Optimizing Nozzle Travel Time in Proton Therapy

Matteo Spezialetti
DISIM Department
University of L'Aquila, Italy

Renata Di Filippo
DISIM Department
University of L'Aquila, Italy

Ramon Gimenez De Lorenzo
Medical Physics Unit, San Salvatore Hospital, ASL1 Abruzzo, Italy

Giovanni Luca Gravina
DISCAB Department
University of L'Aquila, Italy

Giuseppe Placidi
MESVA Department
University of L'Aquila, Italy

Guido Proietti
DISIM Department
University of L'Aquila, Italy

Fabrizio Rossi
DISIM Department
University of L'Aquila, Italy

Stefano Smriglio
DISIM Department
University of L'Aquila, Italy

João Manuel R.S. Tavares
Inst. de Ciência e Inovação em Eng. Mecânica e Eng. Industrial, Departamento de Eng. Mecânica
Faculdade de Engenharia, Universidade do Porto, Portugal

Francesca Vittorini
Medical Physics Unit, San Salvatore Hospital, ASL1 Abruzzo, Italy

Filippo Mignosi
DISIM Department
University of L'Aquila, Italy
ICAR-CNR, Palermo, Italy

***Abstract*—Proton therapy is a cancer therapy that is more expensive than classical radiotherapy but that is considered the gold standard in several situations. Since there is also a limited amount of delivering facilities for this techniques, it is fundamental to increase the number of treated patients over time. The objective of this work is to offer an insight on the problem of the optimization of the part of the delivery time of a treatment plan that relates to the movements of the system. We denote it as the Nozzle Travel Time Problem (NTTP), in analogy with the Leaf Travel Time Problem (LTTP) in classical radiotherapy.**

In particular this work: (i) describes a mathematical model for the delivery system and formalize the optimization problem for finding the optimal sequence of movements of the system (nozzle and bed) that satisfies the covering of the prescribed irradiation directions; (ii) provides an optimization pipeline that solves the problem for instances with an amount of irradiation directions much greater than those usually employed in the clinical practice; (iii) reports preliminary results about the effects of employing two different resolution strategies within the aforementioned pipeline, that rely on an exact Traveling Salesman Problem (TSP) solver, Concorde, and an efficient Vehicle Routing Problem (VRP) heuristic, VROOM.

***Index Terms*—Proton Therapy, Delivery Time Optimization, Generalized Traveling Salesman Problem.**

I. Introduction

Proton therapy uses accelerated particle beams of protons to treat and eliminate various forms of cancer. Unlike x-rays and electron beams which, in classical radiotherapy, release energy along the entire path through the patient's body, the beams used in proton therapy have a characteristic depth/dose profile that allows to concentrate the maximum of the energy at a precise depth position, called Bragg Peak [1]. For this reason, proton therapy allows to obtain the maximum biological effect near the tumor mass or, in general, in the target volume established by the medical team, while minimizing the damage caused to the surrounding healthy tissues [2].

The treatment planning is the process that analyses and selects the particle beams that can bring the most benefit to the therapy, i.e. those that are compliant with the dose requirements of the target and preserve as much as possible the healthy surrounding tissues and organs at risk. It starts from diagnostic information, e.g. the results of magnetic resonance and CT, and produce a treatment plan, i.e. in a set of beams, each with its own orientation, energy and fluence.

In order to deliver the therapy, a particle production equipment, the cyclotron, and a beam transport and release system are required. The typical delivery front-end structure in the case of Intensity Modulated Proton Therapy (IMPT) is composed by a nozzle, from which the beams are irradiated, that is placed in a rotating gantry structure. During the treatment, the patient is immobilized on a bed inside the gantry room, thus allowing the beams to be directed by rotating the nozzle around the patient. The bed itself can rotate around the vertical axis, perform rolling and pitching movements and translations along the three orthogonal axes. These movements are critical, in order to obtain the correct relative positioning of the patient with respect to the nozzle during the treatment: positions must be reproducible with extreme accuracy in order to fully exploit the precision of the proton beams and achieve the expected result of the therapeutic treatment plan [3].

The delivery system is particularly sophisticated and expensive, as well as slow in its movements: consequently the cost of a treatment with protons is higher than the cost of a normal classic radiotherapy irradiation [2]. Completing as much as possible daily treatments is therefore essential to reduce the average costs of therapies with protons and make them more easily accessible to a greater number of patients [4].

The treatment planning presents several computational bottlenecks, such as the accurate computation of beams dose profiles [5], [6] and the development of treatments including the set of irradiation directions, called fields (or angles) and

This work was partially supported by the project "Optimization of computational resources in TPS (Treatment Planning Systems) for hadrontherapy", funded by the Italian Government (PON-AIM 1877124-3) and by the research starting grant "Design and development of an optimization pipeline for proton therapy" (04ATE2021-PROGETTI.SPEZIALETTI) funded by the University of L'Aquila.

fluences, that provide the best outcome for patients [7], [8].

The objective of this work is to address the minimization of the delivery time of a treatment plan once that the set of fields is fixed. In our optimization problem we say that the fields in above set have to be "covered", and we call our problem the Nozzle Travel Time Problem (NTTP), in analogy with the Leaf Travel Time Problem (LTTP) in classical radiotherapy. Notice that inside the NTTP the bed movements are taken into account.

To this aim, an optimization pipeline for solving NTTP instances is proposed. It has been designed to be modular, in the sense that, once the problem instance is pre-processed and the set of configurations that the delivery system can assume to fulfill the requirements is defined, multiple exact and heuristic strategies can be followed to solve it. This work offers a comparison of performances, both in terms of computational time and accuracy of the solution, between two methods: an exact Traveling Salesman Problem solver (Concorde) and an efficient heuristic Vehicle Routing Open-source Optimization Machine (VROOM).

The manuscript is organized as follows: Section II summarizes related work; Section III described the NTTP and provides a formal representation of the delivery system and of the NTTP instances needed for further steps; Section IV concerns the implementation of the solvers and the simulation results; finally, Section V illustrates conclusions and future developments.

II. Related work

The problems described in this work are related to Traveling Salesman Problem (TSP) and its variants. TSP is a very well known optimization problem: given a set of m cities, with integer distance $d(c_i, c_j)$ between each pair of cities (c_i, c_j), find a Hamiltonian cycle of minimum total distance. TSP has been proven to be NPO-complete [9] and the decision version of the problem is NP-complete [10].

In particular NTTP is a special case of the Equality Generalized TSP (E-GTSP), that is, in turn, a variant of Generalized TSP (GTSP), also known as International TSP or Travelling Politician Problem. GTSP was first introduced by Henry-Labordere in 1969 [11]. Given a weighted complete directed graph $G = (V, A)$ and a partition $V_1, \ldots, V_k$ of its vertices, GTSP asks for a minimum weight cycle containing a vertex from each set [12]. The "Equality" variant imposes that each cluster is visited exactly once. Obviously, being TSP a particular case of E-GTSP, where each cluster is a singleton, E-GTSP is at least hard as TSP. More precisely, in NTTP, since all the distances that are taken in account preserve the triangle inequality, there is no difference between the general an the equality version of the problem. A natural approach to deal with GTSP/E-GTSP problems is to reduce it to TSP [13], [12], [14], [15], since scholars put a large effort in developing exact and heuristic algorithms for TSP [16].

From the application point of view, the problem of optimizing the treatment delivery time is strictly related with the Robotic Task Sequencing Problem (RTSP), i.e. optimizing the sequence of movements that a manufacturing robot (e.g. a welding machinery) has to perform to accomplish a task [17], [18]. In particular, the task space (or T-space) that describes position and orientation of the end-effector in RTSP, is a generalization of the concept of "field" described here: the irradiation direction toward the patient's body. In the same way, the configuration space (or C-space), i.e. the set of possible joints position of the robot, corresponds to the configurations described here.

NTTP is related with LTTP in classical radiotherapy, where dose optimization routines generate the intensity maps, that could be delivered with different techniques. Conceptually the IMPT delivery method considered here is analogous to the "step-and-hoot" method considered in classical radiotherapy. For further details the reader can refer to [19], [20] and references therein.

III. The NTTP problem

In order to address the NTTP definition, a formal description of the delivery system has to be provided. Consider, in a 3-dimensional Cartesian coordinate system with origin O, a delivery system for proton therapy consisting of:

- a circumference, henceforth simply called "*ring*" (R), of radius r, centered in the origin and lying in the XY plane, representing the gantry structure around which the nozzle, N, rotates: let be P_{ring} the XY plane itself;
- a *bed*, B, represented by a rectangle lying in the XZ plane (P_{bed}) that can translate in along X and Z axes and rotate around an axis Y_{bed} (parallel to Y) and passing for O_{bed}, the Center of Gravity (COG) of B. On B, the midpoint of the shorter side closest to the patient's head is denoted by H.

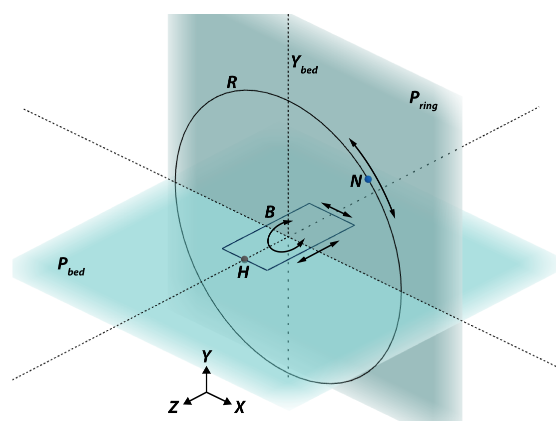

Fig. 1. Geometric representation of the delivery system

Figure 1 shows the aforementioned components. In order to simplify, only bed rotations around gravity (Y_{bed}) are considered. This constraint also reflects the clinical practice, since internal organs and soft tissue geometry have to be preserved. Each position of the nozzle corresponds to a radius of R and identifies the direction of the proton beam. The reciprocal positions of the nozzle and the bed identify a "*field*", i.e. the direction of the beam toward the patient's body. Let a *configuration* represent a specific position of the system by a quadruple $c = (\alpha, t_z, t_x, \beta)$, where:

- α is the angle, expressed in radians, formed by Y and the half line with origin in O and passing for N, i.e. the position of the nozzle with respect to the apex of R;
- tx and tz are the X and Z coordinates of O_{bed}, expressed in meters;
- β is the angle, expressed in radians, of rotation of B around Y_{bed}, that is, β is the angle formed by:
  - the vector with origin in O_{bed} and same direction and sense of Z;
  - the half line with same origin and passing for H.

The *reference configuration* $c_0 = (0, 0, 0, 0)$, i.e. with $N = (0, r, 0)$, $O_{bed} = O$ and H on the positive side of Z, is the initial position of the system. *System features* describe the physical constraints of the delivery hardware, e.g. there are systems that limit the nozzle allowed rotation around R, and the movement speed of the components. System features can be formalized by the tuple

$SF = (\alpha_{rng}, \omega_\alpha, tx_{rng}, v_x, tz_{rng}, v_z, \beta_{rng}, \omega_\beta)$,
where:

- $\alpha_{rng} = (\alpha_{min}, \alpha_{max})$ represents the feasible positions of N (rad);
- ω_α is the angular velocity of N (rad/s);
- $tx_{rng} = (tx_{min}, tx_{max})$ and $tz_{rng} = (tz_{min}, tz_{max})$ indicate the positions O_{bed} can assume (m);
- v_x and v_z are the linear velocity of O_{bed} along X and Z, respectively (m/s);
- $\beta_{rng} = (\beta_{min}, \beta_{max})$ indicates which rotations the bed can perform (rad);
- ω_β is the angular velocity of H around Y_{bed} (rad/s).

Given the system features, for any pair of configurations (c_1, c_2), it is possible to compute the distance $\Delta(c_1, c_2) = \Delta(c_2, c_1)$ that represents the time needed by the system to move from c_1 to c_2 or vice-versa. This metric is, in turn, function of:

- $\delta_\alpha(c_1, c_2) = \omega_\alpha min(|\alpha_{c_1} - \alpha_{c_2}|, |2\pi - \alpha_{c_1} - \alpha_{c_2}|)$
- $\delta_{tx}(c_1, c_2) = v_x|tx_{c_1} - tx_{c_2}|$
- $\delta_{tz}(c_1, c_2) = v_z|tz_{c_1} - tz_{c_2}|$
- $\delta_\beta(c_1, c_2) = \omega_\beta min(|\beta_{c_1} - \beta_{c_2}|, |2\pi - \beta_{c_1} - \beta_{c_2}|)$

that are the time needed to align each element of the configurations: nozzle position, bed X and Z translations and bed rotation (for simplicity, the configurations arguments of the functions have been omitted). If the system constraints allow to move multiple components at the same time, then $\Delta(c_1, c_2) = max(\delta_\alpha, \delta_{tx}, \delta_{tz}, \delta_\beta)$ (L_∞ norm), otherwise, if only one component at time can be moved, then $\Delta(c_1, c_2) = \delta_\alpha + \delta_{tx} + \delta_{tz} + \delta_\beta$ (L_1 norm).

For the proposed application, it is convenient to model a second coordinates system, integral with B, with origin in O_{bed} and axes corresponding to X, Y and Z when the system is in the reference configuration. Now therefore, a *field* can be described by the pair (V, T), with V and T free vectors. Specifically, the irradiation direction can be represented by applying the translation T (in the bed coordinates system) to V, initially with its tail in O_{bed}. The previous definition is suitable with a "patient-centric" field description, since it is independent from the system features and configuration and depends only on the position of the patient on the bed. We say that a configuration $c = (\alpha, t_x, t_z, \beta)$ "covers" a field f, i.e. c allows to irradiate along the direction indicated by f, if, rotating B by $-\beta$ around Y_{bed}, translating it by $(-t_x, 0, -t_z)$ and, finally, rotating it by $-\alpha$ around Z, f has its start point on O and has the same direction and sense of vector $(0, -1, 0)$ in the original coordinates system. It is to note that the last rotation corresponds to the rotation by the angle α of the nozzle. Consistently with the delivery system features, a maximum of two configurations cover the same field: if $c_1 = (\alpha, t_x, t_z, \beta)$ cover f, also $c_2 = (-\alpha, -t_x, -t_z, \beta + \pi)$ does. A pair of configurations such c_1 and c_2 will be henceforth called *twin configurations*: this property of the delivery system has a key role in the problem definition, since, given a set of fields, at least one of the configurations covering each of them has to be assumed by the system.

A. Problem Description

NTTP consists in the following problem: given a set of distinct fields $F = \{f_1, f_2, ..., f_n\}$ and the system features $SF = (\alpha_{rng}, \omega_\alpha, tx_{rng}, v_x, tz_{rng}, v_z, \beta_{rng}, \omega_\beta)$, find a sequence of configurations $C^* = (c_1, ..., c_n)$ that covers each field and minimize $T = \Delta(c_n, c_1) + \sum_{i=1}^{n} \Delta(c_{i-1}, c_i)$, i.e. the time needed to visit the whole sequence and return in c_1.

Notice that, as usually assumed in proton therapy, the order in which the fields are covered is not relevant for the final outcome of the treatment.

NTTP can be formulated as follows: given $F = \{f_1, f_2, ..., f_n\}$, construct a complete undirected graph $G = (V, E)$ where vertices in V correspond to the set C of all configurations that cover a field in F and assign weight $w_ij = \Delta(c_i, c_j)$ to each $(i, j) \in E$. Let $C = \bigcup_{s=1}^{n} K_s$, where K_s is the subset of configurations covering f_s (depending on f_s orientation and the system features, there could be a single configuration ore two twins configurations in K_s). Thus, NTTP asks for a cycle Γ of minimum weight in G, such that it visits exactly one configuration for each K_i. In [21] it is proved that NTTP is hard to solve but due to space constrains we cannot report the proof here.

IV. Optimization Pipeline for NTTP

In order to solve instances for the NTTP, the pipeline depicted in Figure 2 has been followed. The proposed method takes in input the characteristics of the delivery system and the prescribed set of fields that have to be covered and finds an optimal sequence of configurations such that each field is covered by exactly one configuration. In the first part, the field set is processed in order to build the graph $G = (V, E)$. At this point the NTTP, that, as remark, is a special case of GTSP, can be transformed in an Asymmetric TSP (ATSP) instance by applying the algorithm proposed by Noon and

Bean [22]. Then, alternative approaches can be followed. Here we implemented and tested the following:

- consider ATSP as a special (simpler) case of Vehicle Routing Problem (VRP). Here, the VROOM engine[1] is employed as heuristic approach;
- reformulate ATSP as TSP, thanks to the construction proposed by Jonker and Volgenant [23]. In this work, TSP instances have been solved by using the best known TSP solver, Concorde[2].

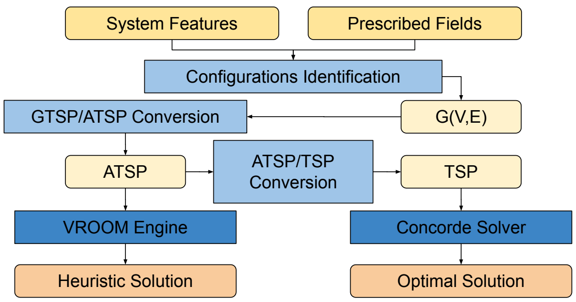

Fig. 2. Solving pipeline for NTTP

A. *Identification of Configurations*

As first step, the fields have to be translated in configurations. Specifically, let $f = (V, T)$ be the considered field: first the angles β_1 and $\beta_2 = \beta_1 + \pi$ are computed (if exist) such that the rotations around Y $rot_{Y\beta_1}$ and $rot_{Y\beta_2}$, applied to V, produce vectors V_1 and V_2 parallel to P_{ring}. In the same way, the angles α_1 and α_2, if exist, are computed, such that $V_3 = rot_{Z\alpha_1}(V_1)$ and $V_4 = rot_{Z\alpha_2}(V_2)$ have the same direction and sense of vector $(0, -1, 0)$. Finally, $T_1 = rot_{Y\beta_1}$ and $T_2 = rot_{Y\beta_2} = -T_1$ are computed and the configurations $c_1 = (-\alpha_1, -T_{1(Z)}, -T_{1(X)}, \beta_1)$ and $c_2 = (-\alpha_2, -T_{2(Z)}, -T_{2(X)}, \beta_2)$ are composed, where $T_{(X)}$ and $T_{(Z)}$ represent the X and Z components of vector T. The starting configuration $c_0 = (0, 0, 0, 0)$ is added to the instance, since the system is expected to start and stop in this position.

In this phase, the distance between each pair of configuration $\Delta(c_1, c_2) = \Delta(c_2, c_1)$ is also computed, in order to assign weights to all edges in E.

By considering a small toy-example, given the following system features:

```
betamin: [-3.14], betamax: [3.14], txmin: -1, txmax: 1,
tzmin: -1.5, tzmax: 1.5, alphamin: [-1.57], alphamax: [3.14],
omegalpha: 0.11, omegabeta: 0.21, v_x: 0.06, v_z: 0.06
```

and the set of prescribed fields:

```
V:[ 0.58,  0.66, -0.47], T:[-0.68,  0.,  0.50]
V:[-0.37,  0.58,  0.73], T:[-0.45,  0., -0.37]
V:[-0.46, -0.49,  0.74], T:[-0.17,  0.,  0.62]
```

the system computes the following set of configurations:

```
(0, 1): [0, 0, 0, 0], (1, 1): [2.29, 0.04, 0.84, -0.68]
(2, 1): [2.18, -0.57, 0.12, 2.04],
(3, 1): [-1.06, -0.19, 0.62, -1.01],
(3, 2): [1.06, 0.19, -0.62, 2.13],
```

that represents a graph composed by 4 clusters, one of which includes 2 twin configurations. Note that values are reported after rounding to help readability. Angle ranges for bed and nozzle are represented as lists (in this case with a single elements) because the system is designed to work with angles represented in the range $[-\pi, \pi)$, thus some intervals have to be described with two ranges, also if continuous. In the proposed example, by assuming the distance computed as L_∞ norm, the system produces the following distance matrix:

```
[[0.,    21.84, 20.80, 10.33, 10.33]
 [21.84, 0.,    13.01, 31.90, 24.39]
 [20.80, 13.01, 0.,    30.86, 12.60]
 [10.33, 31.90, 30.86, 0.,    20.66]
 [10.33, 24.39, 12.60, 20.66, 0.   ]]
```

B. *GTSP to ATSP conversion*

The transformation follows the approach by Noon and Bean [22]. Let $G(V, E)$ be the graph representing the instance, with distance matrix c, the algorithm produces a new matrix c', maintaining the original vertex set by (i) picking an arbitrary tour between vertex of each cluster and set $c'_{i,j} = 0$ if j follows i in the tour; (ii) setting $c'_{i,j} = M + c_{k,j}$, where k is the vertex following i in the in-cluster tour and M is a constant larger than the sum of the $|V|$ largest distances; (iii) set any other $c'_{i,j}$ to $2M$. The algorithm produce a ATSP equivalent instance, in the sense that the optimal solution can be retrieved by choosing the first vertex visited in each cluster and the minimum distance d^* can be computed from the obtained cost d'^* as $d^* = d'^* - nM$ (with n being the number of clusters). In the proposed pipeline, M is computed as $|V| * m$, where m is the larger weight between edges (in this case $M = 5 * 31.9 = 159.5$). Therefore, by considering the previous example, the new distance matrix is:

```
[[0.,     181.34, 180.30, 169.83, 169.83]
 [181.34, 0.,     172.51, 191.40, 183.89]
 [180.30, 172.51, 0.,     190.36, 172.10]
 [169.83, 183.89, 172.10, 0.,     0.    ]
 [169.83, 191.40, 190.36, 0.,     0.    ]]
```

It is to note the presence in the distance matrix of the subtour of cost 0 in the cluster with two nodes, induced by the GTSP-ATSP conversion and the asymmetry that arises in presence of cluster with two configurations.

C. *VROOM solver*

VROOM is an open-source optimization engine designed for solving instances of VRP, that is a generalization of ATSP, thus it can be used to solve the problem in the current asymmetric form. VROOM is implemented in C++ and can be fed with *json* input file. Here, the ATSP is formalized as a VRP with just one vehicle, that must start and return in a single depot and visit all the customers. The following json code represents an example of how the input for VROOM is prepared, starting from the data considered in the toy example. The first job is the starting configuration c_0. In order to feed VROOM with integer values, the weights have been casted

[1]https://github.com/VROOM-Project/vroom

[2]https://www.math.uwaterloo.ca/tsp/concorde.html

after a multiplication by 10, meaning that the system has a precision of 0.1 seconds.

```
{'matrices': {'car': {'durations':
  [[0,    1813, 1803, 1698, 1698],
   [1813, 0,    1725, 1914, 1839],
   [1803, 1725, 0,    1907, 1721],
   [1698, 1839, 1721, 0,    0],
   [1698, 1914, 1904, 0,    0]] } },
 'vehicles': [{'id': 0, 'start_index': 0,
               'end_index': 0}],
 'jobs': [ {'id': 0, 'location_index': 0},
  {'id': 1, 'location_index': 1},
  {'id': 2, 'location_index': 2},
  {'id': 3, 'location_index': 3},
  {'id': 4, 'location_index': 4}  ]}
```

D. ATSP to TSP conversion

In this step, the transformation proposed by Jonker and Volgenant [23] is applied. It reformulates ATSP to TSP by creating a dummy duplicate of each vertex $i \in V$. From the original complete directed graph, a complete undirected one with $2|V|$ vertices is obtained by (i) setting to 0 the cost of the edge $(i, |V|+i)$ for each $i \in V$; (ii) assign the cost $c_{ij} + B$ to edge $(|V|+i, j)$ $\forall i, j \in V$ (where B is a sufficiently large positive value); (iii) setting to $+\infty$ the cost of any other edge [24]. The set of optimal solutions contains for sure a tour in the form $i_1 \to i_{1+|V|} \to i_2 \to i_{2+|V|} \to ... \to i_{|V|} \to i_{2|V|}$. The solution for the original problem can be retrieved from this tour by removing nodes with indices greater than $|V|$. Similarly, the optimal cost of the original problem can be retrieved by subtracting $|V| * B$ from the obtained one.

E. Concorde Solver

In the proposed pipeline the solving of TSP instances has been managed by using Concorde. Concorde is a computer code for the symmetric TSP and some related network optimization problems, written in C. Here, the executable version of the solver for linux, in conjunction with QSopt linear programming solver [3], has been employed. In particular, B is chosen as $2M$ and the value $+\infty$ is replaced by $4M$. As for VROOM, values are casted to integer after being multiplied by 10 to be computed by Concorde.

In the following, it is possible to see the instance obtained from the previous steps for the considered example, written following the TSPLIB [4] format.

```
NAME: NTTP2TSP
TYPE: TSP
COMMENT: conversion from ATSP to TSP
DIMENSION: 10
EDGE_WEIGHT_TYPE: EXPLICIT
EDGE_WEIGHT_FORMAT: FULL_MATRIX
EDGE_WEIGHT_SECTION
6380 6380 6380 6380 6380 0    5003 4993 4888 4888
6380 6380 6380 6380 6380 5003 0    4915 5029 5104
6380 6380 6380 6380 6380 4993 4915 0    4911 5094
6380 6380 6380 6380 6380 4888 5104 5094 0    3190
6380 6380 6380 6380 6380 4888 5029 4911 3190 0
0    5003 4993 4888 4888 6380 6380 6380 6380 6380
5003 0    4915 5104 5029 6380 6380 6380 6380 6380
4993 4915 0    5094 4911 6380 6380 6380 6380 6380
4888 5029 4911 0    3190 6380 6380 6380 6380 6380
4888 5104 5094 3190 0    6380 6380 6380 6380 6380
EOF
```

[3]https://www.math.uwaterloo.ca/ bico/qsopt/

[4]http://comopt.ifi.uni-heidelberg.de/software/TSPLIB95/

F. Simulations

The proposed pipeline has been tested by considering common system features and a variable amount of randomly generated prescribed fields, by employing Concorde and VROOM as solvers. Simulations have been conducted on a desktop PC (HP® Omen 30L) with a AMD™ Ryzen 9 5900x processor, 32 GB of memory, running Python 3.8.10 over Ubuntu 20.04.4 LTS operating system. Simulations take into account three feature sets, namely SF1, SF2 and SF3, that share all the parameters ($\omega_\alpha = 0.105$ rad/s, $tx_{rng} = [-1.0, 1.0]$ m, $tz_{rng} = [-1.5, 1.5]$ m, $v_x = v_z = 0.06$ m/s, $\beta_{rng} = [-\pi, \pi)$ rad, $\omega_\beta = 0.209$ rad/s), but differ with regards to the allowed position of the nozzle, α_{rng} ($[-\pi, \pi)$, $[-pi/2, \pi)$ and $[0, \pi]$ rad, respectively). For each feature set, 5 to 100 (step 5) prescribed fields has been simulated 100 times both using Concorde and VROOM and, for each simulation, the required computation time has been evaluated. Moreover, each simulation has been performed twice, in order to take into account both L_∞ and L_1 norms as distances. Results show differences induced by different system features: since $SF2$ and $SF3$, starting from the same amount of required fields, allow less configurations than $SF1$, their instances are easier to solve and, thus, computation time is lower. However, all the accounted combination of system feature and metric share the same trend in the results. Here, as representative example, the results obtained with $SF1$ and L_∞ norm are reported and discussed. All the considered instances, together with solutions obtained by both solvers and the source code, are available on an online repository[5]. Figure 3 reports the average computation time over the number of prescribed fields both for Concorde and VROOM. Both approaches demonstrated to be largely capable to manage an amount of fields much greater than those usually employed in clinical practice. As it is possible to observe, VROOM showed an average computation time several orders of magnitude lower than Concorde (0.01 seconds versus 29.62 seconds for 100 fields), confirming its efficiency in addressing VRP instances. Moreover, it shows a linear trend in the solving time, suggesting that it would be capable to manage also bigger instances of NTTP. However, being VROOM an heuristic, the provided result are rarely optimal, especially with the increase of prescribed fields. On the other hand, result obtained by VROOM are quite close to the optimum: Figure 3 also shows the estimated travel time computed by Concorde and VROOM and the relative error of the heuristic with respect to the exact solution (expressed in percentage). Here it is possible to observe that the curve draws a plateau between 7% and 8% after 70 fields, suggesting an asymptotic behavior even for larger instances.

V. CONCLUSION

In this work, we addressed the problem of minimizing the nozzle travel time in proton therapy, on one hand by proposing a formalization of NTTP and, on the other hand, by providing a modular pipeline to solve the problem. Obtained results show

[5]https://doi.org/10.5281/zenodo.6686176

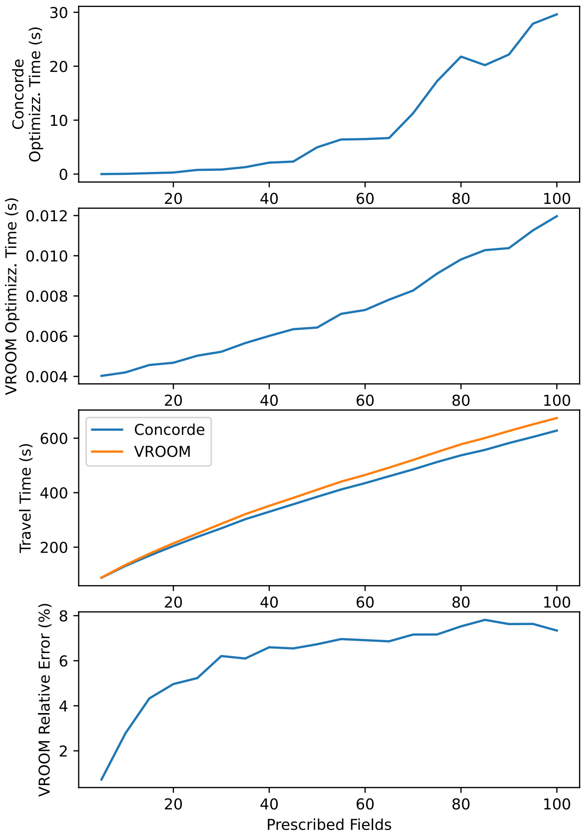

Fig. 3. Simulation results for SF1 and L_∞ norm distance. From top to bottom: average optimization time over the number of prescribed fields for Concorde (i) and VROOM (ii), computed nozzle travel time for both solvers (iii), relative distance from optimum achieved by VROOM (iv).

that both the exact and the heuristic approaches are capable to solve instances with a large number of fields, with respect to those usually employed in actual proton therapy treatment plans. In particular, the heuristic solver based on VROOM has shown to be very effective in terms of computation time, at cost of a little increment of the obtained travel time. Even if the gain in terms of computational time could be negligible when applied to Intensity Modulated Proton Therapy, where the amount of prescribed fields is in general low and the exact solver can efficiently solve the problem, it represents a quite promising advantage if applied to more recent techniques such as Proton Arc Therapy (PAT) [25], where the radiation is continuously on during the movement of the gantry and the complexity of the problem increases.

Future developments will include a theoretical perspective on the complexity of NTTP as well as the development and the test of alternative solving strategy to be included in the pipeline.

References

[1] W. D. Newhauser and R. Zhang, "The physics of proton therapy," *Physics in Medicine & Biology*, vol. 60, no. 8, p. R155, 2015.

[2] T.-Z. Yuan, Z.-J. Zhan, and C.-N. Qian, "New frontiers in proton therapy: applications in cancers," *Cancer Communications*, vol. 39, no. 1, pp. 1–7, 2019.

[3] R. Mohan and D. Grosshans, "Proton therapy–present and future," *Advanced drug delivery reviews*, vol. 109, pp. 26–44, 2017.

[4] H. Paganetti, C. Beltran, S. Both, L. Dong, J. Flanz, K. Furutani, C. Grassberger, D. R. Grosshans, A.-C. Knopf, J. A. Langendijk *et al.*, "Roadmap: proton therapy physics and biology," *Physics in Medicine & Biology*, vol. 66, no. 5, p. 05RM01, 2021.

[5] H. Nystrom, M. F. Jensen, and P. W. Nystrom, "Treatment planning for proton therapy: what is needed in the next 10 years?" *The British Journal of Radiology*, vol. 93, no. 1107, p. 20190304, 2020.

[6] M. Spezialetti, F. Lapenna, P. Caianiello, F. Fracchiolla, F. Muciaccia, G. Placidi, G. Russo, and F. Mignosi, "Using deep learning for fast dose refinement in proton therapy," in *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2021, pp. 1783–1789.

[7] E. Altobelli, M. Amichetti, A. Langiu, F. Marzi, F. Mignosi, P. Pisciotta, G. Placidi, F. Rossi, G. Russo, M. Schwarz *et al.*, "Combinatorial optimisation in radiotherapy treatment planning," *AIMS Medical Science*, vol. 5, no. 3, pp. 204–223, 2018.

[8] M. L. Kessler, D. L. Mcshan, M. A. Epelman, K. A. Vineberg, A. Eisbruch, T. S. Lawrence, B. Fraass *et al.*, "Costlets: A generalized approach to cost functions for automated optimization of imrt treatment plans," *Optimization and Engineering*, vol. 6, no. 4, pp. 421–448, 2005.

[9] P. Orponen and H. Mannila, "On approximation preserving reductions: Complete problems and robust measures (revised version)," *Department of Computer Science, University of Helsinki*, p. 79, 1990.

[10] M. R. Garey and D. S. Johnson, *Computers and Intractability; A Guide to the Theory of NP-Completeness*. USA: W. H. Freeman & Co., 1990.

[11] A. Henry-Labordere, "The record balancing problem: A dynamic programming solution of a generalized traveling salesman problem," *Revue Francaise D Informatique DeRecherche Operationnelle*, vol. 3, no. 2, pp. 43–49, 1969.

[12] M. Fischetti, J.-J. Salazar-González, and P. Toth, "The generalized traveling salesman and orienteering problems," in *The traveling salesman problem and its variations*. Springer, 2007, pp. 609–662.

[13] M. Fischetti, J. J. S. González, and P. Toth, "The symmetric generalized traveling salesman polytope," *Networks*, vol. 26, no. 2, pp. 113–123, 1995.

[14] D. Ben-Arieh, G. Gutin, M. Penn, A. Yeo, and A. Zverovitch, "Transformations of generalized atsp into atsp," *Operations Research Letters*, vol. 31, no. 5, pp. 357–365, 2003.

[15] K. Helsgaun, "Solving the equality generalized traveling salesman problem using the lin–kernighan–helsgaun algorithm," *Mathematical Programming Computation*, vol. 7, no. 3, pp. 269–287, 2015.

[16] G. Gutin and A. P. Punnen, *The traveling salesman problem and its variations*. Springer Science & Business Media, 2006, vol. 12.

[17] S. Alatartsev, S. Stellmacher, and F. Ortmeier, "Robotic task sequencing problem: A survey," *Journal of intelligent & robotic systems*, vol. 80, no. 2, pp. 279–298, 2015.

[18] X. Wang, X. Zhou, Z. Xia, and X. Gu, "A survey of welding robot intelligent path optimization," *Journal of Manufacturing Processes*, vol. 63, pp. 14–23, 2021.

[19] R. A. C. Siochi, "Minimizing static intensity modulation delivery time using an intensity solid paradigm," *International Journal of Radiation Oncology* Biology* Physics*, vol. 43, no. 3, pp. 671–680, 1999.

[20] M. Ehrgott, Ç. Güler, H. W. Hamacher, and L. Shao, "Mathematical optimization in intensity modulated radiation therapy," *Annals of Operations Research*, vol. 175, no. 1, pp. 309–365, 2010.

[21] R. D. Filippo, "Modelli ed algoritmi di ottimizzazione combinatoria in radioterapia adronica," Master's thesis, University of L'Aquila, Department of Information Engineering, Computer Science and Mathematics, L'Aquila, Italy, December 2021, supervisors: Filippo Mignosi, Fabrizio Rossi, Matteo Spezialetti.

[22] C. E. Noon and J. C. Bean, "An efficient transformation of the generalized traveling salesman problem," *INFOR: Information Systems and Operational Research*, vol. 31, no. 1, pp. 39–44, 1993.

[23] R. Jonker and T. Volgenant, "Transforming asymmetric into symmetric traveling salesman problems," *Operations Research Letters*, vol. 2, no. 4, pp. 161–163, 1983.

[24] M. Fischetti, A. Lodi, and P. Toth, "Solving real-world atsp instances by branch-and-cut," in *Combinatorial Optimization—Eureka, You Shrink!* Springer, 2003, pp. 64–77.

[25] X. Li, G. Liu, G. Janssens, O. De Wilde, V. Bossier, X. Lerot, A. Pouppez, D. Yan, C. Stevens, P. Kabolizadeh *et al.*, "The first prototype of spot-scanning proton arc treatment delivery," *Radiotherapy and Oncology*, vol. 137, pp. 130–136, 2019.