A numerical design methodology for optimal pacing strategy in cycling

Asker Friis Bach, Christian Bach Lundgaard, Joe Alexandersen

A numerical design methodology for optimal pacing strategy in cycling

Asker Friis Bach[1*], Christian Bach Lundgaard[2] and Joe Alexandersen[3]

[1*]Independent scientist, 2100 Østerbro, Denmark .
[2]Independent scientist, 2200 Nørrebro, Denmark.
[3]Department of Mechanical and Electrical Engineering, University of Southern Denmark, 5230 Odense, Denmark.

*Corresponding author(s). E-mail(s): asker@friisbach.dk;
Contributing authors: chrlundg@gmail.com; joal@sdu.dk;

Abstract

The individual time trial discipline (ITT) in professional cycling is dubbed the race of truth as the cyclists are racing alone against the clock. These stages can be very close and they often play a decisive role in the Grand Tours. At the first ITT stage of the Tour de France 2022, the first and second places were separated by **5** seconds corresponding to a **0.5%** difference in the finishing times. This makes the discipline highly interesting from an optimization perspective. A pacing strategy defines how a cyclist distributes their power output along a course and hence also their velocity. We have developed a numerical design methodology for computing optimal pacing strategies - i.e. the pacing strategy that minimizes the finishing time for a given cyclist racing on a given course. We apply the method to four hypothetical courses of **2 km** simulating various gradients and wind conditions. The optimized pacing strategies led to improved finishing times ranging from **0.45%** to **2.84%** compared to benchmark pacing strategies. We also show how the method can be applied to compute an optimized pacing strategy for a complex real world course. The result is closely related to a pacing strategy derived from data from professional cyclist and ITT specialist Martin Toft Madsen, with the optimized strategy being **1.2%** faster over **21.3 km**. We believe that the method presented here constitutes a promising framework for computing optimal pacing strategies and with further research and a more accurate physiological model, this could prove an important tool for strategizing in professional cycling.

Keywords: Numerical optimization, optimal pacing strategy, Finite Element Method, adjoint sensitivity analysis, cycling time trial

Competing Interests and Funding: On behalf of all authors, the corresponding author states that there is no conflict of interest and that no funding has been received for the preparation of this article.

1 Introduction

In professional cycling the individual time trial (ITT) is a discipline where competitors take turn to race individually on a fixed course. Known as the race of truth, the ITT involves no team tactics so the race depends only on the individual rider's strength and endurance. It is famous for its salient

Fig. 1: Picture of a power meter (small grey box) measuring the pedal force. The velocity is measured with a GPS and the power output is calculated as the product of velocity and pedal force.

role in major competitions such as the Tour de France and the Olympics. At the first ITT stage of the Tour de France 2022, there was a 5 s difference between the finishing times of the first and second place corresponding about 0.5% over 13.1 km. The closeness and importance of these races motivates the search for marginal gains by optimizing all areas that can improve the performance of riders.

This article is a contribution to the field of optimal pacing strategies for ITT stages. The term *pacing strategy* refers to the way a rider distributes their power output along a course (Abbiss and Laursen, 2008). The pacing strategy of a rider defines both their resulting velocity along the course and the physiological toll of each segment as well as the workout as a whole. An optimal pacing strategy is a rider- and course-dependent pacing strategy that minimizes the finishing time subjected to an upper bound on the physiological toll. In other words, it is the most efficient way for a rider with a certain physique to race a given ITT. The aim of studying optimal pacing strategies is to develop a numerical methodology for efficient computation of physiologically accurate optimal pacing strategies.

With modern power meters (see Figure 1) and cycling computers one can measure the power output and velocity of a rider in real-time. This presents an opportunity for comparing theoretical pacing strategies with experimental data. The long-term goal is to exploit the unique optimal pacing strategy for a given rider when strategising for an ITT.

1.1 Overview of the literature

The field is relatively new with most advances made within the past decade. Early papers showed theoretically and experimentally that varying power output in response to changes in conditions, such as hills and wind, can lead to improved finishing times (Swain, 1997; Atkinson et al, 2007; Cangley et al, 2011; Atkinson and Brunskill, 2000). Although these early papers only considered simple models on piecewise linear courses, they sparked an interest in the field.

1.1.1 The numerical methods

In recent years, there have been a number articles published with various numerical methods (Gordon, 2005; Boswell, 2012; Sundström et al, 2013; Yamamoto, 2018; Fayazi et al, 2013; Wolf et al, 2019; Zignoli and Biral, 2020). Gordon (2005) introduced mathematical optimization with Lagrange multipliers and Boswell (2012) introduced a differential equation model relating the power output of a cyclist to their velocity. This allowed for optimization taking into account the effects of acceleration. Sundström et al (2013) used adjoint sensitivity analysis and the Method of Moving Asymptotes (MMA) albeit with a very coarse discretization.

Recent studies employ libraries with built-in optimization solvers. Examples are MATLAB's function *fmincon()* used by Dahmen (2016), the MATLAB library *FALCON.m* used by Wolf et al (2019) and the Maple package *XOptima* used by Zignoli and Biral (2020). These methods have achieved the impressive goal of optimizing real world courses and the authors have taken further steps such as implementing a control process and considering the effects of corners with a three-dimensional model (Wolf et al, 2019; Dahmen, 2016).

Though these optimization solvers are readily implemented, we believe that there is still more insights and possibly a higher efficiency to be obtained with further research into the methodology for optimizing. It is towards this aim that we present a design methodology for optimization in this article.

1.1.2 The physiological constraint

The question of how to incorporate a physiologically accurate constraint modelling human fatigue when optimizing pacing strategies has been an active topic of research in recent years with several different approaches appearing (Gordon, 2005; Fayazi et al, 2013; Sundström et al, 2013, 2014; Dahmen et al, 2012; Dahmen, 2016; Wolf et al, 2016). These include bioenergetic models which e.g. separately take into account the effects of aerobic and anaerobic work based on various models for fatigue dynamics tracing back to the work by Morton (1986) and Morton and Billat (2004). The arguments are mainly theoretical and all studies conclude that further work is required to develop a more realistic model. For this reason, the concern of this paper will not be the nature of the constraint, but rather the method of optimization employed subsequently. Nevertheless, we will here provide arguments in favour of our choice of constraint.

Within the world of cycling the principle of *Normalized Power®* (NP) is a well established measure developed by Allen et al (2019) in the book *"Training and racing with a power meter"*. It is described in an article by Hurley (2021) as a measure that *"reflects the disproportionate metabolic cost of riding at high intensity, by weighting hard efforts and deemphasizing periods of easy spinning"*. The idea is to take the average of the power output function raised to the fourth power and subsequently take the fourth root. Mathematically, this is equivalent to placing an upper bound on the p-norm, with $p = 4$, of the power function. This penalizes variations in power output whilst placing an upper limit on the average power output. Small variations in power output were reported to have a relatively large effect on the physiological toll in a study by Foster et al (1993). The choice of raising the power output to the power of 4 is based on a regression of blood lactate due to the power output (Coggan, 2017). NP has already been introduced as a constraint when optimizing power output by Yamamoto (2018) and we will similarly base our physiological model on this principle. Furthermore, it is already used as a metric for evaluating the physiological stress of a workout by e.g. Garmin, trainerroad and trainingpeaks (Garmin, 2022; Hurley, 2021; Ganoung, 2022) and Strava uses a similar weighted average method (Strava, 2022).

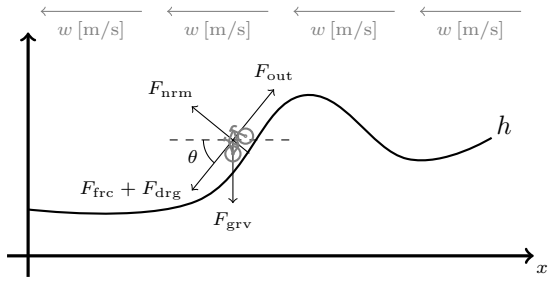


Fig. 2: *Free body diagram* of a cyclist moving on a road defined by the height profile $h(x)$ with wind $w(x)$. θ is the steepness of the course given by $\theta(x) = \arctan\left(\frac{dh}{dx}\right)$

2 Physical model

For a course of distance L we define the following variables and functions:

$$\text{Distance variable: } x \in [0, L] \tag{1a}$$
$$\text{Height function: } h : [0, L] \to \mathbb{R} \tag{1b}$$
$$\text{Wind function: } w : [0, L] \to \mathbb{R} \tag{1c}$$
$$\text{Angle of slope: } \theta = \arctan\left(\frac{dh}{dx}\right) \tag{1d}$$
$$\text{Velocity of cyclist: } v(x) = \frac{dx}{dt} \tag{1e}$$

As shown in Figure 2 we subject the cyclist to to the following forces:

$$\text{Gravity: } F_{\text{grv}} = mg \tag{2a}$$
$$\text{Normal force: } F_{\text{nrm}} = mg\cos(\theta) \tag{2b}$$
$$\text{Rolling resistance: } F_{\text{frc}} = C_r F_{\text{nrm}} \tag{2c}$$
$$\text{Drag: } F_{\text{drg}} = \frac{1}{2}\rho A C_d v^2 \tag{2d}$$

where $m\,[\text{kg}]$ is the total mass of the cyclist and the bike, $g\,[\text{m/s}^2]$ is the constant of gravitational acceleration, $C_r\,[1]$ is a coefficient taking all frictional forces into account, $\rho\,[\text{kg/m}^3]$ is the density of air, $A\,[\text{m}^2]$ is the total frontal surface area of the cyclist and the bike and $C_d\,[1]$ is the total drag coefficient of the cyclist and the bike. This leads to our governing physics equation relating the power

Mass:	$m = 85\,\mathrm{kg}$
Frontal Area:	$A = 0.3\,\mathrm{m}^2$
Coefficient of drag:	$C_d = 0.9$
Rolling resistance:	$C_r = 0.003$
Normalized Power:	$P_c = 450\,\mathrm{W}$
Gravitational constant:	$g = 9.807\,\mathrm{m/s}^2$
Air density:	$\rho = 1.2255\,\mathrm{kg/m}^3$

Table 1: Constants defining a *standard cyclist*.

output of the cyclist to their velocity:

$$\overbrace{P_{\text{out}}(x)}^{\text{Power output of cyclist}} = \overbrace{m\frac{dv}{dx}\,v(x)^2}^{\text{Inertia}} + \overbrace{mg\sin\left(\theta(x)\right)\,v(x)}^{\text{Gravity}} + \underbrace{C_r mg\cos\left(\theta(x)\right)\,v(x)}_{\text{Total rolling resistance}} + \underbrace{\frac{1}{2}\rho A C_d\,\left(v(x)-w(x)\right)^2 v(x)}_{\text{Total drag}} \quad (3)$$

Throughout the article, the term *standard cyclist* will refer to a cyclist modelled with the constants given in Table 1. The mass m includes a cyclist of mass 75 kg and a bicycle of mass 10 kg.

3 Method

Our method can be summarized as the following. First, the governing differential equation (3) is discretized using the Galerkin Finite Element Method (FEM). We define a design vector $\boldsymbol{\rho}$ to tune the power output at a series of nodes. Then we formulate the optimization problem as a discrete minimization of a functional subject to a residual and a functional constraint - all of which depend on the design vector. Finally, we introduce adjoint sensitivity analysis to solve the optimization problem using the Method of Moving Asymptotes (MMA) (Svanberg, 1987).

3.1 Finite Element formulation

The governing differential equation (3) is defined in space on the interval $[0, L]$. We discretize this into n elements Ω^e separated by the mesh nodes $\{x_k\}_{k=0,\ldots,n}$ as shown in the Figure 3. To each node x_k we associate a design variable $\rho_k \in [0,1]$ and a velocity variable $v_k \geq 0$. We collect the design variables in the design vector, $\boldsymbol{\rho} = (\rho_0\ \rho_1\ \cdots\ \rho_n)^T$, and likewise the velocity variables in the state vector, $\boldsymbol{V} = (v_0\ v_1\ \cdots\ v_n)^T$.

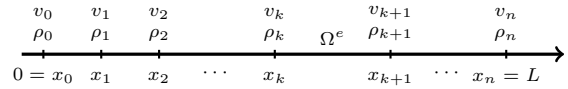


Fig. 3: *Discretization of the course.* The interval $[0, L]$ is divided into elements Ω^e separated by the mesh nodes $\{x_k\}_{k=0,\ldots,n}$. To each node x_k we associate a design variable ρ_k and a velocity variable v_k.

The power output on a given element Ω^e is some interpolation of the design variables at the neighbouring nodes. By tuning the numbers ρ_k and ρ_{k+1} we can change the power output on the element Ω^e in between the nodes x_k and x_{k+1}. In general, we say that the power output function depends on the whole design vector by:

$$P_{\text{out}} = P^h_{\text{out}}\left(\boldsymbol{\rho}\right) \quad (4)$$

where the superscript h indicates discretization. Similarly, the velocity function on the element Ω^e will be some interpolation of the velocity variables v_k and v_{k+1} at the neighbouring nodes. The velocity function therefore depends on the state vector by:

$$v = v^h\left(\boldsymbol{V}\right) \quad (5)$$

The velocity is related to the power output via the governing equation (3). Hence, the state vector will depend on the design vector according to the physics. The continuous strong form of our governing equation (3) is translated into the discrete weak form:

$$\underbrace{\boldsymbol{K}\left(\boldsymbol{V}\right)}_{n\times n}\underbrace{\boldsymbol{V}}_{n\times 1} = \underbrace{\boldsymbol{F}}_{n\times 1} \quad (6)$$

where the stiffness matrix $\boldsymbol{K}$ has contributions from each of the terms in the governing equation (3):

$$\boldsymbol{K} = \boldsymbol{K}_{irt} + \boldsymbol{K}_{grv} + \boldsymbol{K}_{frc} + \boldsymbol{K}_{drg} \quad (7)$$

If we let $\sum_{k\in\mathbb{N}}$ denote the standard finite element assembly, then these are given by the set of equations:

$$\boldsymbol{K}_{irt} = \sum_{k\in\mathbb{N}} \int_{\Omega^e} m\,\boldsymbol{N}\left(\boldsymbol{N}^T\boldsymbol{v}^e\right)^2 \boldsymbol{B}^T\,dx \tag{8a}$$

$$\boldsymbol{K}_{grv} = \sum_{k\in\mathbb{N}} \int_{\Omega^e} mg\sin(\theta)\;\boldsymbol{N}\boldsymbol{N}^T\,dx \tag{8b}$$

$$\boldsymbol{K}_{frc} = \sum_{k\in\mathbb{N}} \int_{\Omega^e} mgC_r\cos(\theta)\;\boldsymbol{N}\boldsymbol{N}^T\,dx \tag{8c}$$

$$\boldsymbol{K}_{drg} = \sum_{k\in\mathbb{N}} \int_{\Omega^e} \frac{1}{2}\rho A C_d\,\boldsymbol{N}\left(\boldsymbol{N}^T\boldsymbol{v}^e - w\right)^2 \boldsymbol{N}^T\,dx \tag{8d}$$

$$\boldsymbol{F}(\boldsymbol{\rho}) = \sum_{k\in\mathbb{N}} \int_{\Omega^e} P_{out}(\boldsymbol{\rho}^e)\,\boldsymbol{N}\,dx \tag{8e}$$

where $\boldsymbol{N}$ contains the shape functions (i.e. the functions that define the interpolation for the velocity function in (5)), $\boldsymbol{B} = \frac{d}{dx}\boldsymbol{N}$ and $\boldsymbol{\rho}^e$ and $\boldsymbol{v}^e$ are the design and state vectors localised to a single element, respectively. For example, on the element Ω^e in Figure 3, they are given by:

$$\boldsymbol{\rho}^e = \begin{pmatrix}\rho_k\\ \rho_{k+1}\end{pmatrix} \qquad \text{and} \qquad \boldsymbol{v}^e = \begin{pmatrix}v_k\\ v_{k+1}\end{pmatrix} \tag{9}$$

3.2 Optimization

The general optimization problem is to minimize an objective functional f subject to three constraints:

$$\begin{aligned} \min_{\boldsymbol{\rho}}\;\; & f(\boldsymbol{V}) \\ \text{s.t.}\;\; & \boldsymbol{R}(\boldsymbol{\rho},\boldsymbol{V}) = \boldsymbol{0}, \\ & g(\boldsymbol{\rho},\boldsymbol{V}) \le P_c, \\ & \boldsymbol{0} \le \boldsymbol{\rho} \le \boldsymbol{1} \end{aligned} \tag{10}$$

In our case, the object functional f will be to minimize the finishing time T, and the constraints are the residual equation $\boldsymbol{R}$, a physiological constraint g along with a chosen upper bound P_c and a standard box constraint on the design vector $\boldsymbol{\rho}$.

The objective functional f is expressed as an integral over space using $dt = dx/v$:

$$f(\boldsymbol{V}) := T = \int_0^L \frac{1}{v^h(\boldsymbol{V})}\,dx \tag{11}$$

The residual $\boldsymbol{R}$ is the discretized governing equation (6) rewritten as:

$$\boldsymbol{R}(\boldsymbol{\rho},\boldsymbol{V}) := \boldsymbol{K}(\boldsymbol{V})\,\boldsymbol{V} - \boldsymbol{F}(\boldsymbol{\rho}) = \boldsymbol{0} \tag{12}$$

As argued for in Section 1, we choose our physiological constraint based on the principle of Normalized Power (NP). It is defined by Coggan (2017) as the 4-norm of the power output function:

$$NP := \left(\frac{1}{T}\int_0^T {P_{out}^h}^4\,dt\right)^{1/4} \tag{13}$$

Weighted averages like this are commonly used (e.g. by Strava and Garmin Strava (2022); Garmin (2022)) to measure the physiological toll of a workout. We rewrite this to an integral over space and define the constraint as:

$$g(\boldsymbol{\rho},\boldsymbol{V}) := \left(\frac{1}{T}\int_0^L \frac{P_{out}^h(\boldsymbol{\rho})^4}{v^h(\boldsymbol{V})}dx\right)^{1/4} \tag{14}$$

The strategy for minimizing the objective function is to compute the sensitivities of the objective function ${}^{df}/_{d\boldsymbol{\rho}}$ and the constraint ${}^{dg}/_{d\boldsymbol{\rho}}$ and then use MMA to update the design vector $\boldsymbol{\rho}$ until an optimal solution is reached.

3.3 Adjoint sensitivity analysis

To compute the sensitivities of the objective functional, ${}^{df}/_{d\boldsymbol{\rho}}$, we define the Lagrangian:

$$\mathcal{L} = f + \boldsymbol{\lambda}^T\mathbf{R} \tag{15}$$

where $\boldsymbol{\lambda}$ is the vector of Lagrange multipliers for the residual equations. Taking the total derivative of the Lagrangian yields:

$$\frac{d\mathcal{L}}{d\boldsymbol{\rho}} = \frac{\partial f}{\partial \boldsymbol{V}}\frac{\partial \boldsymbol{V}}{\partial \boldsymbol{\rho}} + \frac{\partial f}{\partial \boldsymbol{\rho}} + \boldsymbol{\lambda}^T\left(\frac{\partial \mathbf{R}}{\partial \boldsymbol{V}}\frac{\partial \boldsymbol{V}}{\partial \boldsymbol{\rho}} + \frac{\partial \mathbf{R}}{\partial \boldsymbol{\rho}}\right) \tag{16}$$

which can be rewritten to:

$$\frac{d\mathcal{L}}{d\boldsymbol{\rho}} = \frac{\partial f}{\partial \boldsymbol{\rho}} + \boldsymbol{\lambda}^T \frac{\partial \mathbf{R}}{\partial \boldsymbol{\rho}} + \left(\frac{\partial f}{\partial \boldsymbol{V}} + \boldsymbol{\lambda}^T \frac{\partial \mathbf{R}}{\partial \boldsymbol{V}} \right) \frac{\partial \boldsymbol{V}}{\partial \boldsymbol{\rho}} \quad (17)$$

By defining $\boldsymbol{\lambda}$ as the solution to the following adjoint problem, the parenthesis above is made equal to zero:

$$\left(\frac{\partial \mathbf{R}}{\partial \boldsymbol{V}} \right)^T \boldsymbol{\lambda} = - \left(\frac{\partial f}{\partial \boldsymbol{V}} \right)^T \quad (18)$$

The sensitivities are then given by:

$$\frac{df}{d\boldsymbol{\rho}} = \frac{\partial f}{\partial \boldsymbol{\rho}} + \boldsymbol{\lambda}^T \frac{\partial \mathbf{R}}{\partial \boldsymbol{\rho}} = \boldsymbol{\lambda}^T \frac{\partial \mathbf{R}}{\partial \boldsymbol{\rho}} \quad (19)$$

where the last equality is a result of f not depending explicitly on the design vector $\boldsymbol{\rho}$ in (11). We can go through the exact same procedure for the sensitivities of the constraint g using a different Lagrange multiplier vector $\widetilde{\boldsymbol{\lambda}}$ and write the sensitivities as:

$$\frac{dg}{d\boldsymbol{\rho}} = \frac{\partial g}{\partial \boldsymbol{\rho}} + \widetilde{\boldsymbol{\lambda}}^T \frac{\partial \mathbf{R}}{\partial \boldsymbol{\rho}} \quad (20)$$

where $\widetilde{\boldsymbol{\lambda}}$ is given via the corresponding adjoint problem:

$$\left(\frac{\partial \mathbf{R}}{\partial \boldsymbol{V}} \right)^T \widetilde{\boldsymbol{\lambda}} = - \left(\frac{\partial g}{\partial \boldsymbol{V}} \right)^T \quad (21)$$

The procedure for solving the optimization problem numerically is now the following. Start with an initial guess of $\boldsymbol{\rho}$ satisfying the criteria in (10) and then execute the following loop:

- Assemble the matrix $\boldsymbol{K}$ with (7) and Finite Element assemblies (8).
- Solve the non-linear system of equations (6) for $\boldsymbol{V}$.
- Compute the objective functional f using (11) and the constraint g using (14).
- Compute the sensitivities of f and g with (18)-(21).
- Use the sensitivities to update the design vector using the Method of Moving Asymptotes (MMA) Svanberg (1987).
- Break loop if solution is considered converged.

4 Results

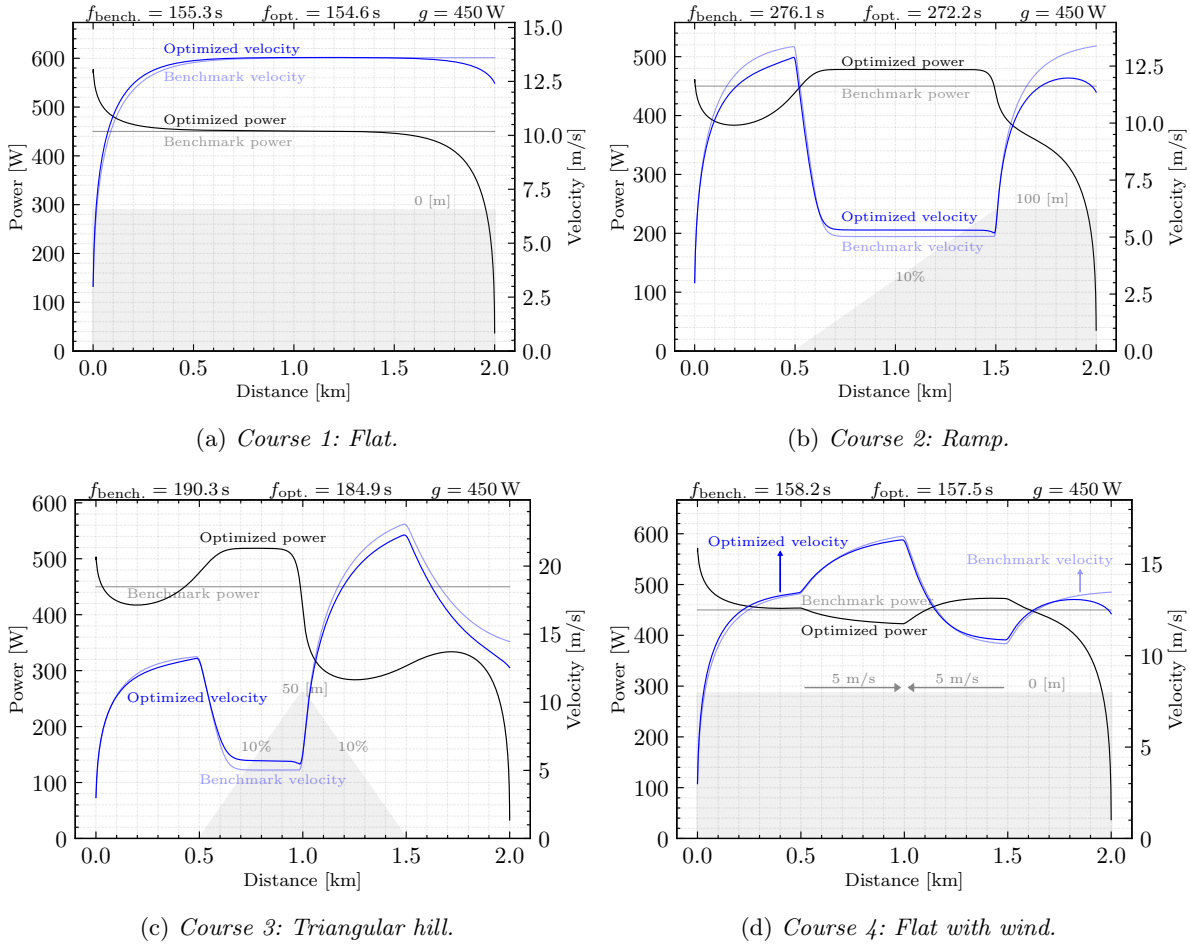
We implemented the method from Section 3 in the COMSOL Multiphysics® software (COMSOL Inc., 2022). We applied it to four hypothetical courses (from here on: "Elementary Courses") as well as one course adapted from a real world route (from here on: "Real Course"), labelled Course 1-5. The hypothetical courses are designed to isolate the most elementary conditions allowing for insights into the basic properties of the optimized strategies. Furthermore, we believe these elementary courses allow future work to readily make comparisons to our method and results.

The Elementary Courses 1-4 are 2 km long and were discretized using 2000 elements (see Figure 3). The Real Course 5 is 21.3 km and was discretized with 20 000 elements. For the numerical optimization to yield a result, a strictly positive initial velocity was required. This was chosen to be 3 m/s in all cases. The initial guess for the optimization algorithm was always a constant pacing strategy. At points where the conditions change (height gradient or wind) we included a transition zone of 20 m such that height profiles and wind functions always have 2 continuous derivatives. We elaborate on these numerical subtleties in Section 5.

4.1 Elementary Courses

Course 1 is a flat course with no wind. Figure 4a shows the benchmark and optimized pacing strategies for a standard cyclist. The optimized pacing strategy is 0.45% faster than the benchmark. The main strategy difference is a power surge in the very beginning and a sharp decrease in power towards the end. The result is a faster acceleration of the cyclist. At 100 m the velocity with a constant pacing strategy is 8.80 m/s and with an optimal pacing strategy the velocity is 9.25 m/s. The optimized pacing strategy at the extremities of the course seems to be a general result, as they appear for all courses considered here.

Course 2 includes a 1 km long climb with a constant gradient of 10% from 0.5 km to 1.5 km. Figure 4b shows the benchmark and optimized pacing strategies for a standard cyclist. The optimized pacing strategy is 1.41% faster than the benchmark. Optimizing leads to increased power



(a) *Course 1: Flat.*

(b) *Course 2: Ramp.*

(c) *Course 3: Triangular hill.*

(d) *Course 4: Flat with wind.*

Fig. 4: Model predictions of a standard cyclist (1) racing on the Elementray Courses employing a benchmark pacing strategy (faded black) with constant power output resulting in a benchmark velocity curve (faded blue) with finishing time $t_{\text{bench.}}$, and the optimized pacing strategy (black) resulting in the optimized velocity curve (blue) with finishing time $t_{\text{opt.}}$. The optimized pacing strategies are 0.45%, 1.41%, 2.84% and 0.44% faster than the benchmark for Course 1-4, respectively. All courses are 2 km long with conditions indicated in the background (dimgray). Course 1 is flat, Course 2 includes a 1 km climb with constant gradient of 10%, Course 3 includes a 0.5 km climb followed by a 0.5 km descend with constant gradients of ±10%, Course 4 is flat with a 0.5 km tailwind segment followed by a 0.5 km headwind segment with wind strengths of ±5 m/s

output on the uphill segment of the course compared to the flat segments in the beginning and end. Despite the very sudden change in height gradient at 500 m (within 20 m), the model yields a smooth increase in power from about 200 m to 650 m. For this course, it is advantageous to accelerate several hundred meters before the climb begins and to not reach maximum power output until after the velocity has dropped sufficiently due to gravity. On the other hand, the fall in power output at the top of the hill is almost instantaneous. To the best of our knowledge, features such as these have not been reported by previous studies.

Course 3 includes a 0.5 km climb with a constant gradient of 10% from 0.5 km to 1 km followed by a 0.5 km descend with a constant gradient of −10% from 1 km to 1.5 km. Figure 4c shows the benchmark and optimized pacing strategies for

a standard cyclist. The optimized pacing strategy is 2.84% faster than the benchmark. This is the largest relative difference in finishing times among the five courses presented. The optimized pacing strategy on the first half of the course is similar to Course 2 in Figure 4b, though the maximum power output on the climb is 40 W higher for Course 3 than Course 2. The optimized power output falls significantly over the top of the hill. Contrary to the climb segment, the power does not stabilize to a constant value on the descend. A minimum in power output is reached at 1.25 km followed by a slow increase until 200 m after the bottom of the hill. We notice that positive changes in the height gradient can result in a slow and dispersed increase in power output whilst negative changes in the height gradient may induce a quick and sharp decrease in power output.

Course 4 in Figure 4d is flat with a 0.5 km tailwind segment from 0.5 km to 1 km followed by a 0.5 km headwind segment. The wind strength is 5 m/s. Contrary to the force of gravity, the force of drag depends on the velocity of the rider. For this reason, we see a more complicated and non-linear response in the optimized pacing strategy to a constant wind compared to a hill with constant height gradient. We also notice that locally a change in wind conditions mainly induces a response in the rate of change of power output and not the power output itself.

Below is a summary of the main takeaways from the optimized pacing strategies on the Elementary Courses:

1. Power outputs are increased on uphill and headwind segments and decreased on downhill and tailwind segments.
2. There is a surge in power output at the very beginning of the course, to accelerate the cyclist, and a sharp decline just before the finish line, where the relative gain is lowest.
3. The increase in power output due to a positive change in height gradient may be slow and dispersed over several hundred meters. A negative change in height gradient can result in a quicker and relatively local drop in power output.
4. The response in power output due to wind is non-linear and velocity dependent. The power output may not stabilize to a constant value.
5. Locally, a change in wind condition mainly affects the rate of change of power output instead of the power output itself.

4.2 Real world course

Course 5 is a real world course in Norway. Professional cyclist Martin Toft Madsen has provided us with the data measured by his cycling computer and power meter while he was racing on this course. Madsen won gold at the Danish National Road Championship in the individual time trial discipline in 2016, 2017 and 2018 and silver in 2019, 2020 and 2021. The data includes GPS and altitude measurements along with his power output. The altitude and power data was imported to COMSOL Multiphysics® and a linear interpolation was used to create a height profile and power function compatible with our model. The resulting power function constitutes Madsen's pacing strategy on the course with a Normalized Power of $P_c = 385$ W.

Figure 5 compares his pacing strategy to the benchmark pacing strategy and the optimized pacing strategy constrained by the same Normalized Power. The finishing times for a standard cyclist (1) are 1814 s, 1810 s and 1788 s for the benchmark, Madsen's and the optimized pacing strategy, respectively. The optimized pacing strategy is 1.2% faster than Madsen's pacing strategy.

There was close to no wind and the course only had a single turn sharp enough so that Martin Toft Madsen had to use his breaks. This occurred at about 9 km and is the reason for the sharp trough followed by a peak in his power data. This makes the course and data obtained close to ideal for the purpose of comparing the predictions of our model to his choice of pacing strategy. For visual purposes, Madsen's pacing strategy is smoothed with a rolling average of about 53 m.

The main takeaways are that: 1) the method allows for efficient computation of optimized pacing strategies on real world courses; 2) the optimized pacing strategy is close to Martin Toft Madsen's pacing strategy. Regarding point 1, the time it takes for the method to converge on a solution for the optimized pacing strategy depends on numerous factors such as the implementation

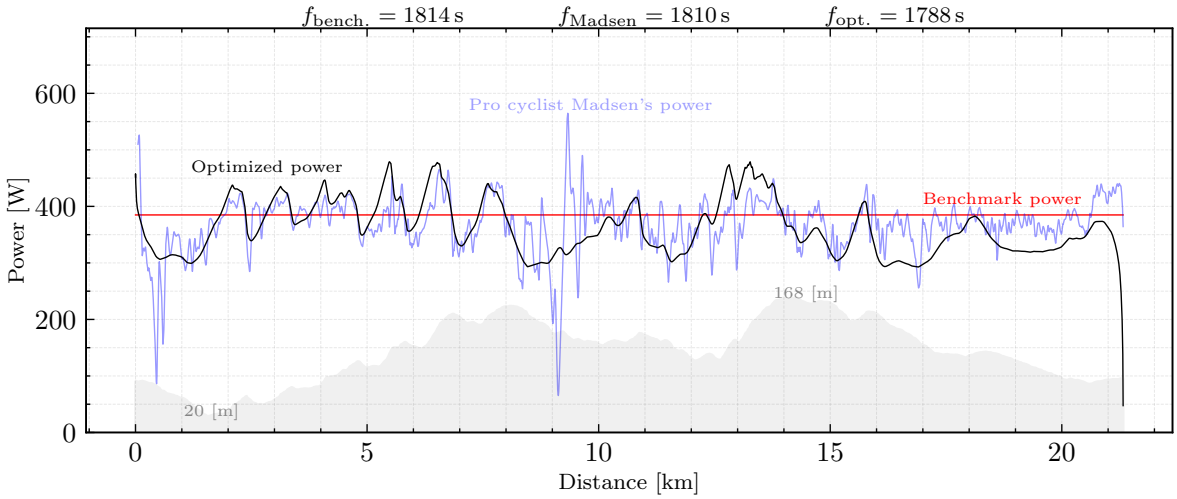


Fig. 5: *Course 5: Real world.* A 21.3 km real world course (dimgray) with the following plots: A benchmark pacing strategy (red) with constant power output, the power output data from professional cyclist Martin Toft Madsen (faded blue) racing on this course and the optimized pacing strategy (black). The Madsen data is smoothed with a rolling average of 53 m for visual purposes solely. The Normalized Power of each strategy is $g = 385\,\text{W}$ and the finishing times for a standard cyclist (1) are 1814 s, 1810 s and 1788 s for the benchmark, Madsen's and the optimized strategy, respectively. The optimized strategy is 1.2% faster than Madsen's strategy

chosen, the mesh for discretization and the criteria one sets for convergence. We noted that using COMSOL Multiphysics® and a mesh of 20 000 elements, a solution with finishing time within 1 s of the optimal pacing strategy could be computed with only 2-3 iterations of the method and within 25 s. Regarding point 2, a pertinent question is how close the optimized pacing strategies, computed within the model presented here, are to the actual optimal pacing strategies in the real world. As for all frameworks for computing optimal pacing strategies, this can only be answered with appropriate experimental studies. However, we consider the comparison to the pacing strategy of Martin Toft Madsen provided here as a first test of the relevance of our method for real world courses. In conclusion, we believe this example provides some credibility towards the capability of our method to efficiently compute relevant optimized pacing strategies for real world courses.

Finally, it is worth mentioning that realizing a 22 s (1.2%) difference in the finishing time by optimizing ones pacing strategy is sufficient to make a significant contribution to performances of cyclists competing in the ITT discipline.

5 Discussion

5.1 Physical model

In this paper, we have chosen a relatively simple equation for the governing physics - e.g. we have not differentiated between different types of frictions. Our focus is on the method of optimization and we believe the method can be adapted to more detailed equations. Though a crucial element that is unaccounted for here is the effects of sharp turns forcing the cyclist to brake. This is clear in Figure 5 where Madsen encountered a sharp turn and had to brake at the distance of 9 km. Zignoli and Biral (2020) considered a three-dimensional road geometry and cornering strategies in their paper, so there is some work on this problem within the field.

Any physical model of bicycle dynamics also faces the problem of constants measurement. The drag coefficient C_d, for example, might be very difficult to measure to a high accuracy. The drag term in the governing equation is therefore difficult to predict, highly dependent on both the rider and course conditions, due to the surface area A and effects of wind w, and at the same time it will have

significant influence on pacing strategies and finish times. Furthermore, the riders maximal available Normalized Power may also be difficult to estimate prior to the race, which in part motivates an optimal control formulation (Wolf et al, 2019; Zignoli and Biral, 2020). A feedback loop can be used to update the constants during a race when the measured power output and velocity do not match the predictions of the chosen physical model.

A third problem concerning the physical model is whether a cyclist can maintain a smooth and consistent power output as opposed to constantly oscillating slightly up and down. In Figure 5, the measured power output of Madsen clearly oscillates much more than the optimized power curve and this is after his data has been smoothed with a rolling average of 53 m. The model should take any oscillations in the power output due to human factors into account before optimization.

5.2 Normalized Power constraint

Our choice of constraint for optimization is based on the principle of Normalized Power and is motivated in Section 1.1.2. The formula is a $p = 4$ norm of the power output with respect to time as given in equation (13). Firstly, the choice of taking the fourth power might not be the most physiologically accurate way to constrain the power output of a cyclist. A higher exponent penalizes variations in power output more and therefore forces the cyclist to maintain a closer to the average power output throughout the course. Hence, a high exponent might be more suitable for longer courses where long-term endurance is imperative. A shorter course might allow for larger variations in power output and therefore require a lower exponent in the constraint. Hence, the constraint itself can be course dependent. Additionally, the physiology of a unique rider may also influence which constraint leads to the actual optimal pacing strategy of that rider.

Secondly, regardless of the choice of exponent, the constraint is solely global as a measure of the physiological toll of the workout as a whole. Consider a course that starts with steep hill followed by a long downhill segment. Our constraint would allow a pacing strategy with a very high power output on the hill as long as the power output on the long descend is sufficiently low. It does not consider the local question of whether or not the cyclist can actually maintain the high power output all the way up the hill. This is where more sophisticated and physiologically accurate constraints that consider the effects of different types of work done during a race, though these demand experimental verification. That being said, we believe our method of optimization can be adapted to more advanced constraints.

5.3 Numerical method

The differential equation model for the governing physics requires a choice of initial velocity. Since, we have discretized the course (Figure 3), the initial velocity affects the whole first element and not just the boundary. Because of this, the numerical method requires a non-zero initial velocity to allow the modelled cyclist to simply accelerate and start cycling. In general, the finer the discretization the lower this initial velocity can be, but it also depends on the course conditions in the beginning. Throughout this article, we have chosen an initial velocity of 3 m/s, which turned out to be a safe choice, in terms of not running into the mentioned problems, when discretizing with approximately 1 element per meter. The initial velocity is also related to initial power output of the optimized pacing strategy. A low initial velocity might lead to the optimized power output on the first few elements being unreasonably high. Hence, there is a trade-off between having closer to realistic initial conditions and the methods ability to yield a reasonable result or even a result at all.

Secondly, the optimization method requires an initial choice of pacing strategy, which amounts to choosing an initial design vector, in order to start the optimization loop. We always chose a constant power output, referred to as the benchmark pacing strategy, as the initial guess. The immediate question is whether or not the resulting optimized pacing strategies depend on this initial guess. That is, would we see different results if other initial pacing strategies were chosen? We will not fully address this problem but nonetheless, mention that for the real world course in Figure 5, the exact same optimized pacing strategy is obtained when using Madsen's power output as the initial guess.

5.4 Application to real world courses

With Figure 5 we conclude that our method may be applied to real world courses and, by comparing the pacing strategy of professional cyclist Madsen, that the optimized pacing strategy seems reasonable in this example. Clearly, this single application and single comparison to experimental data does not validate our methods applicability to real courses. A far larger study of how the methods predictions compare to experimental data is necessary.

We implemented the method with the COMSOL Multiphysics® software (COMSOL Inc., 2022). Because of this, the GPS data used for the real world course in Figure 5 had to be interpreted accordingly. We used a linear interpolation of the altitude measurements to create a height profile suitable for COMSOL. In general, any implementation of the method will require some manipulation of GPS data. A potential problem arises if the GPS measurements are either not accurate enough or too far apart to create an accurate representation of the actual course. We have not yet considered how to construct the most accurate height profile from the available data. For these reasons, a different implementation of the method might be more appropriate for applying the method to real world courses.

6 Conclusion

We have presented a design methodology for computing the optimal pacing strategy for a cyclist competing in the ITT discipline of professional cycling. We used a Finite Element formulation and adjoint sensitivity analysis to minimize the finishing time subjected to a physiological constraint based on the principle of Normalized Power.

We applied the method to four hypothetical courses to simulate various gradients and wind conditions as well as one real world course. With the hypothetical courses, we saw that the method can effectively compute the highly non-trivial optimized pacing strategies, which led to relative time gains between 0.45% to 2.84% compared to a benchmark strategy. Furthermore, we saw how the optimized pacing strategies can provide certain insights into local strategy questions such as how one should distribute power output at the beginning of a hill, over the top of a hill or when the wind changes. Lastly, we saw how the method can effectively be applied to a real world course. The computed optimized pacing strategy was similar to a strategy obtained from data from professional cyclist and ITT specialist Martin Toft Madsen cycling the same route. The optimized pacing strategy was 1.2% faster than the Madsen's strategy and 1.4% faster than the benchmark strategy.

In conclusion, the results show that the method is relevant as an effective optimization methodology for pacing strategies and capable of providing insights that have not been seen in the current literature.

7 Replication of results

The method described in Section 3 can be implemented in a variety of programming languages and software. We employed the COMSOL Multiphysics® software (COMSOL Inc., 2022). The following is available from the corresponding author upon reasonable request: an example of COMSOL file used, the data generated with COMSOL for the Figures 4 and 5, the data from professional cyclist Martin Toft Madsen (raw data and the cleaned data) used in Figure 5 and lastly all python scripts used for creating the plots in Figures 4a to 4d and 5 and the script to clean the data from Martin Toft Madsen for compatibility with COMSOL. Researchers and interested parties are welcome to contact the authors for further explanation, who may also provide further material under request.

References

Abbiss CR, Laursen PB (2008) Describing and understanding pacing strategies during athletic competition. Sports medicine 38(3):239–252

Allen H, Coggan AR, McGregor S (2019) Training and racing with a power meter. VeloPress

Atkinson G, Brunskill A (2000) Pacing strategies during a cycling time trial with simulated headwinds and tailwinds. Ergonomics 43(10):1449–1460

Atkinson G, Peacock O, Passfield L (2007) Variable versus constant power strategies during cycling time-trials: Prediction of time savings using an up-to-date mathematical model. Journal of Sports Sciences 25(9):1001–1009

Boswell GP (2012) Power variation strategies for cycling time trials: a differential equation model. Journal of Sports Sciences 30(7):651–659

Cangley P, Passfield L, Carter H, et al (2011) The effect of variable gradients on pacing in cycling time-trials. International journal of sports medicine 32(02):132–136

Coggan AR (2017) Training and racing using a power meter: an introduction. URL http://www ipmultisport com/ref_lib/Coggan_Power_Meter pdf Accessed 8

COMSOL Inc. (2022) Comsol multiphysics® v. 6.0. URL http://www.comsol.com/products/multiphysics/

Dahmen T (2016) A 4-parameter critical power model for optimal. In: Workshop Modelling in Endurance Sports, p 7

Dahmen T, Wolf S, Saupe D (2012) Applications of mathematical models of road cycling. IFAC Proceedings Volumes 45(2):804–809

Fayazi SA, Wan N, Lucich S, et al (2013) Optimal pacing in a cycling time-trial considering cyclist's fatigue dynamics. In: 2013 American Control Conference, IEEE, pp 6442–6447

Foster C, Snyder A, Thompson NN, et al (1993) Effect of pacing strategy on cycle time trial performance. Medicine and science in sports and exercise 25(3):383–388

Ganoung G (2022) What is normalized power? URL https://www.trainingpeaks.com/blog/what-is-normalized-power/, [Online; accessed 10-07-2022]

Garmin (2022) What is normalized power and how is it calculated on my garmin device? URL https://support.garmin.com/en-US/?faq=8r4llV3DFK5jc13BOHion5, [Online; accessed 10-07-2022]

Gordon S (2005) Optimising distribution of power during a cycling time trial. Sports Engineering 8(2):81–90

Hurley S (2021) Normalized power®: What it is and how to use it. URL https://www.trainerroad.com/blog/normalized-power-what-it-is-and-how-to-use-it/, [Online; accessed 19-04-2022]

Morton RH (1986) A three component model of human bioenergetics. Journal of mathematical biology 24(4):451–466

Morton RH, Billat LV (2004) The critical power model for intermittent exercise. European journal of applied physiology 91(2):303–307

Strava (2022) Strava training glossary for cycling. URL https://support.strava.com/hc/en-us/articles/216917147-Strava-Training-Glossary-for-Cycling, [Online; accessed 10-07-2022]

Sundström D, Carlsson P, Tinnsten M (2013) On optimization of pacing strategy in road cycling. Procedia Engineering 60:118–123

Sundström D, Carlsson P, Tinnsten M (2014) Comparing bioenergetic models for the optimisation of pacing strategy in road cycling. Sports engineering 17(4):207–215

Svanberg K (1987) The method of moving asymptotes—a new method for structural optimization. International journal for numerical methods in engineering 24(2):359–373

Swain DP (1997) A model for optimizing cycling performance by varying power on hills and in wind. Medicine and science in sports and exercise 29(8):1104–1108

Wolf S, Bertschinger R, Saupe D (2016) Road cycling climbs made speedier by personalized pacing strategies. In: icSPORTS, pp 109–114

Wolf S, Biral F, Saupe D (2019) Adaptive feedback system for optimal pacing strategies in road cycling. Sports Engineering 22(1):1–10

Yamamoto S (2018) Optimal pacing in road cycling using a nonlinear power constraint.

Sports Engineering 21. https://doi.org/10.1007/s12283-018-0264-3

Zignoli A, Biral F (2020) Prediction of pacing and cornering strategies during cycling individual time trials with optimal control. Sports Engineering 23(1):1–12