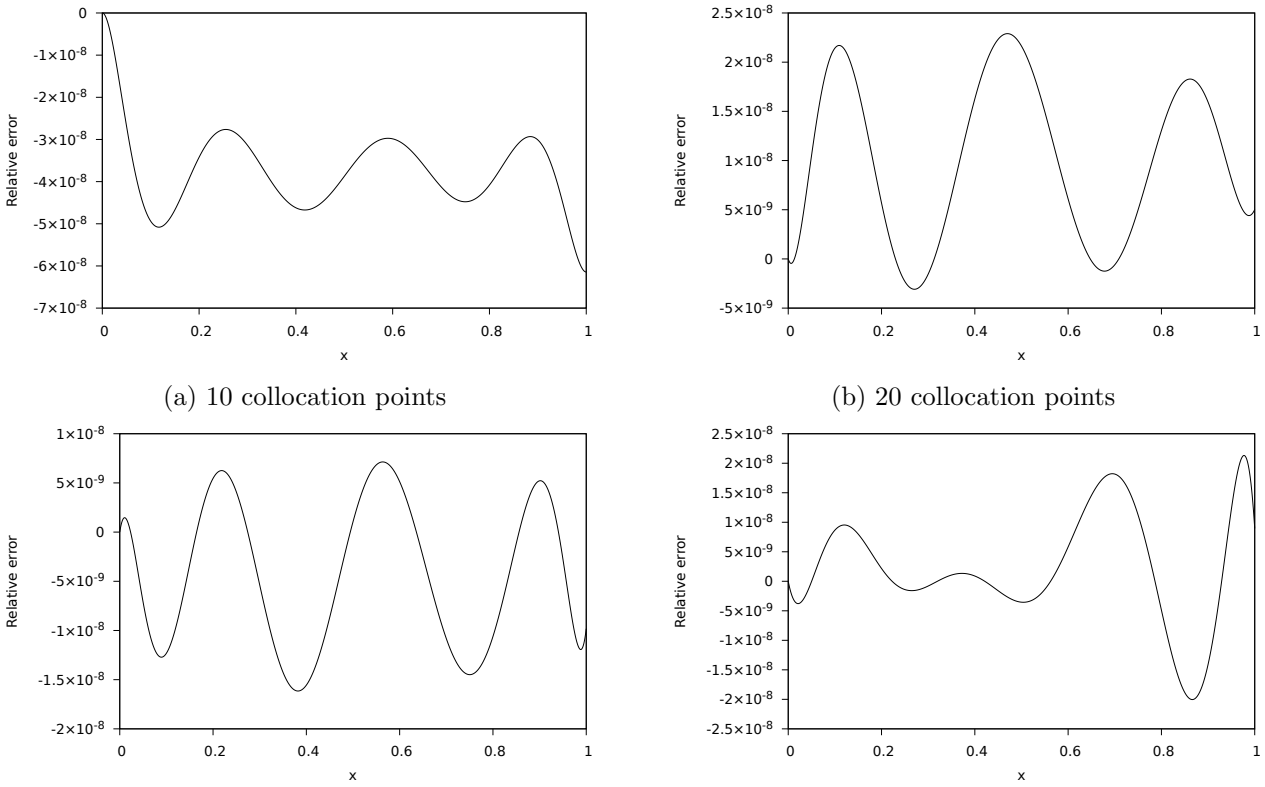
(c) 50 collocation points



(d) 100 collocation points

Figure S12: Best results for ODE-4 with the penalty method (pm).

7.5 ODE-4 + threshold



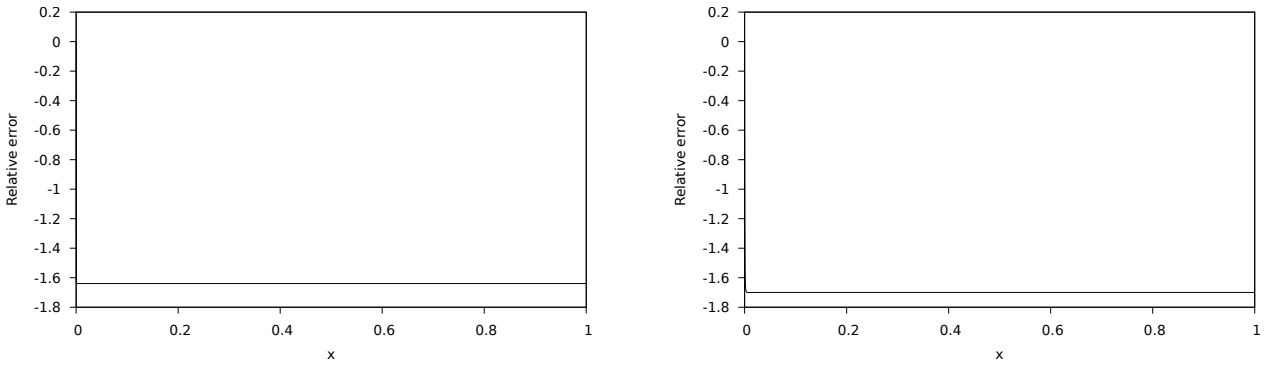
(a) 10 collocation points

(b) 20 collocation points

(c) 50 collocation points

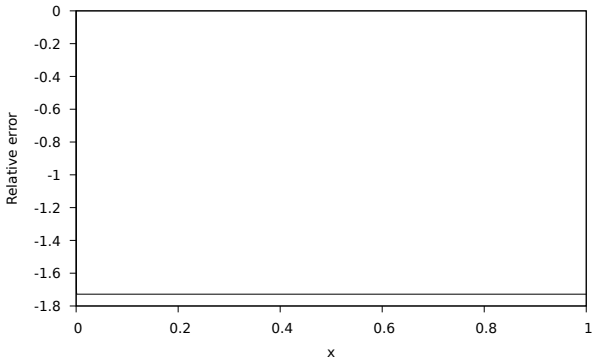
(d) 100 collocation points

Figure S13: Best results for ODE-4_threshold (hc).

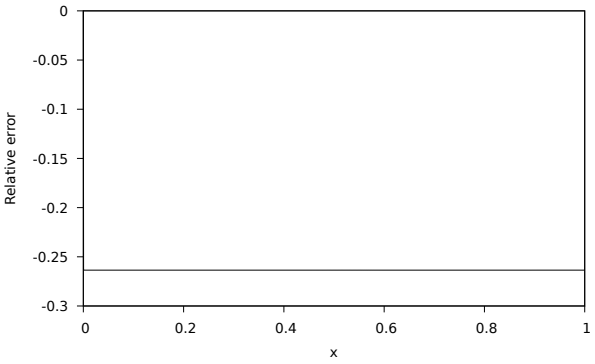


(a) 10 collocation points

(b) 20 collocation points

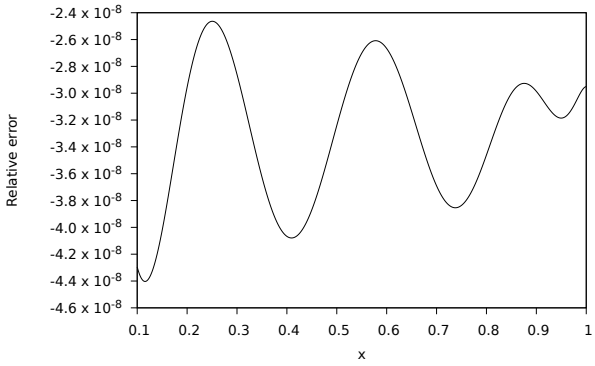


(c) 50 collocation points

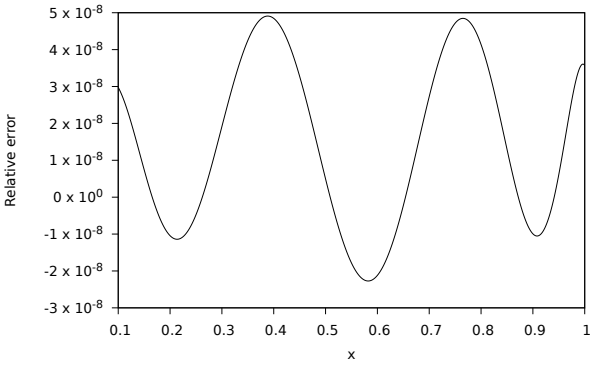


(d) 100 collocation points

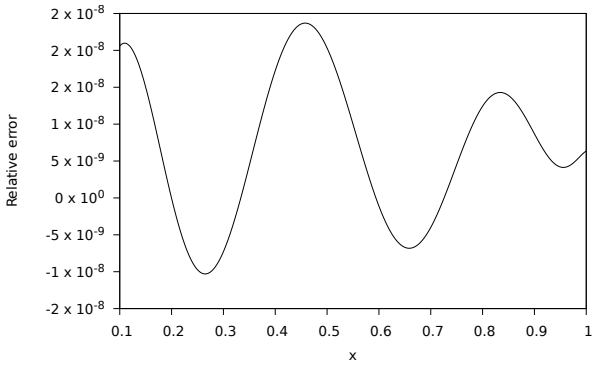
Figure S14: Best results for ODE-4_threshold with death penalty (dp).



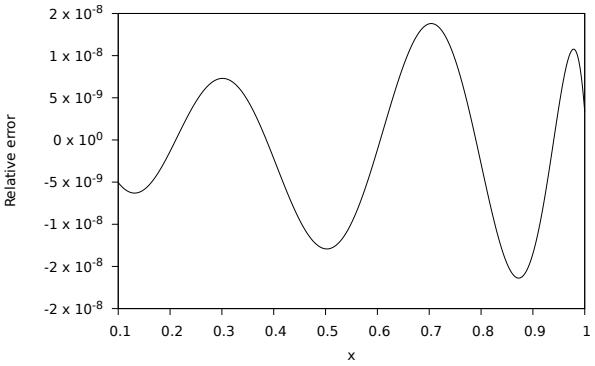
(a) 10 collocation points



(b) 20 collocation points



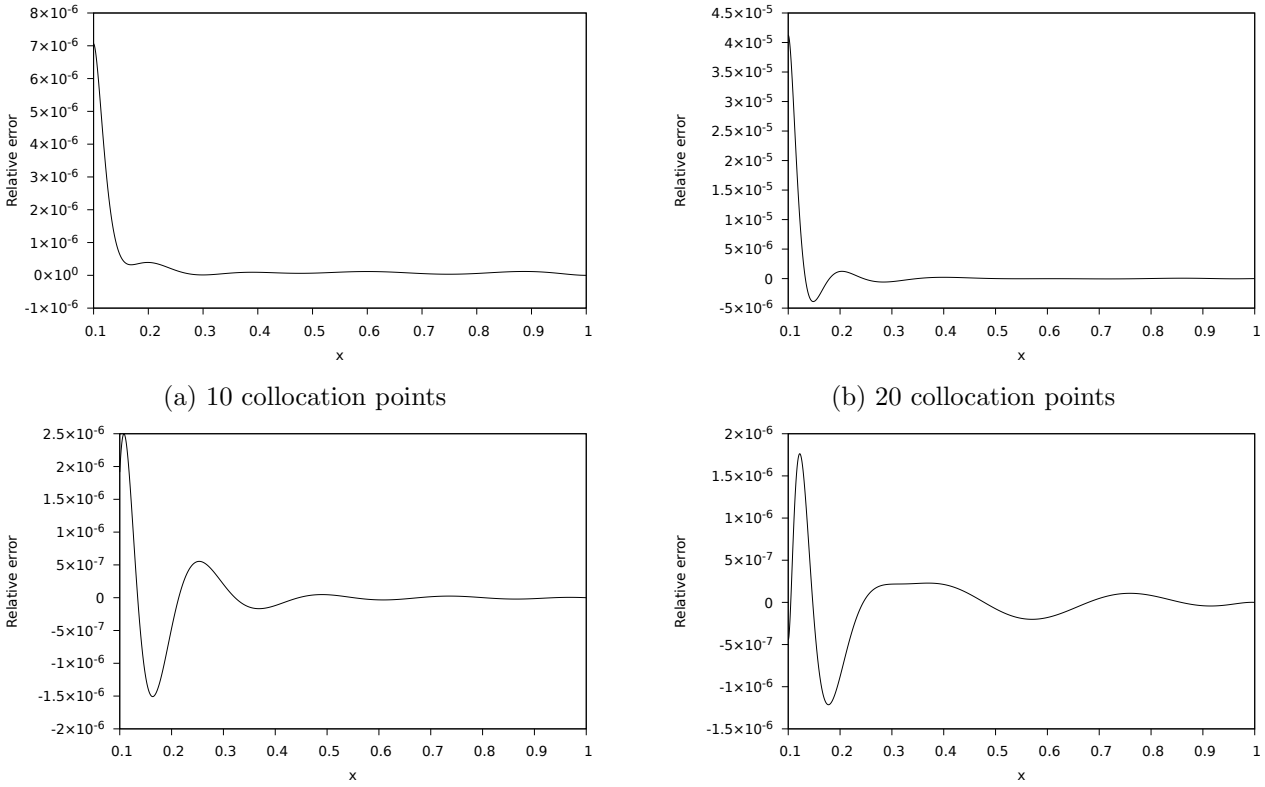
(c) 50 collocation points



(d) 100 collocation points

Figure S15: Best results for ODE-4_threshold with the penalty method (pm).

7.6 ODE-5



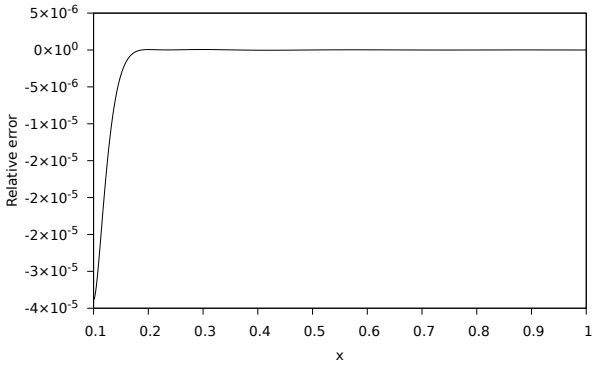
(a) 10 collocation points

(b) 20 collocation points

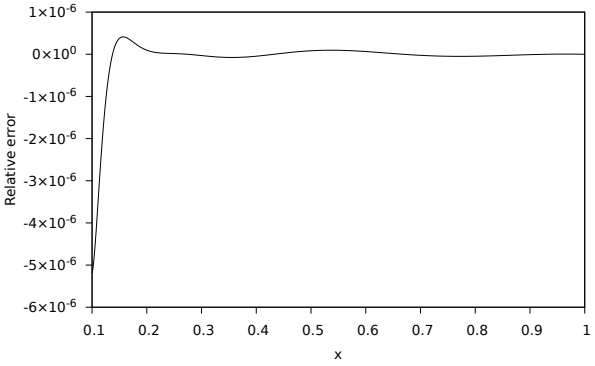
(c) 50 collocation points

(d) 100 collocation points

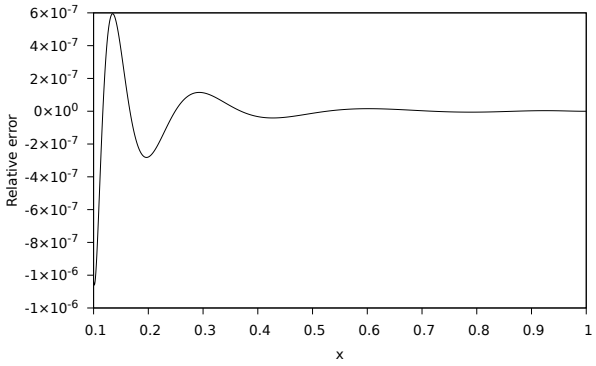
Figure S16: Best results for ODE-5 (hc).



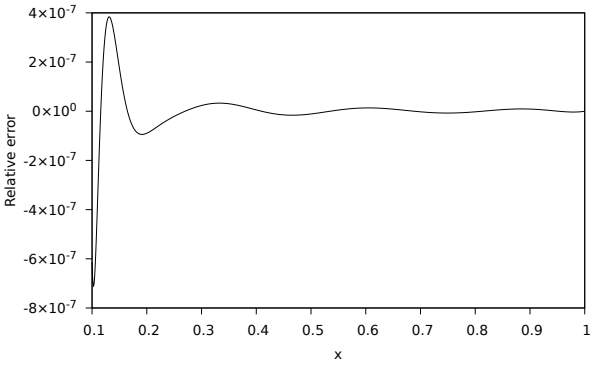
(a) 10 collocation points



(b) 20 collocation points

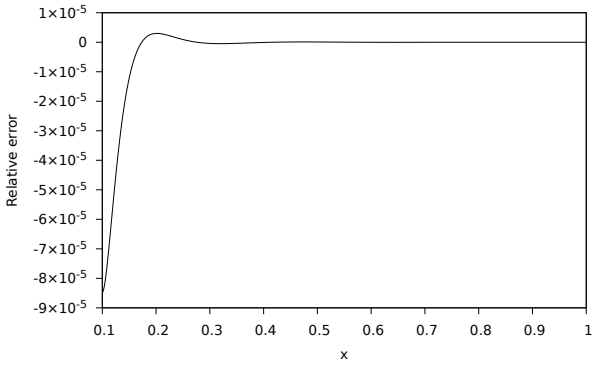


(c) 50 collocation points

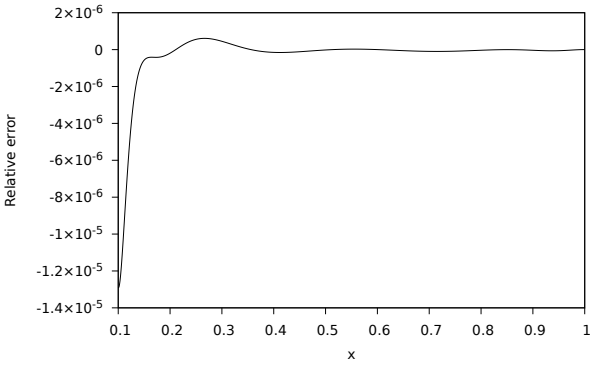


(d) 100 collocation points

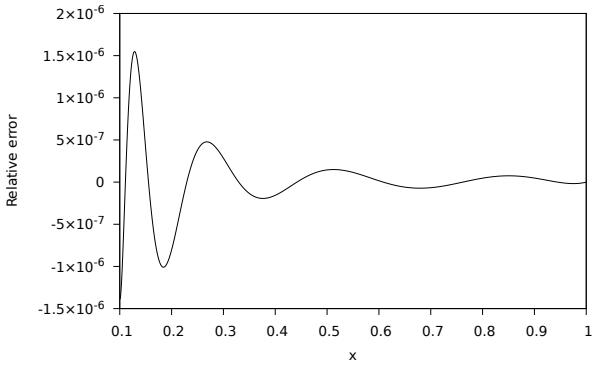
Figure S17: Best results for ODE-5 with death penalty (dp).



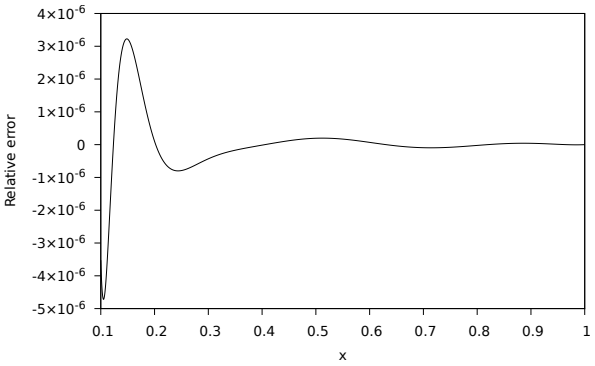
(a) 10 collocation points



(b) 20 collocation points



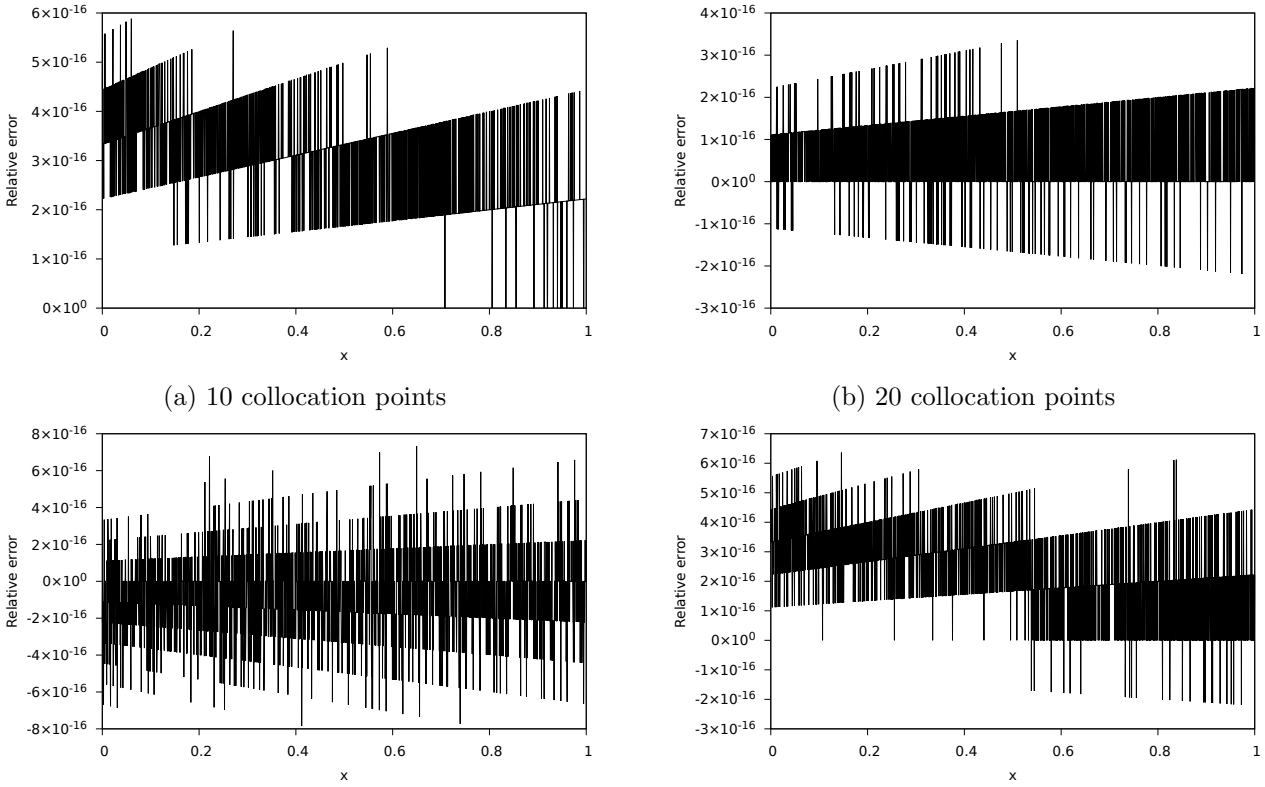
(c) 50 collocation points



(d) 100 collocation points

Figure S18: Best results for ODE-5 with the penalty method (pm).

7.7 ODE-6



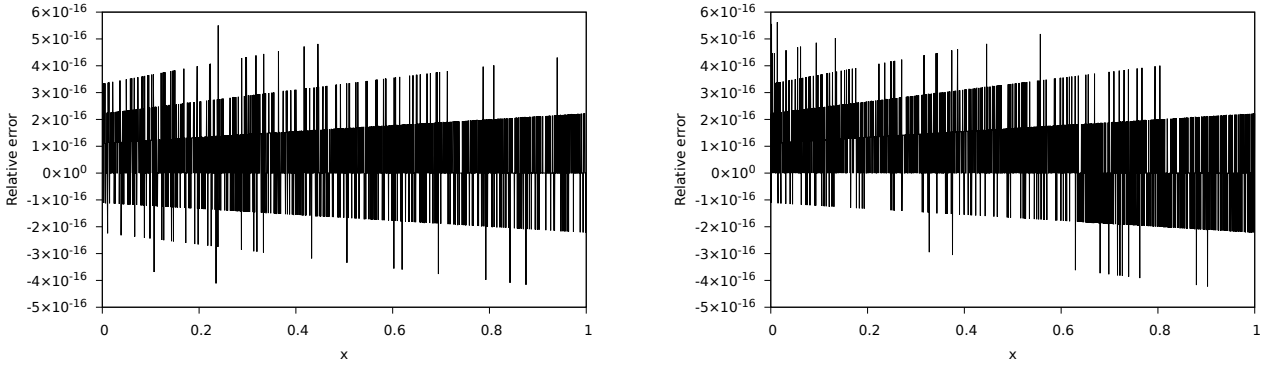
(a) 10 collocation points

(b) 20 collocation points

(c) 50 collocation points

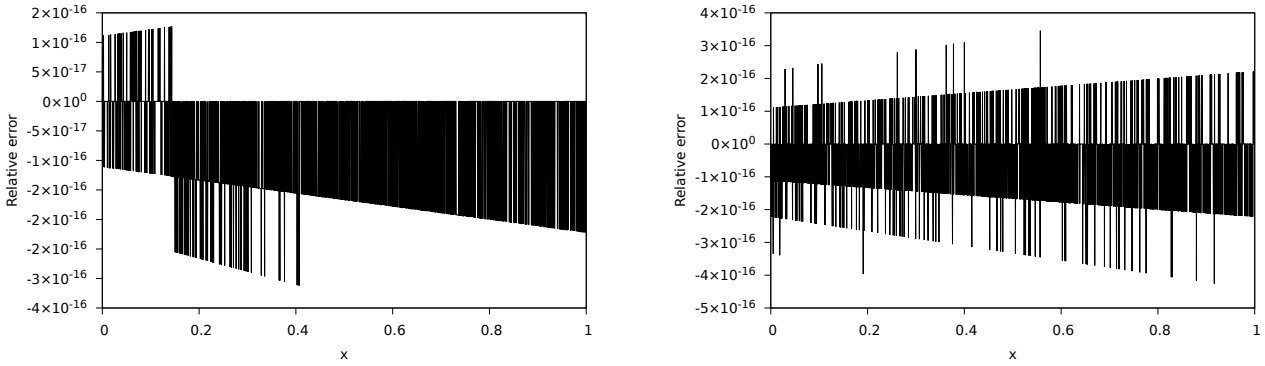
(d) 100 collocation points

Figure S19: Best results for ODE-6 (hc).



(a) 10 collocation points

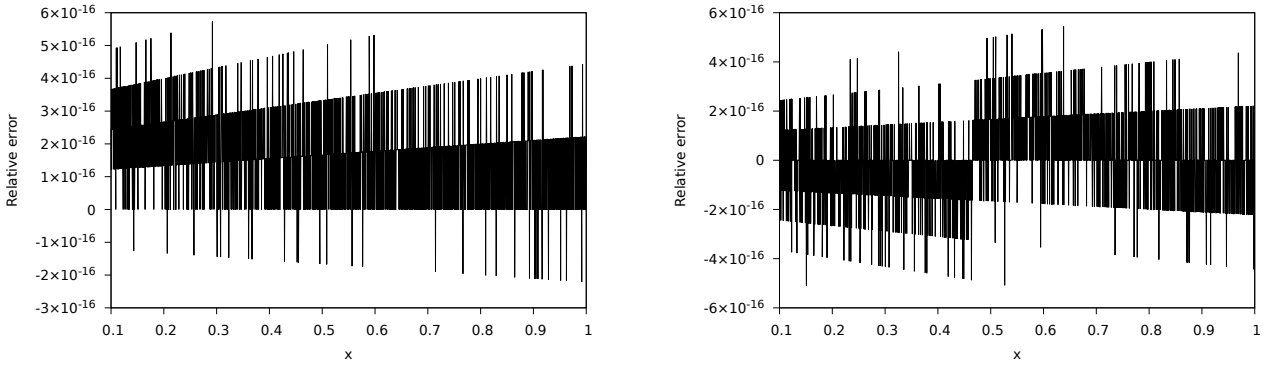
(b) 20 collocation points



(c) 50 collocation points

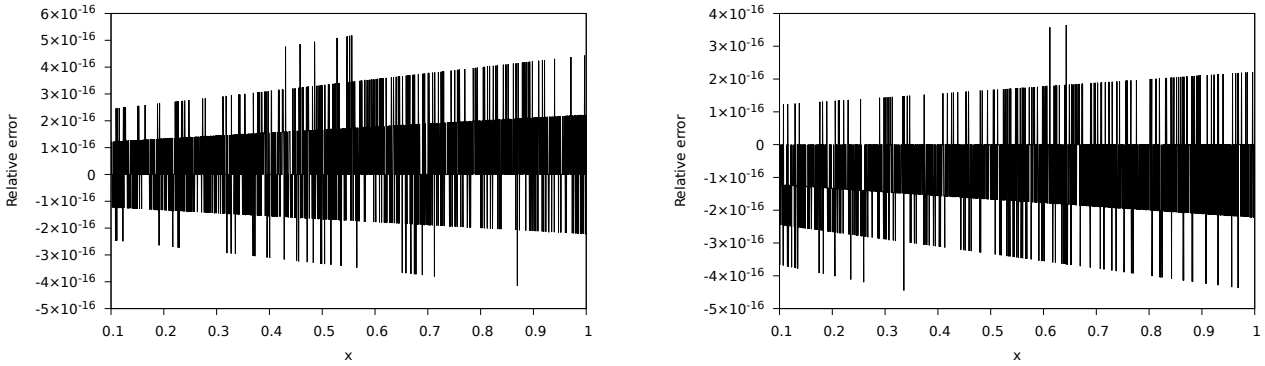
(d) 100 collocation points

Figure S20: Best results for ODE-6 with death penalty (dp).



(a) 10 collocation points

(b) 20 collocation points



(c) 50 collocation points

(d) 100 collocation points

Figure S21: Best results for ODE-6 with the penalty method (pm).

7.8 ODE-7

The results for all 30 runs for ODE-7 with both hard constraints, death penalty and the penalty method all produced trivial solutions regardless of the numbers of collocation points. These plots are therefore omitted.

7.9 ODE-7 + threshold

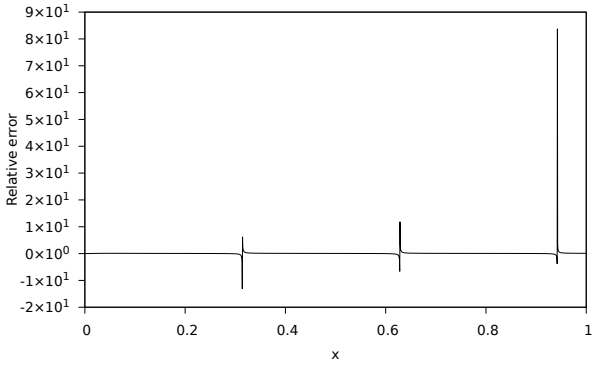
The situation with ODE-7 with a threshold is rather more complicated. The numbers of successful (*i.e.* non-trivial) solutions is shown in Table S75. It is clear that without the addition of the extra threshold constraint, the success rate is very poor whereas either of the hard-constraint or death-penalty methods produces 30 successful runs. Notably, the penalty method has zero success rate without the threshold, and only a very modest success rate with the threshold.

Table S75: Numbers of successful (i.e. non-trivial) solutions out of 30 runs for ODE-7. '#pts' = number of collocation points; '**hc**' = hard constraints; '**dp**' = death penalty; 't' = threshold; '**pm**' = penalty method.

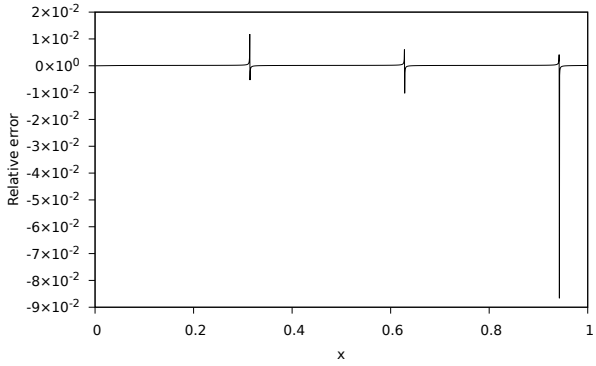
#pts	**hc**	**hc+t**	**dp**	**dp+t**	**pm**	**pm+t**
10	0	30	0	30	0	0
20	0	30	0	30	0	1
50	0	30	0	30	0	1
100	0	30	0	30	0	6

The plots of relative error for ODE-7 with hard and threshold constraints are shown in Figure S22 from which a very clear 'spiking' behavior can be observed; the relative errors in the notionally 'zero' sections of the plots are mostly around 5×10^{-5}. To take the most conspicuous spike at $x = 0.9425$ from Figure S22d (for 100 collocation points), the absolute error is 2.13×10^{-6} whereas the true function value is 2.22×10^{-4} leading to a relative error of 9.59×10^{-3}. In fact, the range of *absolute* errors over the domain is $[-8.55 \times 10^{-5}, 8.47 \times 10^{-5}]$, a fairly small span. The spiking behavior thus seems to be a consequence of the small denominator values used for calculating the relative error rather than the function approximation itself. This would seem to highlight the shortcoming of using relative error to assess the approximation of a function that can have very small absolute values.

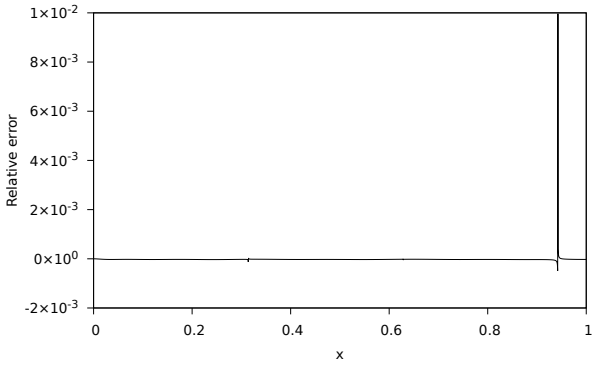
Since it is clear from Table S75 that the penalty method is rather unsuccessful with ODE-7, we omit the plots of the best-of-run errors for this method.



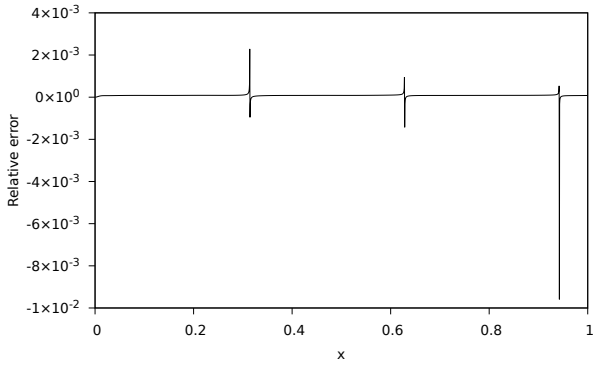
(a) 10 collocation points



(b) 20 collocation points

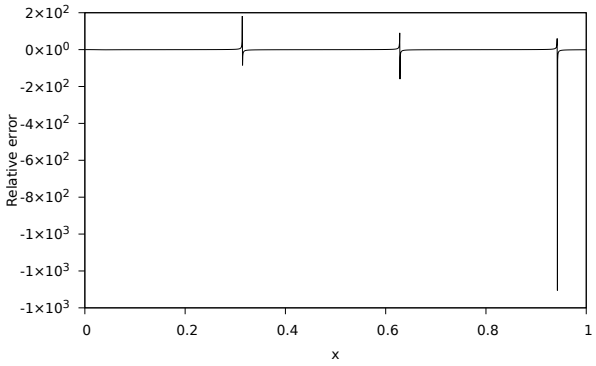


(c) 50 collocation points

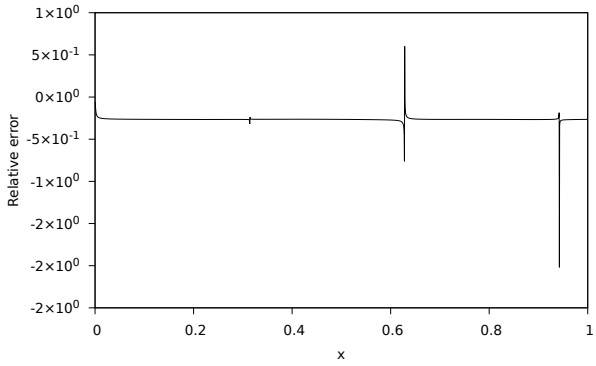


(d) 100 collocation points

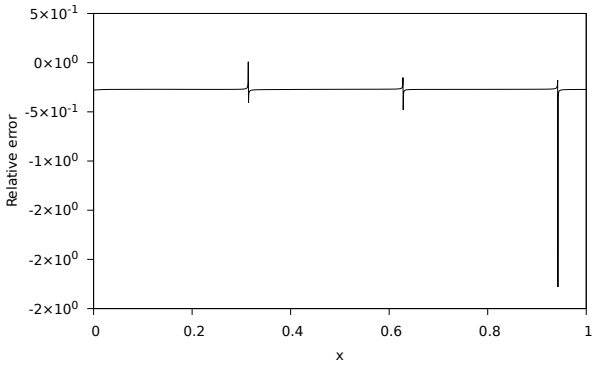
Figure S22: Best results for ODE-7_threshold (hc).



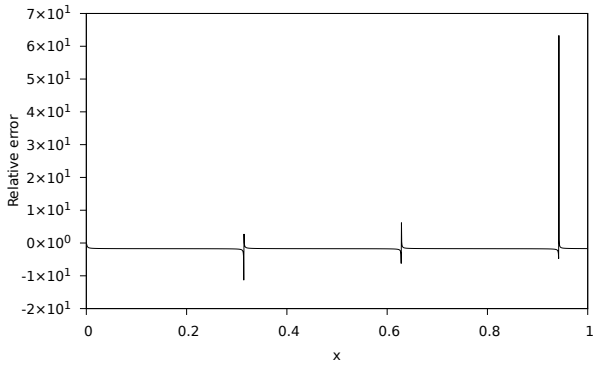
(a) 10 collocation points



(b) 20 collocation points



(c) 50 collocation points



(d) 100 collocation points

Figure S23: Best results for ODE-7_threshold with death penalty (dp).

7.10 ODE-8

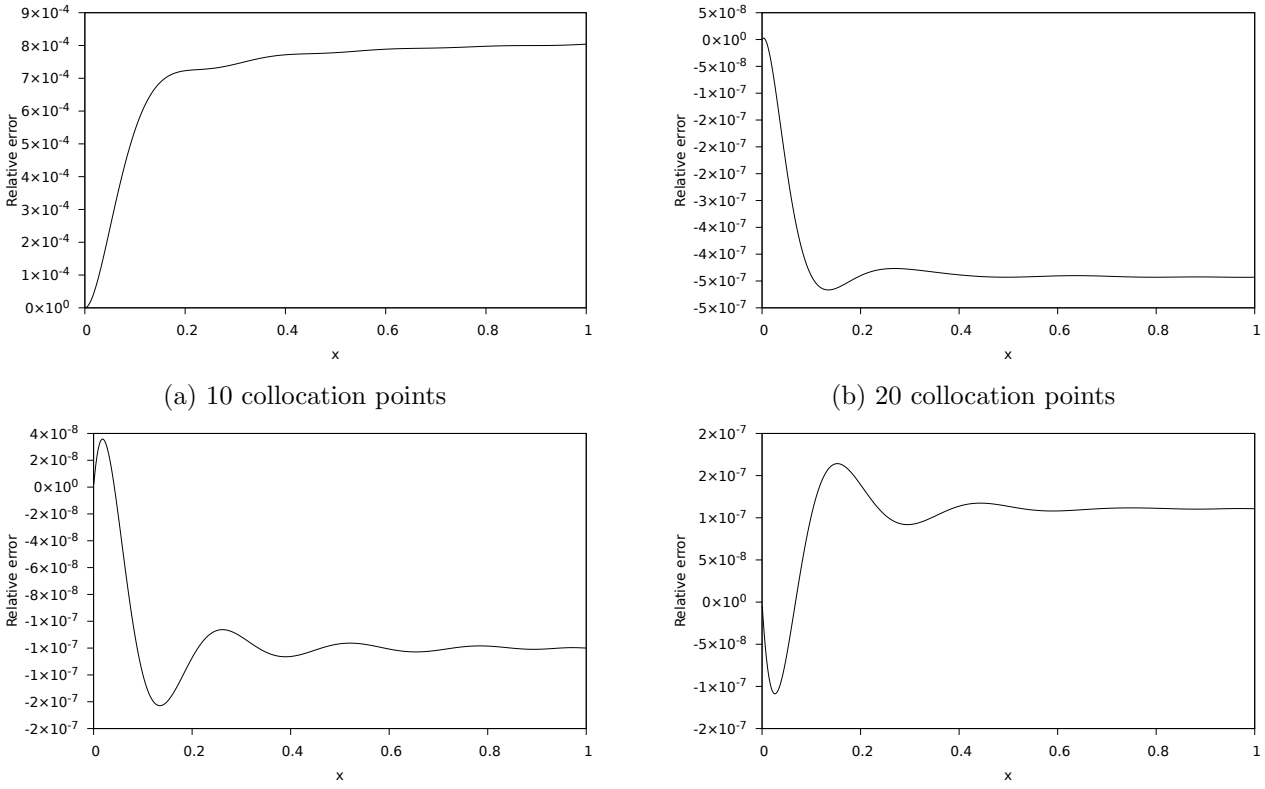
As with ODE-7 (see Section 7.8), all 30 runs for ODE-8 with both hard constraints and death penalty all produced trivial solutions so these plots are therefore omitted.

The penalty method, on the other hand had some minor success as shown in Table S76. While this approach produced some non-trivial solutions, the smallest training losses were all associated with trivial solutions—in fact, the *largest* training losses were produced by the non-trivial solutions, As these non-trivial solutions could not, in practice, be identified without knowing the true answer, we omit the plots, and conclude that all three methods are unsuccessful for ODE-8 without the threshold constraint.

Table S76: Numbers of non-trivial solutions out of 30 runs for ODE-8 with the penalty method. '#pts' = number of collocation points

#pts	**#successes**
10	0
20	0
50	4
100	9

7.11 ODE-8 + threshold



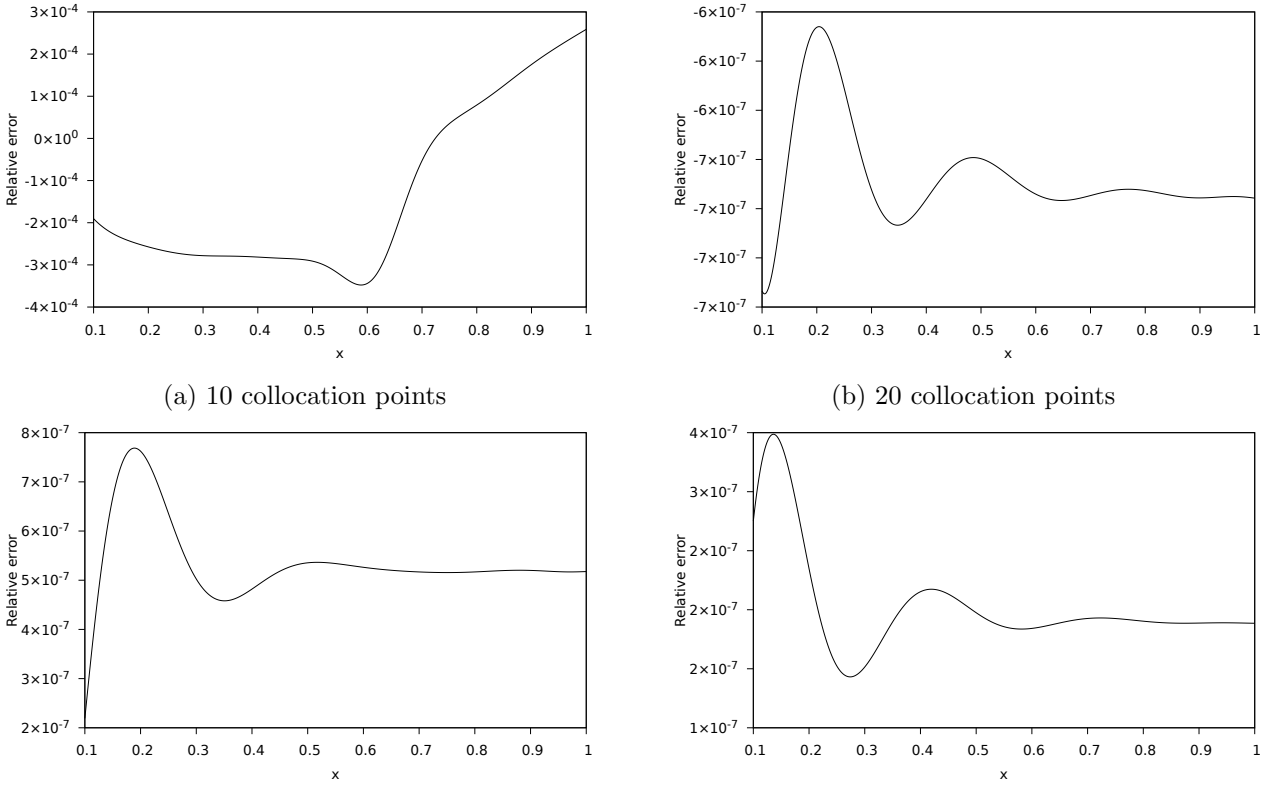
(a) 10 collocation points

(b) 20 collocation points

(c) 50 collocation points

(d) 100 collocation points

Figure S24: Best results for ODE-8_threshold (hc).



(a) 10 collocation points

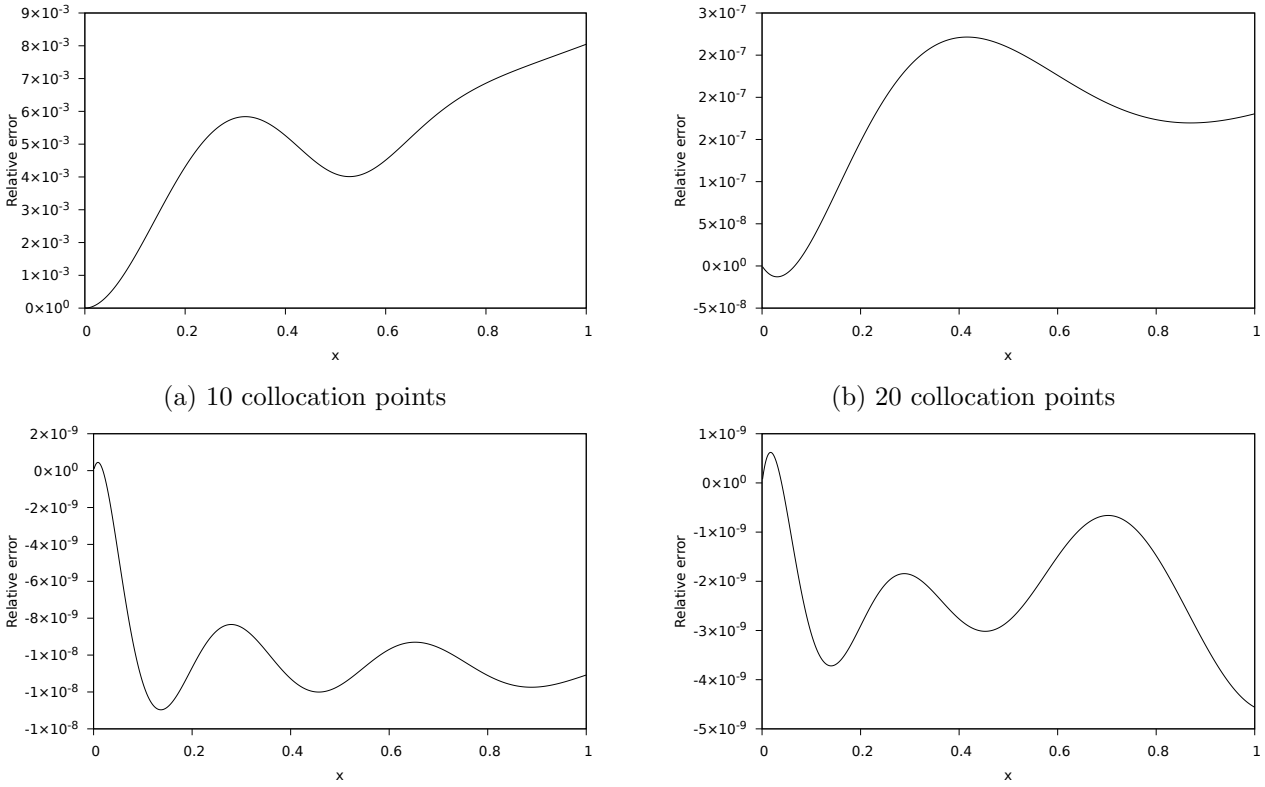
(b) 20 collocation points

(c) 50 collocation points

(d) 100 collocation points

Figure S25: Best results for ODE-8_threshold with the penalty method (pm).

7.12 ODE-9



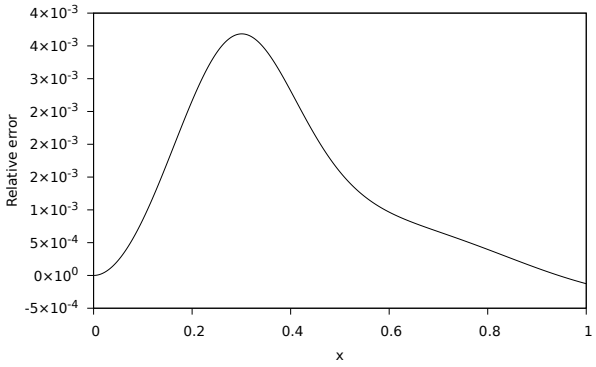
(a) 10 collocation points

(b) 20 collocation points

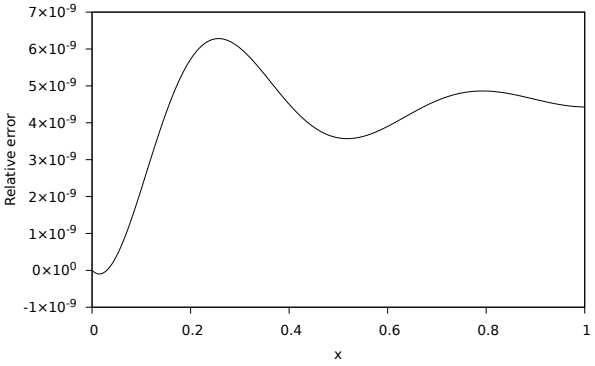
(c) 50 collocation points

(d) 100 collocation points

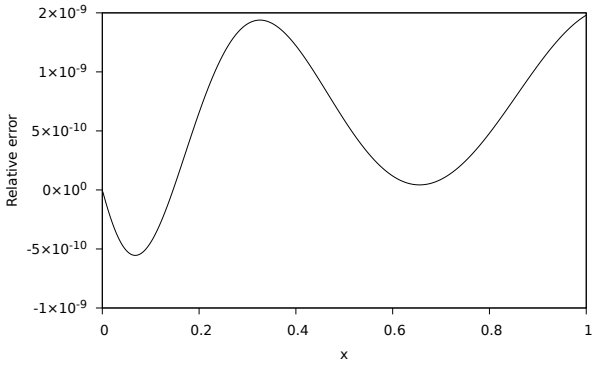
Figure S26: Best results for ODE-9 (hc).



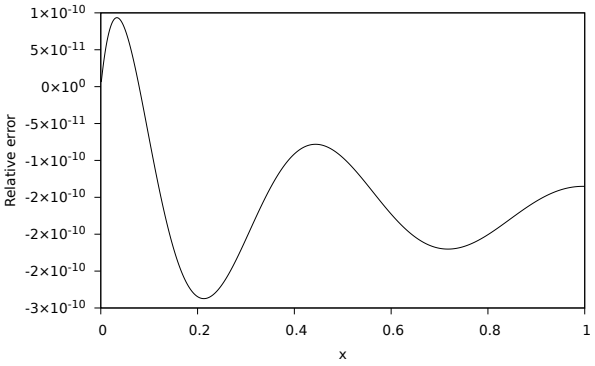
(a) 10 collocation points



(b) 20 collocation points

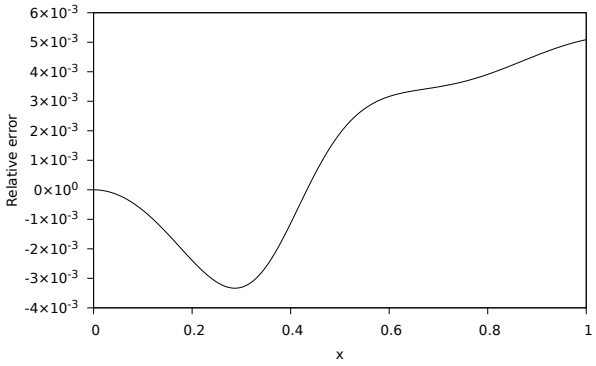


(c) 50 collocation points

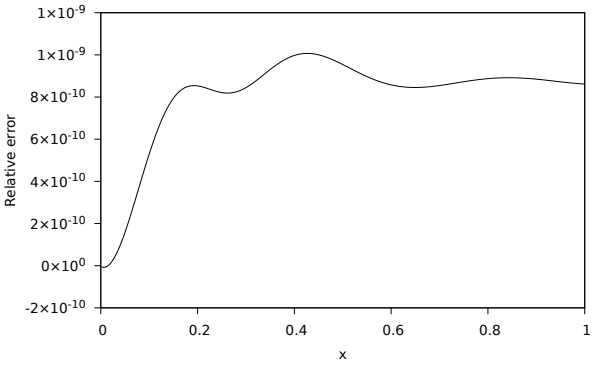


(d) 100 collocation points

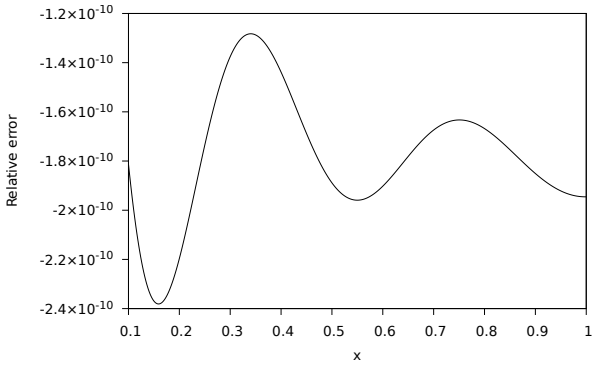
Figure S27: Best results for ODE-9 with death penalty (dp).



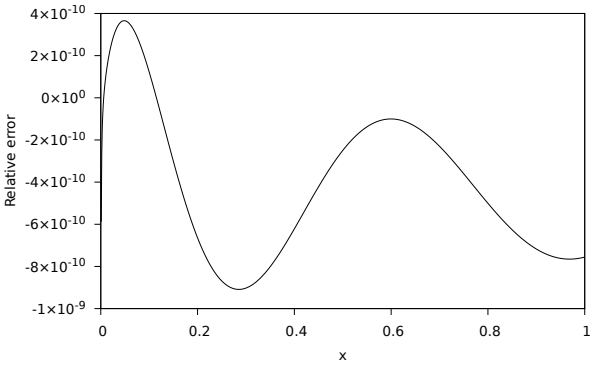
(a) 10 collocation points



(b) 20 collocation points



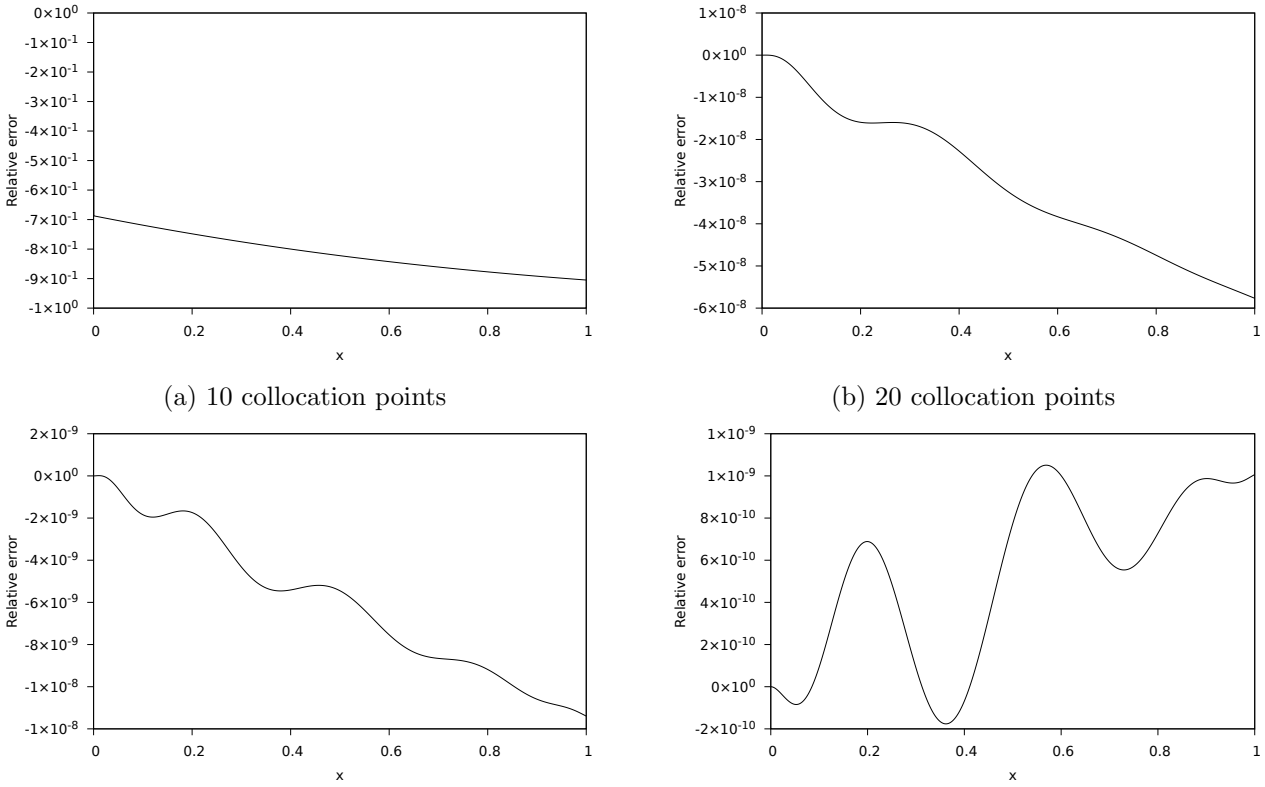
(c) 50 collocation points



(d) 100 collocation points

Figure S28: Best results for ODE-9 with the penalty method (pm).

7.13 ODE-10



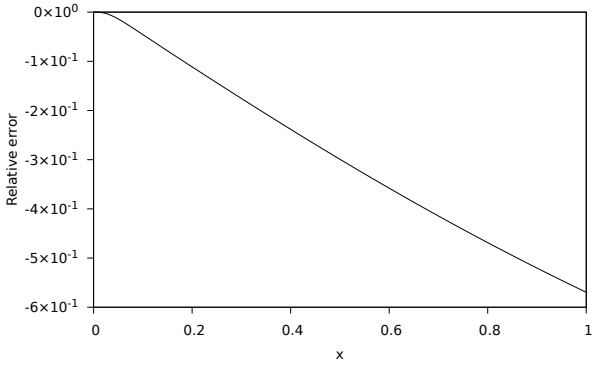
(a) 10 collocation points

(b) 20 collocation points

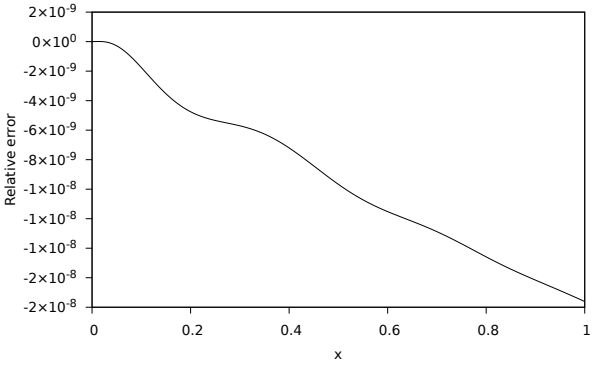
(c) 50 collocation points

(d) 100 collocation points

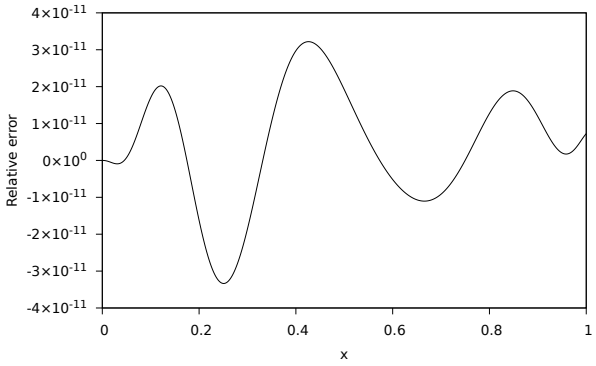
Figure S29: Best results for ODE-10 (hc).



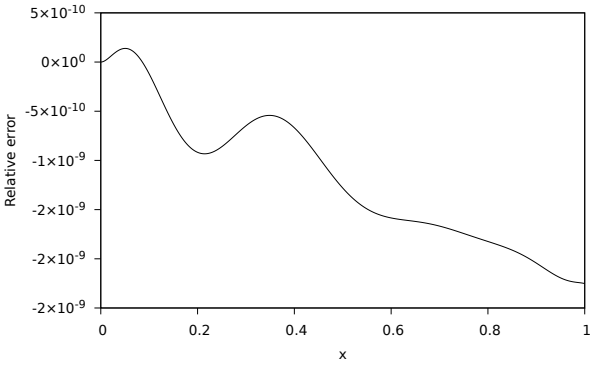
(a) 10 collocation points



(b) 20 collocation points

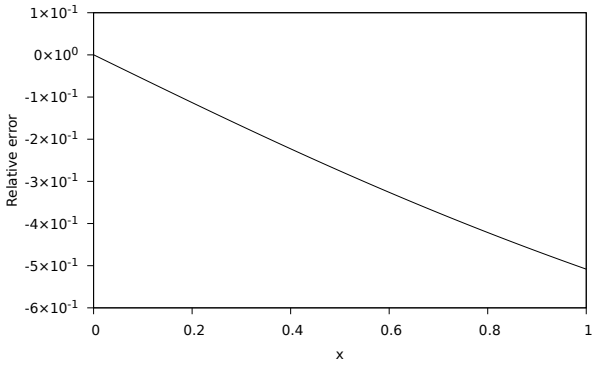


(c) 50 collocation points

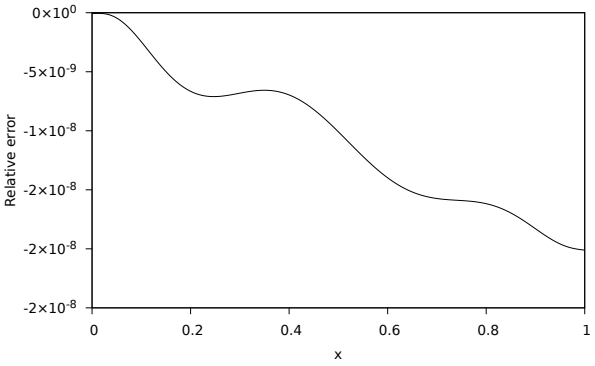


(d) 100 collocation points

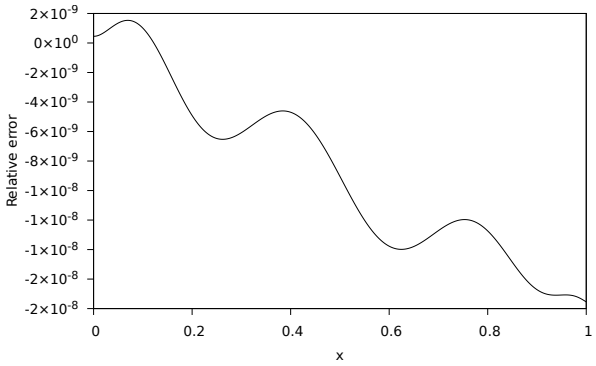
Figure S30: Best results for ODE-10 with death penalty (dp).



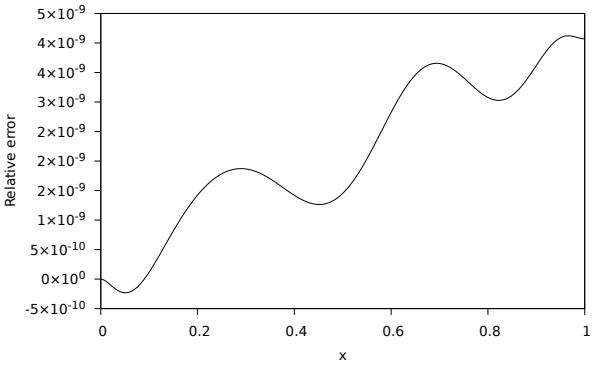
(a) 10 collocation points



(b) 20 collocation points



(c) 50 collocation points



(d) 100 collocation points

Figure S31: Best results for ODE-10 with the penalty method (pm).

8 Training loss analysis

As pointed out above, the present setting is unusual from a machine learning standpoint in that we are solving an equation, not inducing a function from a noisy dataset. We can thus directly use the training loss in the model selection phase—we have no need to consider overfitting to a particular data sample. The corollary of this is that for practical situations where the analytical solution to the ODE is not known, we can use the training loss a measure of solution quality: it follows that a small training loss at any point in the domain implies that the ODE is being well satisfied at that point. Figure S32 shows both the relative error and the training loss over the domain for the same result as shown in Figure S26a.

Figure S33 on the other hand shows the corresponding results for the same (much better) solution in Figure S26d. It is clear from these two figures that a better training loss implies a better solution. An area of future work is to propagate the training loss values through to errors on the evolved solution thereby providing an error bound when the ODE's solution is not known a priori.

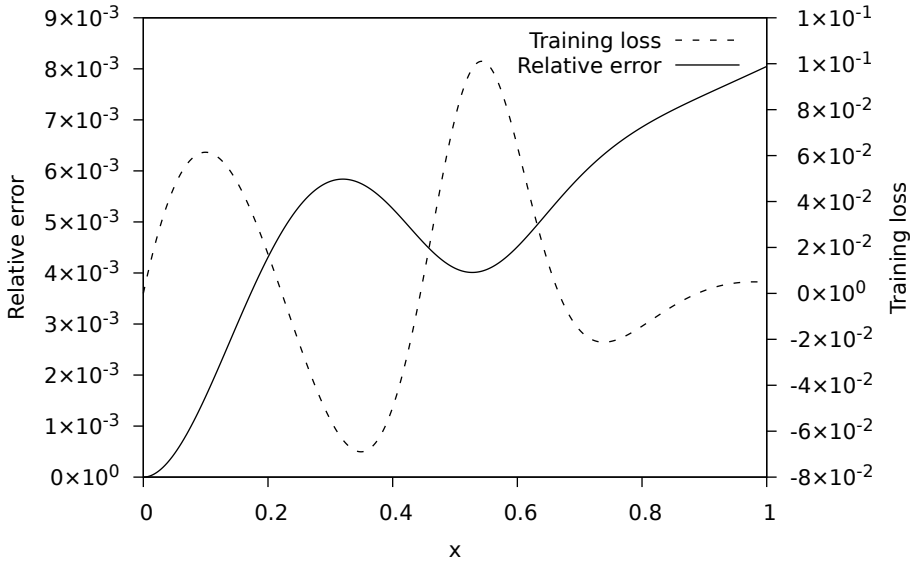


Figure S32: Relative error (left ordinate scale) and training loss (right ordinate scale) over the domain for ODE-9 with hard constraints with 10 collocation points.

It is also instructive to examine the self-diagnostic information available for ODE-7. Figure S34 shows the same result as Figure S22d. Here the largest training loss occurs near $x \approx 0$ and thereafter takes on much smaller values. This figure reinforces the point that the spikes in the previously presented relative error plot are a numerical issue with calculating the relative error,

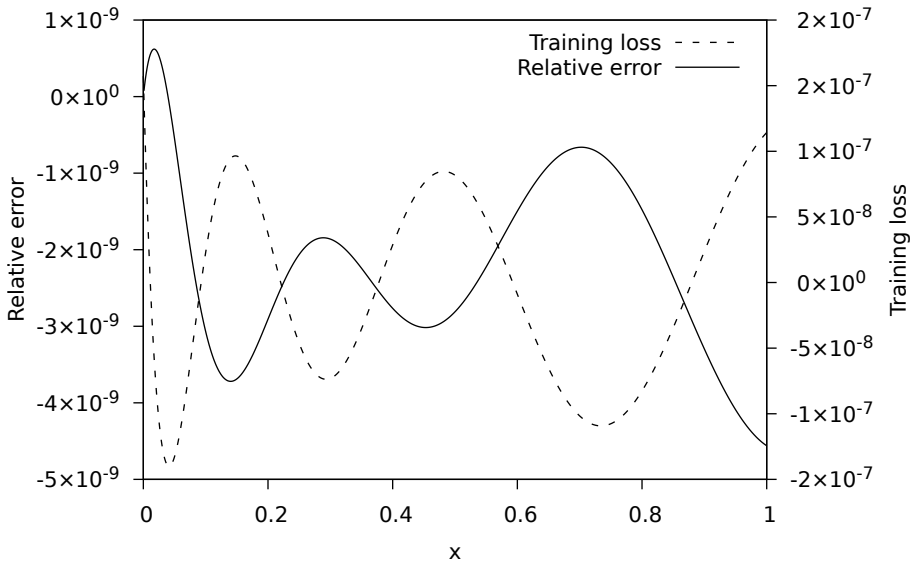


Figure S33: Relative error (left ordinate scale) and training loss (right ordinate scale) over the domain for ODE-9 with hard constraints with 100 collocation points.

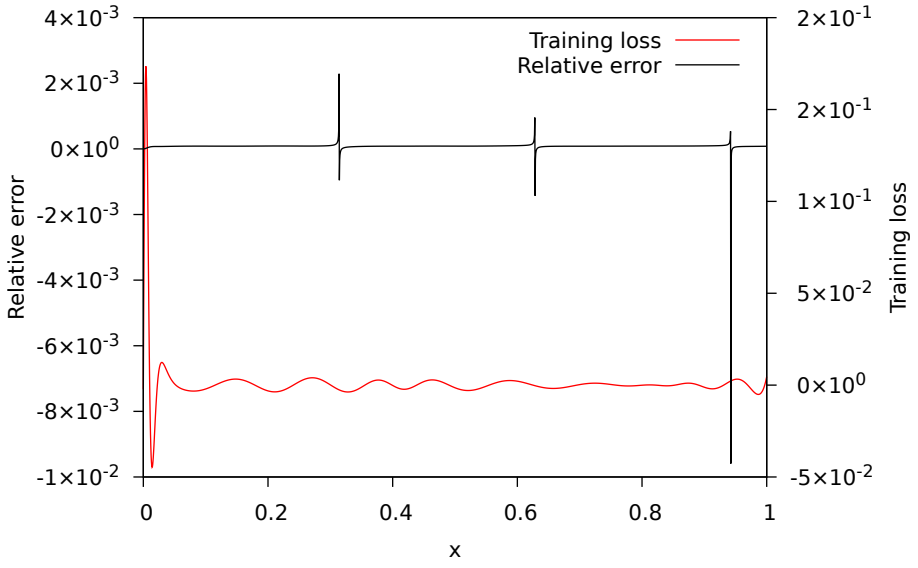


Figure S34: Relative error (left ordinate scale) and training loss (right ordinate scale) over the domain for ODE-7 with threshold/hard constraints with 100 collocation points.

References

P. Rockett. The solution of ordinary differential equations using genetic programming with constant tuning. In *24th UK Workshop on Computational Intelligence*, Edinburgh, UK, 3-5 September 2025.

T. Seaton, G. Brown, and J. F. Miller. Analytic solutions to differential equations under graph-based genetic programming. In A. I. Esparcia-Alcázar, A. Ekárt, S. Silva, S. Dignum, and A. S. Etaner-Uyar:, editors, *13th European Conference on Genetic Programming (EuroGP'10)*, pages 232–243, Istanbul, Turkey, 7-9 April 2010. doi: 10.1007/978-3-642-12148-7_20.