

# OPEN Individualized optimal strategy in team pursuit for track cycling

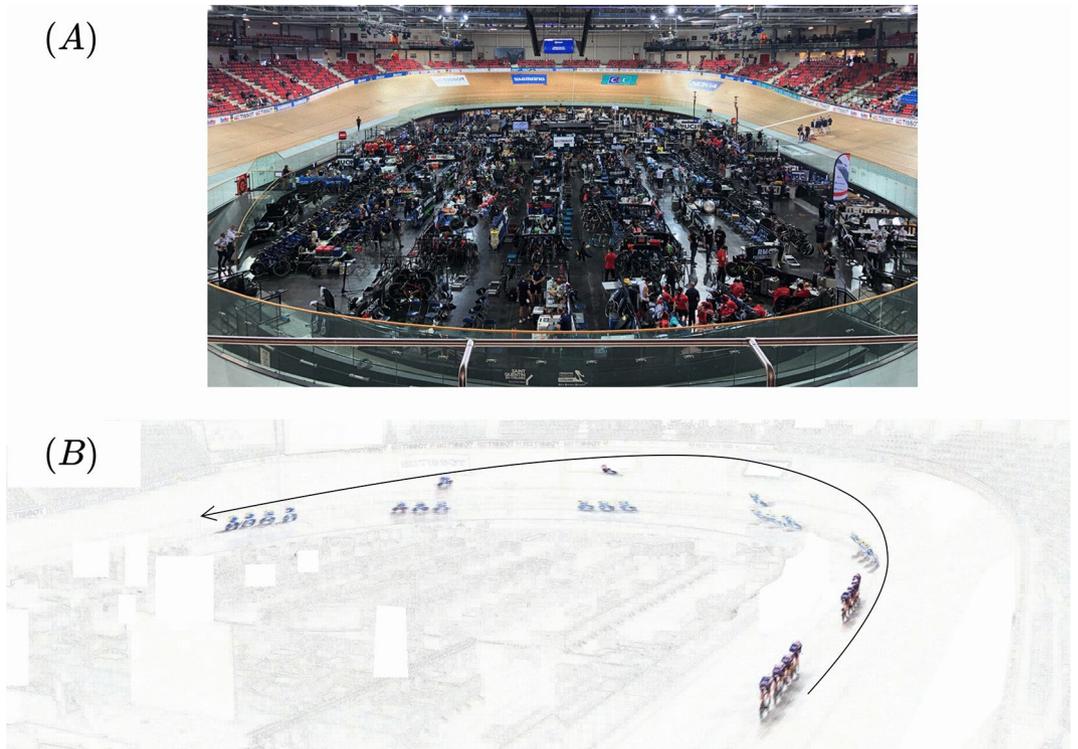
Alice Boillet<sup>1✉</sup>, Maxence Noble<sup>2</sup>, Iris Sachet<sup>3</sup>, Laurent A. Messonnier<sup>4,5</sup> & Caroline Cohen<sup>1</sup>

In track cycling, performance in the team pursuit depends on the mechanical and physiological abilities of each member of the team, but also on the choice of racing strategy. Athletes must cover the 4000 m of the race, sharing the effort between them in successive relays. This raises the question of the optimum strategy. We propose a method for resolving this question by coupling a mechanical model of the race to physiological models (digital twins) of the athletes. The mechanical model enables one to predict a theoretical finishing time for a given strategy, while the physiological model enables one to determine whether or not a given strategy is feasible. By coupling the two models and using numerical optimization, an optimal strategy for a given team can then be predicted. Simplified team composition case studies are explored. For each case studied, an optimal strategy to maximize performance is obtained and composed of a set of three variables: relay lengths, power values for each relay, and the starting order of cyclists. The proposed method can be used for real athletes and extended to other disciplines.

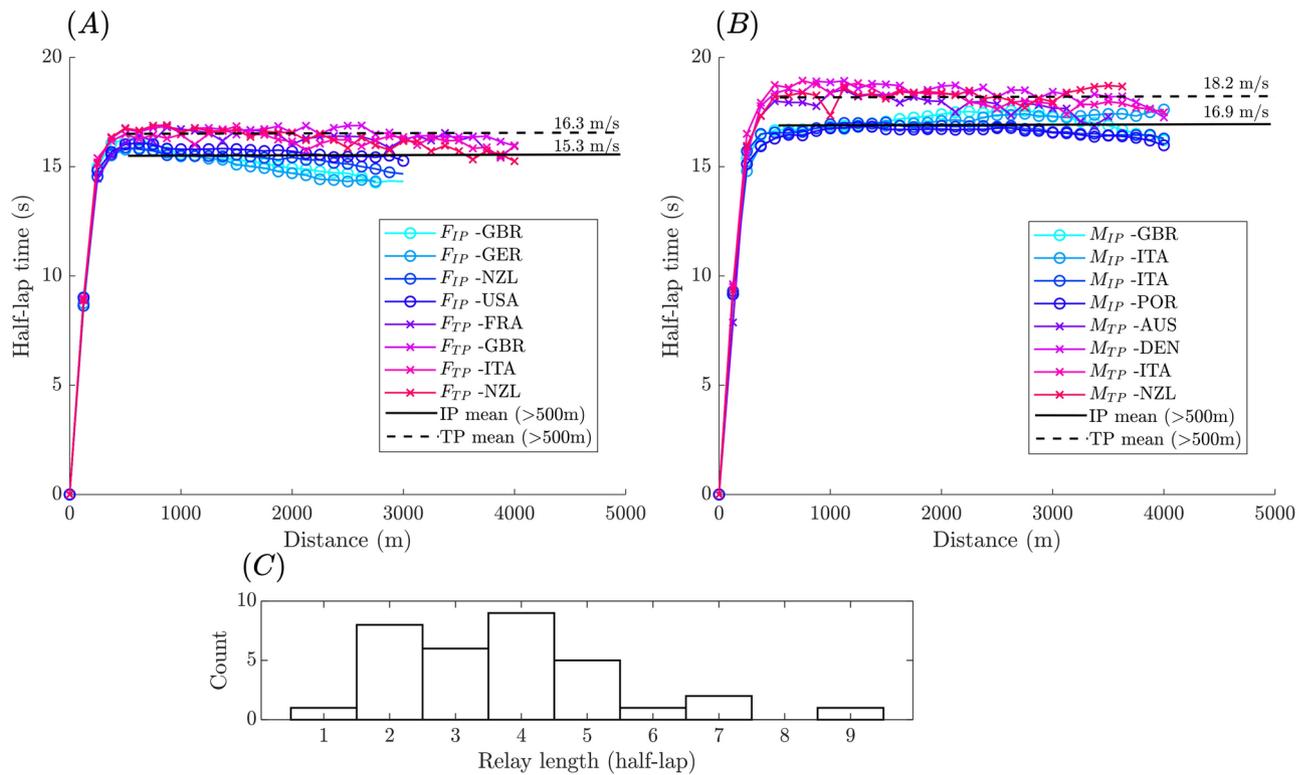
Track cycling is one of the original 43 sports events contested (by 280 male athletes from 13 nations) in the modern Olympic Games when they were restored in 1896 (Athens, Greece). Track cycling disciplines have appeared in all but one Olympic Games since, except for Stockholm 1912 when only a time trial was held. The team pursuit was first an Olympic event for men in 1908 with the women added to the program only in 2012.

The principle of the event is to cover 16 laps on a velodrome (Figure 1 (A)), namely 4000 m, as its length is 250 m at the black line. Each team starts with four cyclists. After the start, the cyclists line up in a peloton to take advantage of aerodynamic drafting. After some time in front, the cyclist in first position moves off to get the back of the peloton, thereby “passing the relay” to the second cyclist who becomes first. Several relay passages take place during the event, often with each cyclist performing and passing two relays. They pass the relay on a whole number of half-turns as they use the geometry of the track in the turns to pass the relay, converting kinetic energy to potential energy climbing and then getting back kinetic energy going down, once passed by the team as it can be seen on Figure 1 (B). Their order of passage is dictated by the order adopted at the start. In championships, the aim in qualifying races is to minimize the time taken by the third cyclist to reach the 4000 m mark, so it is not uncommon for the race to finish with only three cyclists if one “cracks” during the race (ie. stops before the end of the 4000 m). The teams with the fastest times then compete in duels. Each team starts from the opposite position on the track. The aim is either to obtain a shorter time than the opponent for the third cyclist on the 4000 m or to catch up with the other team, hence the term “pursuit”. In any configuration, cases, where a team is caught up by the opponent, are relatively rare, and the overall aim of the teams remains to minimize the time taken for the third cyclist to complete the distance. For a given team (four cyclists), several race configurations are conceivable. The variables on which a team can play are of three kinds: i) the starting order of the cyclists, which conditions the order of the relays, ii) the length of the relays (number of half-laps), iii) the speed maintained by each cyclist when in first position, which is directly related to the power injected by the first cyclist and their mechanical characteristics via their equation of motion. So, the following cyclists necessarily ride at the same speed to remain aerodynamically sheltered. However, only strategies that can be sustained by athletes can be considered. The search for the optimum strategy is therefore a problem of optimization (minimization of running time) under constraints (athletes’ physiological capacities). A similar race exists for individual athletes (3000 m for women, 4000 m for men). As expected, speeds in these individual events are lower than in their team counterparts as can be seen in Figure 2 (A) and (B)<sup>1</sup>. For both women (A) and men (B), the average speed to cover a half-lap from the 500 m mark to the end of the race is smaller for the individual event (by 6.1 % for women and by 7.1 % for men). However, passing a relay presents a technical risk (if the cyclist does not reposition well) and causes the peloton to lose a bike length (around 2 m). It therefore

<sup>1</sup>LadHyX, UMR 7646 du CNRS, Ecole polytechnique, 91120 Palaiseau, France. <sup>2</sup>CMAP - Centre de Mathématiques Appliquées - Ecole polytechnique, 91120 Palaiseau, France. <sup>3</sup>Fédération Française de Cyclisme (FFC), Saint Quentin en Yvelines, Montigny-le-Bretonneux, France. <sup>4</sup>Université Savoie Mont Blanc, Laboratoire Interuniversitaire de Biologie de la Motricité, 73000 Chambéry, France. <sup>5</sup>Institut universitaire de France (IUF), Paris, France. ✉email: alice.boillet@polytechnique.edu



**Fig. 1.** The team pursuit event (A) Track cycling velodrome; (B) Chronophotography of a passing of a relay (using the geometry of the track), the arrow indicates the chronology.



**Fig. 2.** Observed strategies in the 2023 World Track Championship (Glasgow) finals: (A) Best four female team and individual pursuits half-lap speed; (B) Best four male team and individual pursuits half-lap speed ; (C) Histogram of observed length of relays in team pursuits (F-FRA, F-GBR, M-NZL, M-DEN). (A) and (B) data comes from the official timing website<sup>1</sup>; (C) data from observation on race days.

appears that racing as a team is advantageous, as it allows effort to be shared between team members. Cyclists not in the lead tire less and/or recover more than the frontrunner. They are indeed aerodynamically sheltered and benefit from drafting. This well-known effect in road cycling peloton leads to drag reduction of up to 90% of the following cyclists in triangular-shaped pelotons of more than a hundred cyclists<sup>2</sup>, when compared to a cyclist riding alone. In the track time-trial case, where cyclists are aligned one after the other and limited to four, this drag reduction is less important but can lead to a drag reduction of up to 40%<sup>3–6</sup>. In other words, at a given speed, a sheltered cyclist in a team pursuit would have to generate as much as 40% less power (and save the same proportion of energy) if they are riding in the peloton rather than in the first position. As a result, the rider in first position can generate more power on average than in the case of a single race, around 47%<sup>3</sup>. Empirically, the most successful teams use relays of around 4 half-laps as can be seen in Figure 2 (C). Cyclists do not all perform relays of the same length, and when they take several relays in a race, they are not necessarily of the same length either. The dichotomy between fatigue at the front and protection in the peloton, the approximative 2 m cost of the relay, and the benefits of sharing the effort suggest the existence of a strategic optimum. This raises the question of knowing this optimum and how to determine it. Previous work has looked at optimizing performance for team pursuit, particularly in terms of aerodynamic interactions between cyclists<sup>5,7</sup>. Others have explored the question of team cycling strategies, either for track<sup>8</sup> or road cycling<sup>9</sup>, but in a general way, without attempting to model specific athletes.

To tackle more specifically the optimal strategy question, a coupled mechanical and physiological approach is appropriate. This is also the approach taken by Keller in his pioneering work on the optimal effort management strategy for an individual runner as a function of race distance<sup>10</sup>. His method and the simplified models of running mechanics and energy producible by an athlete that he uses enable him to solve the problem analytically. For the optimal team pursuit strategy problem, we propose a similar approach in terms of methods, but where the models are more realistic and thus, complex. The mechanical equation is not linearly simplifiable in the case of cyclists, unlike in running. Indeed, the main resistive force in running is taken to be proportional to speed by Keller, which can be interpreted as a linear decrease of exercisable force with speed by the athlete or friction proportional to running pace and thus speed. In cycling, the main resistive force is aerodynamic and therefore proportional to speed squared. As far as the physiological model is concerned, the simplified one proposed by Keller is attractive for its simplicity, but it does neither take into account the anaerobic lactic pathway nor the potential recovery (for following cyclists) and it does not allow metabolic predictions. Therefore we propose to tackle the optimal strategy problem in track cycling team pursuit using the same kind of approach, but with models of higher complexities, and using numerical optimization as an analytical one is no longer possible.

In the first part of this paper, the mechanical model of team pursuit and the physiological model are presented. The latter is based on the creation of digital twins of athletes. The principle of the numerical optimizations using these coupled models to obtain optimal strategy results is also explained. In the second part - the case study section -, results obtained by the method presented in the first part for simplified team configurations are analyzed.

## Optimization of a coupled model of the race

### Mechanical modeling

The mechanical model must be able, for a given tested strategy, to predict a finishing time and the power that each athlete would have to generate. An energetic approach is therefore used. It will also take into account the mechanical characteristics of each athlete.

Intra-cycle power or speed variation in the cycling motion is not considered. Indeed, left-right alternation of the pedaling movement (no recovery phase) and no relative movement of the cyclist with respect to their equipment make those variations negligible in the overall speed prediction. The motion equation of a single cyclist can be stated as follows (no height variation as they stays on the “black line”, the shorter line of the velodrome), as developed and validated by Martin<sup>11</sup>:

$$Mv_{cycl} \frac{dv_{cycl}}{dt} = P_{mec} - 1/2\rho_{air}Sc_Dv_{cycl}^3 - \mu Mgv_{cycl} \quad (1)$$

Where resistive terms in  $v^3$  are gathered together and those in  $v$  too. Where  $M$  is the mass of the system {cyclist + equipment},  $v_{cycl}$  is the speed of the cyclist,  $\rho_{air}$  is the density of air,  $S$  is the frontal area of the cyclist,  $c_D$  is the global drag coefficient (for all terms proportional to  $v^3$ ),  $\mu$  is the coefficient of friction of the wheels,  $g$  is gravity on Earth. The left-hand side of the equation corresponds to the system's energy variation. The right-hand side shows the mechanical power supplied by the athlete  $P_{mec}$  and the power dissipated by the resistive forces. The latter come from, in the order shown: aerodynamic drag resistance linked to the system's movement in a fluid; friction from the wheels rolling on the track.

The single cyclist equation can be used to compute the speed of cyclists, considering always the leading cyclist. The features of a strategy now have to be stated. To do so, the power generated by the first cyclist is considered constant on a given relay. Indeed, the average power over a lap remains relatively constant. Observed oscillations of speed and power are due to the geometry of the velodrome but do not influence the prediction of the finish time. Indeed, on the velodrome's curves, cyclists lean inwards, and the speed perceived by their bodies (the main source of aerodynamic drag) decreases and becomes smaller than the forward speed of the wheels. This phenomenon causes an artificial drop in their  $Sc_D$  if the speed used to calculate it is the forward speed of the bicycle<sup>12,13</sup>. Although the geometry of the velodrome is not explicitly included in the mechanical model, this effect is taken into account since the  $Sc_D$  used for the cyclists are those derived from experimental power and

speed data on velodromes. Used drag coefficients are therefore the average of the measured ones on straights and curves.

Concerning relays, passing one on the straight sections would be technically and physiologically more challenging. It would require the cyclist to slow down (that is difficult on a track bike due to the fixed gear and lack of brakes) and then to re-accelerate to keep up with the peloton (which is energy consuming and therefore less energetically efficient). Therefore, we assumed that relays would always be made over whole half-lap distances, implicitly accounting for the velodrome's geometry. Moreover, it is assumed that the cyclists always pass relay in a "standard" manner, i.e., the leading cyclist moves to the back of the peloton once their relay is over. There are, however, various alternative approaches, such as having multiple cyclists (e.g., the first and second riders) move apart simultaneously, or allowing the cyclist passing the relay to fall back to a position other than last (for instance, falling into second position while the third and fourth cyclists slow down slightly to create space). This has been considered and shown to be preferable in some cases for the team time trial in road cycling<sup>14</sup>. However, on track, even the "standard" relay technique is challenging and not always perfectly executed, even at the highest levels. Since the presented work focused on strategies that are practically feasible without requiring significant changes in technique or new skill acquisition, we chose to only consider the "standard" approach. Therefore, the initial order of the cyclists gives the order at which they will do their relays.

The strategy can then be stated using the three following variables:

- The starting order of the athletes:  $[A, B, C, D]$
  - The list of the numbers of half-laps per relay:  $n_{list} = [n_1, \dots, n_N]$
  - The list of the values of average mechanical power produced by the leading cyclist per relay:  $P_{list} = [P_1, \dots, P_N]$
- Where  $N$  is the total number of relays done in the race. With these notations,  $n_i$  is the number of half laps of the  $i^{th}$  relay,  $P_i$  is the power in Watts produced by the cyclist in front of the peloton during the  $i^{th}$  relay. Note that  $N$  always verifies  $\sum_{i=1}^N n_i = 32$ , since the total distance (4000 m) is covered by 32 half-laps. Each cyclist achieves a known and repeatable start when placed in the first position during the early part of the race: characterized by a peak power  $P_{max}$  reached at a given time  $t_{max}$ . Therefore for the start, the generated power of the front cyclist is modeled as follows :

$$P_{mec}(t) = \mathbf{1}_{t < t_{max}} [(1 - e^{3t/t_{max}}) P_{max} / (1 - e^{-3}) - P_1] + \mathbf{1}_{t > t_{max}} [(P_{max} - P_1) e^{-(t-t_{max})/(1.3t_{max})}]$$

Where  $\mathbf{1}$  is the identity function equal to one if the condition in indices is met, and equal to zero otherwise. Of note, for the first 125 meters, the cyclists ride standing on pedals. During this phase, the higher  $S_{CD}$  of cyclists in this position are also taken into account in the equation. From those three items describing the strategy and the start parameters, the  $P_{mec}$  of the leading cyclist at each time of the race can be determined. Reusing Equation 1 their speed can be computed and therefore the speed of the peloton (since all cyclists are assumed to be at the same speed). Moreover, relay changes have to be taken into account. For each relay passage, a distance  $d_{pen}$  is lost by the peloton (one cyclist length, approximately 2 m) and the cyclist who is passing the relay has to continue producing the power they were producing in the first position during the transition from first to last position in the peloton (approximately 3 seconds, as measured on elite team pursuit data recorded thanks to an instrumentalized drive, such as shown on figure 3(A)).

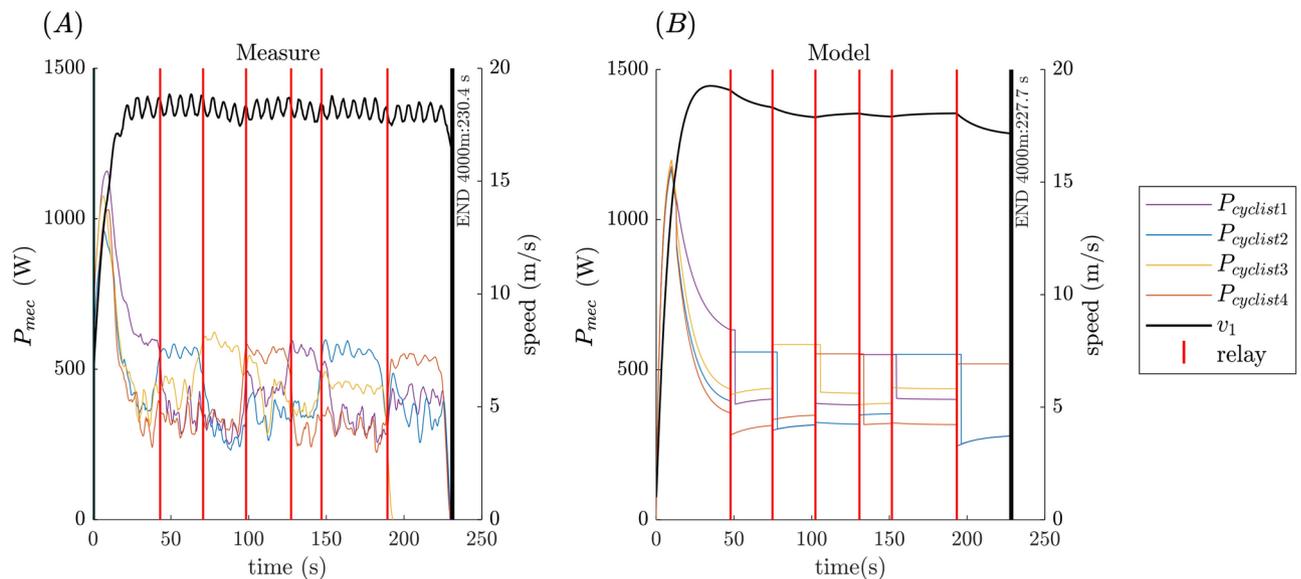
Each time a cyclist takes the lead, they remain in the first position until they have covered the corresponding distance specified in  $n_{list}$ . For example, if the two first values of  $n_{list}$  are 4 and 3 ( $n_{list} = [4, 3, \dots]$ ): the cyclist starting in the first position at the beginning of the race will cover 4 half-laps in their first relay; the cyclist starting the race in second position will then cover a relay of 3 half-laps. In this case, the mechanical characteristics of the first cyclist are used in Equation 1 to calculate the speed of the peloton for the first 4 half-laps, and the associated power is imposed on their digital twin. Once this cyclist has covered the distance of 500 meters ( $4 \times 125$  m, calculated using the modeled speed), they pass the lead. Either due to exhaustion or normal relay transition, the cyclist moves back (to the last position in the peloton), and the second cyclist becomes the leader. This second cyclist then leads the peloton until they have covered the next relay length of 375 meters ( $3 \times 125$  m).

Then, the power that each following cyclist has to generate can be determined. To do so, the equality of speeds and the knowledge of the mechanical characteristics of each cyclist (mass, drag coefficient  $S_{CD}$ ) are used. When a cyclist is not in front, their drag coefficient is reduced. The average of the reduction over the three sheltered positions is taken:  $35\%^{3-6}$ . This value is coherent with data from team pursuit races (power and velocity records). Let  $Y$  be the leading cyclist during the  $i^{th}$  relay, producing the power  $P_i$  and riding at speed  $v_1$ . A following cyclist  $X$  will have to generate the power expressed as in Equation 2.

$$P_{X, follow} = m_X / m_Y P_i - 1/2 \rho_{air} v_1^3 \left( \frac{S_{CD,Y}}{m_Y} - \frac{0.65 S_{CD,X}}{m_X} \right) \quad (2)$$

Where  $S_{CD,X}$  and  $S_{CD,Y}$  are the drag coefficients of  $X$  and  $Y$  respectively,  $m_X$  and  $m_Y$  are their respective masses.

Based on all these considerations, a given race can be modeled, and the theoretical finish time for a given strategy can be deduced numerically, as well as the mechanical power that each cyclist has to produce for this strategy, as schematized in Figure 5. All simulations are performed with MATLAB R2021b software using a forward Crank-Nicolson process. An example from a team at a 2023 UCI Track Nations Cup team pursuit event is given in Figure 3. The power records and the speed signals were obtained using an instrumented drive (Phyling, Palaiseau, France). Experimental data are shown in Figure 3 (A), with a measured arrival time of 230.4 seconds.



**Fig. 3.** Power generated by the four cyclists and team speed for a 2023 UCI Track Nations Cup team pursuit event: **(A)** Experimental data extracted from an instrumentalized drive (Phylog). **(B)** Modeled race: for each relay after a typical start, the power of the leading cyclist is averaged. The speed of the team  $v_1$  is obtained by solving Equation 1 and the powers generated by the following cyclists are calculated with Equation 2.

The cyclist starting in the first position makes a start with peak power at 1200 W at 10 seconds. The succession of different relays is highlighted with vertical red lines, marking the transitions, between which the cyclist in the first position produces a higher power than the followers (aerodynamically sheltered). The oscillations of the speed caused by the track geometry are also visible. Figure 3 (B) exhibits the corresponding modeled race. The starting order and  $n_{list}$  are taken as done with the team.  $P_{list}$  is determined by averaging the measured power of the leading cyclist on each relay. The obtained three strategy variables are then injected into the mechanical model to predict the speed of the peloton, the power that each cyclist has to generate to be at the same speed as the peloton and, a theoretical arrival time. The predicted 4000 m finish time is 227.7 seconds, which is 1.1% smaller than what is experimentally measured. Although acceptable, several parameters may explain the error obtained. Firstly, there may be variations in the values of the mechanical parameters used to predict speed:  $SC_D$  of the athletes if their position on the bike varies a little, variation in density and friction coefficients with pressure, temperature, or humidity conditions. Next, the distance to be covered in the model is  $4000 + N \times 2 \text{ m}$ . This is true if cyclists stay perfectly on the “black line”. In practice, however, cyclists may deviate slightly from it, often covering around 1% (2.5 m) more distance per lap, which would correspond to the observed error. In any case, all the sources of error mentioned are systematic errors. While the model does not provide a second-readiness prediction of performance, it is still suitable for comparing different strategies.

For any strategy, an arrival time can be predicted. However, it lacks a feasibility dimension. For this, the mechanical model presented has to be coupled with a physiological model.

### Physiological modeling

Using the mechanical model described above, the theoretical power each cyclist should generate for a given strategy can be determined. To find out whether the effort is feasible for each cyclist in the team, digital twins of the cyclists are used. These twins correspond to a hydraulic energy model inspired by the Margaria-Morton model and are created following the procedure described in a previous article<sup>15</sup>. A scheme of the model is available in Figure 4. The hydraulic analogy models the different energy pathways involved in efforts lasting from 1 to 15–20 minutes as tanks of fluids: the oxidative (O), non-oxidative lactic (G), and non-oxidative alactic pathways (P). This choice is appropriate since these are the main metabolic pathways involved in the type of effort under consideration<sup>16</sup>.

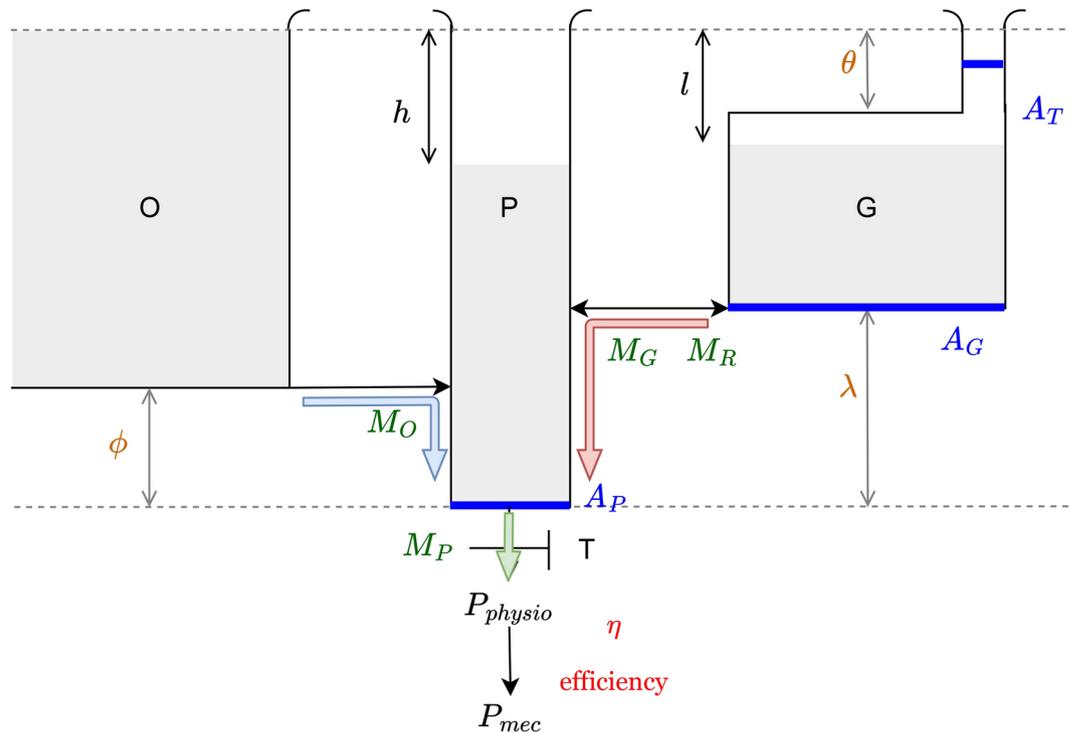
The mechanical power produced is linked to the physiological power generated by an athlete via an efficiency coefficient:  $P_{mec} = \eta P_{physio} \cdot P_{physio}$  corresponds to the flow leaving P via a tap T whose flow rate can be controlled as long as it remains below the maximum flow rate  $P_{max}$  (depending on the model drain state). When full (i.e. before the start of an exercise), this maximum flow rate is  $M_P$ . P is of finite capacity ( $A_P$ ) and is recharged by O, of infinite size, but the flow between O and P is limited ( $M_O < P_{max}$ ). P is also connected to G, via a conduit of maximum flow  $M_G$  in the direction from G to P and  $M_R$  in the direction from P to G. G is also of finite capacity ( $A_G$  and  $A_P$ ), but the fluid/power that can be drawn from it is not limiting ( $M_G > P_{max}$ ).  $P_{max}$  depends on how much fluid remains in G. Thus, a net flow from G to P corresponds to energy expenditure (which leads to fatigue and a decrease in  $P_{max}$ ), while a net flow from P to G corresponds to recovery (increase in  $P_{max}$ ). The bases of O and G are respectively  $\phi$  and  $\lambda$  higher than the base of P. The upper part of G is reduced in size. The wide part of G is lowered by  $\theta$  to the top of P. The evolution of the system is governed by the evolution of flows between the

**Model parameters:**

Maximum flux (W)  
Capacity - Area (J)  
Height (-)

**Metabolic pathways:**

→ : Oxidative pathway  
→ : Non-oxidative lactic pathway  
→ : Non-oxidative alactic pathway



**Fig. 4.** Modified M-M energetical model used to create digital twins of athletes (from a previous article<sup>15</sup>).

reservoirs and the open state of T. Using the physiological characteristics of each athlete, their digital twin can be parameterized. In particular, the model takes into account two important variables in endurance performance: a critical power  $P_{crit}$  and the ability to generate energy above this power  $E_{non-ox}$ <sup>17,18</sup>.

For ease of understanding, duplicates of a single digital twin model and variations of it will be used in the following. This twin corresponds to a female track cycling athlete of international level for the year 2022 (age = 23 years old,  $\dot{V}O_{2,max} = 3.97$  L/min,  $P_{crit} = 285$  W, weight = 54 kg). This study was approved by the French National Ethics Committee, and all participants provided written informed consent for the study. All experiments were conducted in accordance with the Declaration of Helsinki guidelines (ClinicalTrial : NCT05314543). The required incremental and 3-minute all-out tests were performed on a cycling ergometer (Excalibur Sport; Lode, Groningen, the Netherlands). The power record of maximal effort was obtained using an instrumented drive (Phyling, Palaiseau, France). Her resulting model parameters are as follows:  $\phi = 0.3$ ,  $\theta = 0.476$ ,  $\lambda = 0.18854$ ,  $M_O = 1.38 \cdot 10^3$ ,  $M_P = 4.86 \cdot 10^3$ ,  $M_G = 1.036 \cdot 10^4$ ,  $M_R = 4.14 \cdot 10^3$ ,  $A_P = 3.147 \cdot 10^4$ ,  $A_G = 3.4017 \cdot 10^5$ ,  $A_T = 1.145 \cdot 10^4$ ,  $\eta = 0.20205$ . We can impose a theoretical power signal to be supplied by a digital twin and predict i) whether or not this is feasible, ii) if not feasible, when exhaustion is reached (data used later in the search for optimal solutions).

### Coupling and optimization

Consider a given team of athletes and a given strategy for them, i.e., we are given their mechanical characteristics ( $SC_D$ , masses, typical start) and a certain choice for the starting order,  $n_{list}$ , and  $P_{list}$ . First, the mechanical model of a team pursuit race given in the “Mechanical modeling” section generates a prediction for the theoretical arrival time and the power that each cyclist should generate throughout the race to perform that strategy; note that these values are obtained independently of the true capacities of the athletes. To assess how realistic this strategy is, the power data obtained at the first stage are then imposed on the digital twins created as described in the “Physiological modeling” section. For a given strategy (defined by  $n_{list}$ ,  $P_{list}$ , and initial order) applied to four cyclists, it takes approximately 1.5 sec. to simulate the predicted arrival time and the time evolution of the states of the corresponding digital twins. The simulations were conducted on a computer with the following specifications: Intel Core i7 CPU 11850H, 8 cores @ 2.50 GHz. Two outcomes are possible: either the obtained data is realistic, i.e., the strategy is feasible, or it is unrealistic, and the strategy is discarded. Our goal is to find a strategy that minimizes the arrival time (output from the mechanical model) while producing a truthful description of physiological data (constraint from the physiological model), which can be seen as a specific

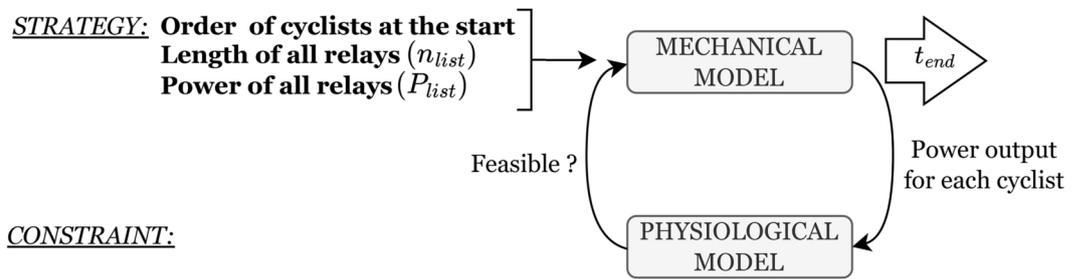
constrained optimization problem. The two models (mechanical and physiological) are thus coupled, as depicted in Figure 5.

Given the complexity of the approach, it is unclear if the considered optimization problem admits a global optimum, and how to find it in the case where it exists. Since testing all the possible strategies is unrealistic, a heuristic approach is proposed to find local optima. Ideally, we seek to find a strategy described by three variables (starting order,  $n_{list}$ ,  $P_{list}$ ) that minimizes the target quantity  $t_{end}$ , which corresponds to the arrival time of the third cyclist generated by the mechanical model, while verifying hard constraints imposed by the physiological model. In this paper, a relaxation of this complicated problem is proposed by minimizing on the set of all strategies a new target quantity defined by  $t_{end} + T_{penalty}$ , where  $T_{penalty}$  is a positive penalty that is taken proportional to the “infeasibility” of the given strategy. To be more specific, for any strategy that verifies the constraint (i.e., it is validated by the physiological model, see Figure 5), we take  $T_{penalty} = 0$ ; on the other hand, for any unrealistic strategy, i.e., in the case where  $K \geq 1$  cyclists are exhausted before  $t_{end}$ , we take  $T_{penalty} = \sum_{i=1}^K (t_{end} - t_{exh,i})$  where  $t_{exh,i}$  is the predicted time of tiredness of the  $i$ -th exhausted cyclist. Defined as such, the model constraint is directly added to the target quantity, which thus favors feasible strategies in the minimization. This results in an unconstrained optimization problem for which a numerical approach can be conducted, as we detail below.

Consider a fixed value for the variable describing the initial order of the cyclists. Our numerical optimization procedure operates on two levels:

1. Assume for now that the variable  $n_{list}$  is fixed. Following the method proposed by Wagner<sup>19</sup>, the variable  $P_{list}$  is optimized via an evolutionary search algorithm named *Covariance Matrix Adaptation Evolution-Strategy* (CMA-ES)<sup>20</sup>. This algorithm is stopped either when the iterations no longer prove to bring any improvement in minimizing the target quantity or if the maximal settled number of iterations is reached (here, we consider at most 40 iterations). This optimization stage is conducted with three different random seeds, and the best of the three results is kept.
2. Optimizing  $n_{list}$  in our framework is not an easy task since it takes values in a discrete but large space. To this aim, we design a specific optimization algorithm, which follows a greedy paradigm, and with restrained values in [2, 9]. We now detail the algorithm.  $n_{list}$  is taken as a list of constant size 16 (since at least 2 half-laps are performed per relay); at the end of this optimization step, only the first part of  $n_{list}$  is kept such that the total number of half-laps does not exceed 32. The algorithm proceeds by sequentially generating proposals for  $n_{list}$  that produce a decreasing value for the target quantity. At each iteration, the current proposal for  $n_{list}$  gives birth to new proposals, namely ‘neighbors’ (this stage is called mutation), from which only the proposal with the minimal value for the target quantity is kept. This approach is greedy in the sense that we iteratively explore the best direction among all the possibilities for  $n_{list}$ , which does not guarantee finding a global optimum. The mutation gives birth to 20 neighbors, by independently modifying  $n_i$  with probability 0.4 for any  $i \in [1, N]$ . This mutation consists of (i) either incrementing or decrementing  $n_i$  with step 1 (with uniform probability 1/2) when  $n_i \notin \{2, 9\}$ , (ii) incrementing when  $n_i = 2$  or (iii) decrementing when  $n_i = 9$ . Multiple settings for the initialization of  $n_{list}$  are considered, where the number of half-laps per relay is taken constant through the race. In practice, this value varies from 2 to 8. The greedy algorithm naturally stops when the target quantity does not decrease anymore among the neighbors obtained after the mutation. This optimization stage is conducted with 3 different random seeds for each initialization setting, and we keep the best of the obtained results. From a mathematical standpoint, it would be ideal to explore many different initializations. However, practical observations and physiological constraints suggest that the search should be confined to a specific range. For the variable  $n_{list}$ , initializing with equal relay lengths for all cyclists is a reasonable approach, given that we are considering four identical or nearly identical athletes. This ensures an equitable distribution of effort. The initial values for  $n_{list}$ , ranging from 2 to 9, correspond to relay lengths commonly observed in practice (see Figure 2). For the power variable  $P_{list}$ , a range of 350 to 450 watts is selected, as it reflects the maximal sustainable power output over time intervals associated with relay lengths of 2 to 9 half-laps. Multiple initializations and simulations are crucial to the optimization process. While the initial  $n_{list}$  values remain the same, variations in relay lengths are introduced by adjusting  $n_{list}$  values within the defined range. Likewise,  $P_{list}$  is randomly initialized between 350 and 450 watts, using three different random seeds for each iteration. This method ensures a broad exploration of potential solutions during optimization, improving the robustness of the results. Despite the consistency of outcomes for a given simulation with a fixed random seed, using different random seeds in the optimization process may yield slightly different optimal times. This variability is inherent to the process and helps in exploring different potential solutions, thereby enhancing the overall effectiveness of the optimization.

Hence, for a given order of the cyclists at the departure, an optimal racing strategy is obtained by the procedure described above, which finds local optimal values for  $n_{list}$  and  $P_{list}$ . We acknowledge that this method of exploration to obtain solutions to the unconstrained minimization problem is not exhaustive, since we have no guarantee to reach a global optimum. Nonetheless, it allows us to obtain plausible results (namely, local optima) in a relatively short time. For a given initial order of cyclists and an initialized  $n_{list}$  (where all relays are of equal length), it takes approximately 20 to 25 minutes to complete both parts of the optimization process and obtain optimal  $n_{list}$  and  $P_{list}$ . For example, computing for  $n_{list}$  initialized with constant relay lengths between 2 and 9 (integer only), it takes approximately 2.5 to 3 hours per given starting order to obtain the final result. Finally, remains the optimization of the order of the cyclists. In theory, one could test the  $4! = 24$  combinations. For simplicity’s sake and ease of understanding, a simplified approach is hereafter taken. First, we assume that the four cyclists have the same athletic profile (hence, the departure order does not matter). Then, we assume that the cyclists have only two different athletic profiles (including one profile for three cyclists), which leads to 4



**Fig. 5.** Evaluation of the strategy by model coupling.

different combinations of the initial order. In practice, coaches often consider a limited number of possible initial orders. Therefore, one could restrain the calculations to the considered orders.

### Case study

In the following section, the optimization results for simplified cases and hypothetical teams are presented. This case study approach aims first at providing insights into the potential of such an approach and the use of the coupled physio-mechanical model. Then, in the context of the team pursuit event, the simplified cases allow us to draw some general conclusions in terms of generic teams rather than specific athletes' combinations. First, the optimum for a "homogeneous" team (four identical athletes) will be studied, focusing on the relay plan (lengths and power). Then the optimal position of an athlete with varied physiological capacities will be addressed.

#### Four identical cyclists

The first studied case corresponds to a combination of four identical athletes, which frees from the consideration of order and allows to focus on the optimal strategy in the case of four athletes of similar levels and physiological profiles. The goal was to create a generalized model that could produce broadly applicable conclusions, with variations between cyclists kept within realistic bounds. In addition, the use of specific data from athletes (of team currently competing) raises confidentiality issues. Instead, we preferred to focus on a generalized scenario.

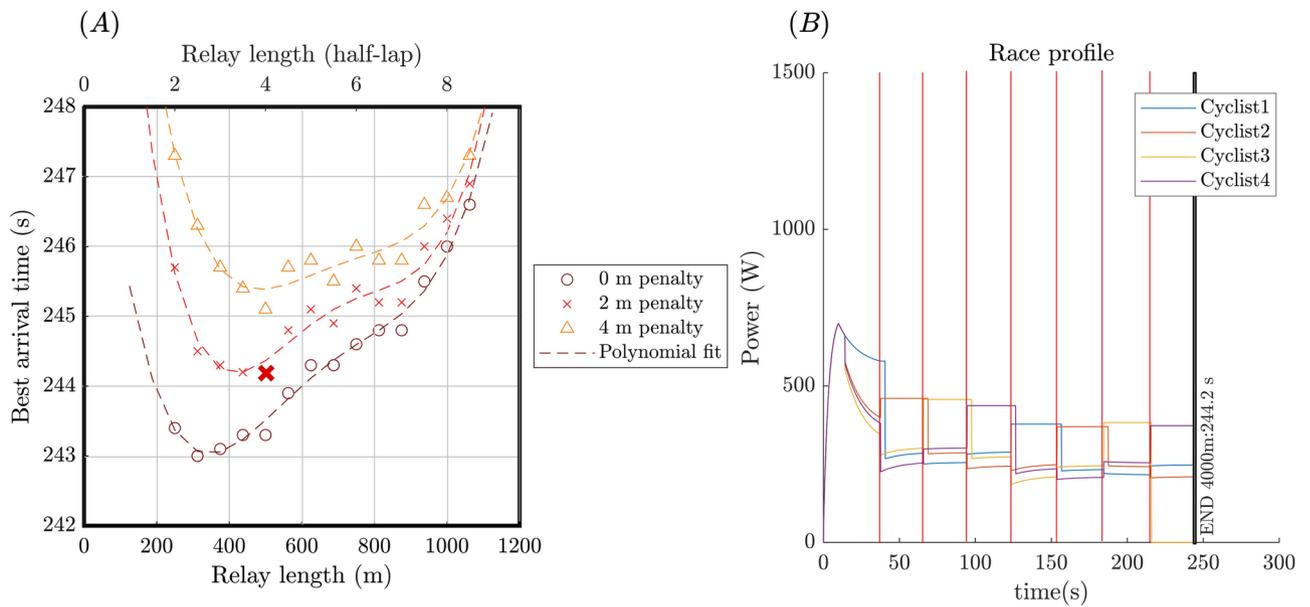
##### Identical length for all relays

To further reduce the problem, all relays are first imposed to be the same length. In this way, one can observe whether an optimal length can be found, resulting from the compromise between the costs of making more relays (lost distances and reduced recovery time for sheltered cyclists) and the benefits (reduced effort time in first position). In this case, for each  $n_{list}$ , only part 1. of the optimization described in the "Coupling and Optimization" section is performed. A fixed relay length is imposed in terms of the number of half-turns, ranging from 2 to 9 half-laps (corresponding to distances ranging from 250 m to 1125 m). To gain resolution, the optimization is also performed with a pitch corresponding to quarter relays. This approach is extended by varying the distance penalty  $d_{pen}$  associated with passing a relay, to highlight the compromise described earlier. The distance lost by passing a relay is taken successively as 0, 2, and 4 meters (remembering that in practice  $d_{pen} \approx 2$  meters). The results are available in Figure 6. To further emphasize the importance of initializing the optimization with different random seeds for exploration and validation, a graph showing the arrival times for the three different random seeds in each scenario (fixed relay length and varying  $d_{pen}$ ) is included in Supplementary Figure S1. This figure illustrates both the proximity of the obtained solutions (demonstrating the robustness of the optimization process) and the slight variations between seeds (highlighting the benefit of varying the initialization).

For ease of interpretation, a polynomial fit (degree 4) is added to each of the cases studied in dashed lines. First focusing on the classic case (2 m penalty, cross markers), a minimum is observed for relays of 3.5 half-turns (437,5 m). Since it is only possible to make relays of a whole number of half-turns, the result corresponding to 4 half-turns (bold crosses on (A)) is plotted on panel (B), and yields an arrival time of 244.2 seconds. The passage of a relay is indicated by a red vertical line, and the finish by a black bold vertical line. The optimum race profile in this case corresponds to a perfectly even distribution of effort in terms of the number of half-turns per cyclist: two relays of 4 half-turns for each cyclist. Power decreases between the first relay and the second one. One cyclist does not finish the race (Cyclist3 - see her power after her second relay that is equal to "zero" in Figure 6 (B)). If we increase the penalty associated with passing a relay ( $d_{pen} = 4$  m, triangles), the optimum shifts to longer relays: 4 half-turns (500 m). This confirms that the distance penalty shifts the optimum towards a lower number of relays, which was expected. Conversely, if  $d_{pen}$  is reduced to 0, the optimum is observed at 2.5 half-turns (312.5 m). However, this optimum is not obtained for the minimum possible (2 half-turns, 250 m). So it means that there is also a metabolic advantage in allowing protected cyclists to recover long enough. Indeed, the shorter the relays, the less recovery time cyclists have. This first simplified approach with a constant number of relays validates the existence of an optimum in this case.

##### Sensitivity analysis

Keeping the relay length constant between athletes (to limit calculation times and keep the approach simple), a "sensitivity" analysis to the parameters of the model can also be carried out. To do this, the relay length is set at 4



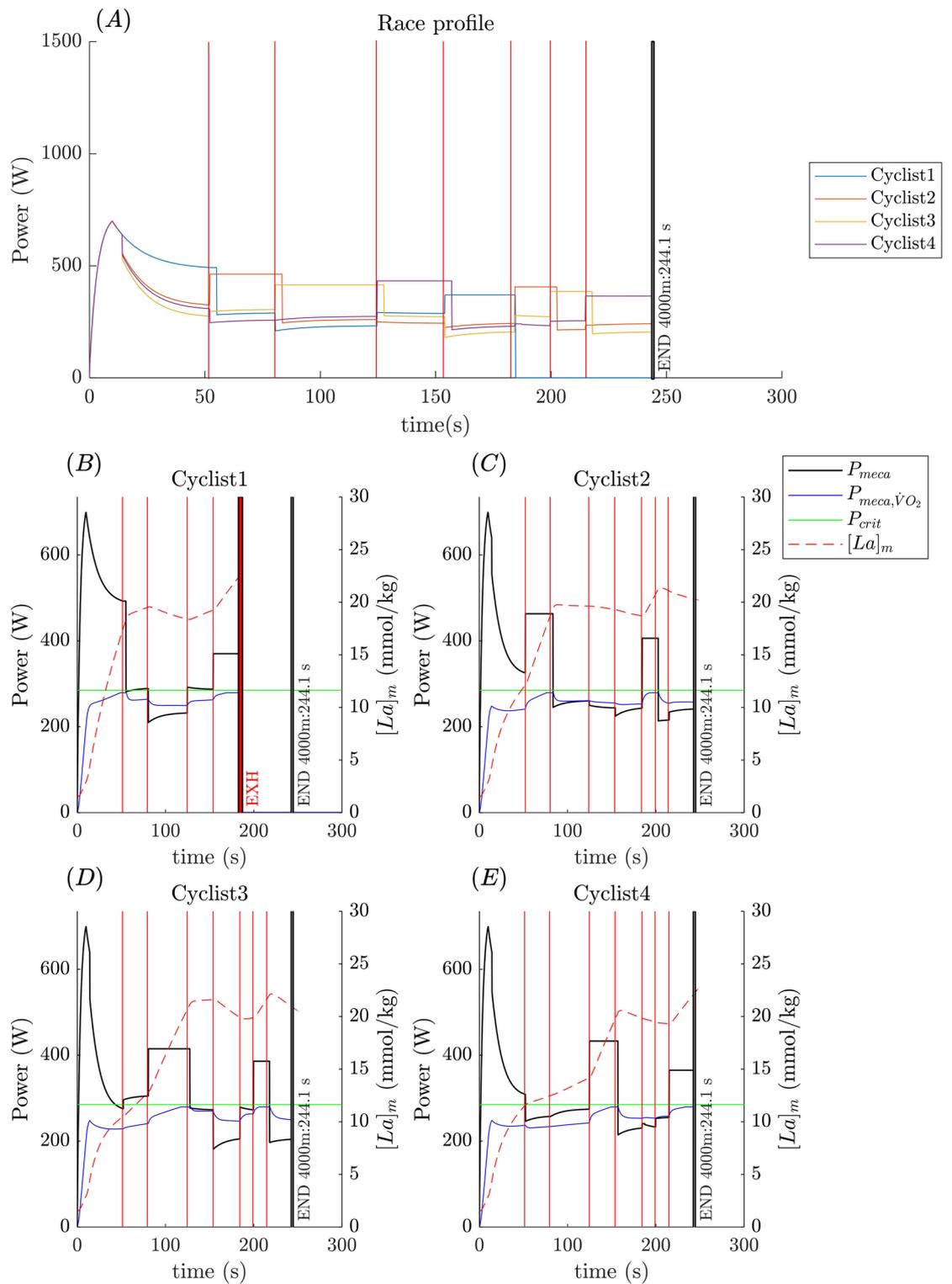
**Fig. 6.** Optimization results when taking four identical cyclists, imposing a constant relay length: **(A)** Arrival times in function of the relay length, varying the cost in length of passing a relay ( $d_{pen}$ ); **(B)** Optimal relay for  $d_{pen} = 2\text{ m}$  (bold red cross in (A)), relay length of 4 half-laps.

$\Pi$ (studied parameter)	$\rho_{air}$	$Sc_D$	$\mu$	$D_{end}$	$P_{start}$	$P_{max}$	$E_{non-ox}$	$P_{LS1}$	$P_{crit}$	$\dot{V}O_{2,max}$
slope of $\frac{dt_{end}}{t_{end}} = f(\frac{d\Pi}{\Pi})$	+0.31	+0.31	+0.21	+1.03	-0.02	-0.02	-0.10	-0.05	-0.20	-0.24
$R^2$	0.9986	0.9984	0.8776	0.9936	0.8980	0.8462	0.9913	0.9720	0.9975	0.9403
typical $\frac{d\Pi}{\Pi}$ (%)	+5	+20	+5	+1	+10	+10	+10	+10	+10	+5
corresponding change in $t_{end}$ (s)	+3.8	+15.2	+2.6	+1.7	-0.4	-0.4	-2.4	-1.2	-4.8	-2.9

**Table 1.** Sensitivity of the performance to various model parameters for four identical cyclists (with characteristics described in “Physiological modeling section”), imposing a constant length of 4 half-laps for each relay.  $\Pi$  denotes the parameters at stake, a linear function is fitted to the graph  $\frac{dt_{end}}{t_{end}} = f(\frac{d\Pi}{\Pi})$  and the

half-laps (500 m optimum from the previous section). The previously used team of four identical athletes is used as a reference point. Each parameter  $\Pi$  of the model is varied (all others remaining constant) by [-10; -5; -2; -1; 0; +1; +2; +5; +10] % and the minimum arrival time at  $n_{list}$  imposed (relay length of 4 half-laps) is calculated for each optimized case. If  $\Pi$  concerns the cyclists (for example their  $Sc_D$ ), the value is modified for the whole team. The results obtained can be represented in the plane  $(\frac{d\Pi}{\Pi}, \frac{dt_{end}}{t_{end}})$ . For the parameters for which the effect is significant on the performance, a linear function is fitted to the data, and the values of the slopes are calculated. To confirm the validity of using a linear fit, the coefficient of determination ( $R^2$ ) is provided in Table 1 to support this approach. These slopes can be used to assess the model’s sensitivity to the input parameters, and they are reported in Table Table 1. Hereafter two examples of how to interpret those results are given. First, for  $\rho_{air}$ , the slope is of +0.31. This means that if  $\rho_{air}$  increases by 5% (all other parameters remaining constant), the arrival time is predicted to increase by  $+0.05 \times +0.31 = 1.5\%$ , thus the performance is worsened by 1.5% (+3.7 secs). Conversely, if  $\rho_{air}$  decreases, the performance improves (decrease in  $t_{end}$ ). This was expected, as the aerodynamic drag is directly proportional to the air density. This is also a well-known effect by the athletes, coaches, and staff, explaining why many records in track cycling are established in velodromes higher than sea level (for example the one in Mexico at 1887 m) where the air density is lower than at sea level. Our analysis allows a quantification of the phenomenon. For an other example we can focus on  $P_{crit}$ . If it were to be increased by 5% for all cyclists, the model predicts a decrease in  $t_{end}$  of  $+0.05 \times -0.20 = -1\%$ , thus an increase in performance (-2.4 secs).

This sensitivity analysis highlights certain performance factors. A large slope value marks a big impact on performance. One can note the relative importance of  $\rho_{air}$ ,  $Sc_D$ , and  $D_{end}$  (the latter can artificially increase by bad piloting) in “mechanical” terms; and  $P_{crit}$ ,  $E_{non-ox}$ , and  $\dot{V}O_{2,max}$  in “physiological” terms. However, it should be borne in mind that these slopes indicate the variation in  $t_{end}$  as a function of the variation in the parameter under consideration. So, for example, variations of the order of 10% in  $\rho_{air}$  are sometimes measured. However, variations of 10% in  $D_{end}$  are highly unrealistic. This would mean that athletes were covering 10% more distance due to poor piloting. In practice, variations of  $D_{end}$  are more in the order of a percent. To highlight



**Fig. 7.** Optimization results for four identical cyclists. Race profile: (A) Power that each cyclist has to generate; (B) to (E) Individual time-courses of physiological parameters of the digital twins. For (A) to (E), passing of a relay is marked by a vertical red line, the end of the race by a bold vertical black line. For (B) to (E) the black line represents the mechanical power that has to be produced; the blue line the part of this power that comes from oxidative sources, the green horizontal line is the  $P_{crit}$ . The dashed red line is the predicted  $[La]_m$  in working muscles.

this effect, the orders of magnitude of variation in the parameters empirically observed and the time variations it would correspond to in the studied case (4 identical female cyclists with the chosen numerical twin) are reported in Table 1.

Now, since the team pursuit event allows it, it is time to add a level of complexity and allow variable lengths between relays.

#### Varying relay's length

In this section, all the steps described in the optimization section are carried out for four identical cyclists. The resulting optimal relay plan is available in Table 2 and shown in Figure 7 (A). The minimum predicted time is 244.1 seconds. In this case, too, each cyclist completes two relays. For all cyclists, the first relay is at a higher power and lasts longer than the second. The starting cyclist does not finish the race and stops after her second relay. An individual focus on each athlete can be used to interpret the predicted optimum. To do this, the values of physiological markers of the digital twins throughout the imposed race are computed. The results are shown in panels (B) to (E) of Figure 7. For each cyclist, the solid line represents powers: mechanical power produced (black), power generated by oxidative sources (blue), and the cyclist's critical power (green). If the power coming from oxidative sources is greater than the produced power, the cyclist is "recovering" (i.e., replenishing some of its non-oxidative capacity). For details concerning the interpretation and the computation of the time-courses of plotted variables, refer to the article describing the creation of the digital twins<sup>15</sup>. Relays passing are always represented by vertical red lines. The dashed red line represents the simulation of the time-course of muscle lactate concentration predicted by the model. An increase in lactate concentration models fatigue (in the sense of a decrease of the maximal power instantaneously producible), while a decrease in lactate concentration models recovery (in the sense of an increase of the maximal power instantaneously producible). Finally, a bold red vertical line framed in black marks potential predicted exhaustion. For Cyclist1 (B) in Figure 7, the start is the most demanding (6 half-laps plateauing at 484 W) and induces the greatest fatigue ( $[La]_m$  of 17.5 mmol/kg after her initial relay). She partially recovers during the Cyclist3's first long relay. She then completes a second shorter and less intense relay (4 half-laps at 370 W) but stops after this one. Cyclist2 ((C) in Figure 7) also undergoes fatigue during the start in the second position ( $[La]_m = 12.5$  mmol/kg after the first Cyclist1's relay). She does not recover before taking her first relay (4 half-laps at 463 W) and recovers slightly (drop in  $[La]_m$ ) before taking her second, shorter relay (2 half-laps at 406 W). Cyclist3 undergoes a similar "fatigue" during the start as Cyclist2. Then, she completes a long relay as her first one (6 half-laps at 415 W). Interestingly, this is not preceded by any recovery. Indeed, before taking her first relay, the  $[La]_m$  modeled does not rise high enough when compared to the demanded power to observe a drop in it. Cyclist3 then recovers during Cyclist2's second relay, before taking a second relay on her own (2 half-laps at 386 W). Cyclist4 makes a similar effort to Cyclist3. She does not recover before her first relay, and recovers between her two relays. Her first relay is more demanding (4 half-laps at 433 W) than the second one (4 half-laps at 365 W).

A slight performance gain of 0.04% is obtained when allowing for variation in relay lengths, compared to the case of identical-length relays. While 0.04% is indeed a small margin, such differences can often be decisive in competitions between teams. However, for four identical cyclists, the potential benefit of varying relay lengths is not clearly apparent. In real-life scenarios, where cyclists have differing strengths and abilities, the advantage could be more pronounced.

### Three identical cyclists and one different cyclist

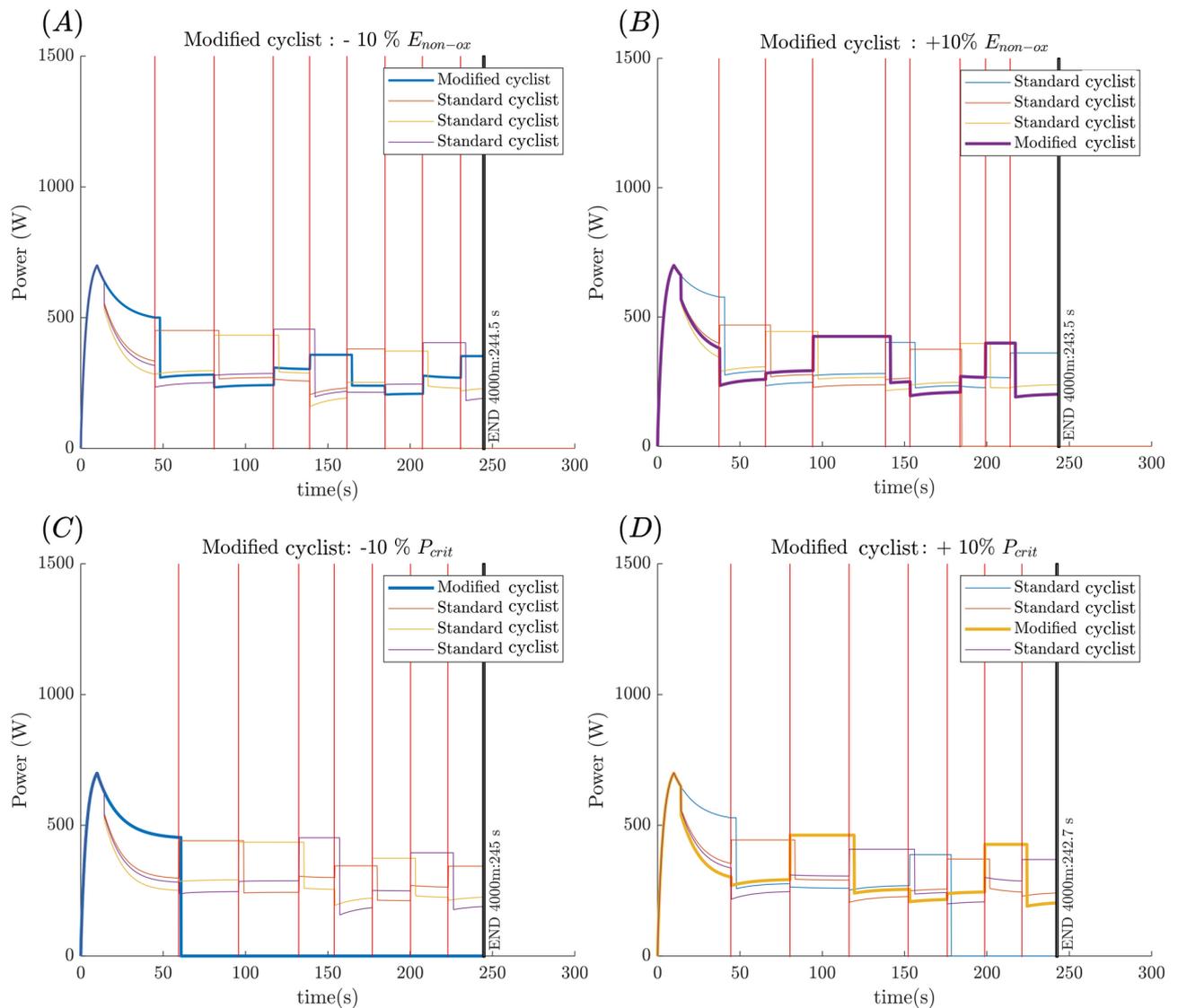
To investigate the cyclists' order parameter in more detail, the physiological characteristics of one of the cyclists are varied ("modified cyclist"). The others cyclists are maintained as before ("standard cyclists"). The global optimization is then conducted for each of the four possible positions for the "modified cyclist". The position and relay plan that will achieve the minimum finish time is then determined. Here, only the variation of physiological characteristics will be explored, in particular, energy-production capacities. The limiting factor for oxidative energy production in the model's parameterization is  $P_{crit}$ . The limiting factor for non-oxidative energy production is the amount that can be generated above  $P_{crit}$  ( $E_{non-ox}$  in the model's parameterization). One of those two variables is varied for the "modified cyclist". A variation of  $\pm 10\%$  is chosen. This corresponds to the order of magnitude of the maximum variation (to have a significant difference in profile) observed between athletes competing in the same team.

#### Modified cyclist: -10% $E_{non-ox}$

In the first case, the  $E_{non-ox}$  of a cyclist's digital twin is reduced by 10% (all other physiological and mechanical characteristics remain constant). The optimal strategy is predicted when the modified cyclist is in first position, as shown in Figure 8 (A). The minimum predicted time is 244.5 seconds. The drop in performance was to be expected (-0.2 %) since one cyclist is less "strong" than in the baseline case (four standard cyclists). The overall relay plan remains similar in idea: each cyclist completes at least two relays, the first being more demanding (in power and/or length) than the second. The cyclist starting in second position does not finish the race, she stops after her second relay. Surprisingly, the modified cyclist completes three relays, including a second relay

$n_{list}$ (half-laps)	6	4	6	4	4	2	2	4
$P_{list}$ (W)	484	463	415	433	370	406	386	365

**Table 2.** Optimal strategy for four identical cyclists (the starting order is of no interest in case of identical cyclists).



**Fig. 8.** Race profile of the optimal strategy for a team composed of one modified cyclist and three standard cyclists. Modified cyclist corresponding to a variation of the standard cyclist where: (A) her  $E_{non-ox}$  is reduced by 10%; (B) her  $E_{non-ox}$  is increased by 10%; (C) her  $P_{crit}$  is reduced by 10%; (D) her  $P_{crit}$  is increased by 10%.

relatively low in power (compared with her teammates' relay). One possible interpretation of this relay is that it benefits the following cyclists. The slower the cyclist in first position, the lower the power required from the protected cyclists. This weaker relay is therefore of interest in this case, as it enables the following cyclists to deplete less of their stocks of  $E_{non-ox}$  or to recover more.

#### Modified cyclist: +10% $E_{non-ox}$

Now we increase the  $E_{non-ox}$  of the modified cyclist by 10% compared with a standard cyclist. The optimal solution is reached when the modified cyclist is in fourth position at the start. This is associated with a time saving of 0.2% compared with the baseline situation. The associated optimal relay plan can be seen in Figure 8 (B). In this relay plan, once again, the starting cyclist makes three relays, while all the others make two. The cyclist in second position at the start stops after her second relay. Each cyclist also completes relays of decreasing intensity (length and/or power). The modified cyclist completes two relays, the first of which is very long (6 half-laps at 425W, the longest of all relays). This relay allows her teammates to recover during a relatively long phase. Moreover, her second relay, although shorter remains quite high (2 half-laps at 399W), here increased  $E_{non-ox}$  allows her to have a greater recovery potential when she is in the peloton riding below her  $P_{crit}$ .

#### Modified cyclist: -10% $P_{crit}$

Now the  $P_{crit}$  is modified, starting by lowering it by 10% for the modified cyclist. The optimal result is obtained when the cyclist is placed in the first position. The optimal relay plan is shown in Figure 8 (C). Performance drops by 0.4% compared with the baseline. The drop was expected, but it is greater than for a drop of the same

percentage of  $E_{non-ox}$ . This was expected given the higher sensitivity of the model to  $P_{crit}$  as reported in Table 1 for the case of four identical cyclists. In the optimal relay plan obtained, the modified cyclist stops after her first long single relay (of 7 half-turns). The other cyclists finish the race according to a plan similar to the ones above, completing relays of 5 or 3 half-laps of decreasing intensity. One explanation for the single relay of the modified cyclist is that since critical power is reduced, she would recover little or not at all, even protected in the peloton (power to be supplied above her  $P_{crit}$ ).

#### *Modified cyclist : +10% $P_{crit}$*

Now the  $P_{crit}$  of the modified cyclist is increased by 10%. The optimal result is obtained when the modified cyclist is placed in the third position, the optimal relay plan is shown in Figure 8 (D). Performance improved by 0.6% compared with the baseline. The improvement is greater than for an increase of the same percentage of  $E_{non-ox}$ . Again, this was expected given the higher sensitivity of the model to  $P_{crit}$  (Table 1). Each cyclist completes two relays. The starter cyclist stops after her second relay. The same overall pattern of decreasing relay intensity from first to second remains. The modified cyclist can complete a second relay with a lower intensity than her teammates. This is explained by her higher  $P_{crit}$ , which enables her to recover better when sheltered in the peloton.

## Discussion and limitations

In the previous section, we explored the problem of optimizing the team pursuit strategy. The studied simplified cases allowed us to explore the different variables of a team pursuit strategy.

By first taking four identical cyclists, the focus was set on the relays' lengths and powers, keeping the order variable for later. For four identical cyclists and exploring configurations with identical lengths for all relays (hypothetical simplified case), we have seen the benefits of distributing effort among the different cyclists. In this case, this observation takes the form of an optimum relay length, which is neither the shortest possible (minimizing the number of relay passing) nor the highest possible (minimizing the duration of the effort bouts). This highlights the advantage of splitting each cyclist's effort into two relays. This is due firstly to the distance cost of passing a relay, as shown in Figure 6. But also to the compromise between the effort time of the cyclist in the first position in the peloton and the recovery time of the followers. The optimum is found around 4 half-lap relays, perfectly splitting the effort in distances between cyclists (8 half-laps for each cyclist). For a team of four identical cyclists, we slightly improve the result by exploring configurations where relay lengths can vary. In this case, the effort is distributed with two relays per cyclist. The second relay is shorter in length and/or lower in power than the first. However, it should be noted that this scenario applies to the hypothetical case studied and may vary according to team composition. First, for four identical cyclists, the predicted optimum values of relay length may vary if the physiological or mechanical characteristics of the modeled cyclist were to be changed. However, we expect the general observations and conclusion to be the same: advantages of splitting the efforts between cyclists with several relays for each and of decreasing length and/or intensity. Then, the situation in which all four cyclists are identical is unrealistic and predictions should vary if the team is composed of more diverse cyclists. Namely, the benefit of varying relay lengths might be more significant in teams composed of cyclists with differing abilities.

To study the order parameter, one of a cyclist's metabolic parameters ( $E_{non-ox}$  or  $P_{crit}$ ) was modified and the predicted optimal position and relay plan was obtained. The plans (shown in Figure 8) all follow (except for (C)) the principle of at least two relays per cyclist, with the second relay less long and/or power-intensive than the first one. A cyclist whose one of the modified physiological parameters is artificially lowered (Figure 8 (A) and (C)) is placed in the first or second position. Indeed, her ability to recover is diminished, and it seems better to limit the time she spends recovering since it will be less "exploited" than if one of her teammates recovers. In contrast, a stronger cyclist is placed in the third or fourth position (Figure 8 (B) or (D)). She has a higher recovery potential than her teammates, so it is in the team's interest to let her do relays at the end of the race (as she will have been able to recover relatively more). Here also, these results are only true for the simplified studied case and are presented to illustrate the potential of the method. They show in a team with varying capacities, there is an optimal order they should adopt to maximize performance. Additionally, it is worth noting that in the scenario where one cyclist differs from three identical ones, the performance variations do not simply amount to one-fourth of those observed in the sensitivity analysis for the same parameters ( $E_{non-ox}$  and  $P_{crit}$ ), where the parameters of all four cyclists were modified. Moreover, these variations are not always symmetrical-modifying the performance. I.e. a parameter can have different magnitudes of impact depending on whether the parameter is increased or decreased (see  $P_{crit}$ ). These two observations highlight the interactions between cyclists: changes in one cyclist's capacity can influence the performance of the entire team. For instance, a change in one cyclist's ability may lead to differences in relay length or power output, thereby altering the recovery conditions for the sheltered cyclists.

The developed method can be applied for races or competition. For live feedback during a race, speed is monitored rather than exerted power, as speed is easier to measure and is directly related to power. Additionally, for delayed feedback and post-race analysis, instrumented drives on the bikes provides exploitable data. Consequently the method can be used at three key stages of a race or competition. First, in the preparation phase, it can be used with parameterized digital twins to generate optimal relay strategies. They can then be tested and practiced during training, allowing the cyclists to familiarize themselves with and validate the proposed plan. Secondly, during the competition, the method can be used between race rounds. For instance, after qualification rounds, the race plan (i.e.,  $n_{list}$ ,  $P_{crit}$ , and the starting order of cyclists) can be simulated and compared to real-world race data. If any discrepancies are observed, such as unexpected exhaustion, the digital twins can be adjusted-particularly by modifying parameters like  $E_{non-ox}$  or  $P_{crit}$ -to better reflect the cyclists' actual performance. The updated model can then be used to rerun simulations and adapt the strategy for the next

round. Finally, after the race or competition, the actual strategies employed during the event can be simulated to analyze how well the digital twins represented the cyclists' performance. This post-race analysis helps identify potential areas for improvement through further optimization and provides insights into key observations, such as why a cyclist fatigued at a particular point or whether there was underperformance or overperformance compared to the predictions. Any discrepancies between the simulated outcomes and real-life performance at any stage can help highlight limitations in the model, offering feedback for refining future strategies.

The tool presented here is powerful, but some of its limitations must be mentioned. On the whole, the results presented are based on a very specific case, and general trends should be retained rather than precise numerical results. Secondly, as far as the mechanical model is concerned, average parameters were considered (relay cost, cyclists' drag coefficient) which may vary throughout the race. Concerning the physiological model, the limitations are those of the digital twin described in the paper used to create them<sup>15</sup>. Without listing them exhaustively, it is worth mentioning that physiological parameters are determined at a given moment and may vary for an athlete, according to their fitness, but also during the race as a function of pedaling rate or fatigue for instance. Additionally, in the context of numerical optimization, despite efforts to explore strategies as comprehensively as possible, there remains the possibility of missing a global optimum. Moreover, the initialization process may also constrain the solution space. For example, when considering cyclists with significantly different characteristics, the assumption of an equal distribution of effort (i.e., initializing  $n_{list}$  with equal relay lengths) or initializing  $P_{list}$  within a uniform power range might be less appropriate. This computational considerations could be investigated further in future research. Moreover, another limitation arises from the consideration of only the "standard" relay-passing technique, where the leading cyclist moves to the back of the peloton. If cyclists of very different performance levels were included, more advanced relay-passing strategies might be beneficial. However, incorporating such techniques would not only require changes in the optimization process and mechanical model but would also have significant practical implications, necessitating cyclists to learn new skills and techniques. Investigating this numerically before introducing such changes in training could be judicious.

## Conclusion

The work presented here presents a method and a case study of race strategy optimization for team pursuit in track cycling. Without being exhaustive, we have highlighted several results that can be generalized to all teams (existence of optimal strategies, interest in sharing effort, importance of race parameters) and the method presented can be applied to specific teams or to other disciplines.

All examples illustrate the application of the coupled mechanical and physiological model. This method can be applied to real-life athletes (and with more complicated cases than the ones illustrated here) or to other events and/or disciplines by adapting the input parameters and the mechanical model. For application to other sports events, it will be necessary to adapt the mechanical model and ensure that the metabolic pathways involved in effort are similar, if the same physiological model is to be used. Examples include pursuit events in speed skating (similar problem with first position and recovery phases), or rowing races (where athletes row simultaneously and therefore deliver a similar effort throughout the race).

## Data availability

The datasets generated and/or analysed during the current study are not publicly available as it corresponds to national athlete data. Although the data were collected in 2021, the French Cycling Federation prefers to keep the data confidential. However, data are available from the corresponding author on reasonable request.

Received: 18 April 2024; Accepted: 9 October 2024

Published online: 25 October 2024

## References

1. Official Website | Tissot Timing - <https://www.tissottiming.com>.
2. Blocken, B. et al. Aerodynamic drag in cycling pelotons: New insights by CFD simulation and wind tunnel testing. *Journal of Wind Engineering and Industrial Aerodynamics* **179**, 319–337. <https://doi.org/10.1016/j.jweia.2018.06.011> (2018).
3. Broker, J. P., Kyle, C. R. & Burke, E. R. Racing cyclist power requirements in the 4000-m individual and team pursuits. *Medicine and Science in Sports and Exercise* **31**, 1677–1685. <https://doi.org/10.1097/00005768-199911000-00026> (1999).
4. Defraeye, T. et al. Cyclist drag in team pursuit: influence of cyclist sequence, stature, and arm spacing. *Journal of Biomechanical Engineering* **136**, 011005. <https://doi.org/10.1115/1.4025792> (2014).
5. Torre, A. & Iñiguez, J. Aerodynamics of a cycling team in a time trial: Does the cyclist at the front benefit?. *European Journal of Physics* **30**, 1365. <https://doi.org/10.1088/0143-0807/30/6/014> (2009).
6. Underwood, L. Aerodynamics of track cycling. *Publisher: University of Canterbury. Mechanical Engineering*. <https://doi.org/10.26021/3548> (2012).
7. Heimans, L., Dijkshoorn, W. R., Hoozemans, M. J. M. & de Koning, J. J. Optimizing the Team for Required Power During Track-Cycling Team Pursuit. *International Journal of Sports Physiology and Performance* **12**, 1385–1391. <https://doi.org/10.1123/ijssp.2016-0451> (2017).
8. de Koning, J. J., Bobbert, M. F. & Foster, C. Determination of optimal pacing strategy in track cycling with an energy flow model. *Journal of Science and Medicine in Sport* **2**, 266–277. [https://doi.org/10.1016/S1440-2440\(99\)80178-9](https://doi.org/10.1016/S1440-2440(99)80178-9) (1999).
9. Overtoom, M. *Optimisation of Changing Strategy using a Mathematical Performance model*. Master's thesis, TU Delft Mechanical, Maritime and Materials Engineering (2013).
10. Keller, J. B. A theory of competitive running. *Physics Today* **26**, 43–47. <https://doi.org/10.1063/1.3128231> (1973).
11. Martin, J. C., Gardner, A. S., Barras, M. & Martin, D. T. Modeling sprint cycling using field-derived parameters and forward integration. *Medicine & Science in Sports & Exercise* **38**, 592–597. <https://doi.org/10.1249/01.mss.0000193560.34022.04> (2006).
12. Laigret, S. *Stratégies de cours en cyclisme sur piste*. Master's thesis, Ecole polytechnique, Palaiseau (2020).
13. Slawinski, M. A., Slawinski, R. A. & Stanoev, T. On modelling bicycle power for velodromes: Part I: Formulation for individual pursuits, <https://doi.org/10.48550/arXiv.2005.04691> (2020). [ArXiv:2005.04691](https://arxiv.org/abs/2005.04691) [physics].

14. Ashtiani, F. & Vahidi, A. Dynamic optimization of formation and pace in a cycling team time trial considering effects of drafting and cyclists' fatigue and recovery. *Control Engineering Practice* **137**, 105563. <https://doi.org/10.1016/j.conengprac.2023.105563> (2023).
15. Boillet, A., Messonnier, L. A. & Cohen, C. Individualized physiology-based digital twin model for sports performance prediction: a reinterpretation of the Margaria-Morton model. *Scientific Reports* **14**, 5470, <https://doi.org/10.1038/s41598-024-56042-0> (2024). Publisher: Nature Publishing Group.
16. Craig, N. & Norton, K. Characteristics of Track Cycling. *Sports medicine (Auckland, N.Z.)* **31**, 457–68, <https://doi.org/10.2165/00007256-200131070-00001> (2001).
17. Pugh, C. F. et al. Critical Power, Work Capacity, and Recovery Characteristics of Team-Pursuit Cyclists. *International Journal of Sports Physiology and Performance* 1–8, <https://doi.org/10.1123/ijsp.2021-0478> (2022).
18. Craig, N. P. et al. Aerobic and anaerobic indices contributing to track endurance cycling performance. *European Journal of Applied Physiology and Occupational Physiology* **67**, 150–158. <https://doi.org/10.1007/BF00376659> (1993).
19. Wagner, M., Day, J., Jordan, D., Kroeger, T. & Neumann, F. Evolving Pacing Strategies for Team Pursuit Track Cycling. In Di Gaspero, L., Schaefer, A. & Stützle, T. (eds.) *Advances in Metaheuristics*, vol. 53, 61–76, [https://doi.org/10.1007/978-1-4614-6322-1\\_4](https://doi.org/10.1007/978-1-4614-6322-1_4) (Springer New York, New York, NY, 2013). Series Title: Operations Research/Computer Science Interfaces Series.
20. Hansen, N. The CMA evolution strategy: a comparing review. *Towards a new evolutionary computation: Advances in the estimation of distribution algorithms* 75–102 (2006).

## Acknowledgements

The authors would like to thank the staff and athletes of the French Cycling Federation. Without their questioning and collaboration, this method would not have emerged. Without the valuable data and participation of the athletes, the models could not have been parameterized. Special thanks to Emmanuel Brunet, the federation's scientific manager and project facilitator. This work was funded by a grant from the PIA (Programme d'Investissements d'Avenir) via the ANR (Agence Nationale de la Recherche) with the reference ANR-20-STHP-0006 and funding from an AMX grant (École polytechnique, Institut polytechnique de Paris).

## Author contributions

A.B., L.A.M., and C.C. conceived the physiological model. A.B., L.A.M., C.C., and I.S. conceived the mechanical model. A.B. and M.N. adapted and tuned the optimization method. A.B. implemented the coupled model and conducted the simulations. All authors analyzed the results and reviewed the manuscript.

## Declarations

### Competing interests

The authors have no competing interests as defined by Nature Research, or other interests that might be perceived to influence the results and/or discussion reported in this paper.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-75963-4>.

**Correspondence** and requests for materials should be addressed to A.B.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024