

Doses Relations

Contents

1 Distance Between Doses	2
A non-trivial task	2
Dose Evaluation	2
1.1 Method	2
Naive Doses Comparison	2
Pathological example	2
1.1.1 Doses Samples	4
1.1.2 Distances Between Doses	4
Comparing Doses Voxel-wise	4
Comparing Dose Volume Histogram Curves	4
Discrete DVH Approximation	4
Distance Measure	4
Fréchet Distance	4
Hausdorff Distance	5
Wasserstein Distance	6
Kolmogorov-Smirnov Distance	6
Total Variation Distance	6
1.2 Results	8
1.2.1 Dose Distances Comparison	8
1.2.2 Link between Total Variation and Wasserstein	8
1.2.3 Bounding of Total Variation and Voxel Distance	9
Sorting lists	9
Proof Outline	9
Conclusion	9
1.2.4 Distances Distribution Comparison	10
Comparing Total Variation and Voxel-wise	10
Comparing Total Variation and Wasserstein	10
1.3 Discussion	10
Stop Criterion	11
2 Network of Doses	11
2.1 Methods	11
Multiple Plans Generation	11
Objectives' Weights Generation	11
Dose Normalization	12
Phantom Patient	12
Dose Clustering Techniques	12
Dose Distance	12
Community Detection	12
Evaluating Communities Split	12
2.2 Results	12
Visual Clustering Evaluation	12
Graph Plots	13
DVH Plot	13

Numerical Clustering Evaluation	13
2.3 Discussion	15
3 Novel Dosimetry Automation Approach with Graph Theory	15
3.1 Challenges in Current Practices	15
3.2 Proposed Framework	16
3.3 Discussion	16

Abstract

Fluence Map Optimization (FMO) requires selecting importance factors for each clinical constraint, which reflect the relative priority of achieving specific dose goals. These importance factors significantly influence the resulting dose distribution, as varying their values can lead to diverse treatment outcomes. This chapter investigates the relationship between the chosen importance factors and the resulting dose distributions. We aim to understand how different configurations affect the trade-offs between conflicting clinical objectives. This analysis provides insight into optimizing importance factors to achieve the most clinically effective treatment plans.

1 Distance Between Doses

In this section, we aim to establish a robust metric for quantifying the distance between different dose distributions. Such a distance should provide a numerical comparison that reflects the clinical discrepancies between two dose distributions. By developing a distance measure that captures these nuances, we can better evaluate and compare treatment plans.

A non-trivial task Quantifying the difference between two dose distributions, particularly regarding clinical impact, is inherently challenging. This complexity arises because not all patient anatomy regions contribute equally to treatment outcomes. Dose variations in critical structures may significantly influence clinical effects, while similar variations in less critical areas may have negligible impact. Moreover, the potential for dose compensation (underdosing in one region counterbalancing overdosing in another) further complicates the development of a reliable metric for comparing dose distributions. This compensatory effect is only sometimes applicable, making establishing a standardized method for assessing dose distribution clinical differences challenging.

Dose Evaluation To assess the quality of a dose distribution, dosimetrists primarily focus on DVHs as the key information. While they also consider aspects of the three-dimensional (3D) dose distribution, such as inter-structure dose gradients and the presence, number, and location of hot spots, their primary attention is directed towards the analysis of DVHs, which provide a comprehensive overview of dose coverage and sparing of organs at risk.

1.1 Method

Naive Doses Comparison The most straightforward method for comparing two dose distributions, thus defining a distance metric, is to perform a voxel-by-voxel comparison of the dose values. However, this approach overlooks the inherent anatomical structure of the human body and the fact that not all voxels have the same clinical significance. Consequently, even if the voxel-wise distance between two dose distributions is considerable, their overall clinical effects may still be similar.

Pathological example We constructed a simplified example, as illustrated in Figure 1. This hypothetical scenario involves a phantom model consisting of a homogeneous water-equivalent material containing a cubic planning target volume (PTV) and a cubic organ-at-risk (OAR). Although this model lacks anatomical realism, it effectively highlights the limitations of using basic voxel-wise comparisons for dose evaluation. It emphasizes the need for more sophisticated techniques to capture clinically relevant differences in dose distributions accurately.

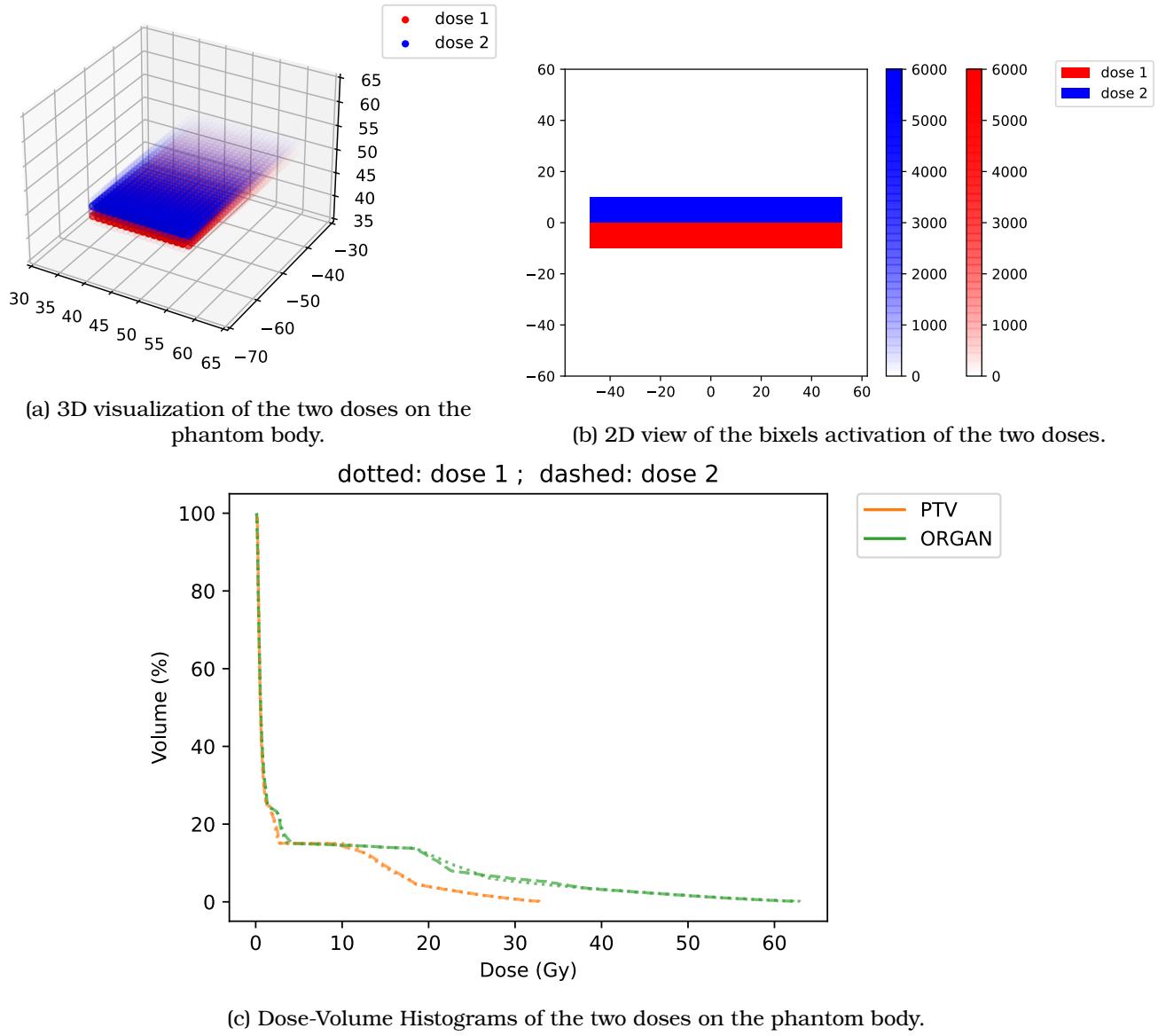


Figure 1: Example of two doses that have the same clinical effect (measured from the DVHs), but very different voxel-wise dose values.

1.1.1 Doses Samples

We assessed the efficacy of our proposed method for comparing radiation doses using the TG-119 Prostate case, a well-established benchmark for evaluating radiation therapy plans [MYH⁺12]. The TG-119 dataset includes predefined dose objectives, which we utilized to formulate our cost function. We performed optimizations with different weight assignments applied to each constraint to generate varying treatment dose distributions for the same patient case under identical constraints.

1.1.2 Distances Between Doses

Comparing Doses Voxel-wise When two radiation dose distributions are closely aligned, the voxel-wise comparison is an effective measure, as it can be assumed that the global distribution is similar. This approach allows for a detailed comparison of local dose variations. Mathematically, given the voxel-wise dose \mathbf{d} , the distance between two dose distributions, \mathbf{d}^1 and \mathbf{d}^2 , is defined as the norm of their difference: $\sum_{v \in \mathcal{V}} |\mathbf{d}_v^1 - \mathbf{d}_v^2|$, where v represents the voxels in the set of interest \mathcal{V} , and \mathbf{d}_v^i is the dose value at voxel v for dose distribution \mathbf{d}^i ¹. However, voxel-wise distance can become misleading if two regions of equal volume within the same anatomical structure have their dose values swapped. In such cases, the voxel-wise difference would appear large despite the clinical equivalence of the two doses. Furthermore, this method is limited to comparing doses within the same patient, as it requires a direct correspondence between the dose voxels in both distributions.

Comparing Dose Volume Histogram Curves We propose comparing the Dose Volume Histogram (DVH) curves. We have one curve for each structure; we define the distances between doses for each structure, and in the end, we sum up all structures to end up with a single scalar distance between two doses. We can quantify the variation between the two dose distributions in aggregated forms, using the structures.

Discrete DVH Approximation The DVH is obtained after sorting the voxel-wise dose of the structure: Let $\mathbf{d}[s]$ be the voxel-wise dose of the structure s (therefore, a list, of length $n[s]$, the number of voxels that belong to the structure). Let $\mathbf{d}[s]$ be the list above in descending order (i.e. $\mathbf{d}[s]_i > \mathbf{d}[s]_j$ if $0 < i < j \leq n[s]$). Then, the DVH of s can be approximated by the continuous line composed of the segments linking the following points: $(\mathbf{d}[s]_i, i/n[s]) \quad 0 < i \leq n[s]$. Since we compute the dose voxel-wise, we may only have an approximation of the DVH. However, in practice, most structures of interest have more than a hundred voxels, which makes the DVH approximation very precise.

Since we draw one curve per structure of interest, this capture some of the importance of voxel over others. In fact, when analyzing a dose, doctors look at the dose volume (voxel-wise), but they also take a close look at the DVHs; this is an incentive that DVHs should contain meaningful information.

Distance Measure To measure how different two DVH curves, we can imagine several techniques:

- Fréchet distance (treating DVHs as curves in a 2D space)
- Hausdorff distance (treating DVHs as 1D manifolds in a 2D space)
- Wasserstein distance (treating DVHs as probability distributions)
- Kolmogorov-Smirnov test (treating DVHs as probability distributions)
- Total variation between curves (treating DVHs as functions)

We evaluated all the aforementioned distance metrics and propose to retain only the one that yields the most clinically meaningful results.

Fréchet Distance DVH (Dose-Volume Histogram) curves can be interpreted as lines in \mathbb{R}^2 ². In this context, the Fréchet distance is a well-known metric for assessing the similarity between two curves, particularly useful for comparing poly-lines [EMO2]. It measures the minimum distance a particle would

¹This is often written as $|\mathbf{d}_1 - \mathbf{d}_2|$, with the summation over voxels implied.

²In the case of voxel-wise dose approximations, they are represented as poly-lines in \mathbb{R}^2 .

travel when simultaneously traversing both curves. In this study, we apply the Fréchet distance to compare the DVH curves of two radiation dose distributions.

Formally, let P and Q represent the curves being compared, with γ denoting a parametrization defined on the interval $[0, 1]$. The positions of a particle moving along curves P and Q at time t are given by $P(\gamma(t))$ and $Q(\gamma(t))$, respectively. The Fréchet distance is defined as:

$$d_{\text{Fréchet}}(P, Q) = \inf_{\gamma} \max_{t \in [0, 1]} d(P(\gamma(t)), Q(\gamma(t)))$$

When applied to DVH curves, let \mathcal{C}_A and \mathcal{C}_B denote the discrete DVH curves of two dose distributions. These curves consist of line segments connecting a series of points $\{\mathcal{C}_A(i) = (d_i, v_i), 1 \leq i \leq n_A\}$ and $\{\mathcal{C}_B(j) = (\tilde{d}_j, \tilde{v}_j), 1 \leq j \leq n_B\}$; where d_i and \tilde{d}_j denote the dose levels³, v_i and \tilde{v}_j represent the corresponding volumes, and n_A and n_B are the number of points forming \mathcal{C}_A and \mathcal{C}_B ⁴.

The Fréchet distance, in this case, is defined as the infimum over all possible traversal times. Given that the curves are discrete line segments, the Fréchet distance can be expressed as:

$$d_{\text{Fréchet}}(\mathcal{C}_A, \mathcal{C}_B) = \min_{\substack{j: \llbracket 1, n_A \rrbracket \rightarrow \llbracket 1, n_B \rrbracket \\ j \nearrow (j \text{ increasing})}} \sum_{i=1}^{n_A} \text{dist}(\mathcal{C}_A(i), \mathcal{C}_B(j(i)))$$

$$\text{where } \text{dist}(\mathcal{C}_A(i), \mathcal{C}_B(j(i))) = \sqrt{(d_i - \tilde{d}_{j(i)})^2 + (v_i - \tilde{v}_{j(i)})^2}$$

Here, j represents a (discrete) a parametrization, and $\text{dist}(\mathcal{C}_A(i), \mathcal{C}_B(j(i)))$ is the distance between points $\mathcal{C}_A(i)$ and $\mathcal{C}_B(j(i))$.

One drawback of the Fréchet distance is its computational expense, particularly for structures with a large number of voxels. To mitigate this, we applied the Ramer–Douglas–Peucker algorithm for curve simplification [PLQC12]. We employed this algorithm with $\epsilon = 0.05$, and after testing on a subset of DVH curves, it was found to accelerate computations by a factor of 3-5, while the calculated Fréchet distance deviated by less than 0.5 %. This method was therefore used in the results presented below.

Hausdorff Distance The Hausdorff distance is another commonly used metric for measuring the similarity between two curves [Hen99]. It is defined as the greatest of the shortest distances between any point on one curve and the closest point on the other. Formally, let X and Y be two non-empty sets; the Hausdorff distance between X and Y , denoted $d_{\text{Hausdorff}}(X, Y)$, is given by:

$$d_{\text{Hausdorff}}(X, Y) = \sup_{x \in X} \inf_{y \in Y} \text{dist}(x, y)$$

where $\text{dist}(x, y)$ represents the distance between points x and y (typically, the Euclidean distance).

In this study, we treat DVH (Dose-Volume Histogram) curves as sets of points in a two-dimensional space \mathbb{R}^2 , using the Hausdorff distance to quantify their difference. Using the same notation for the DVH curves \mathcal{C}_A and \mathcal{C}_B as previously defined, the discrete Hausdorff distance is computed as:

$$d_{\text{Hausdorff}}(\mathcal{C}_A, \mathcal{C}_B) = \max_{i \in \llbracket 1, n_A \rrbracket} \min_{y \in \mathcal{C}_B} \text{dist}(\mathcal{C}_A(i), y)$$

where \mathcal{C}_B is represented by the set of points

$$\left\{ \left((1 - \lambda)\tilde{d}_j + \lambda\tilde{d}_{j+1}, (1 - \lambda)\tilde{v}_j + \lambda\tilde{v}_{j+1} \right) \mid \lambda \in [0, 1], j \in \llbracket 1, n_B - 1 \rrbracket \right\}.$$

³Derived from \mathbf{d} after selecting voxels of the structure of interest, and sorting voxels.

⁴Here we are constantly comparing two DVH curves of the same structure on the same patient, so we always have $n_A = n_B$.

Wasserstein Distance The Wasserstein distance, also known as the Earth Mover's Distance, is a metric used to quantify the difference between two probability distributions [OP82]. Formally, given two probability distributions μ and ν defined on a metric space X , the Wasserstein distance, denoted $d_{\text{Wasserstein}}(\mu, \nu)$, represents the infimum cost of transporting the mass of distribution μ to match distribution ν , where the transportation cost is determined by the distance metric $dist$ on X . It is defined as:

$$d_{\text{Wasserstein}}(P, Q) = \inf_{\gamma \in \Gamma(\mu, \nu)} E_{(x, y) \sim \gamma} [dist(x, y)]$$

where $\Gamma(\mu, \nu)$ represents the set of all possible joint distributions $\gamma(x, y)$ with marginals μ and ν .

In our analysis, we treat DVH (Dose-Volume Histogram) curves as probability distributions and employ the Wasserstein distance to assess their differences. This metric has the distinct advantage of capturing both local and global variations between the curves, offering a more comprehensive comparison. However, it can be computationally demanding, particularly when dealing with DVHs of large anatomical structures.

Kolmogorov-Smirnov Distance Another distance metric commonly used to compare DVH curves is the Kolmogorov-Smirnov (KS) distance [Ste74]. The KS distance measures the maximum vertical separation between two curves and is particularly well-suited for comparing non-parametric distributions, such as DVH curves.

Mathematically, let the two DVH curves be represented by functions f and g , mapping dose levels to volume ratios. The KS distance, d_{KS} , is then defined as:

$$d_{KS} = \sup_{x \in \mathbb{R}^+} |f(x) - g(x)|.$$

In the case of discrete DVH data, f and g are piecewise linear, continuous functions from \mathbb{R}^+ to $[0, 1]$, with their values set to zero beyond the maximum dose level.

Total Variation Distance We propose a distance metric that computes the integral of the absolute difference between two DVH (dose-volume histogram) curves. This metric is straightforward to compute and provides a balanced measure of local and global differences between the curves [Cha]. Additionally, it is computationally efficient and well-suited for analyzing large structures with many voxels. This approach yielded the most consistent and clinically relevant results among the metrics tested. As such, we selected this distance measure for our analysis.

Traditionally, the total variation distance is defined as the integral of the absolute difference between two DVH curves. While the dose domain is theoretically unbounded, the volume domain is bounded between 0% and 100%. To avoid integrating over an unbounded dose domain, we opted to reverse the axes, placing dose on the y -axis and volume on the x -axis and subsequently integrating the absolute difference in dose over the volume range $[0, 1]$.

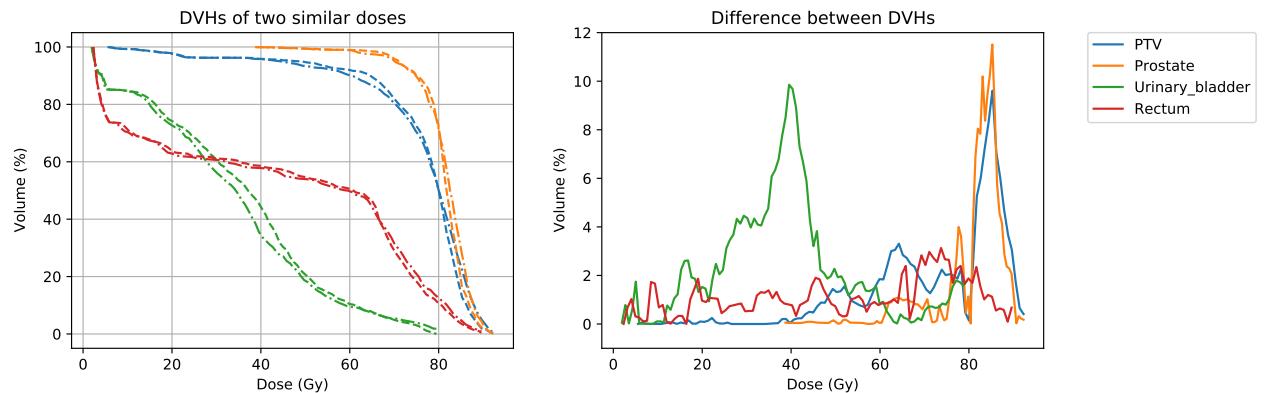
Mathematically, standard DVHs are described by $V : \mathbb{R}^+ \rightarrow [0, 1]$. For two DVHs $V(d)$ and $\tilde{V}(d)$, the total variation distance is given by:

$$d_{\text{TotalVariation}} = \int_0^{+\infty} |V(d) - \tilde{V}(d)| dd$$

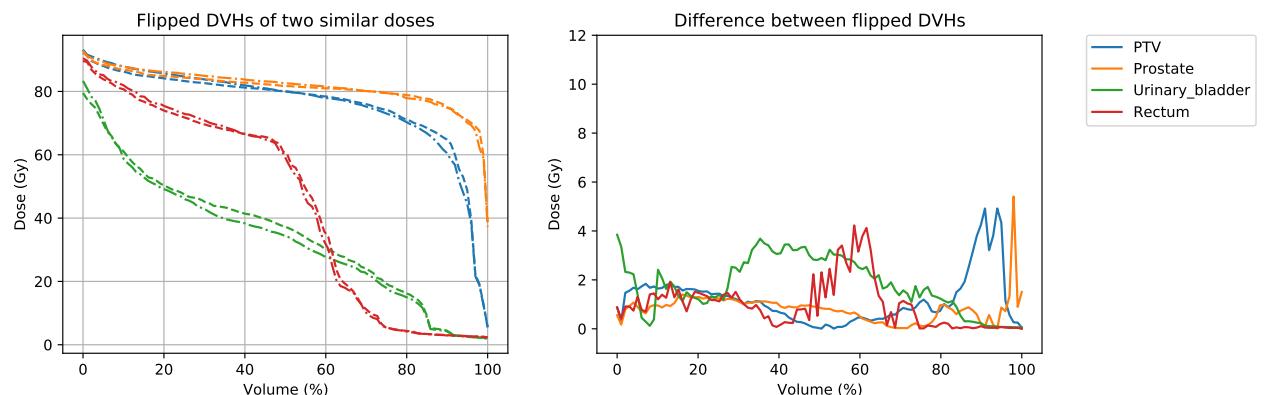
However, in our approach, we express DVHs with dose as a volume function, denoted $D : [0, 1] \rightarrow \mathbb{R}^+$. Thus, for two DVHs $D(v)$ and $\tilde{D}(v)$, the total variation distance becomes:

$$d_{\text{TotalVariation}} = \int_0^1 |D(v) - \tilde{D}(v)| dv$$

While the theoretical value of the integral remains unchanged, we prefer integrating over the finite volume domain $[0, 1]$ instead of the unbounded dose domain $\mathbb{R}^+ = [0, +\infty[$. An illustration highlighting the differences between the classical DVH and the version with swapped x - and y -axes is presented in Figure 2. The two compared doses were optimized on the TG-119 phantom prostate case, using different weights (1 and 3) for the PTV objective.



(a) Classical DVH (dose on the *x*-axis).



(b) Flipped axes DVH (volume on the *x*-axis).

Figure 2: DVHs: Comparison of classical and flipped axes styles.

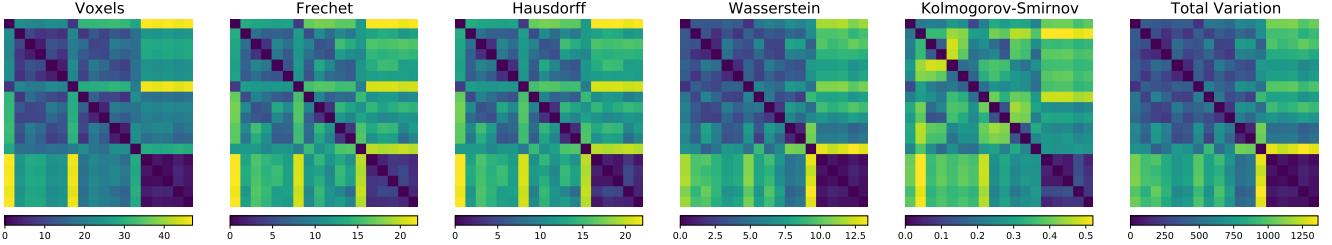


Figure 3: Pairwise distances between doses
(with different distances calculation method)

As Figure 2 shows, the difference between DVHs exhibits less noise (fewer fluctuations) when the dose is on the x -axis. This observation suggests a reduction in numerical error, providing additional motivation to place the volume on the x -axis.

Computing the total variation distance is computationally efficient, requiring only $\mathcal{O}(n_s)$ operations per structure, where n_s represents the number of voxels in the structure of interest, s . Overall, this method achieves a good balance between capturing local and global differences in DVH curves.

1.2 Results

1.2.1 Dose Distances Comparison

We optimized each constraint with all weights initially set to 1 but sequentially increased one to 3, 10, and 100. This process resulted in 18 distinct dose distributions, which were compared using the distance metrics described earlier. We calculated the pairwise distances for each pair of doses, effectively constructing the adjacency matrix of a fully connected graph, where each optimized dose corresponds to a node. See Figure 3 for comparing the adjacency matrices.

Ideally, the distance metric should satisfy the following criteria:

- It should match the voxel-wise distance when the voxel-wise difference is small.
- It should remain small in cases where the voxel-wise distance is significant. However, the clinical significance of the two doses is similar, even if the doses are voxel-wise different.

From the pairwise distances shown in Figure 3, we make the following observations:

- The Fréchet and Hausdorff distances behave similarly to the voxel-wise distance, indicating that they are too sensitive. Thus, they are not suitable for our purpose.
- The Kolmogorov-Smirnov distance appears to degenerate, likely capturing noise due to numerical approximations in the DVH calculations. Therefore, it is also not suitable for our purpose.
- The Wasserstein and Total Variation distances produce more clinically relevant results. As a result, we chose to focus further analysis on these two metrics.

1.2.2 Link between Total Variation and Wasserstein

The adjacency matrices for the Wasserstein and Total Variation distances exhibit substantial similarity. This similarity is expected, as the two metrics are equivalent in this context, given that we employed the Earth Mover’s Distance (Wasserstein distance with $p = 1$). The Total Variation distance can be regarded as a particular case of the Wasserstein distance.

The Wasserstein distance, also known as the Earth Mover’s Distance, provides a metric for quantifying the distance between two probability distributions. Let X and Y be two distributions with cumulative distribution functions (CDFs) F and G , respectively. The Wasserstein distance between them is formally defined as:

$$W_p(F, G) = \inf_{\pi \in \Pi(F, G)} \left(\iint_{x, y \in \mathbb{R}^2} |x - y|^p d\pi(x, y) \right)^{1/p}$$

where $\Pi(F, G)$ represents the set of all possible joint distributions with F and G as marginals.

In contrast, the Total Variation distance between the two curves F and G is defined as:

$$\text{TotalVariation}(F, G) = \int_{x \in \mathbb{R}} |F(x) - G(x)| dx$$

When the Wasserstein distance is computed with $p = 1$, it becomes equivalent to the Total Variation distance:

$$W_1(F, G) \equiv \text{TotalVariation}(F, G).$$

Thus, the only expected differences between these two distance metrics in our analysis should arise from numerical errors.

1.2.3 Bounding of Total Variation and Voxel Distance

The Voxel Distance can bound the Total Variation distance. However, the reverse is impossible, as illustrated by the example in the introduction, where two doses exhibit nearly identical dose-volume histograms but significantly different voxel-wise distances.

In the following, we provide a bound for the Total Variation distance of a single DVH, which can be generalized to the sum of the Total Variation distances across all DVHs.

This proof demonstrates that while the Voxel Distance constrains the Total Variation distance, the converse does not hold, especially when voxel-wise variations do not translate to clinically meaningful differences in the global dose distribution.

We aim to compare two dose distributions on a structure, d and \tilde{d} (we suppose that d and \tilde{d} are lists containing only the values on the voxels of the structure).

Sorting lists

Lemma 1. Let $\dot{l}, l^* \in \mathbb{R}^n$. Let \dot{l} be sorted and $\dot{l}^* \in \mathbb{R}^n$ be sorted version of l^* . Then, we have:

$$|\dot{l} - l^*| \geq |\dot{l} - \dot{l}^*|$$

Proof. Suppose $a < b$ and $c < d$, and WLOG, $a \leq c$.

We have $|a - d| = |a - c| + |c - d|$ so $|a - d| + |b - c| = |a - c| + |c - d| + |b - c|$ using triangle inequality ($|c - d| + |b - c| \geq |b - c|$): $|a - d| + |b - c| \leq |a - c| + |b - d|$.

Thus, with \dot{l} sorted, swapping elements l_i and l_j ($i < j$) of l^* decreases $|\dot{l} - l^*|$ if $l_i \geq l_j$. Applying, bubble sort on l^* , we obtain \dot{l}^* doing only permutations satisfying the condition just stated.

Hence, we obtain $|\dot{l} - l^*| \geq |\dot{l} - \dot{l}^*|$ at the end of the bubble sort. \square

Corollary 1. Let $l, l^* \in \mathbb{R}^n$. Let $\dot{l}, \dot{l}^* \in \mathbb{R}^n$ be sorted version of l, l^* . Then:

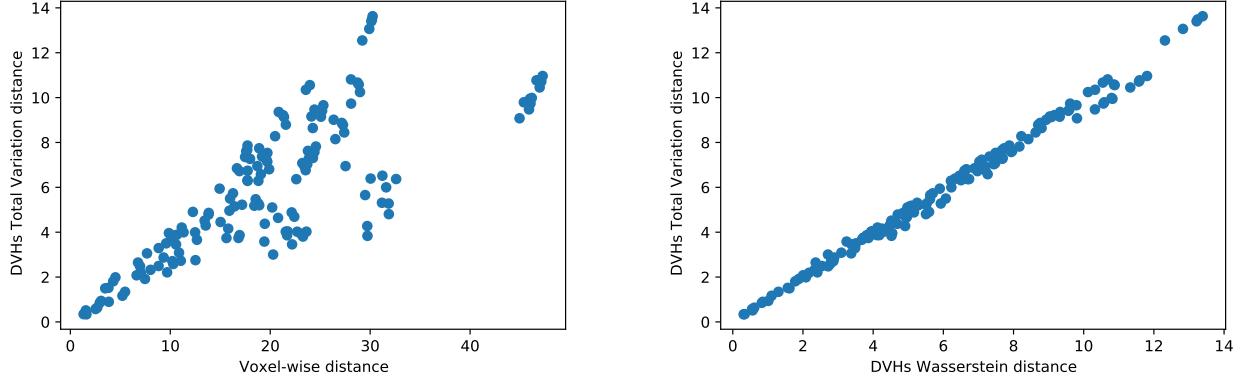
$$|l - l^*| \geq |\dot{l} - \dot{l}^*|$$

Proof. The order in which we perform $|l - l^*| = \sum_{k=1}^n |l_k - l_k^*|$ can be chosen, so $|l - l^*| = \sum_{k=1}^n |l_{\sigma(k)} - l_{\sigma(k)}^*|$ (with σ a permutation of $\llbracket 1, n \rrbracket$). Taking σ such that $l_{\sigma(i)} \leq l_{\sigma(j)}$ for $i < j$ and using lemma finishes the proof. \square

Proof Outline Suppose the voxel-wise difference is ε -small (i.e. $|d_i - \tilde{d}_i| < \varepsilon$). Then, the total variation of the unsorted vector doses is $|d - \tilde{d}| < n_S \varepsilon$. Let \dot{d} be sorted d and $\dot{\tilde{d}}$ be sorted \tilde{d} . Then, by Corollary 1, we have: $|\dot{d} - \dot{\tilde{d}}| \leq |d - \tilde{d}| < n_S \varepsilon$.

Therefore, if d and \tilde{d} are sufficiently close, $\varepsilon \rightarrow 0$ and $|\dot{d} - \dot{\tilde{d}}| \rightarrow 0$.

Conclusion Thus, voxel-wise very close doses distributions will also have close DVHs distances, which ensure DVHs distances are non-degenerative.



(a) DVHs Total Variation vs Voxel-wise.

(b) Total Variation vs Wasserstein.

Figure 4: Comparing Distances

1.2.4 Distances Distribution Comparison

Comparing Total Variation and Voxel-wise The bounding of the total variation DVH distance in terms of voxel-wise distance is clearly illustrated in Figure 4a, where a linear upper bound can be observed in the scatter plot. However, specific pairs of doses are closer regarding DVH distance than initially anticipated based solely on voxel-wise comparisons. This observation underscores the need for a more nuanced analysis beyond voxel-wise comparison, as it may overlook clinically relevant similarities between dose distributions.

Comparing Total Variation and Wasserstein Figure 4b shows that the two DVH distances are nearly perfectly proportional. This result aligns with expectations, given that they are mathematically equivalent. The only difference lies in the integration axis in the total variation distance, which accounts for the small fluctuations observed, likely due to accumulated numerical error.

1.3 Discussion

In this section, we introduce a novel metric for comparing radiation doses. This metric offers the advantage of being insensitive to dose changes in certain regions, provided they are compensated in other regions, thus achieving the intended objective. This property makes the metric particularly useful in various applications, including dose mimicking and determining early stopping criteria for fluence map optimization.

Despite the advantages, this distance metric has certain limitations. A notable drawback is its inability to capture spatial dose distribution, which may pose challenges in specific cases. Pathological examples exist where two DVHs appear similar, but the clinical interpretation differs significantly. Other factors, such as the spatial distribution of the dose within the target volume or surrounding tissues, can play a pivotal role in the treatment's effectiveness. To cover this, incorporating hot points in the metric could help. However, it remains to be seen how to do it in a clinically meaningful way.

For instance, two dose distributions might deliver the same high dose, with one distributed across several small regions and the other concentrated in a single large region. While the DVHs may appear identical, clinicians would interpret these two dose distributions differently. Such edge cases, however, are sporadic in clinical practice. Nonetheless, for critical cases, we recommend complementing this metric with voxel-invariant approaches and other techniques to evaluate the radiation doses comprehensively.

When comparing two distinct doses, a considerable distance between them may indicate a significant difference in the intensity or frequency of the treatment. However, this does not necessarily imply that one dose is superior. The effectiveness of a dose depends on several other factors, such as the individual patient's characteristics, medical history, and treatment response.

The distance between two doses does not assess which is more effective or clinically appropriate in a specific case. It only acknowledges that they are clinically equivalent or not, and in the latter case, a dosimetrist needs to make a choice. It is essential to account for all relevant factors when evaluating the efficacy of a treatment dose to ensure a comprehensive understanding.

Overall, the proposed dose comparison technique presents a promising tool for radiation dose evaluation. While it has certain limitations, it can serve as a valuable addition to the repertoire of methods employed by radiation oncologists and medical physicists for optimizing treatment plans and improving patient outcomes. Complementing existing techniques offers an additional layer of analysis, contributing to more informed decision-making in clinical practice.

Stop Criterion Defining an adequate stopping criterion for the fluence map optimization process is a critical challenge in radiotherapy dose optimization. In clinical practice, dosimetrists often guide optimization, who may terminate the process when they are satisfied with the outcome. However, the need for fully automated optimization processes requires the establishment of systematic and objective stopping criteria. One potential approach is to compare the clinical effects of two dose distributions and stop when one optimization step does not change the clinical effect. This method can help evaluate different solutions and determine the optimal point to terminate the optimization process.

2 Network of Doses

In this section, we aim to construct a clinically meaningful network of dose distributions, where each node represents a distinct dose distribution, and the edges quantify the relationships between them. By creating such a network, we can identify clusters of similar dose distributions and uncover patterns that reflect clinical relevance. This network-based approach will enable us to visualize better, analyze, and interpret the relationships between various treatment plans, ultimately improving the comparison and optimization of radiotherapy strategies.

Numerous efforts have been made to automate the treatment optimization process in radiation therapy. One promising avenue is the exploration of the Pareto frontier, as discussed in [KGC⁺23] and [OKB10], which seeks to identify a set of treatment plans that balance conflicting objectives, such as maximizing tumor control while minimizing damage to surrounding healthy tissue. Another approach proposed [CX03] consists of directly extracting leaf movements from patient data to enhance the automation of dose delivery.

Despite these advancements, fully automated approaches have yet to be widely adopted in clinical practice, primarily due to practical limitations and the complexity of translating these methods into routine use. Using doses network analysis, we will propose a hybrid approach that integrates both manual and automated treatment optimization. This method will combine the computational efficiency of automation with the expert judgment of clinicians, ensuring that the final treatment plans are both optimal and clinically relevant.

2.1 Methods

Multiple Plans Generation We employed the same dose optimization process as previously described. The cost function utilized in this study is designed to be convex, ensuring that, irrespective of the specific weight assignments given to the objectives, minimizing this cost function will consistently converge to an optimal radiotherapy plan. To generate a diverse set of treatment doses for a given patient case and set of constraints, we optimized the cost function with varying weight assignments for each constraint.

This approach mirrors the current practice in which dosimetrists adjust the weights associated with different dose objectives to guide the optimization engine toward a clinically acceptable solution. By altering these weights, it becomes possible to explore different trade-offs and prioritize certain aspects of the treatment plan, such as sparing healthy tissue or enhancing tumor coverage, according to the patient's specific clinical needs.

Objectives' Weights Generation The weights assigned to each optimization objective dictate the relative importance of each objective in the trade-offs made by the optimization engine. Initially, we begin

with a typical weight assignment used in clinical practice. To generate a diverse set of weights, we perturb these initial values by adding random normal noise, resulting in a unique new set of weights.

By repeating this process, we can explore a wide range of potential treatment plans. This thorough exploration provides a nuanced understanding of the trade-offs between competing clinical objectives. The random computational exploration of irradiation strategies extends beyond the capabilities of manual exploration, enabling clinicians to make more informed and tailored decisions based on a broader array of treatment possibilities.

Dose Normalization For consistency across cases, we normalized the doses using the " $D_{50\%}$ " normalization method, a common practice in radiation therapy. This method normalizes the dose such that the median dose delivered to the PTV equals the prescribed dose, ensuring comparability across treatment plans.

Phantom Patient Our proposed method for clustering radiation doses was evaluated using the TG-119 Prostate case [MYH⁺12], a well-established benchmark dataset commonly employed to assess the quality of radiation therapy plans. The TG-119 dataset provides predefined dose objectives, which were incorporated into the formulation of our cost function. This benchmark allows for a robust evaluation of our clustering approach in the context of clinically relevant treatment goals, ensuring that the method aligns with widely accepted standards in radiation therapy planning.

Dose Clustering Techniques

Dose Distance We first needed to establish a method for defining the distance between individual doses to perform the dose clustering. We employed the Euclidean distance between voxel-wise dose distributions as our distance metric. The weight of each edge between two doses was then defined as the inverse of this distance as we sought to maximize the edge weights between similar doses. Defining edge weights as the inverse of the distance between nodes is a common practice in graph theory, as noted in previous works [?] [?].

Community Detection We employed Louvain's method for community detection to cluster the radiation doses. Louvain's method is a greedy optimization algorithm that aims to maximize the modularity of the graph. Modularity is a metric used to assess the quality of a graph partition by quantifying how well the graph is divided into distinct communities. For our analysis, we utilized the implementation of Louvain's method available in the Python library NetworkX [?], which facilitated partitioning the dose similarity graph into meaningful clusters.

Evaluating Communities Split The clustering quality was evaluated using DVHs derived from the different doses. We computed the mean and standard deviation of the DVHs curves for each cluster and the entire dataset.

We analyzed the relative volume doses across four distinct anatomical structures. One hundred one dose values were sampled for each structure, corresponding to volume percentages ranging from 0 % to 100 % in equal 1 % intervals. These values were aggregated into a single vector, resulting in a vector of length 404 for each dose.

To assess the variability within these dose vectors, we calculated the standard deviation for each of the 404 elements in the vector. This measure provides insight into the dispersion of dose values within the structures. By averaging the calculated standard deviations, we derived a scalar metric that quantifies the degree of separation among the clustered doses, reflecting the consistency or variability of the doses within each group.

2.2 Results

Visual Clustering Evaluation

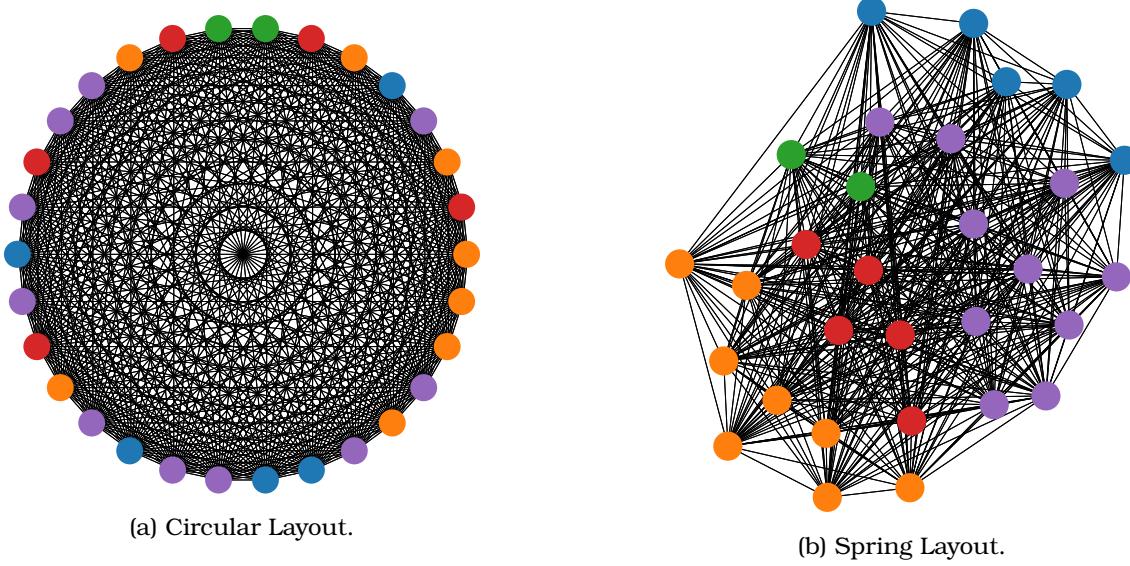


Figure 5: Plot of the Network
 Edges width \propto edge weight $\propto 1/\text{distance}$
 Node's color reflects community attribution

Graph Plots In figure 5, each node represents a dose. The communities are attributed using Louvain method and are identified by colors. Since the graph is clearly not planar, we choose to plot it in a circular layout (fig. 5a) and in a spring layout (fig. 5b).

In order to obtain a more precise understanding of the edge weights in the network, one can refer to the adjacency matrix of the edge weights, as depicted in Figure 6.

DVH Plot Figure 7 illustrates the DVHs, with the colors of the plots corresponding to the communities identified in the network analysis. To ensure clarity and prevent confusion, each structure is presented on a separate plot. Notably, our observations align with expectations, as doses assigned to nodes within the same community (indicated by the same color) exhibit nearly overlapping DVHs.

The visualization of DVHs offers critical insights into the distribution of radiation doses across different anatomical structures. The strong alignment between observed dose patterns and community assignments highlights the potential of network analysis to reveal significant patterns within radiation therapy data. By mapping the colors of the DVH plots to the communities identified in the network, we gain a deeper understanding of the relationship between dose assignments and structural characteristics.

This analysis supports the hypothesis that nodes (representing doses) within the same community exhibit similar dose profiles. As a result, these doses could be consolidated, given that they are likely to have comparable clinical outcomes. This method provides an efficient approach to identifying groups of treatment plans that share equivalent therapeutic effects, enabling a more streamlined decision-making process.

Numerical Clustering Evaluation As explained in subsection 2.1, we use the mean standard deviation of doses at 101 equispaced values of volume to obtain a scalar value of how far apart are a set of doses; see table 1 for results.

The mean standard deviation values for the clusters — 7.30, 4.51, 0.95, 2.17, and 7.92, with an average of 4.57 — are significantly lower (nearly four times) than the mean, standard deviation of the entire network, which is 15.54 (see Table 1). This substantial reduction in standard deviation provides clear quantitative evidence that our clustering method yields favorable results.

This quantitative improvement is further corroborated by qualitative insights, as shown in Figure 7, where the dose clusters effectively characterize the distribution of radiation doses. The strong alignment

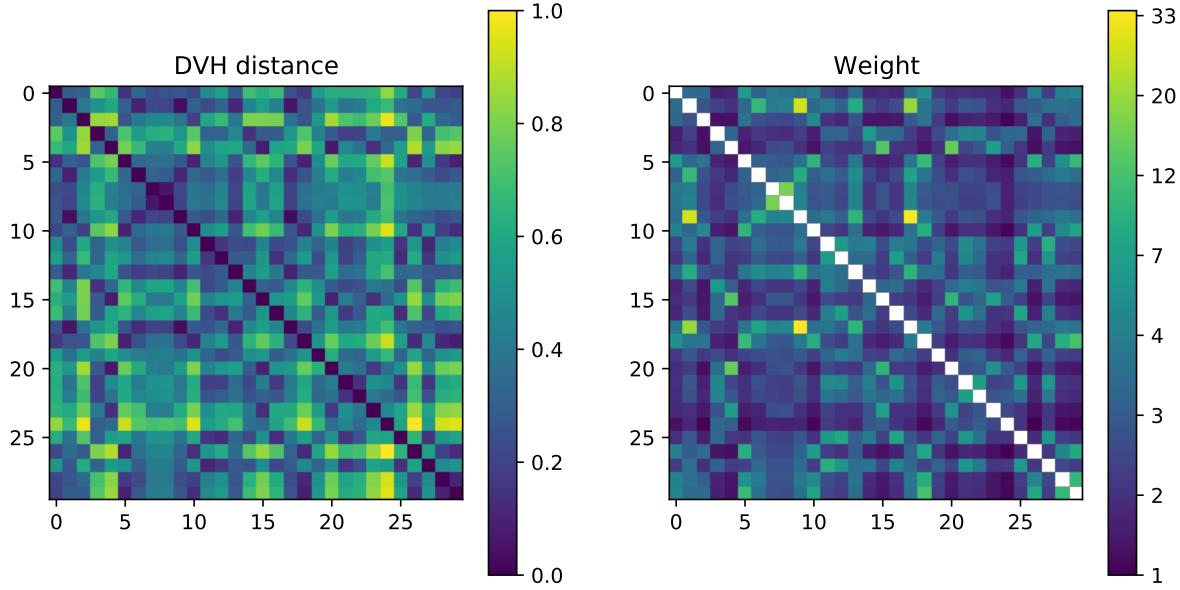


Figure 6: Weight Adjacency Matrix of the Network

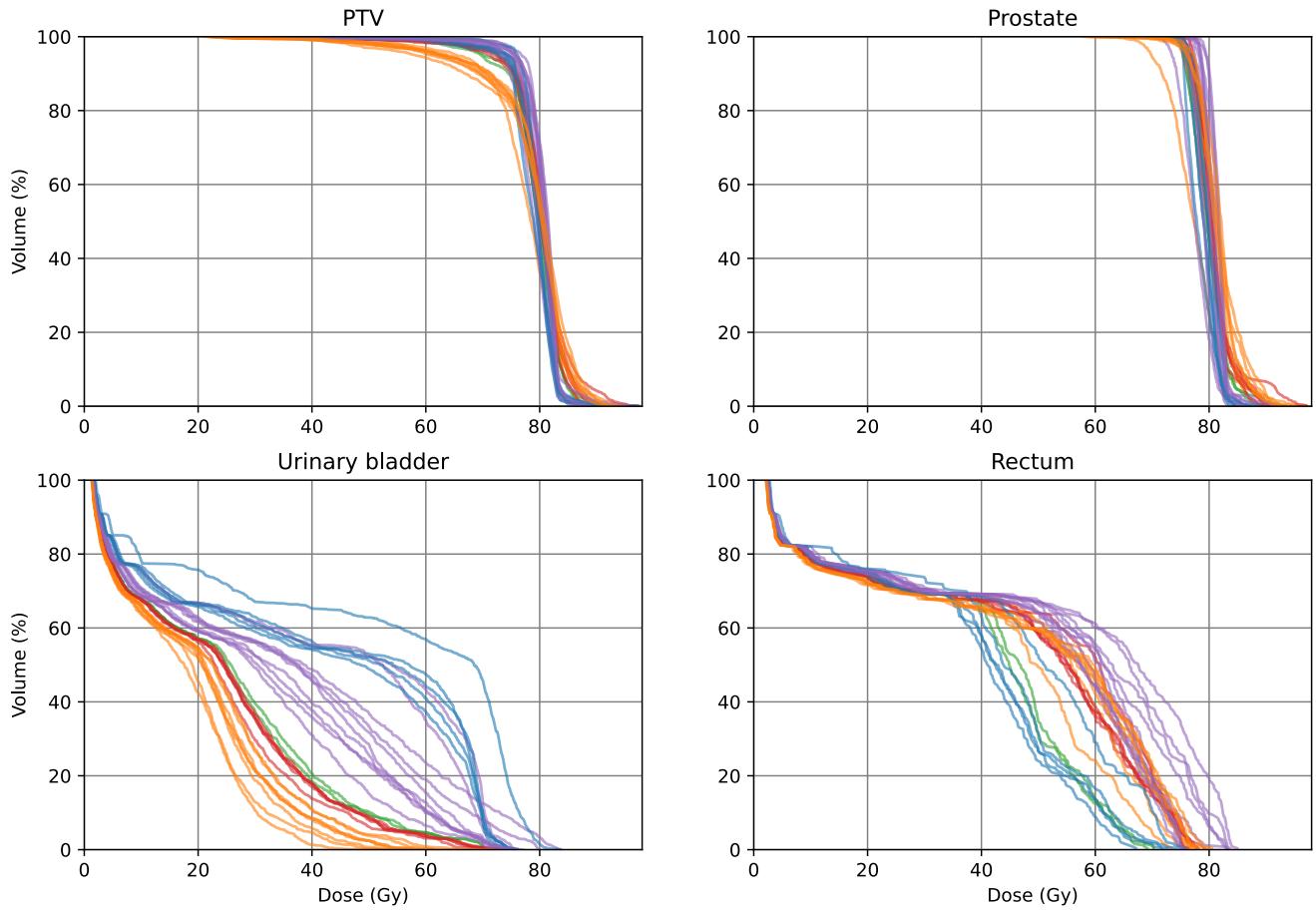


Figure 7: Dose-Volume Histogram

Set	Set Size	Mean Standard Deviation
Cluster 1 (blue)	5	7.30
Cluster 2 (orange)	8	4.51
Cluster 3 (green)	2	0.95
Cluster 4 (red)	5	2.17
Cluster 5 (purple)	10	7.92
<i>All</i>	19	15.54

Table 1: Clustering Quality

between the dose clusters and the DVH patterns reinforces the validity and relevance of our clustering strategy. Our proposed clustering approach performs well in qualitative and quantitative evaluations, underscoring its potential utility in optimizing radiation therapy plans.

2.3 Discussion

This study presents a novel approach to clustering doses based on their distributions, providing both quantitative and qualitative evidence of the significance of such clustering. However, clustering doses remains a complex task due to the high dimensionality of dose space and the irregularities in dose distributions. While the results presented in this study are promising, they are not yet ready for clinical implementation. The results suggest that doses within the same cluster likely have indistinguishable clinical effects, which opens up several practical applications.

Regrouping doses into clusters offers valuable insights into clinical practices and introduces new possibilities for optimizing radiotherapy treatment plans. One potential application of dose clustering is to use the identified clusters as a similarity measure for early stopping in optimization processes, leading to potential time savings and computational efficiency. Specifically, if the dose distribution shows convergence with the previous n dose distributions, this could serve as an indication to halt further optimization, thereby improving efficiency.

Moreover, expanding this clustering approach to include doses from multiple patients could provide insight into the treatment practices of different clinical centers, facilitating comparisons of outcomes for similar anatomies.

In the following section, we propose a clinical application of the dose clustering technique that can potentially streamline the optimization process for dosimetrists. This approach could significantly reduce the number of manual interactions, or "clicks," required within the Treatment Planning System, thereby enabling faster and more efficient dose optimization.

3 A Novel Framework for Multi-Objective Optimization and Robust Plan Selection Using Graph Theory⁵

3.1 Challenges in Current Practices

In recent years, the complexity of treatment planning in radiotherapy has increased substantially, with planners having to balance multiple objectives, such as maximizing tumor control while minimizing damage to surrounding healthy tissue. Despite advances in radiotherapy planning, current practices still face limitations. Traditional approaches often rely on manual fine-tuning of parameters, making it difficult to achieve a globally optimal solution. Multi-objective optimization balances competing goals and is a manual, time-consuming process. Planners must evaluate and adjust multiple treatment plans to account for uncertainties and variations in patient anatomy. This approach is labor-intensive and prone to inconsistencies due to human factors, which can lead to suboptimal patient outcomes. We present a novel framework for multi-objective optimization and robust treatment plan selection based on graph theory to address these challenges. This framework introduces significant automation into the planning process, allowing for more efficient and reliable plan generation.

⁵Presented at ESTRO 2024.

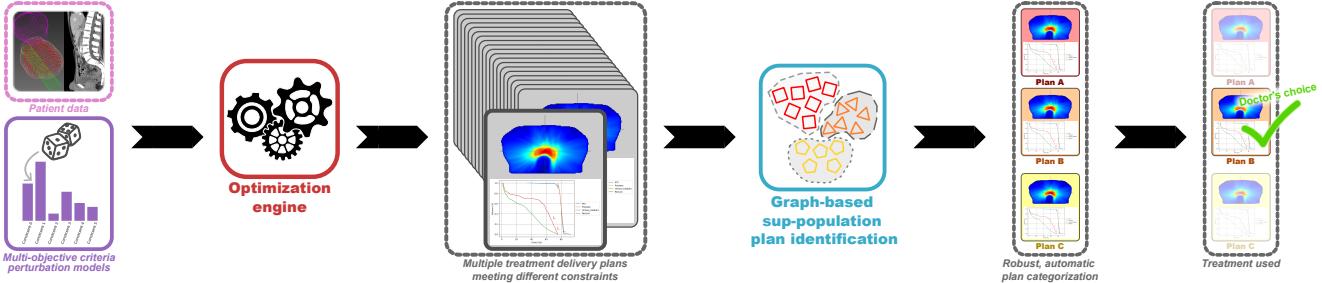


Figure 8: Graphical Abstract of the Proposed Process.

3.2 Proposed Framework

The first step in the framework is the construction of the graph, which is based on the systematic random perturbation of importance factors assigned to clinical constraints. By treating this set of plans as a graph, we can explore the landscape of possible solutions more systematically. It would not be practical to ask doctors and dosimetrists to evaluate dozens of treatment plans.

We use clustering algorithm described in the previous section to regroup doses with similar effects. Clusters of plans, or "communities," emerge based on their proximity in the graph, allowing us to narrow down the set of candidates for more detailed evaluation. By focusing on these robust communities, we reduce the need for manual intervention in plan selection. Instead of planners having to sift through dozens of plans, the framework automatically highlights the most promising options, significantly reducing both time and effort (see figure 8).

This framework represents a shift from the traditional "many-clicks" approach—where planners must manually adjust and evaluate numerous treatment plans—to a "few-clicks" workflow, where the system presents only the most relevant, robust plans for final review. This automation not only saves time but also reduces the risk of human error, leading to more consistent, higher-quality outcomes for patients.

3.3 Discussion

The proposed framework transitions from a traditional N -click approach⁶ to a more efficient n -click solution⁷. While this shift significantly reduces the time required for plan selection, it still requires dosimetrists and doctors to adapt to a new workflow. This change in habit—moving from manually reviewing and adjusting plans to relying on semi-automated plan suggestions—may pose a barrier to clinical implementation.

For a solution to be adopted in everyday practice, it must minimize the disruption to current workflows. Ideally, this means evolving the framework into a proper 1-click solution, where the system autonomously generates and selects the optimal, robust treatment plan without requiring manual intervention. Such a system would eliminate the need for dosimetrists and doctors to alter their habits and ensure consistent, high-quality outcomes with minimal effort. Achieving this level of automation is essential for widespread clinical adoption. Therefore, future work in this manuscript will focus on refining the framework to meet these demands, bringing it closer to a fully automated, practical tool for radiotherapy planning.

References

- [Cha] Sourav Chatterjee. *Distances between probability measures*.
- [CX03] Cristian Cotrutz and Lei Xing. Segment-based dose optimization using a genetic algorithm. *Physics in Medicine & Biology*, 48(18):2987, sep 2003.

⁶with $N \gg 10$

⁷with $10 > n > 1$

- [EM02] Har-Peled Sariel Mitchell Efrat, Guibas and Murali. New similarity measures between polylines with applications to morphing and polygon sweeping. *Reports of Practical Oncology and Radiotherapy*, 28(4):535 – 569, 2002.
- [Hen99] Jeffrey T. Henrikson. Completeness and total boundedness of the hausdorff metric. 1999.
- [KGC⁺23] Sander Kuipers, Jérémie Godart, Anouk Corbeau, Abdul Wahab Sharfo, Sebastiaan Breedveld, Jan Willem Mens, Stephanie de Boer, Remi Nout, and Mischa Hoogeman. The impact of bone marrow sparing on organs at risk dose for cervical cancer: a pareto front analysis. *Frontiers in Oncology*, 13, 2023.
- [MYH⁺12] Dinesh Kumar Mynampati, Ravindra Yaparpalvi, Linda Hong, Hsiang-Chi Kuo, and Dennis Mah. Application of aapm tg 119 to volumetric arc therapy (vmat). *Journal of Applied Clinical Medical Physics*, 13(5):108–116, 2012.
- [OKB10] R O Ottosson, A Karlsson, and C F Behrens. Pareto front analysis of 6 and 15 mv dynamic imrt for lung cancer using pencil beam, aaa and monte carlo. *Physics in Medicine & Biology*, 55(16):4521, jul 2010.
- [OP82] I. Olkin and F. Pukelsheim. The distance between two random vectors with given dispersion matrices. *Linear Algebra and its Applications*, 48:257–263, 1982.
- [PLQC12] Dilip K. Prasad, Maylor K.H. Leung, Chai Quek, and Siu-Yeung Cho. A novel framework for making dominant point detection methods non-parametric. *Image and Vision Computing*, 30(11):843–859, 2012.
- [Ste74] M. A. Stephens. Edf statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association*, 69(347):730–737, 1974.