IPEM
Institute of Physics and
Engineering in Medicine

PAPER

Optimizing radiotherapy plans for cancer treatment with Tensor Networks

View the article online for updates and enhancements.
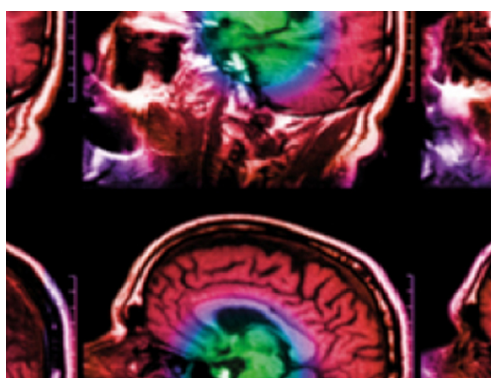
IPEM | IOP

Series in Physics and Engineering in Medicine and Biology

Your publishing choice in medical physics, biomedical engineering and related subjects.

Start exploring the collection–download the first chapter of every title for free.

Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

PAPER

Optimizing radiotherapy plans for cancer treatment with Tensor Networks

Samuele Cavinato[1,2,*] , Timo Felser[1,3,4,5] , Marco Fusella[2] , Marta Paiusco[2] and Simone Montangero[1,3,6]

1 Dipartimento di Fisica e Astronomia 'G. Galilei', Università degli Studi di Padova, I-35131 Padova, Italy
2 Medical Physics Department, Veneto Institute of Oncology IOV-IRCCS, I-35128 Padua, Italy
3 INFN, Sezione di Padova, I-35131 Padova, Italy
4 Theoretische Physik, Universität des Saarlandes, D-66123 Saarbrücken, Germany
5 Tensor Solutions, Institute for Complex Quantum Systems, University of Ulm, D-89069 Ulm, Germany
6 Padua Quantum Technologies Research Center, Università degli Studi di Padova, I-35131 Padova, Italy
* Author to whom any correspondence should be addressed.

E-mail: samuele.cavinato@gmail.com

Abstract

We present a novel application of Tensor Network methods in cancer treatment as a potential tool to solve the dose optimization problem in radiotherapy. In particular, the intensity-modulated radiation therapy technique—that allows treating irregular and inhomogeneous tumors while reducing the radiation toxicity on healthy organs—is based on the optimization problem of the beamlets intensities that shall result in a maximal delivery of the therapy dose to cancer while avoiding the organs at risk of being damaged by the radiation. The resulting optimization problem is expressed as a cost function to be optimized. Here, we map the cost function into an Ising-like Hamiltonian, describing a system of long-range interacting qubits. Finally, we solve the dose optimization problem by finding the ground-state of the Hamiltonian using a Tree Tensor Network algorithm. In particular, we present an anatomical scenario exemplifying a prostate cancer treatment. A similar approach can be applied to future hybrid classical–quantum algorithms, paving the way for the use of quantum technologies in future medical treatments.

1. Introduction

Radiotherapy is one of the techniques used to treat solid tumors by means of ionizing radiation. The radiation dose released into the cancer tissue damages the DNA of the tumor cells leading to their death or slowing down the growth of the tumor (Baskar *et al* 2014, Wang *et al* 2018a). Radiotherapy is often used in combination with other therapies like surgery, chemotherapy or immunotherapy to improve their global efficacy (Ma *et al* 2010, Formenti and Demaria 2013, Brunner 2016, Wang *et al* 2018b).

One of the hardest challenges encountered while treating patients with ionizing radiations is to deliver an optimal dose to the targeted tumor while keeping the radiation as low as possible in the surrounding healthy tissues. One of the most frequently used techniques nowadays is the *intensity-modulated radiation therapy* (IMRT) (Taylor and Powell 2004, Bortfeld 2006, Zahra *et al* 2018, Elith *et al* 2011). In IMRT, the applied radiation is modulated to reach an optimal dose distribution inside the patient. This optimal modulation is obtained by solving a highly non-trivial numerical optimization problem with a high number of optimization parameters and numerous constraints on the final radiation dose distribution. Over the past years, several numerical techniques have been developed to address this challenge (Webb 1995, Ezzell 1996, Cotrutz *et al* 2001, Bortfeld 2006, Censor and Unkelbach 2012). In 2015, Nazareth and Spaans proposed to solve the IMRT beam fluence optimization problem using the D-wave annealer (Nazareth and Spaans 2015), and more recently Pakela *et al* proposed an approach exploiting the simulated quantum tunnelling effect (Pakela *et al* 2020).

Despite these remarkable novel approaches paved the way for future applications of quantum computation in medicine, their application is still limited. On the one hand, this is due to the lack of scalable quantum hardware. On the other hand, it is not straightforward to extend what is done on classical computers to quantum ones due to the lack of a robust and clear strategy to map the classical IMRT problem to quantum hardware. To try to overcome these limitations as well as to further investigate the applicability of quantum-inspired techniques to the solution of classical optimization problems and foster future applications of quantum technologies to medicine, here we apply Tensor Networks (TNs) to an IMRT dose optimization problem.

TNs are one of the most successful algorithms for simulating quantum many-body systems on classical computers. Indeed, whenever possible, they efficiently represent quantum many-body wavefunctions in a compact form on classical computers (Evenbly and Vidal 2009, Schollwöck 2011, Orús 2014, Montangero 2018, Silvi *et al* 2019). In the last few decades, TNs have proven their effectiveness in the research and analysis of quantum many-body systems, especially for low-dimensional ground-state (Felser *et al*, McCulloch 2007, Singh and Vidal 2013, Dalmonte and Montangero 2016, Gerster *et al* 2017, Bañuls *et al* 2019, Felser *et al* 2019, Bañuls and Cichy 2020). In addition to that, thanks to the properties they share with quantum hardware, TNs may play the role of test benches for the development of quantum algorithms (Kim and Swingle 2017, Huggins *et al* 2019, Zhou *et al* 2020).

Hereafter, we show how to solve an IMRT optimization problem with TNs. We first introduce how the classical cost function is mapped into an Ising-like Hamiltonian, where the optimization variables are represented as a set of long-range interacting spins. Finally, we solve the classical optimization problem by finding the ground-state for this Hamiltonian using TNs. We present the application of TNs to two different toy models and to a more realistic anatomical scenario simulating a prostate cancer treatment. We show that TNs results are compatible with other classical techniques, quadratic programming (QP) and simulated annealing (SA). Our results pave the way to the application of TNs methods to more complex and realistic clinical scenarios, contributing to building solid foundations for future applications of quantum computation to medicine. In the midterm, we foresee the development and application of classical TN methods to the solution of the IMRT problem.

The manuscript is structured as follows. In section 2.1 we provide a brief introduction to the dose optimization problem in IMRT together with its basic mathematical description. In section 2.2 the mapping procedure of the classical cost function to the Ising-type Hamiltonian is described in more detail. Furthermore, this section includes a brief introduction to TNs as well. Finally, we present and discuss the main results obtained from this study in section 3.
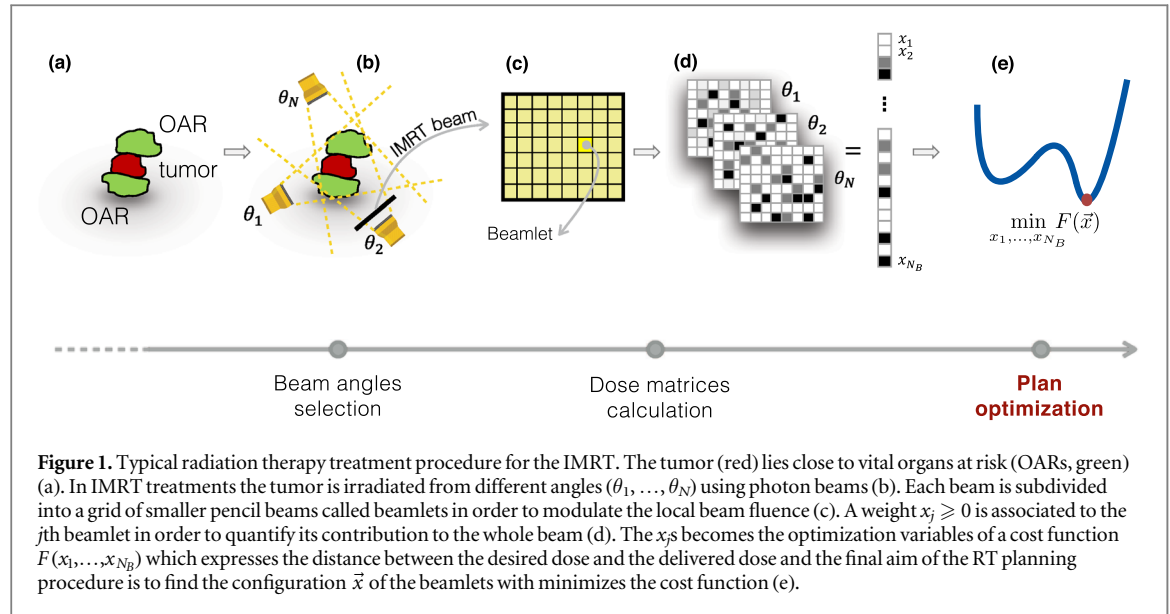
2. Methods

In what follows, we describe the classical optimisation problem behind the IMRT in more detail, followed by a description of the procedure we implemented to map the classical cost function to a classical Hamiltonian. In the last part, we provide a brief description of TNs, focusing on how they can be used to solve the initial IMRT optimization problem.

2.1. Radiotherapy optimization problem

In radiotherapy, cancer cells are treated by releasing a certain amount of radioactive dose inside the tumor. Modern radiotherapy offers various techniques for treating tumors (Teoh *et al* 2011, Alongi *et al* 2012, Baskar *et al* 2012, Jaffray 2012, Chargari *et al* 2019) and the choice of one of them depends on factors like the site of the disease, type of cancer and overall patient's conditions (Delaney *et al* 2017, Adeberg *et al* 2018, Zhang *et al* 2020). One of the most impacting techniques in terms of improvement of treatments quality is the IMRT which became clinically available in its first implementations in the early 2000s (Bortfeld 2006, 2001 *Int. J. Radiat. Oncol. Biol. Phys.* **51** 880).

The goal of IMRT is to create a personalized dose distribution for each patient's anatomy that ensures the appropriate dose for the tumor while saving the organs at risk (OARs) as much as possible. The choice of IMRT treatment is nowadays mandatory in very challenging cases where a high dose is required for very irregular tumor shapes surrounded by critical OARs. In order to effectively treat this kind of diseases, the IMRT benefits from a non-uniform intensity distribution of the radiation beams. In particular, the radiation beams are modulated by dividing the fluence of each treatment beam into a certain number of smaller pencil-beams called *beamlets* that can be delivered through the movement of the multileaf collimator.

However, finding the ideal intensity of every single beamlet for a desired treatment that optimises the relation between fields arrangement and dose distribution inside the patient is a highly complex problem. This optimisation problem is typically solved as an *inverse problem* which is encoded into a cost function to be minimized in order to find the optimal beamlet configuration.

Figure 1. Typical radiation therapy treatment procedure for the IMRT. The tumor (red) lies close to vital organs at risk (OARs, green) (a). In IMRT treatments the tumor is irradiated from different angles ($\theta_1, \ldots, \theta_N$) using photon beams (b). Each beam is subdivided into a grid of smaller pencil beams called beamlets in order to modulate the local beam fluence (c). A weight $x_j \geqslant 0$ is associated to the jth beamlet in order to quantify its contribution to the whole beam (d). The x_js becomes the optimization variables of a cost function $F(x_1, \ldots, x_{N_B})$ which expresses the distance between the desired dose and the delivered dose and the final aim of the RT planning procedure is to find the configuration $\vec{x}$ of the beamlets with minimizes the cost function (e).

In the past years, different approaches have been proposed, among them, the analytic transform method (Bortfeld and Boyer 1995), algebraic solutions (Censor and Unkelbach 2012), gradient descent (Cotrutz *et al* 2001), genetic algorithms (Ezzell 1996) and SA (Webb 1995). Although these powerful methods have been used in everyday clinical practice with increasing success, finding patient-specific plans is still an open problem since it should account for many factors, increasing the complexity of the problem. Therefore, new mathematical and physical solutions need to be developed in order both to spare precious computational time and to improve the quality of the treatments delivered, with the aim of enhancing our capability to save human lives in the fight against cancer.

In what follows, we present the main elements of radiotherapy, explaining the general planning procedure for an IMRT treatment, focusing on the underlying numerical optimization problem.

Figure 1 illustrates a typical treatment procedure. The targeted tumor (red), located between two OARs (green), is irradiated by several photon beams from different angles θ_k (see figure 1(b)). The underlying geometry is defined in the dosimetric plans created by the medical physicist. The desired dose to the target and the OARs are defined in the treatment plan, following the goal of sparing healthy organs while the proper dose is delivered to the target.

To achieve this treatment goal, the IMRT optimization process begins with the splitting of each beam fluence into a grid of beamlets as shown in figure 1(c). Only beamlets that traverse the target can be optimized, and the intensity of each jth beamlet can be modulated independently with a weight x_j (figure 1(d)). The volume of the patient considered for the optimization can be divided in 3D finite-size elements called *voxels*. The contribution of each beamlet to each voxel depends on their intensity as well as on the geometry and physical properties of each beam and patient's characteristics.
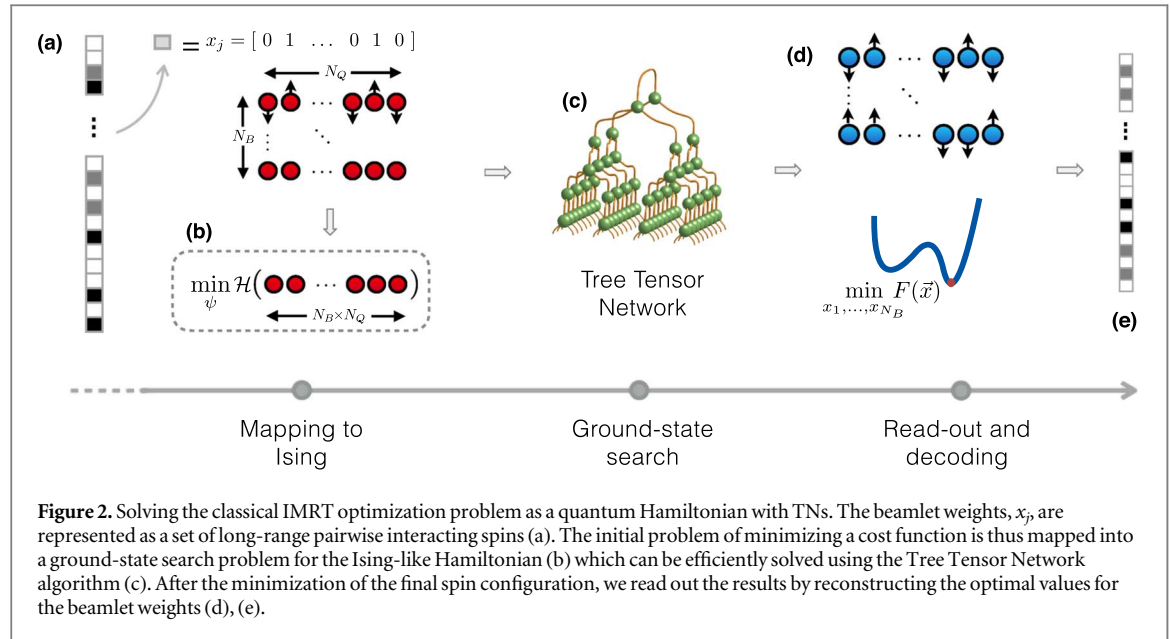
All these information are gathered in the *influence matrix A* which maps every beamlet to each voxel. The influence matrix is usually provided by the treatment planning system used in the clinical environment, and a complete description of its computation goes far beyond the aim of this work.

The sum of all beamlets produces the total dose within a given voxel i and contributes to the definition of a global dose distribution $D(\mathbf{x})$. Thus, the total dose delivered to the voxel i can be expressed as:

$$D_i(\mathbf{x}) \equiv D_i(x_1, \ldots, x_{N_B}) = \sum_{j=1}^{N_B} a_{ij} x_j \,, \tag{1}$$

where N_B is the total number of beamlets and a_{ij} describes the so-called *influence matrix*, A, giving the unmodulated contribution of the jth beamlet to the ith voxel.

The IMRT planning procedure is then solved through an iterative inverse planning process: the intensity of the applied photon beams are optimized towards the prescribed dose distribution $D^{(P)}$ inside the patient. The dose criteria are typically defined using dose-volume histograms (DVH). In particular, we minimize the distance between the delivered dose $D(\mathbf{x})$ and the desired dose in the patient $D^{(P)}$ in the discretized volume with the beamlet weights $\mathbf{x}$ being the optimization variables (see in figure 1(e)). We will describe the process through the following quadratic cost function:

Figure 2. Solving the classical IMRT optimization problem as a quantum Hamiltonian with TNs. The beamlet weights, x_j, are represented as a set of long-range pairwise interacting spins (a). The initial problem of minimizing a cost function is thus mapped into a ground-state search problem for the Ising-like Hamiltonian (b) which can be efficiently solved using the Tree Tensor Network algorithm (c). After the minimization of the final spin configuration, we read out the results by reconstructing the optimal values for the beamlet weights (d), (e).

$$F(\mathbf{x}) = \sum_{r=0}^{R}\sum_{i=1}^{\mathcal{V}_r} \gamma_i [D_i(\mathbf{x}) - D_i^{(P)}]^2, \tag{2}$$

where r is an index running over all the volumes (the targeted tumor and the OARs), and γ_i a weight assigned to the ith voxel in order to prioritize certain volumes during the treatment. $\mathcal{V}_r$ gives the total number of voxels belonging to the object r.

2.2. Solving classical problems with TNs

2.2.1. Mapping the problem to a classical Hamiltonian

In the following, we describe how to rewrite the optimization problem in equation (2) into a ground-state search of a quantum many-body Hamiltonian. In particular, we here propose a procedure based on the binary-decimal conversion to map the cost function $F(\mathbf{x})$ into an Ising-type Hamiltonian, $\mathcal{H}_{IMRT}$. The procedure is summarized in figure 2.

We first discretize the weights x_j for each beamlet by a set of N_Q bits. N_Q is called the *bit-depth*. Thus, we represent

$$x_j \approx \frac{1}{(B/2)}\sum_{n=1}^{N_Q} 2^{n-1} b_n^{(j)} \tag{3}$$

with the bits $b_n^{(j)} = \{0,\ 1\}$, and introducing a normalization constant B to set the range of the beamlets weights such that $x_j \in \left[0;\ \frac{2^{N_Q}-1}{(B/2)}\right]$. With an increasing number of bits N_Q we can increase the resolution of the discretization. Then, we map the binary values $b_n^{(j)}$ into spin variables $s_n^{(j)} = \{-1;\ +1\} = 2b_n^{(j)} - 1$ for each site j. Consequently, we construct a $N_B \times N_Q$-dimensional many-body Hamiltonian with the first dimension running over the different beamlets and the second representing the discretized space for each beamlet. Inserting the spin variables together with equation (3) into the cost function of equation (2), we obtain this Ising-type Hamiltonian with the following general expression:

$$\mathcal{H}_{IMRT} = \mathcal{H}_{SP} + \mathcal{H}_{INT}^{(a)} + \mathcal{H}_{INT}^{(b)}, \tag{4}$$

where the first term describes the single-particle terms with

$$\mathcal{H}_{SP} = \sum_{j,n}\left[\sum_i \frac{\gamma_i}{B}\left(a_{ij}\sum_k a_{ik} - 2D_i^{(P)}a_{ij}\right)2^{n-1}\right]s_n^{(j)}, \tag{5}$$

the second term captures spins interacting in the same beamlet

$$\mathcal{H}_{INT}^{(a)} = \sum_j \sum_{m \neq n}\left[\sum_i \gamma_i \frac{a_{ij}^2}{B^2}2^{n-1}2^{m-1}\right]s_n^{(j)}s_m^{(j)}, \tag{6}$$

and the last term represents the interactions between different beamlets

$$\mathcal{H}_{INT}^{(b)} = \sum_{j \neq k} \sum_{n,m} \left[\sum_i \gamma_i \frac{a_{ij} a_{ik}}{B^2} 2^{n-1} 2^{m-1} \right] s_n^{(j)} s_m^{(k)}. \tag{7}$$

We point out, that this Hamiltonian describes a two-dimensional fully connected lattice of long-range interacting spins with N_B sites in one direction and N_Q in the other. At this point, we can solve the initial problem of minimizing equation (2) by finding the ground-state for the classical Ising-type Hamiltonian in equation (4).

The ground state of the system is given by the classical spin-configuration which provides the lowest energy E_0, that is the configuration that minimizes equation (2). The case $E_0 = 0$ corresponds to the optimal beamlet setup which results in exactly the desired dose distribution within the patient. Anyhow, in practice, the exact desired dose distribution is not always achievable for the optimization. Thus, we can find the solution to the initial optimization problem by going through all possible spin-configurations of the Ising-type Hamiltonian. However, this classical search rapidly becomes unfeasible, as the number of spin-configurations grows exponentially with increasing system size. And since the number of beamlets required in a radiotherapy treatment can easily get in the order of a few thousand, the number of spins required to represent them is practically too high to solve the problem efficiently in the classical regime. For this reason, in the following section, we introduce an approach based on TNs to address this complex optimization task.

2.2.2. Description of Tree Tensor Networks (TTN) for solving quantum many-body problems
Considering the classical Hamiltonian $\mathcal{H}_{IMRT}$, we further allow each spin to be a quantum variable by representing $s_n^{(j)}$ with the Pauli matrix σ_z. In this way, the quantum Ising-type Hamitonian $\mathcal{H}_{IMRT}$ is a diagonal matrix with its entries corresponding to the energies of all possible spin combinations of the classical Hamiltonian. Here, TNs are a vital tool for finding ground states and their physical properties despite the exponentially growing Hilbert space (Nielsen and Chuang 2000, Schollwöck 2011, Montangero 2018). In the following, we introduce the idea of TNs to address this challenging task of investigating complex quantum many-body systems. For a more in depth introduction on TNs we refer to sophisticated literature (Schollwöck 2011, Orús 2014, Montangero 2018, Silvi *et al* 2019).

TNs are used to efficiently represent quantum many-body wavefunctions $|\psi\rangle$, which live in the tensor product $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \mathcal{H}_N$ of N local Hilbert spaces $\mathcal{H}_k$, each assumed to be of finite dimension d. Expressing such a state in terms of a product basis in real-space means decomposing the wavefunction as

$$|\psi\rangle = \sum_{i_1,\ldots i_L = 1}^{d} c_{i_1,\ldots,i_L} |i_1\rangle_1 \otimes |i_2\rangle_2 \otimes \ldots \otimes |i_L\rangle_L , \tag{8}$$

where $\{|i_k\rangle\}_i$ is the canonical basis of site k, spanning $\mathcal{H}_k$. The exact description of such a general state by all possible combinations of local states requires d^N coefficients $c_{i_1,\ldots,i_N}$. Thus, the number of coefficients increases exponentially with the system size N in the exact representation of the wave-function.

TNs offer a more efficient representation by decomposing the complete rank-N tensor (containing all d^N coefficients) into a set of local tensors with smaller rank, connected with auxiliary indices. We control the dimension of the auxiliary indices with the bond-dimension χ and thereby the amount of captured information. Thus, tuning this parameter χ, TNs interpolate between a product state, where quantum correlations are neglected, and the exact, but inefficient representation.

The decomposition of the complete rank-N tensor can be executed in several ways giving rise to different TN geometries. The most prominent TN representations are the matrix product states (MPS) for 1D systems (Östlund and Rommer 1995, Verstraete and Cirac 2006, Schollwöck 2011), which addresses all sites with one corresponding tensor and their two-dimensional variant, the projected entangled pair states (PEPS) (Verstraete and Cirac 2004, Verstraete *et al* 2006, Orús 2014). TTN (Shi *et al* 2006, Silvi *et al* 2010, Gerster *et al* 2014, 2017) with their hierarchical structure can in principle be defined in any lattice dimension. As an illustration of the representation power, the MPS, for instance, reduces the number of parameters to the upper bound of $Nd\chi^2$ controlled by the chosen bond-dimension χ, leading to a linear dependence on the system size N rather than an exponential one.

In our analysis, we use the latter, a TTN. The TTN offers better connectivity between long-range interactions (with a logarithmic distance through the network), while for the simpler MPS the distance by connecting tensors within the network is linear. In contrast to the PEPS, the TTN can be optimized with a lower computational complexity as well ($\mathcal{O}(\chi^4)$ for the TTN versus $\mathcal{O}(\chi^{10})$ for the PEPS).

Due to the bond-dimension $\chi > 1$, we perform a quantum ground state search within the subspace limited by χ. Thus, in contrast to the classical optimization routines, we explore within one optimization step several classical solutions as they are superposed in the quantum representation of the TTN. This allows us further to tunnel through higher, but reasonably thin, potentials within the optimization landscape.

After the TTN analysis, the obtained ground state of the quantum Hamiltonian might, in general, not be classical but rather a superposition of classical states. However, we know, that all solutions to the initial IMRT problem are classical, thus we expect that each of the superposed classical solutions separately has the same ground state energy E_0 once the TTN algorithm is fully converged to the quantum ground state. Therefore, we enforce the 'classicality' of the solution as a constraint after the TTN minimisation problem by selecting one of the classical solutions from the converged TTN by truncating the bond-dimension to $\chi = 1$, resulting in a separable, product-state solution

$$|\Psi_{\chi=1}\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \otimes ... \otimes |\psi_L\rangle \,,$$

in which the only superposition can be local (such as a local site being $|\psi_n\rangle = 1/\sqrt{2} \, (|\uparrow\rangle + |\downarrow\rangle)$). From here on, we measure the quantum observable $\langle \sigma_n^z \rangle$ for each site n, resulting in $\langle \sigma_n^z \rangle = -1$ for a spin down, $\langle \sigma_n^z \rangle = 1$ for a spin up and in between those in case of a local superposition. In the case of the latter, we project the spin to the classical one with the highest probability by using the sign. The resulting spin configuration is further mapped back to the binary encoded solution for each voxel as described in the previous section. In fact, our analysis confirms that most of the times the resulting TTN ground state is indeed a classical state ($\chi = 1$), or in some scenarios that the obtained state is, as expected, a superposition of degenerate classical solutions, each of them with the same energy E_0.

In the case of the IMRT optimization problem, we are dealing with a non-trivial quantum spin-glass Hamiltonian type with over 32 000 long-range interactions. Thus, the optimization is a highly non-trivial task and the TTN-algorithm can be quite sensitive to the initialization procedure. Therefore, for each run, we randomly initialize several samples of the TTN from which we start the optimization. In the end, we can verify the best simulation by comparing the resulting energies.

3. Results

In the following, we show the results of TTN for two different toy models to validate our approach, followed by a more realistic anatomical scenario simulating a prostate IMRT treatment. We first compare the TTN algorithm with the analytical solution for a 3D box toy-model scenario. For the second, more complex two-sphere model and the more realistic IMRT phantom, we lack a general analytical solution. Therefore, we test the results of our TTN approach by comparing it with two different optimization methods, QP and SA. In particular, QP refers to a set of methods and algorithms used to solve quadratic optimization problems subject to linear constraints and it was exploited to address the initial optimization problem in equation (2). On the other hand, the problem in the Ising-type formulation in equation (4) is addressed using both TTN and SA. QP and SA are used to validate the results obtained with the TTN approach. For further details on the QP and SA, we refer to appendix A.

We point out, that both SA and TNs algorithms contain elements of randomness: samples of independent and randomly initialized simulations are always collected and the best solution considered. This also allows us to calculate the standard deviation of the samples and have a quantitative idea of the general behaviour of the algorithms.
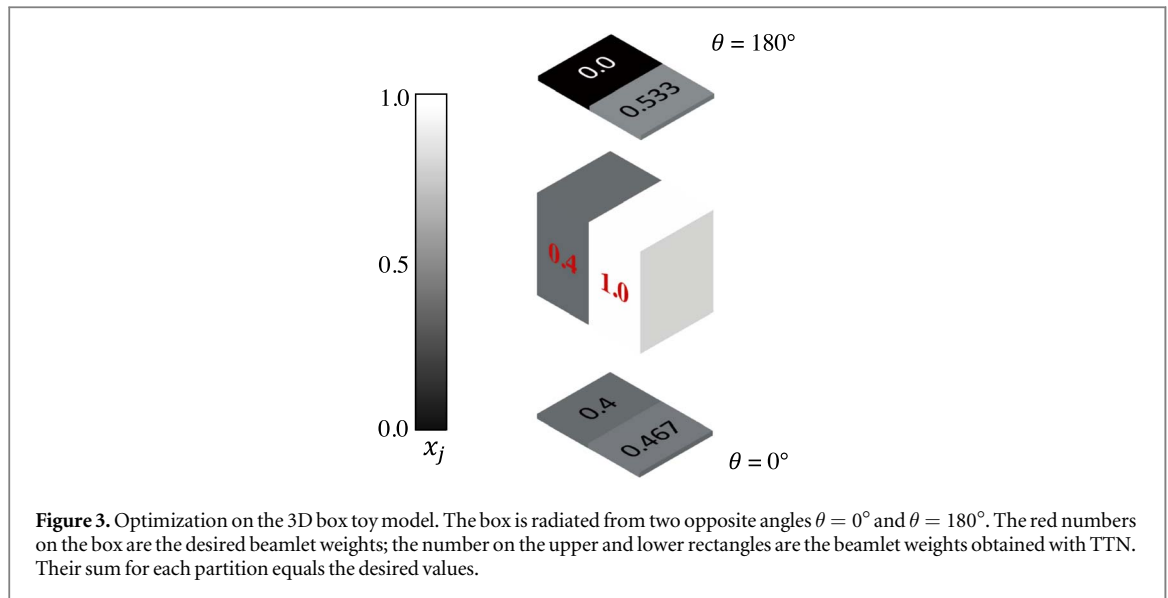
3.1. Toy models

3.1.1. Bipartite box

In this section, we describe a simplified analytical model used to validate the correctness of the mapping discussed in section 2.2.1 and further our TN approach.

The model consists of a 3D box subdivided into two different regions for which we assign specific dose prescriptions as shown in figure 3. The red number on the front of the box corresponds to the desired beamlet weights for each bipartition. We radiate the cube with two beams from two opposite directions (i.e. $\theta_1 = 0°$ and $\theta_2 = 180°$, here represented as two rectangles, with a variable number of beamlets N_B for each beam, two in this example. The radiation beam is modelled as an *ideal beam* that releases the same amount of dose to each voxel. Moreover, we neglect scattering effects limiting the interactions among different beamlets to those acting geometrically on the same voxels only. The dose prescriptions and the number of voxels are chosen in a way to ensure the exact ground state with energy $E_0 = 0$. Therefore, for instance, we choose dose values that are compatible with the number of discretization levels 2^{N_Q} used in equation (3).

We point out, that these conditions are not fulfilled in general cases. Anyhow, we introduce them to obtain an exactly solvable model for the sake of validation. Later on, we present a more realistic anatomical scenario of cancer treatment. The model is then described by the influence matrices A_r for each region r, the priorities γ_i and the dose prescriptions $D_i^{(P)}$ for each voxel. These are the same information that can be extracted from a real therapy planning system, as we will see also in the following.

The number of bits used to represent each beamlet is set to $N_Q = 4$ and will be kept constant throughout the whole study. This number is choosen following the IMRT literature where the effect of beamlet intensities

Figure 3. Optimization on the 3D box toy model. The box is radiated from two opposite angles $\theta = 0°$ and $\theta = 180°$. The red numbers on the box are the desired beamlet weights; the number on the upper and lower rectangles are the beamlet weights obtained with TTN. Their sum for each partition equals the desired values.

discretization has been widely studied in the past (Sun *et al* 2004). The target dose prescriptions for the two partitions are $D_{left}^{(P)} = 6.0$ and $D_{right}^{(P)} = 15.0$, which are arbitrary number chosen according to the number of discretization levels for the x_j, 16 in this case. The influence matrix, mapping the beamlets intensities to the voxels, is defined to have a uniform dose release into the box. Thereby, as we simplified the physical effects of the beam, we end up with the influence matrix consisting of either zero and otherwise constant entries mapping the beamlet intensities to the voxels.

The values contained inside the upper and lower and rectangles correspond to the beamlet weights obtained with the TTN algorithm. Wo observe that their sum equals the desired values for each bipartition, showing a perfect agreement between the analytical and numerical solution. This proves the well functioning of the algorithm as well as the correctness of the mapping procedure.

We point out, that, in general, there can be more than one configuration satisfying the constraints: indeed, depending on the system parameters, the ground state of the Ising–Hamiltonian can be degenerate. Anyhow, for the treatment, we are satisfied obtaining anyone of the degenerate ground states.

3.1.2. Sphere

In this paragraph, we illustrate the application of the TTN algorithm to a more realistic clinical scenario. In this case, instead of assuming an *ideal beam*, we model the photon beam using the Matlab-based software CERR (Apte *et al* 2010). This software generates radiotherapy plans and can be used to obtain the influence matrixes a_{ij}. It allows working with physical effects introduced by the beams such as scattering and further allows for higher freedom in choosing the geometry, the number of beams and other typical model parameters. Thereby, the underlying medical images are in the standardised DICOM format.

In figure 4(a) we illustrate the model analysed in this paragraph. The model described consists of a cubic box dimensions $(50.0 \times 50.0 \times 100.0)\,cm^3$ filled with water in which two spherical regions, shown in red and in orange respectively, of diameter $d = 3\,cm$ placed, one considered as the targeted tumor and the other one as an organ at risk. We irradiate the box using four beams at $0°$, $120°$, $180°$ and $240°$ and a total amount of $N_B = 64$ beamlets (16 beamlets/beam). The influence matrices (A_{red} and A_{orange}) are obtained using the CERR's dose calculation algorithm QIB (Hardiyanti *et al* 2016) with the default settings; the dimensions of the beamlets are set to $(1.0 \times 1.0)\,cm^2$.

The optimization goals are set to $D_{red}^{(P)} = 50.0$ Gy for the red sphere and $D_{orange}^{(P)} = 0.0$ Gy for the orange one, considering the first as the targeted tumor and the second as an OAR. Each sphere is weighted equally in the cost function with $\gamma = 1.0$. During the optimization procedure, the influence matrixes are always normalized to keep the final beamlets intensities in the interval $[0,\,1]$.

In the mapping to the discrete problem, the bit-depth was fixed to $N_Q = 4$ bits, resulting in a fully connected lattice of 256 sites for the underlying Hamiltonian. The total number of non-zero interaction terms was $n_{int} = 32\,640$. Thus, the underlying quantum many-body system is a challenging long-range spin-glass Ising model to be solved with the TTN algorithm.

In practice, it is unfeasible to obtain $E_0 = 0$ and thereby to reach exactly the prescribed dose distribution. Thus, in this example, the optimization balances the different goals for each organ according to their priorities γ. For this reason, to evaluate the quality of the results returned by the TTN algorithm, the same optimization task
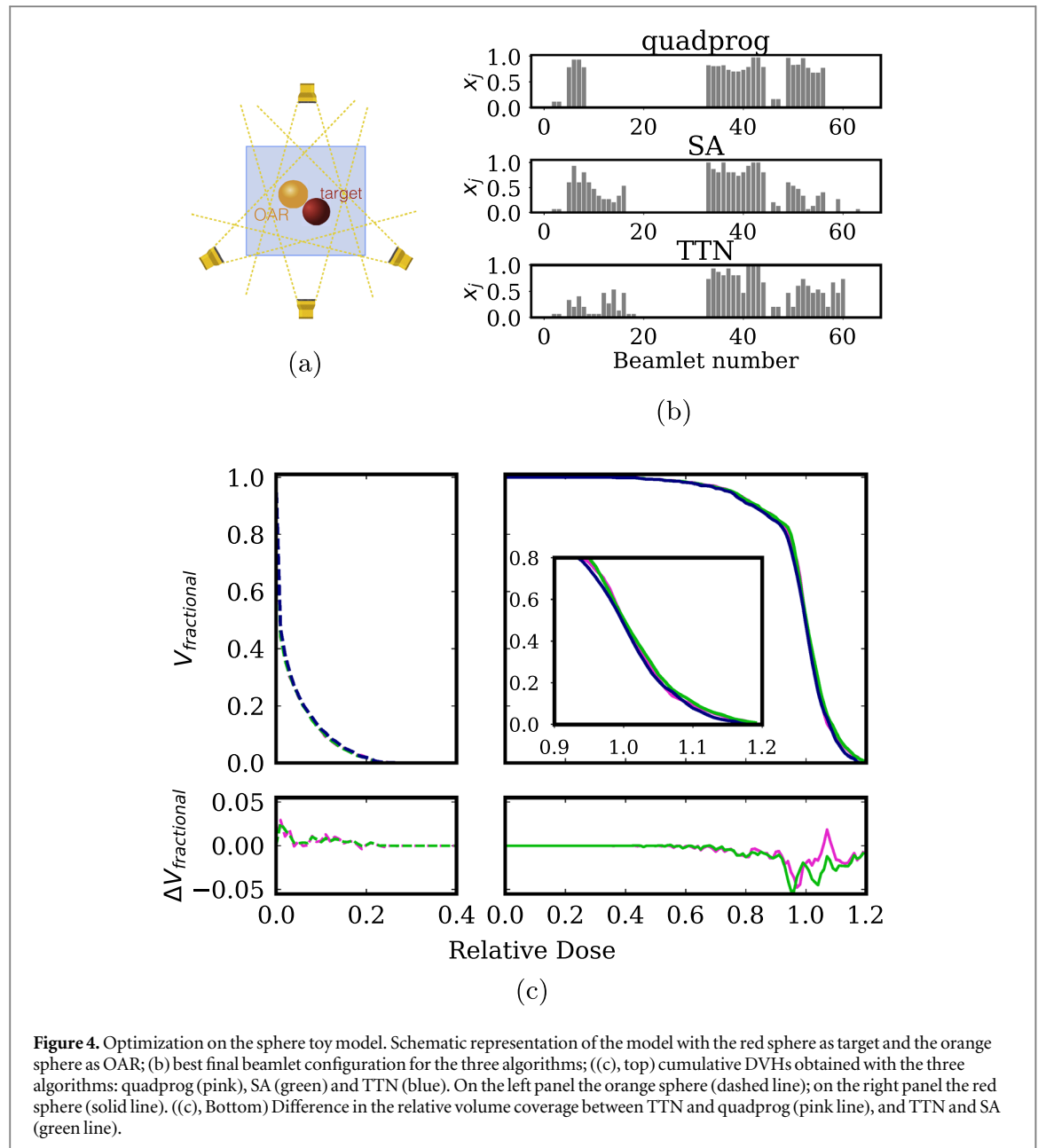
Figure 4. Optimization on the sphere toy model. Schematic representation of the model with the red sphere as target and the orange sphere as OAR; (b) best final beamlet configuration for the three algorithms; ((c), top) cumulative DVHs obtained with the three algorithms: quadprog (pink), SA (green) and TTN (blue). On the left panel the orange sphere (dashed line); on the right panel the red sphere (solid line). ((c), Bottom) Difference in the relative volume coverage between TTN and quadprog (pink line), and TTN and SA (green line).

was attacked using the Matlab build-in function *quadprog*, which exploits QP to optimize the cost function and SA. We recall that the optimization with *quadprog* was directly performed on the function in equation (2), while SA and the TTN were applied to the discretized problem in equation (4).

A standard method used for plans quality evaluation is the cumulative DVH histogram, which shows a 2D projection of the 3D dose distribution inside a given volume. It represents the fractional volume receiving at least a given value of dose. Given a generic volume, r, we can easily build the dose vector $\mathbf{D}_r(\mathbf{x})$ as described in equation (1) by applying its influence matrix, A_r, to the beamlets vector $\mathbf{x}$. The resulting vector contains the total dose delivered to each one of the voxels in the volume r. By subdividing the dose interval $[0, \ D_i^{(max)}]$ into n_b (dose) bins, for each of them, we can count how many voxels receive a dose greater or equal than the corresponding dose value. In other words, the number of entries in the kth bin indicates the number of voxels receiving at least the corresponding dose. The obtained distribution results in a cumulative DVH with the fractional volume represented on the y-axis and the dose values on the x-axis. We point out, that in this representation of the dose distribution we lose the spatial information of the problem.

On the top panels of figure 4(c), we show the DVHs obtained with the three methods. It is clearly visible, that the three methods show a very good qualitative agreement in the resulting DVHs for each organ. This agreement is further quantitatively confirmed by the energy E_0—or cost—after the minimization: within the statistical uncertainty, all three methods result in a final energy $E_0 = 0.0181(8)$. By looking at the bottom panels of

figure 4(c) we see that the difference between TTN and the other two methods in the relative volume coverage is globally very close to zero, with only a few peaks at about 2%–5%.

In figure 4(b), we present the final beamlets configuration for each of the three methods. Despite their global consistency, we observe that local differences arise. First, this is due to the fact the final configurations have slightly different energies, despite are all consistent. However, in general, there may exist more than one configuration satisfying the constraints and minimizing the energy E_0 for the underlying system and this effect is further amplified when comparing the optimization on the discrete space (SA and TTN) to that on the continuous one (quadrog). This happens because the energy landscape may be altered by the discretization procedure. Thus, in this case, we have that the ground state of the underlying Hamiltonian is either degenerate or its energy gap is reasonably small. In fact, the underlying quantum spin-glass Hamiltonian has a highly non-trivial spectrum with many local optima and depending on the system parameters degenerate ground-states.

3.2. Prostate cancer treatment with TG119 IMRT phantom

We now show the results obtained on a standard IMRT phantom provided by the American Association of Physicists in Medicine Task Group 119 for use in institutional IMRT commissioning (Craft *et al* 2014, Villaggi *et al* 2019). This dataset contains several segmented structures and we chose the following with the aim of simulating a prostate cancer case: prostate as the targeted tumor, bladder and rectum as the OARs.

The geometry we used is characterized by two beams (33 and 31 beamlets) placed at 90°, 270° as shown in figure 5(a), resulting in a total amount of $N_B = 64$ beamlets. The dimensions of each beamlet were $(0.9 \times 0.9)\,\text{cm}^2$. The dose prescriptions were set to $D_{prostate}^{(P)} = 50$ Gy and $D_{bladder}^{(P)} = D_{rectum}^{(P)} = 0.0$ Gy and the priority assigned to the different structures was $\gamma = 1.0$ for all of them. The dose calculation algorithm used was CERR's QIB algorithm in the default settings. The fraction of non-zero elements in the influence matrixes was 0.55 for the $A_{prostate}$, 0.28 for $A_{bladder}$ and 0.45 for A_{rectum}. For the discrete problem, the system obtained is again characterized by 256 fully connected long-range interacting spins and a total amount of $n_{int} = 32\,640$ interaction terms.

The top panels of figure 5(c) show the comparison between the DVHs for the three optimization methods. The bottom panels show that the difference between TTN and the other two methods in the volume coverage for the rectum and the prostate is within the 4%, proving a very good quantitative agreement between them. The results for the bladder are not shown since the differences were negligible. This agreement is additionally confirmed by the obtained energy E_0—or cost—after the minimization: within the statistical uncertainty, the three methods result in a final energy $E_0 = 0.043(1)$. This result further confirms what found in the study on the toy models.

4. Conclusions

In this manuscript, we presented a new approach based on TNs to optimize the dose distribution for an IMRT cancer treatment. We showed a feasibility study on three different cancer treatment scenarios. First, we provided a proof-of-principle TN analysis by successfully investigating an analytically solvable toy-model of a radiated box. Then, we compared the TNs with the classical approaches of QP and SA in the case of a spherical cancer and a spherical OAR. Finally, we illustrated the successful application of the TNs approach to a more realistic anatomical scenario simulating a prostate cancer.

The main goal of our work was to show the applicability of TNs to the IMRT optimization problem, fostering new applications of quantum-inspired techniques to the solution of classical optimization problems. Along the road, we defined a clear strategy to map the classical problem to simulated quantum-like hardware.

Our results indicate that the TN approach can achieve results compatible with other optimization techniques such as QP and SA. We stress that for this feasibility study we (i) used a reduced number of beamlets to reduce the complexity in the models, (ii) kept the cost function convex for sake of simplicity and (iii) used the TNs code 'out-of-the-box' originally engineered for typical quantum systems with significantly fewer interactions but higher entanglement. Further software developments will allow addressing these three points, increasing the TN approach efficiency. In particular, extending this study to non-convex and non-differentiable functions and by further specializing and parallelizing the TNs code for this particular field of application, further significant steps forward can be made towards real-world scenarios and their use in every-day medical care.

Finally, we point out that TNs are particular examples of quantum circuits, thus, this study opens the way to the application of quantum computation to cancer treatment, for example through the application of hybrid quantum–classical optimization algorithms (McClean *et al* 2016). Once quantum computer hardware will be scaled up, one could replace the TNs simulation with actual quantum computation, possibly further enhancing our capabilities of fighting cancer via IMRT.
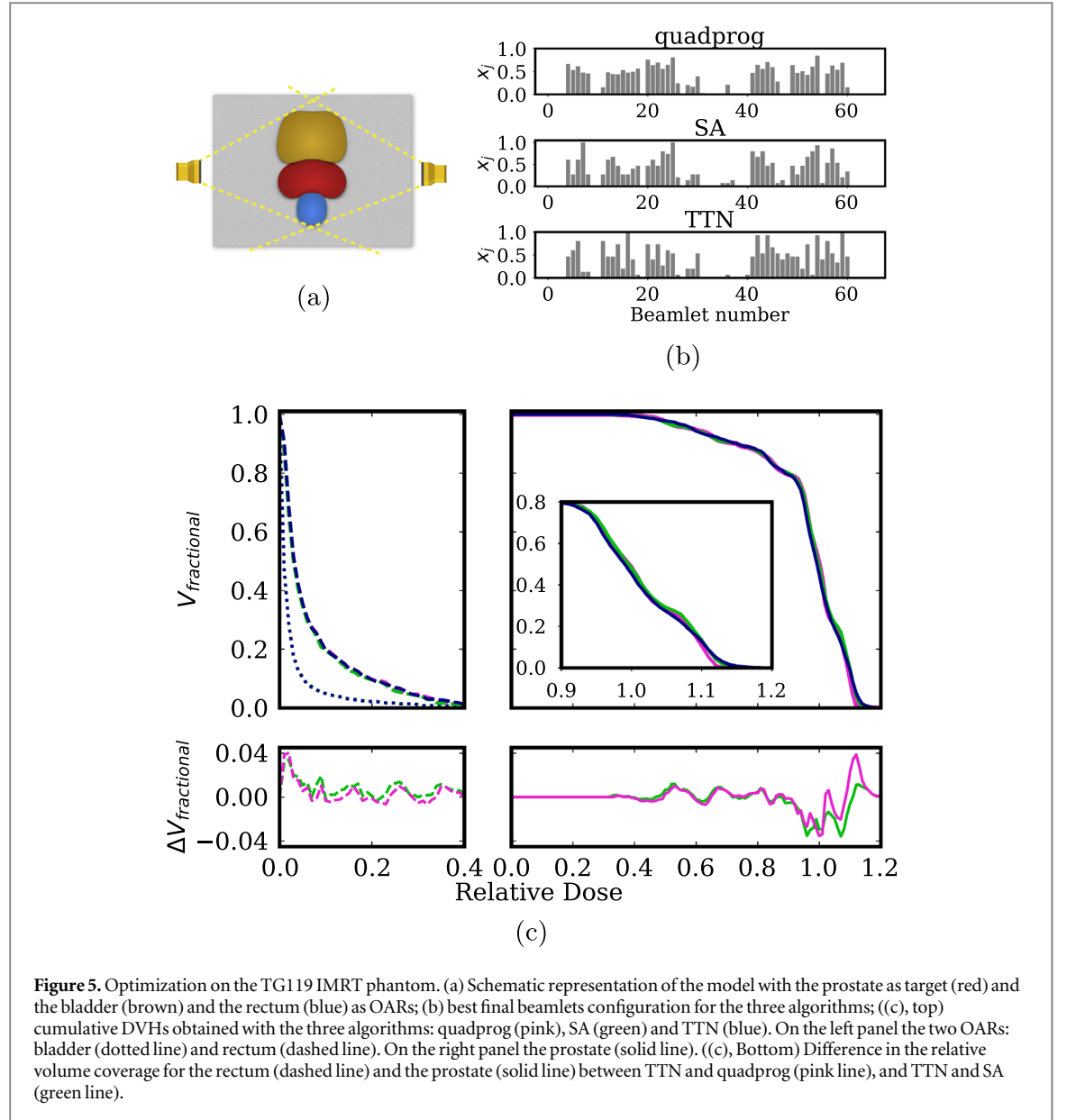
Figure 5. Optimization on the TG119 IMRT phantom. (a) Schematic representation of the model with the prostate as target (red) and the bladder (brown) and the rectum (blue) as OARs; (b) best final beamlets configuration for the three algorithms; ((c), top) cumulative DVHs obtained with the three algorithms: quadprog (pink), SA (green) and TTN (blue). On the left panel the two OARs: bladder (dotted line) and rectum (dashed line). On the right panel the prostate (solid line). ((c), Bottom) Difference in the relative volume coverage for the rectum (dashed line) and the prostate (solid line) between TTN and quadprog (pink line), and TTN and SA (green line).

Acknowledgments

Appendix A. Classical optimization

A.1. Quadratic programming

QP refers to a set of widespread methods for solving (nonlinear) quadratic optimization problems subject to linear constraints. A general QP problem can be formulated as follows:

$$\min_x \mathbf{q}^T \mathbf{x} + \frac{1}{2} \mathbf{x}^T Q \mathbf{x} \quad s.t. \quad \begin{cases} A\mathbf{x} = \mathbf{a}, \\ B\mathbf{x} \leqslant \mathbf{b} \\ \mathbf{x} \geqslant 0 \end{cases} \tag{A1}$$

where $f(\mathbf{x})$ is the target function, $\mathbf{q}$ its gradient and Q its Hessian matrix. The equation $A\mathbf{x} = \mathbf{a}$ contains all the equality constraints, while $B\mathbf{x} \leqslant \mathbf{b}$ all the inequality constraints.

In the QP formulation, the initial IMRT optimization problem in equation (2) can be written as follows:

$$\min_{\mathbf{x}} 2\sum_r (\mathbf{x}^T \tilde{A}_r^T \tilde{A}_r \mathbf{x} - \mathbf{D}_r^{(P)} \tilde{A}_r \mathbf{x}), \ 0 \leqslant \mathbf{x} \leqslant 1, \tag{A2}$$

where r is an index running over the different volumes or organs (the targeted tumor and the OARs), $\tilde{A}_r$ is the influence matrix for the rth volume whose entries are weighted by the pre-assigned priorities $\gamma_{i(r)}$ to each voxel, i, and $\mathbf{D}_r^{(P)}$ a vector containing the dose prescriptions for the voxels in the rth volume.

Matlab's *Optimization Toolbox*$^{\text{TM}}$ provides the function *quadprog* which exploits different solvers to attack a wide class QP problems. In this work, this methods is used to solve the initial quadratic problem and produce results to be compared to those obtained with TTN.

A.2. Simulated annealing

SA is a widespread combinatorial optimization method based on randomization techniques (Aarts and van Laarhoven 1987), in particular, based on the Metropolis–Hastings algorithm. By varying a control parameter T, called *temperature*, it is possible to explore the landscape of a target free-energy function in order to find its global minima. SA is typically applied to large-scale non-convex optimization problems where the number of local minima in the energy landscape is very high. The algorithm always requires a starting point which in practice is either a random one or the best one known for a specific problem. From the initial starting point, a rule to generate new configurations is given: configurations with an energy lower than the previous configuration are always accepted (with probability one). On the other hand, a move towards configurations with a higher energy is accepted with a certain probability which significantly reduces the risk of getting stuck into a local minimum. In this case, the acceptance probability decreases dynamically with the temperature T, which itself decreases during the optimization from a given value $T_{\max}$ to $T_{\min}$ according to a pre-defined annealing schedule.

In this work, SA is applied to find the ground state of the classical Ising spin-glass problem representing the initial IMRT optimization problem. We always start the SA from a random configuration for the spins in the lattice. New configurations are generated by flipping a randomly chosen spin in the prior configuration. Due to the intrinsic stochastic nature of SA, we perform a statistical sampling of N_{runs} independent and randomly-initialized optimization runs for the same problem. For both, the sphere and the prostate cancer case, N_{runs} is set to 100.

The code used in this study to perform SA is based on the Python library in Wagner (2019).

Appendix B. Ground-state search via TTN

B.1. Setting an optimization problem on TTN

In the following, we try to provide a list of the main parameters which need to be tuned in order to set an IMRT optimization problem on the TTN algorithm as well as a brief explanation of their effect on the performance of the algorithm.

1. *Bit-depth N_Q*: it is the number of qubits used to represent each beamlet in the discrete form. This number multiplied by the total number of beamlets N_B determines the total number of qubits N in the lattice. N influences two different parameters: (1) the number of tensors, T, in the network and (2) the number of interactions terms, n_I, in the system. The former shows a linear dependency on $T = N - 2$. The second one, in the case of a fully connected interaction scheme as the one presented here, shows a quadratic dependency with $n_I = \frac{N(N-1)}{2}$. The total time Γ scales linearly in both, $\Gamma \sim \mathcal{O}(T)$ and $\Gamma \sim \mathcal{O}(n_I)$. The latter dependency is explicitly shown in figure B3 of appendix B.3 when the reduction of the number of interaction terms in the system is discussed.

2. *Maximum bond dimension χ*: it corresponds to the dimension of the auxiliary indexes of the TTN and represents an upper limit for the amount of entanglement encoded in the TN representation of the ground-state. Changing the bond dimension allows interpolating between the exact representation of a state and the mean-field one. Increasing the bond dimension allows to explore a wider portion of the underlying Hilbert space but affects the computational time which scales asymptotically as $\mathcal{O}(\chi^4)$. In the present work, the χ was set to 5.

3. *Threshold for convergence ϵ*: the optimization of the parameters in the TTN is done by iterating many times over all tensors in the network. After every complete iteration (also called a *sweep*) the energy is computed. ϵ introduces a stop criterion that terminates the optimization if the absolute (or relative) decrease in energy over a certain number of sweeps does not exceed ϵ. This affects the total number of sweeps I_S performed before reach the stop criterion is triggered. In this work, a good choice of the absolute convergence threshold
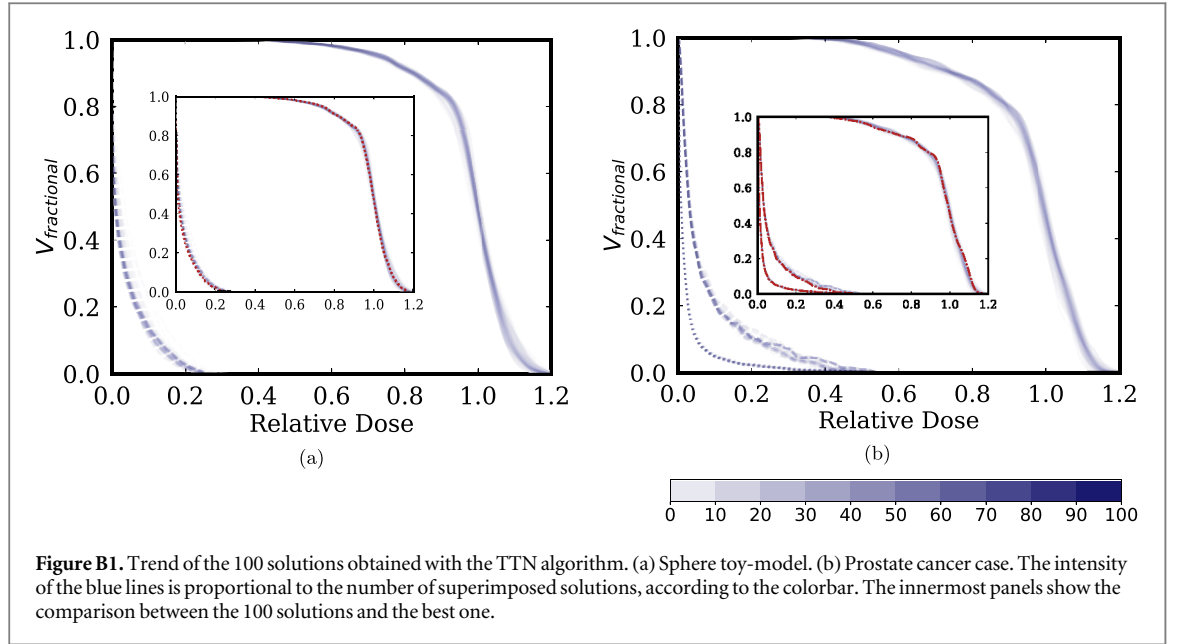
Figure B1. Trend of the 100 solutions obtained with the TTN algorithm. (a) Sphere toy-model. (b) Prostate cancer case. The intensity of the blue lines is proportional to the number of superimposed solutions, according to the colorbar. The innermost panels show the comparison between the 100 solutions and the best one.

was found to be $\epsilon_{abs} = 10^{-5}$. Additionally, we set a maximum number of iterations $I_{\max}$ after which the optimization is stopped even if the convergence criterion is not met. The computational time scales linearly in $I = \min(I_{\max}, I_S)$ with $I \cdot \mathcal{O}(Tn_I\chi^4)$.

We mention, that the numerical complexity can be reduced by parallelising the TTN computation which we did not exploit at the current stage.

B.2. Trend of the solutions

The TTN algorithm explores a corner of the full many-body Hilbert space $\mathcal{H}$ which becomes wider as the bond dimension χ increases. In this work, the bond dimension was fixed to $\chi = 5$ throughout the whole study. Since we are dealing a classical problem we know that the final solution is not entangled. However, the introduction of a bond dimension $X > 1$ increases the probability for the algorithm to converge to the global minimum, as it thereby explores a greater solution space.

In section 3 we have shown the best results obtained with the TTN both for the prostate and the sphere over samples of $N_{runs} = 100$ runs. We recall that the entries of the tensors are randomly initialized at the beginning of each new optimization run.

In order to see which is the general behaviour of the algorithm throughout the whole sampling, we compare in figures B1(a) and (b) the results over all $N_{runs} = 100$ runs. We observe that they are globally very closely distributed around the best solution and this proves the precision of the TTN algorithm. However, making precise clinical considerations about the consistency between the different solutions goes far beyond the purpose of this feasibility study since many different factors should be considered depending on the specific case.

B.3. Reduction of interaction coefficients

The model we have considered so far is a fully connected lattice system of N interacting spins, with the number of interactions terms given by $N \times (N - 1)/2$. The density of the interaction scheme has an impact on the computational time required to the TTN to converge. Thereby, it is interesting to investigate the behaviour of the accuracy of the solution when neglecting some of the coefficients. In particular, we fix a threshold, $\lambda_{\min}$ such that $(|J_{ij}| < \lambda_{\min}) = 0$, with J_{ij} the pairwise interaction term between the ith and the jth spin in the lattice, to cut away the interaction terms with the lowest coefficients.

In the following, we considered four different cuts, keeping respectively the 10%, 40%, 70% and 100% of the coefficients. The result for the sphere cancer is shown in figure B2(a). For each of the four cases, samples of $N_{runs} = 100$ runs were collected. Each histogram shows the average position of each bin (we set $N_{bin} = 100$ bins) with the associated standard deviation both on the relative dose and the factional volume axes. For the red sphere, (the most-right lines) 70%- and 100%-results are very well superimposed, while the 40%-line lays within the standard deviation of the former ones. Bigger differences arise between these three and the 10% line. No significant differences can be found for the orange sphere (most-left lines) between the 100%, 70% and 40% lines. Even in this case, the main differences arise between the 10% line and the others.
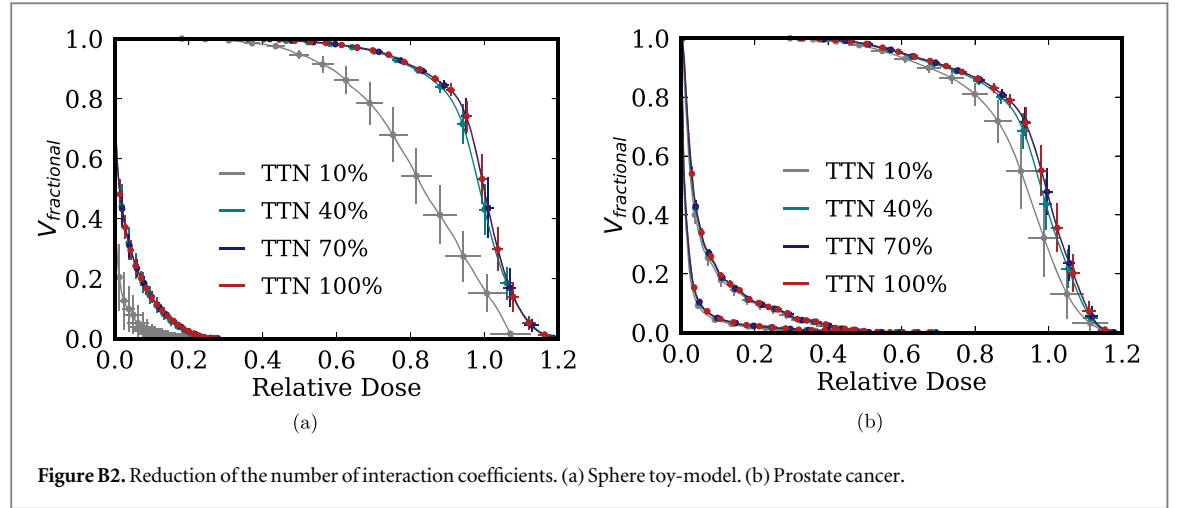
Figure B2. Reduction of the number of interaction coefficients. (a) Sphere toy-model. (b) Prostate cancer.
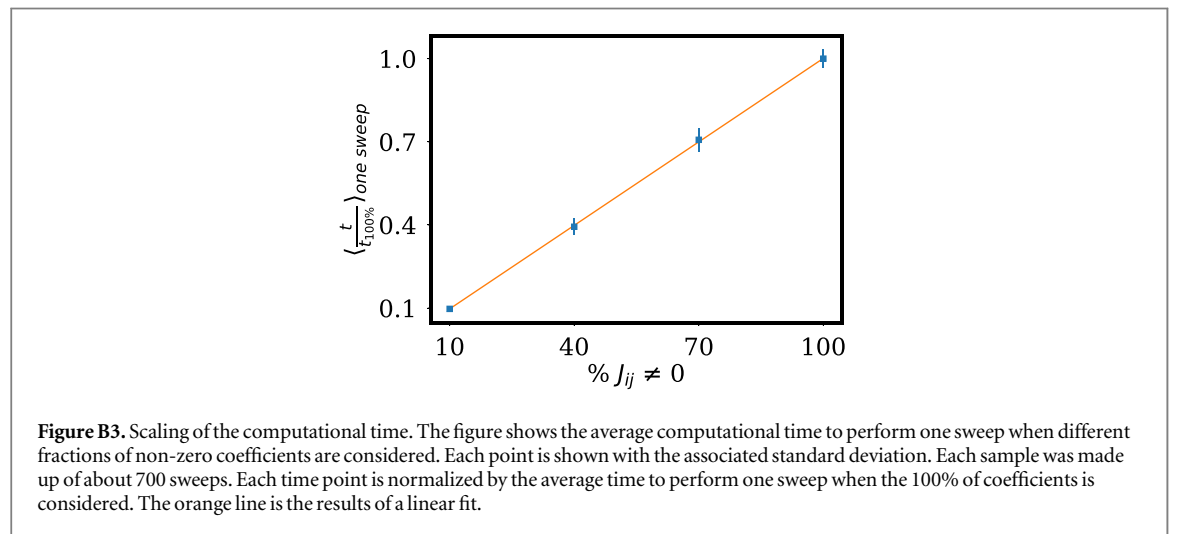

Figure B3. Scaling of the computational time. The figure shows the average computational time to perform one sweep when different fractions of non-zero coefficients are considered. Each point is shown with the associated standard deviation. Each sample was made up of about 700 sweeps. Each time point is normalized by the average time to perform one sweep when the 100% of coefficients is considered. The orange line is the results of a linear fit.

Interestingly, these results confirm that accurate solutions can be obtained even considering only 40% of the coefficients. The study on the prostate cancer case (see figure B2(b)) further confirms this insight. However, in this case, the 10% line is closer to the others than in the previous case.

This result becomes even more interesting if we consider the scaling of the computational time when reducing the number of interaction coefficients. Figure B3 shows the result for the prostate cancer case. We observe that the scaling is linear. In conclusion, using this procedure we should be able to obtain very accurate solutions spending only the 40% of the initial computational time.

Consequently, one question arises naturally: How to find the optimal balance between speeding up the simulations by cutting terms versus keeping an adequate amount of information to accurately describe the system? We therefore propose the following heuristic procedure:

- Build the fully connected interaction matrix of the initial problem, J_{ij}^0 and compute its eigenvalues s_0.

- Choose a cut in order to keep the $\eta\%$ of the coefficients and compute the new eigenvalues s_{red} of the reduced interaction matrix, J_{ij}^{red}.

- Check the statistical difference between the two samples of eigenvalues s_0 and s_{red} (before and after the cut). This can be performed using a hypothesis testing procedure, where the null-hypothesis H_0 is that the two sets of eigenvalues are sampled from the same population, hence no statistical differences arise between them. Within this framework, the *p*-value of the test can be used to predict the accuracy of the optimization. In particular, the smaller the *p*-value is, the more the two sets of eigenvalues are likely to come from different populations and thus the accuracy of the optimization on the reduced model to be poor. We point out, that the choice of the statistical test in this procedure is not unique. Our trials exploited the Wilcoxon non-parametric test for dependent samples.

ORCID iDs

Samuele Cavinato ⬤ https://orcid.org/0000-0002-4706-1551
Timo Felser ⬤ https://orcid.org/0000-0002-4059-8065
Marco Fusella ⬤ https://orcid.org/0000-0002-5361-8387
Marta Paiusco ⬤ https://orcid.org/0000-0001-8647-911X
Simone Montangero ⬤ https://orcid.org/0000-0002-8882-2169

References

Intensity Modulated Radiation Therapy Collaborative Working Group 2001 Intensity-modulated radiotherapy: current status and issues of interest *Int. J. Radiat. Oncol. Biol. Phys.* **51** 880–914

Aarts E H L and van Laarhoven P J M 1987 *Pattern Recognition Theory and Applications* ed P A Devijver and J Kittler (Berlin: Springer) pp 179–92

Adeberg S, Harrabi S B, Bougatf N, Verma V, Windisch P, Bernhardt D, Combs S E, Herfarth K, Debus J and Rieken S 2018 Dosimetric comparison of proton radiation therapy, volumetric modulated arc therapy, and three-dimensional conformal radiotherapy based on intracranial tumor location *Cancers* **10** 401

Alongi F, Arcangeli S, Filippi A R, Ricardi U and Scorsetti M 2012 Review and uses of stereotactic body radiation therapy for oligometastases *Oncologist* 11 **17** 1100–7

Apte A, Khullar D, Alaly J and Deasy J O 2010 CERR https://github.com/cerr/CERR

Bañuls M C and Cichy K 2020 Review on novel methods for lattice gauge theories *Rep. Prog. Phys.* **83** 024401

Bañuls M C *et al* 2019 Simulating lattice gauge theories within quantum technologies *Eur. Phys. J.* D **74** 165

Baskar R, Dai J, Wenlong N, Yeo R and Yeoh K-W 2014 Biological response of cancer cells to radiation treatment *Frontiers Mol. Biosci.* **1** 24

Baskar R, Lee K A, Yeo R and Yeoh K-W 2012 Cancer and radiation therapy: current advances and future directions *Int. J. Med. Sci.* **9** 193–9

Bortfeld T 2006 IMRT: a review and preview *Phys. Med. Biol.* **51** R363–79

Bortfeld T R and Boyer A L 1995 The exponential radon transform and projection filtering in radiotherapy planning *Int. J. Imaging Syst. Technol.* **6** 62–70

Brunner T B 2016 The rationale of combined radiotherapy and chemotherapy – Joint action of castor and pollux *Best Pract. Res. Clin. Gastroenterol.* **30** 515–28

Censor Y and Unkelbach J 2012 From analytic inversion to contemporary IMRT optimization: radiation therapy planning revisited from a mathematical perspective *Phys. Med.* **28** 109–18

Chargari C *et al* 2019 Brachytherapy: an overview for clinicians *CA: Cancer J. Clin.* **69** 386–401

Cotrutz C, Lahanas M, Kappas C and Baltas D 2001 A multiobjective gradient-based dose optimization algorithm for external beam conformal radiotherapy *Phys. Med. Biol.* **46** 2161–75

Craft D, Bangert M, Long T, Papp D and Unkelbach J 2014 Shared data for intensity modulated radiation therapy (IMRT) optimization research: the CORT dataset *Giga Sci.* **3** 2047–217X-3-37

Dalmonte M and Montangero S 2016 Lattice gauge theory simulations in the quantum information era *Contemp. Phys.* **57** 388–412

Delaney A R, Dahele M, Tol J P, Kuijper I T, Slotman B J and Verbakel W F A R 2017 Using a knowledge-based planning solution to select patients for proton therapy *Radiother. Oncol.* **124** 263–70

Elith C, Dempsey S E, Findlay N and Warren-Forward H M 2011 An introduction to the intensity-modulated radiation therapy (IMRT) techniques, tomotherapy, and VMAT *J. Med. Imaging Radiat. Sci.* **42** 37–43

Evenbly G and Vidal G 2009 Algorithms for entanglement renormalization *Phys. Rev. B* **79** 144108

Ezzell G A 1996 Genetic and geometric optimization of three-dimensional radiation therapy treatment planning *Med. Phys.* **23** 293–305

Felser T, Notarnicola S and Montangero S 2021 Efficient tensor network ansatz for high-dimensional quantum many-body problems *Phys. Rev. Lett.* **126** 170603

Felser T, Silvi P, Collura M and Montangero S 2020 Two-dimensional quantum-link lattice quantum electrodynamics at finite density *Phys. Rev. X* **10** 041040

Formenti S C and Demaria S 2013 Combining radiotherapy and cancer immunotherapy: a paradigm shift *J. Natl Cancer Inst.* **105** 256–65

Gerster M, Rizzi M, Silvi P, Dalmonte M and Montangero S 2017 Fractional quantum Hall effect in the interacting Hofstadter model via tensor networks *Phys. Rev. B* **96** 195123

Gerster M, Silvi P, Rizzi M, Fazio R, Calarco T and Montangero S 2014 Unconstrained tree tensor network: An adaptive gauge picture for enhanced performance *Phys. Rev. B* **90** 125154

Hardiyanti Y, Haekal M, Waris A and Haryanto F 2016 The comparison study of quadratic infinite beam program on optimization instensity modulated radiation therapy treatment planning (IMRTP) between threshold and exponential scatter method with CERR® in the case of lung cancer *J. Phys.: Conf. Ser.* **739** 012144

Huggins W, Patil P, Mitchell B, Whaley K B and Stoudenmire E M 2019 Towards quantum machine learning with tensor networks *Quantum Sci. Technol.* **4** 024001

Jaffray D A 2012 Image-guided radiotherapy: from current concept to future perspectives *Nat. Rev. Clin. Oncol.* **9** 688–99

Kim I H and Swingle B 2017 Robust entanglement renormalization on a noisy quantum computer arXiv:1711.07500

Ma Y *et al* 2010 Chemotherapy and radiotherapy: cryptic anticancer vaccines *Semin. Immunol.* **22** 113–24

McClean J R, Romero J, Babbush R and Aspuru-Guzik A 2016 The theory of variational hybrid quantum-classical algorithms *New J. Phys.* **18** 023023

McCulloch I P 2007 From density-matrix renormalization group to matrix product states *J. Stat. Mech: Theory Exp.* **2007** P10014

Montangero S 2018 *Introduction to Tensor Network Methods* (Berlin: Springer)

Nazareth D P and Spaans J D 2015 First application of quantum annealing to IMRT beamlet intensity optimization *Phys. Med. Biol.* **60** 4137–48

Nielsen M and Chuang I (ed) 2000 *Quantum Computation and Quantum Information* (Cambridge: Cambridge University Press)

Orús R 2014 A practical introduction to tensor networks: matrix product states and projected entangled pair states *Ann. Phys.* **349** 117–58

Östlund S and Rommer S 1995 Thermodynamic limit of density matrix renormalization *Phys. Rev. Lett.* **75** 3537–40

Pakela J M, Tseng H-H, Matuszak M M, Ten Haken R K, McShan D L and Naqa I El 2020 Quantum-inspired algorithm for radiotherapy planning optimization *Med. Phys.* **47** 5–18

Schollwöck U 2011 The density-matrix renormalization group in the age of matrix product states *Ann. Phys.* **326** 96–192

Shi Y-Y, Duan L-M and Vidal G 2006 Classical simulation of quantum many-body systems with a tree tensor network *Phys. Rev.* A **74** 022320

Silvi P, Giovannetti V, Montangero S, Rizzi M, Cirac J I and Fazio R 2010 Homogeneous binary trees as ground states of quantum critical hamiltonians *Phys. Rev.* A **81** 062335

Silvi P, Tschirsich F, Gerster M, Jünemann J, Jaschke D, Rizzi M and Montangero S 2019 The tensor networks anthology: simulation techniques for many-body quantum lattice systems *SciPost Phys. Lect. Notes* **8**

Singh S and Vidal G 2013 Global symmetries in tensor network states: symmetric tensors versus minimal bond dimension *Phys. Rev.* B **88** 115147

Sun X, Xia P and Yu N 2004 Effects of the intensity levels and beam map resolutions on static IMRT plans *Med. Phys.* **31** 2402–11

Taylor A and Powell M E B 2004 Intensity-modulated radiotherapy—what is it? *Cancer Imaging* **4** 68–73

Teoh M, Clark C H, Wood K, Whitaker S and Nisbet A 2011 *Br. J. Radiol.* **84** 967–96

Verstraete F and Cirac J I 2004 Valence-bond states for quantum computation *Phys. Rev.* A **70** 060302(R)

Verstraete F and Cirac J I 2006 Matrix product states represent ground states faithfully *Phys. Rev.* B **73** 094423

Verstraete F, Wolf M M, Perez-Garcia D and Cirac J I 2006 Criticality, the area law, and the computational power of projected entangled pair states *Phys. Rev. Lett.* **96** 220601

Villaggi E *et al* 2019 Plan quality improvement by DVH sharing and planner's experience: results of a SBRT multicentric planning study on prostate *Phys. Med.* **62** 73-82

Wagner R J and Perry M T 2019 Python module for simulated annealing https://github.com/perrygeo/simanneal

Wang J-S, Wang H-J and Qian H-L 2018a Biological effects of radiation on cancer cells *Mil. Med. Res.* **5** 20

Wang Y, Deng W, Li N, Neri S, Sharma A, Jiang W and Lin S H 2018b *Frontiers Pharmacol.* **9** 185

Webb S 1995 Optimizing radiation therapy inverse treatment planning using the simulated annealing technique *Int. J. Imaging Syst. Technol.* **6** 71–9

Zahra J U R, Ahmad N, Khalid M, ul Huda Khan Asghar H N, Gilani Z A, Ullah I, Nasar G, Akhtar M M and Usmani M N 2018 Intensity modulated radiation therapy: a review of current practice and future outlooks *J. Radiat. Res. Appl. Sci.* **11** 361–7

Zhang Q, Liu J, Ao N, Yu H, Peng Y, Ou L and Zhang S 2020 Secondary cancer risk after radiation therapy for breast cancer with different radiotherapy techniques *Sci. Rep.* **10** 1220

Zhou Y, Stoudenmire E M and Waintal X 2020 What limits the simulation of quantum computers? *Phys. Rev.* X **10** 041038