2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Numerical Comparison of Collocation vs Quadrature Penalty Methods

Martin P. Neuenhofen[1], Eric C. Kerrigan[2], Yuanbo Nie[3]

***Abstract*— Direct transcription with collocation-type methods (CTM) is a popular approach for solving dynamic optimization problems. It is known that these types of methods can fail to converge for problems that feature singular-arc solutions, high-index differential-algebraic equations and overdetermined constraints. Recently, we proposed the use of quadrature penalty methods (QPM) as an alternative numerical approach to collocation-type methods. In contrast to the concept of collocation, which requires constraint-residuals to equal zero at individual points (e.g. at collocation points), the main idea of QPM is to simply oversample this number of points and use their respective quadrature weights in a quadratic penalty term, coining the name of quadrature penalty. In this paper, we provide numerical case studies and a broad numerical comparison on a wide range of problems, highlighting the benefits of QPM over CTM not only in difficult problems, but also in solving problems competitively to CTM. These results show that QPM can be considered an attractive first go-to method when solving general dynamic optimization problems.**

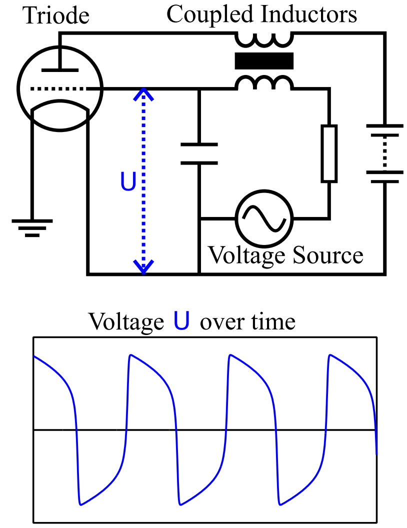

Fig. 1. Top: Diagram of an electric circuit. Bottom: Dynamics of measured voltage reveals a Van-der-Pol oscillation.

I. INTRODUCTION

Many dynamical processes in engineering can be described as differential-algebraic equation (DAE) systems. For example, the famous Van-der-Pol electrical circuit oscillator [1] is illustrated in Figure 1 and given by

$$\dot{y}_1(t) = y_2(t)\,, \tag{1a}$$
$$\dot{y}_2(t) = -y_1(t) + y_{[2]}(t)\cdot\left(1 - y_1(t)^2\right) + u(t)\,, \tag{1b}$$
$$-1 \leq u(t) \leq 1\,, \tag{1c}$$
$$y_1(0) = 0\,, \tag{1d}$$
$$y_2(0) = 1\,. \tag{1e}$$

These equations are from [2], augmented with a control voltage u that can be used to dampen the oscillation in y_1.

Optimal Control describes the discipline of intervening in dynamical systems, using controls, to achieve some mission goals (e.g. to steer the system's states towards a target). In the Van-der-Pol problem, optimal control can be used to construct a function $u(\cdot)$ that minimizes some objective, such as, e.g.,

$$\Phi(y,u) = \int_{t_0}^{t_f} f(\dot{y}(t), y(t), u(t), t)\,\mathrm{d}t\,.$$

For the example of the Van-der-Pol controller, we can use

$$\Phi(y,u) = \int_0^4 \left(y_1(t)^2 + y_2(t)^2\right)\mathrm{d}t\,. \tag{2}$$

For simple problems like the above, optimality criteria, such as Pontryagin's minimum principle, can be applied to analytically determine the optimal solution for $u(t)$. However, for many practical engineering models of large sizes, numerical algorithms may be required.

In this work, we numericallysolve the optimal control problem arising from such dynamical systems with collocation type methods (CTM) and quadrature penalty methods (QPM). We work out differences in both methods' solution processes. We then discuss sources in the problem that result in challenges in numerical solutions, related to singular arcs. We describe how QPM alleviates these issues. Section IV considers another source of numerical issues for CTM: consistent over-determination, and again illustrates how QPM can resolve related issues. We also compare the performance of CTM and QPM in Section V for a wide range of optimal control problems in the

[1]Martin P. Neuenhofen was with the Department of Electrical & Electronic Engineering, Imperial College London, SW7 2AZ London, UK, e-mail: martinneuenhofen@googlemail.com, www.MartinNeuenhofen.de

[2]Eric C. Kerrigan is with the Department of Electrical & Electronic Engineering and Department of Aeronautics, Imperial College London, SW7 2AZ London, UK, e-mail: e.kerrigan@imperial.ac.uk, www.imperial.ac.uk/people/e.kerrigan

[3]Yuanbo Nie is with the Department of Automatic Control and Systems Engineering, University of Sheffield, S1 3JD Sheffield, UK, e-mail: y.nie@sheffield.ac.uk, www.sheffield.ac.uk/acse/department/people/academic/yuanbo-nie

979-8-3503-0123-6/23/$31.00 ©2023 IEEE

literature, before presenting the concluding remarks in Section VI to recommend the use of QPM as the preferred method for solving dynamic optimization problems.

II. Numerical Optimal Control via Direct Transcription

The usual approach to numerically solving optimal control problems operates by fitting piecewise interpolation polynomials for $y(\cdot), u(\cdot)$ of some polynomial degree $p \in \mathbb{N}$, known as finite elements. The polynomials are denoted by $y_h(\cdot), u_h(\cdot)$, depending on a finite element mesh size parameter $h > 0$. The nodal values of $y_h(\cdot), u_h(\cdot)$ are computed as the design vector of a nonlinear programming problem (NLP). In describing this process, for simplicity of exposition we imply treatment of general inequality constraints by the usage of slack variables [3, Chap. 2].

1) Collocation Solution: In order to yield convergence of any constraints $c(\dot{y}(t), y(t), u(t), t) = 0$ for t over the domain $[t_0, t_f]$, CTM (e.g. [3]) uses a set $\mathcal{Q}_h$ of pairs of quadature abscissae τ and weights α over the finite element mesh, and require

$$c\big(\dot{y}_h(\tau), y_h(\tau), u_h(\tau), \tau\big) = 0 \qquad \forall(\tau,\alpha) \in \mathcal{Q}_h\,. \tag{3}$$

This principle is called collocation, and related methods are called collocation methods. It is known that polynomial collocation methods are a special sub-class of the famous Runge-Kutta methods. When replacing the implicit DAE $c(\dot{y}(t), y(t), u(t), t) = 0$ with explicit differential equations $\dot{y} = \tilde{c}(y(t), u(t), t)$, general Runge-Kutta methods can be used instead. Both kinds of methods follow the spirit of solving dynamic equations by virtue of a finite set of nonlinear equations. The NLP reads

$$\begin{aligned}
\min_{y_h,u_h} \quad & \Phi_h(y_h,u_h) := \sum_{(\tau,\alpha)\in\mathcal{Q}_h} \alpha \cdot f(\dot{y}_h(\tau), y_h(\tau), u_h(\tau), t) \\
\text{s.t.} \quad & c(\dot{y}_h(\tau), y_h(\tau), u_h(\tau), \tau) = 0 \qquad \forall(\tau,\alpha) \in \mathcal{Q}_h\,, \\
& u_L(\tau) \le u_h(\tau) \le u_R(\tau) \qquad \forall(\tau,\alpha) \in \mathcal{Q}_h\,, \\
& y_h(t_0) = y_0\,.
\end{aligned}$$

with u_L and u_R the lower and upper bounds of the input.

2) Penalty Solution: The idea of using quadratic penalty functions to solve optimal control problems with explicit initial conditions was studied in [4], focusing on the maximum principles that arise from the penalty function and their connections to the Pontryagin's maximum principle. This work lays the groundwork for an indirect solution to the penalty function, however a numerical method is not introduced. In recent work [5], we have proposed an efficient numerical framework named QPM for the direct solution of optimal control problems of generic forms.

In comparison to CTM, QPM work in a conceptually different way, in that they do not translate the DAE constraints $c(\dot{y}(t), y(t), u(t), t) = 0$ into a finite set of nonlinear equations. Instead, QPMs form a penalty term P, which is composed from quadrature points τ and

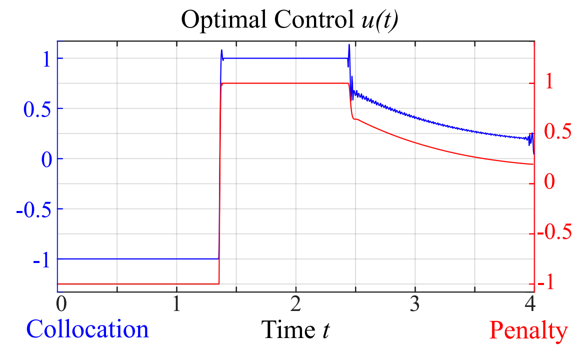

Fig. 2. Comparison of numerical solutions for the Van-der-Pol example, using CTM (blue) and QPM (red). Both methods use finite elements of polynomial degree $p = 4$ on an equidistant mesh of 100 intervals ($h = 0.25$).

weights $\alpha > 0$, given by a set $\mathcal{Q}_h$, with respect to the finite element mesh:

$$P_h(y_h,u_h) := \sum_{(\tau,\alpha)\in\mathcal{Q}_h} \alpha \cdot \|c\big(\dot{y}_h(\tau), y_h(\tau), u_h(\tau), \tau\big)\|_2^2\,.$$

The NLP reads

$$\begin{aligned}
\min_{y_h,u_h} \quad & \Phi_h(y_h,u_h) + \frac{P_h(y_h,u_h)}{2\cdot\omega_h} \\
\text{s.t.} \quad & u_L(\tau) \le u_h(\tau) \le u_R(\tau) \quad \forall(\tau,\alpha) \in \mathcal{Q}_h\,, \\
& y_h(t_0) = y_0\,,
\end{aligned}$$

where $\omega_h > 0$ is a small penalty parameter that approaches zero as h approaches zero.

III. Computational Results for the Van-der-Pol Example

Optimal control problems usually feature solutions that are piecewise smooth. This is also the case for Example 1: The optimal control follows a bang-bang structure, meaning it switches from one of its bounds instantly to another. On the final sub-arc, the solution smoothly approaches zero in an exponential fashion.

Figure 2 shows the numerical solution $u_h(\cdot)$ of CTM and QPM. In particular, both methods use finite elements of polynomial degree $p = 4$ on an equidistant mesh of 100 intervals, resulting in a mesh size of $h = 0.25$. CTM uses the Legendre-Gauss-Radau points of degree p. QPM uses the Gauss-Legendre quadrature method of degree $2 \cdot p$ and a penalty parameter of $\omega = 10^{-6}$.

As the figure shows, the CTM solution significantly overshoots the bound constraints $-1 \le u(t) \le 1$ in the bang-bang region, whereas the QPM features no visually noticeable overshoot. (The exact overshoot for QPM is limited to 0.0031.) In addition, the CTM solution oscillates on the last sub-arc. This phenomenon, known as *ringing*, is common for CTMs when solving certain classes of optimal control problems, known as *singular-arc problems*.

Singular-arc problems are optimal control problems where a certain kind of singularity in the equations that result from the Pontryagin minimum principle occurs. While the exact analytic conditions are well-analyzed, identifying and eliminating singular arcs from complex real-world optimal control problem can be a very challenging task.

Singular arcs can appear in any problem of any size (e.g. number of states, controls) or kind (e.g. linear, non-linear, convex, non-convex). Since singular arcs often only occur on some parts of the solution trajectories and can be affected by many factors, such as the activation status of a constraint, it may not always be possible to eliminate the singularity as a prior measure. Desirably, with QPM we have a method at hand that can solve optimal control problems efficiently, regardless of whether singular arcs are present.

IV. Large-Scale Compound Models

The examples that we give in this paper are relatively small problems for demonstration purposes, but in practical applications, dynamical system models often consist of several hundreds or thousands of states, controls, and are described by a DAE. Frameworks like Modelica or COMSOL are commonly used to automatically generate these models by composing them from smaller sub-modules. In particular when linking modules, high-index DAE or consistently overdetermined equations can result.

For example, consider framework models that use hinges instead of ball joints, or any consistently over-determining type of connector. General models may consist of hundreds of sub-modules and thousands of variables and equations, hence it is practically hard to identify cases like these manually and altering the model using mathematical expertise so as to assert the resulting DAE system is well-determined.

To illustrate the challenges of an overdetermined model, we use a small toy example to illustrate the numerical issues that result from consistent over-determination. The following example studies the $d = 1$-dimensional rotational degree of freedom for a quaternion $q(t)$. For d dimensions, a quaternion is a $d+1$-dimensional vector. Consider

$$\begin{aligned} \min_{q,u} \quad & \int_0^5 u(t)^2 + \dot{u}(t)^2 \, \mathrm{d}t \\ \text{s.t.} \quad & q(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad q(5) = \begin{bmatrix} -0.96 \\ 0.28 \end{bmatrix}, \\ & \dot{q}(t) = q(t) \cdot u(t), \quad \|q(t)\|_2^2 = 1 \,. \end{aligned}$$

The quaternion q rotates with a controlled angular speed u that minimizes some functional, subject to rotating the quaternion from a specified initial angular position into a final angular position within a prescribed amount of time. With input u and the initial conditions, q would already be well determined from the differential equation for q. Thus, the optimal control problem is over-determined due the constraint $\|q(t)\|_2^2 = 1$ and the boundary conditions on q.

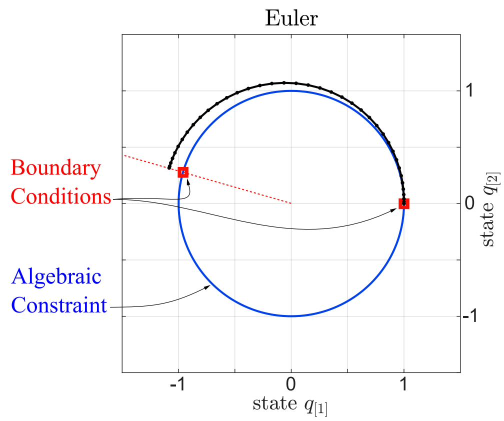

Fig. 3. Explicit Euler solution in state space. The Euler solution features a spiral shape, thus violating the circular constraint $\|q(t)\|_2^2 = 1$.

Elimination of overdetermination for this example can be non-trivial. One approach would be to replace q with an angle φ, but for $d > 1$ this approach would not work; our reason for using $d = 1$ here is to simplify the example in order to illustrate the problems that could occur. We thus intend to keep q and the path constraint. At each time t, one of the differential equations would have to be removed in order for the problem to be well-determined and avoid singularity (e.g. when $q_1 = 0$ and we remove the differential equation for q_2). Finally, we would have to introduce a parameter to neutralize the scaling of the end condition. Otherwise, both the path constraint and the end-condition imply $\|q(5)\|_2^2 = 1$, thus again consistently overdetermining the problem.

Figures 3 and 4 show three collocation solutions, where — to allow the problem to be solved — we removed the pathconstraint and replaced the end condition with $0.28 \cdot q_{[1]}(t_f) + 0.96 \cdot q_{[2]}(t_f) = 0$. The figures show numerical solutions of Explicit Euler, Trapezoidal, and Legendre-Gauss-Radau degree 2 collocation. We see that all three methods violate the path constraint (denoted in black).

QPM faces no issues when solving overdetermined problems, because it treats any equality constraints in a least-squares manner. Thus, since the overdetermination is consistent, the least-squares solution's residual will be in $\mathcal{O}(h^p)$, where h is the mesh size and p depends on the finite element method's order of consistency.

V. Numerical Experiments

In previous sections, we presented a couple of challenging cases that cause CTM to struggle. Although problem reformulations and ad-hoc fixes are available in the CTM literature, they often require the identification of the root causes beforehand, which can be practically prohibitive for real-world engineering problems. In contrast, QPM has the ability to reliably yield accurate and piecewise-smooth solutions for these difficult problems without the need for special considerations.

In this section, we compare computational results

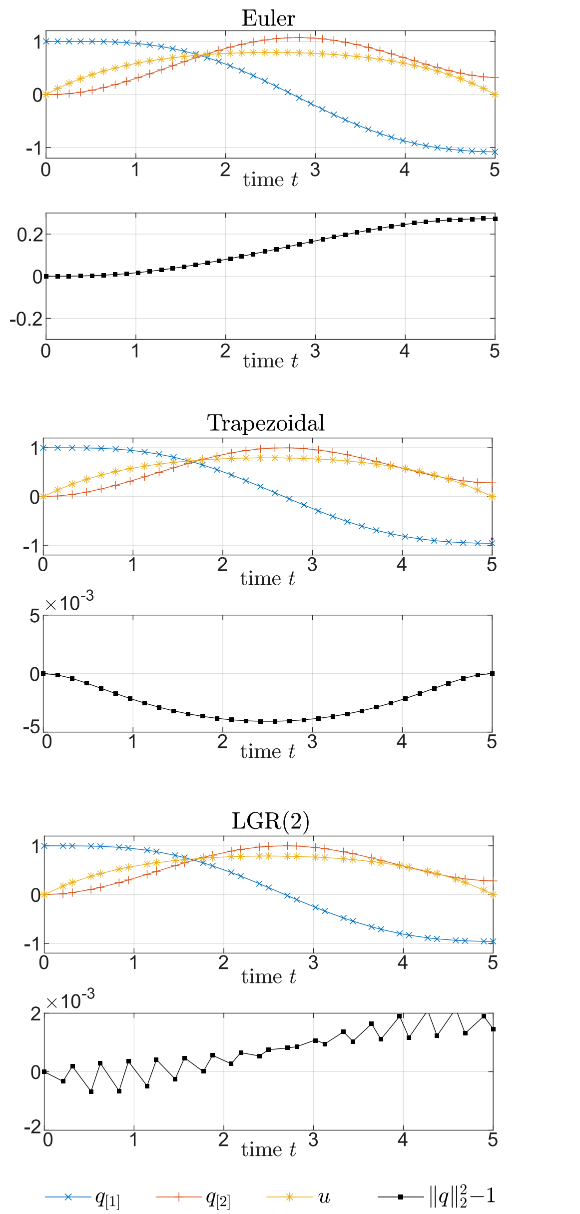

Fig. 4. Trapezoidal and Legendre-Gauss-Radau ($p = 2$) collocation solution in trajectory space. Both methods violate the equation $\|\boldsymbol{q}(t)\|_2^2 = 1$.

between CTM and QPM on a wide selection of benchmark problems from the literature, focusing on the differences in solution accuracy and computational cost. Specifically, we will show that QPM could be used as a reliable first go-to method and as a competitor to CTM.

A. Experimental Setting

Our presentation comprises 20 problems. Each problem is solved with CTM and QPM of polynomial degree 4 on three meshes; a coarse mesh of N intervals, a medium mesh of $4 \cdot N$ intervals, and a fine mesh of $16 \cdot N$ intervals. We use equidistant meshes because different refinement strategies may favour either method, resulting in unfair comparison. In terms of quadrature order $\mathcal{Q}_h$, for collocation, it matches with the polynomial degree hence we select the Legendre-Gauss-Radau (LGR) points of degree 4 (i.e. we use LGR collocation with polynomial degree 4). For QPM we select for $\mathcal{Q}_h$ the Gauss-Legendre of degree 8 in conjunction with a fixed penalty parameter of size $\omega = 10^{-6}$.

As convergence measures, we use the smallest non-negative scalars δ, ρ, γ which satisfy the following bounds.

$$\begin{aligned} \left|\Phi(y_h, u_h) - \Phi(y, u)\right| &\le \delta\,, \\ \sqrt{\int_{t_0}^{t_f} \|c(\dot{y}_h(t), y_h(t), u_h(t), t)\|_2^2 \,\mathrm{d}t} &\le \rho\,, \\ u_L(t) - \gamma \le u_h(t) \le u_R(t) + \gamma \qquad &\forall t \in [t_0, t_f]\,. \end{aligned}$$

The measures δ, ρ, γ are called the *optimality gap*, *feasibility residual* and *bound violation* [5]. In addition, we measure the computation time to solve each problem with each method.

B. Test Problems

Table I depicts the test problems with their respective properties from left to right: active inequality constraints on y, u; smoothness properties of the solution in the literature $y^\star, u^\star$; properties of the minimizer, and numerical properties such as stiffness (see definition in [6]) of the optimality system, scaling issues due to large discrepancy in magnitude of variables, and long timespans $[t_0, t_f]$.

The problems are sorted into categories. Some models permit analytic solutions while others stem from actual engineering applications. Finally, there are two classes of challenges, commonly seen in nonlinear optimal control. These challenges are explained in the following.

The first class of challenges deals with non-strictness and non-uniqueness of solutions: We compute two distinct minimizers to the same problem (one global and one local), to confirm that both methods are able to converge to both minimizers. We also compute non-strict minimizers for a landing-abortion problem that features a family of equally good solutions with regard to how the plane escapes from the wind-shear.

The second class of challenges deals with irregular constraints: The constrained brachistochrone problem features a singular Jacobian [3]; the pendulum determines the beam force implicitly from a differential-algebraic equation of varying index and eventually also imposes a bound on the beam forces. This results in singular optimality conditions. Details on each problem are given in the references in Table I.

C. Results and Discussion

Figure 5 presents the convergence measures on each mesh of each problem from Table I for both CTM and QPM. Both direct transcriptions succeed on all problems in the sense that they generate reasonably good numerical solutions.

TABLE I

List of numerical test problems. Abbreviations: "ineq."=inequalities; "cont."=continuity.

problem		properties										reference
		ineq.		cont.		minimizer		conditioning				
index	name	bound y	bound u	jump u	edge	singular	kind	stiff	bad scale	long span		reference
Analytic solution available												
1	Hager Problem						unique					[7]
2	Bryson-Denham Problem	✓			✓		unique					[8]
3	Singular Regulator		✓	✓	✓	✓	unique					[9]
Applications												
Robotics												
4	Two-Link Robot Arm		✓	✓	✓		strict					[10]
5	Container Crane		✓		✓		strict					[11]
Aircrafts												
6	Alp-Rider		✓				strict	✓	✓			[3]
7	Dynamic Soaring	✓			✓		strict	✓				[12]
Rockets												
8	Goddard Rocket Max Height	✓	✓	✓	✓	✓	strict					[3]
9	Spaceship Control		✓				strict					[13]
10	Spaceshuttle Reentry	✓	✓		✓		strict	✓	✓	✓		[3]
Satellites												
11	Orbit Raising		✓				strict					[14]
12	Low-Thrust MEO-GEO Transfer		✓				strict			✓		[15]
Biochemistry												
13	Tuberculosis Treatment	✓	✓		✓		strict	✓	✓	✓		[16]
14	Batch Fermentation	✓	✓	✓	✓	✓	strict	✓				[17]
15	Kiln Heating PDE		✓		✓		strict	✓				[3]
Challenges												
due to non-uniqueness												
16a	Obstacle Avoidance below		✓		✓		strict					[5]
16b	above		✓		✓		strict					
17a	Free-Flying Robot book		✓	✓	✓		strict					[3]
17b	asymmetric		✓	✓	✓		strict					
18a	Landing Abortion low Exit	✓	✓	✓	✓	✓	non-strict		✓			[3]
18b	high Exit	✓	✓	✓	✓	✓	non-strict		✓			
due to non-regularity												
19a	Brachistochrone unconstr.						unique					[3]
19b	constr. $h = 0.1$		✓		✓		unique					
19c	constr. $h = 0$		✓		✓	✓	unique					
20a	Pendulum Index 1					✓	strict					[18]
20b	Index 2					✓	strict					
20c	Index 3					✓	strict					
20d	Index 3 + constr.		✓		✓	✓	strict					[5]

1) General Observations: Comparing solutions of CTM and QPM on the same mesh, the optimality gap δ is similar in magnitude for QPM and CTM. QPM is a bit slower than CTM, but in return QPM yields on average three orders of magnitude smaller equality constraint residuals ρ and QPM yields on average one order of magnitude smaller inequality constraint residuals γ.

In many problems (index 3, 4, 5, 6, 8, 11, 16, 18, 19, 20), we can observe that the equality constraint residuals and inequality constraint residuals obtained by QPM with the coarse mesh are already smaller than those obtained by CTM with the finest mesh, but using less time. Therefore, when comparing solution time based on a given accuracy level, QPM has the potential to outperform CTM as it can roughly achieve the same accuracy on a much coarser mesh.

2) Exceptions: The three biochemistry problems (index 13–15) are very stiff, with some states' derivatives have very large values. Therefore, the equality feasibility residual ρ is large in these cases. Nonetheless, the solution arcs of both methods match those of the solution in the literature.

On the problems with index 6, 17, 20d (Alp-Rider, Free-Flying Robot, constrained Pendulum), the optimality

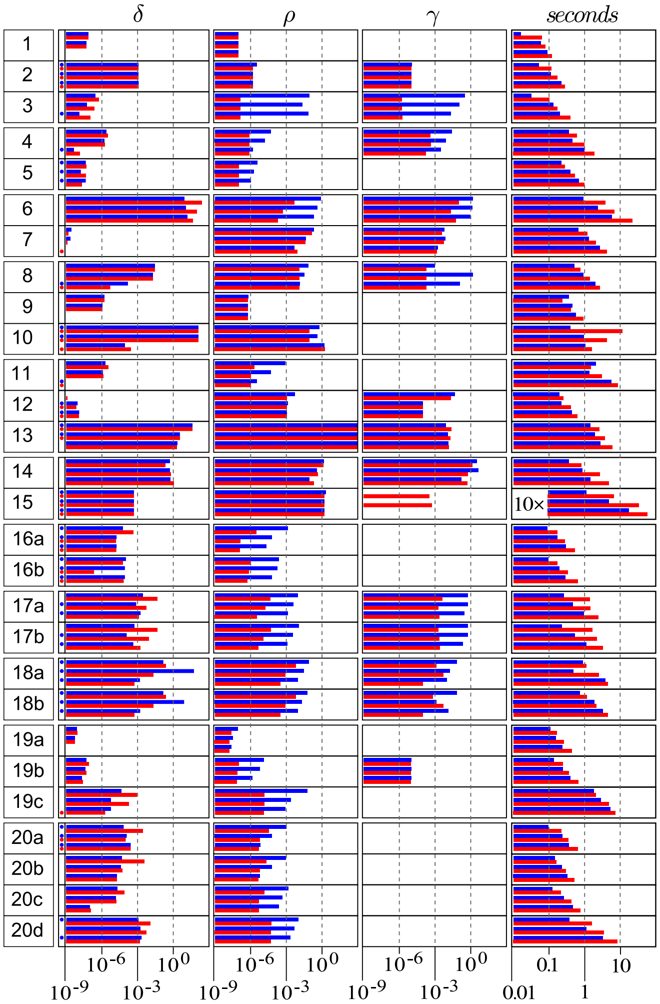

Fig. 5. Convergence measures for CTM (blue) and QPM (red). Three bars per method per cell give the measure from coarsest (top) to finest mesh (bottom). Problem 15 took 10× as many seconds as plotted because the dynamics equation is a PDE (significantly more states).

gaps of QPM are significantly larger than of CTM. In all of these problems, QPM converges to very accurate solutions (in terms of feasibility) whereas CTM converges to inaccurate solutions. However, a comparison of computation time and optimality gap makes sense only when both methods have similarly feasible solutions which is not the case for these problems.

VI. Conclusions

In this paper, we presented a comprehensive numerical study comparing two types of direct transcription methods: the current state-of-the-art direct collocation approach (CTM) and the quadrature penalty method (QPM). This numerical study contains a wide range of representative optimal control problems covering different engineering fields with varying degrees of difficulty. The experiments show that both methods can solve a variety of applications successfully, yet QPM has favourable convergence performance and can reliably solve challenging problems such as singular arc problems and problems with consistently overdetermined constraints.

References

[1] L. Glass and M. C. Mackey, *From clocks to chaos.* Princeton University Press, Princeton, NJ, 1988, the rhythms of life.

[2] H. Maurer, "Theory and applications of bang-bang and singular control problems," in *Conference slides, 2007, 23rd IFIP TC 7 Conference on System Modelling and Optimization Krakow*, 2007.

[3] J. T. Betts, *Practical Methods for Optimal Control and Estimation Using Nonlinear Programming*, 2nd ed. New York, NY, USA: Cambridge University Press, 2009.

[4] A. V. Balakrishnan, "On a new computing technique in optimal control," *SIAM J. Control*, vol. 6, pp. 149–173, 1968.

[5] M. P. Neuenhofen, "Quadratic integral penalty methods for numerical trajectory optimization," Ph.D. dissertation, Imperial College London, 2022.

[6] E. Hairer and G. Wanner, *Solving ordinary differential equations. II*, ser. Springer Series in Computational Mathematics. Springer-Verlag, Berlin, 2010, vol. 14, stiff and differential-algebraic problems, Second revised edition, paperback. [Online]. Available: https://doi.org/10.1007/978-3-642-05221-7

[7] S. Kameswaran and L. T. Biegler, "Convergence rates for direct transcription of optimal control problems using collocation at Radau points," *Comput. Optim. Appl.*, vol. 41, pp. 81–126, 2008.

[8] A. E. Bryson, *Dynamic optimization.* Prentice Hall, 1999.

[9] G. M. Aly and W. C. Chan, "Application of a modified quasilinearization technique to totally singular optimal problems," *International Journal of Control*, vol. 17, pp. 809–815, 1973.

[10] R. Luus, *Iterative dynamic programming.* Chapman and Hall/CRC, 2019.

[11] D. Augustin and H. Maurer, "Sensitivity analysis and real-time control of a container crane under state constraints," in *Online Optimization of Large Scale Systems.* Springer, 2001, pp. 69–82.

[12] Y. J. Zhao, "Optimal patterns of glider dynamic soaring," *Optimal control applications and methods*, vol. 25, no. 2, pp. 67–89, 2004.

[13] R. H. Battin, *An introduction to the mathematics and methods of astrodynamics.* Aiaa, 1999.

[14] A. E. Bryson, Jr. and Y.-C. Ho, *Applied Optimal Control: Optimization, Estimation, and Control.* Taylor & Francis Group, New York (NY), 1975.

[15] C. A. Kluever, "Optimal feedback guidance for low-thrust orbit insertion," *Optimal Control Applications and Methods*, vol. 16, no. 3, pp. 155–173, 1995.

[16] E. Jung, S. Lenhart, and Z. Feng, "Optimal control of treatments in a two-strain tuberculosis model," *Discrete & Continuous Dynamical Systems-B*, vol. 2, no. 4, p. 473, 2002.

[17] J. E. Cuthrell and L. T. Biegler, "Simultaneous optimization and solution methods for batch reactor control profiles," *Computers & Chemical Engineering*, vol. 13, no. 1-2, pp. 49–62, 1989.

[18] S. Campbell and P. Kunkel, "Solving higher index dae optimal control problems," *Numerical Algebra, Control & Optimization*, vol. 6, no. 4, p. 447, 2016.