Trading Robustness? Introducing total variance objectives to Pareto surface approximation techniques for radiotherapy

Remo Cristoforetti[1,2,3], Philipp Süss[4], and Niklas Wahl[1,2]

[1]Department of Medical Physics in Radiation Oncology, German Cancer Research Center (DKFZ), Heidelberg, Germany
[2]National Center for Radiation Research in Oncology (NCRO), Heidelberg Institute for Radiation Oncology (HIRO), Heidelberg, Germany
[3]Department of Physics and Astronomy, Heidelberg University, Heidelberg, Germany
[4]Optimization division, Fraunhofer Institute for Industrial Mathematics (ITWM), Kaiserslautern, Germany

Abstract This work introduces a novel approach to introduce robustness into multicriteria optimization (MCO) in proton therapy. Our approach exploits a formulation of expected value optimization relying on precomputed total variance and expected dose influence. This facilitates merging robust optimization with Pareto front approximation while avoiding explicit calculation of scenarios during optimization. The method also introduces robustness as an objective, enabling interactive management of robustness trade-offs.
We demonstrate MCO including robustness trading with a single field proton plan on the TG119 phantom and a two field proton plan on a liver case. Trade-offs between robustness and dose metrics can be effectively approximated and visualized. Our approach allows planners more explicit control over robustness, and is directly translatable to other MCO approaches like lexicographic optimization.

1 Introduction

Radiotherapy treatment planning is most commonly modeled as a multicriteria optimization (MCO) problem, as trade-offs between multiple organs' toxicities and target coverage need to be managed. Approaches to solve this MCO problem have evolved from the manual weighted-sum approach, which involves repetitive manual tweaking of objective weights in consecutive optimization runs, to Pareto surface approximation techniques or automated lexicographic solvers.
In parallel, techniques like robust optimization were introduced to personalize the management of treatment uncertainties, particularly in particle therapy. In robust optimization, uncertainties are explicitly modeled through computation of error scenarios. These are then aggregated during optimization to minimize a robustness metric of the objective function, like its maximum over all scenarios or its expected value.
While the combination of robust optimization and modern MCO approaches has been suggested before [1], merging both is often limited by incompatible problem formulations or combination of both methods' increased computational demands. For MCO, computational complexity depends on the number of objectives. Performance of robust optimization depends on the number of error scenarios used during optimization. Introducing new robustness objectives computed from a large number of error scenarios to MCO is therefore limited by computational resources.
To address this challenge, we introduce a novel approach that incorporates total variance objectives into Pareto surface approximation techniques. Our method stands out as it formulates robustness objectives whose evaluation cost during optimization is independent of the number of the calculated error scenarios, thereby significantly enhancing computational efficiency without compromising on the robustness of the treatment plans. This independence is achieved by precomputing influence matrices of expected dose and total variance. Then, during approximation of the Pareto front, scenarios can be discarded or stored on disk and only these lightweight matrices are needed. For visualization or navigation of the solutions, the scenarios can be reloaded and the robustness trade-off can be visualized.
We presents the theoretical foundation of our method, the resulting algorithm and their implementation, and demonstrate the possible trade-offs in planning problems.

2 Materials and Methods

Our approach exploits a specific form of robust treatment planning, called probabilistic optimization or expected value optimization [2, 3], and combines it with Pareto surface approximation and navigation techniques.

2.1 Probabilistic Optimization

A common formulation of a probabilistic optimization problem would minimize the expected value $\mathbb{E}[f(d(x)]$ of an objective function f depending on beamlet intensities x [3]:

$$x^* = \arg\min_x \mathbb{E}[f(d(x))]\,. \tag{1}$$

The objective function f is usually formulated depending on dose d and chained to the beamlet intensities x by precomputation of a dose influence matrix D such that $d = Dx$.
For practical purposes, $\mathbb{E}[f]$ in eq. (1) is computed by averaging over a limited number of S samples from the assigned probability distributions. This, in turn, requires computation of S dose influence matrix samples D_s to be able to compute the dose during each iteration of an optimization run. This means that during optimization, the S samples of the dose influence matrix need to be kept in memory while keeping S as small as possible to guarantee computationally efficient treatment plan optimization.

2.2 Total Variance and expected dose influence

For the penalized least-squares objective function $f^{lsq}(d) = (d-d^{ref})^T P(d-d^{ref})$, where d^{ref} is the prescription dose and P a diagonal matrix encoding voxel penalty factors, it was previously shown [4] that its expectation can be separated into a total variance term (written as the trace over the doses' covariance matrix Σ) and an expected value term:

$$\mathbb{E}[f^{lsq}(d)] = \mathrm{tr}(P\Sigma) + (\mathbb{E}[d]-d^{ref})^T P(\mathbb{E}[d]-d^{ref}) \quad (2)$$
$$= x^T\Omega x + \mathbb{E}[D]x - d^{ref})^T P(\mathbb{E}[D]x - d^{ref}) \; . \quad (3)$$

Equation (3) introduces the "total variance" influence matrix Ω together with the expected dose influence matrix $\mathbb{E}[D]$. While the latter can easily be estimated by taking the elementwise mean over all D_s, computing Ω requires estimating the expected value of elementwise combinations of D via

$$\mathrm{tr}(P\Sigma) = \sum_{jk} x_j x_k \underbrace{\sum_i p_i \left(\mathbb{E}\left[D_{ij}D_{ik}\right] - \mathbb{E}\left[D_{ij}\right]\mathbb{E}\left[D_{ik}\right]\right)}_{\Omega_{jk}} \; , \quad (4)$$

with voxel penalties p_i in voxel i and intensity vector x.
Under the assumption that penalties are equal in certain voxels, i. e., voxels i belonging to the same volume of interest (VOI) V, corresponding Ω_V can be computed after pulling the penalty out of the sum, simplifying to

$$\Omega_V = \mathbb{E}[D_V^T D_V] - \mathbb{E}[D_V]^T\mathbb{E}[D_V] \; , \quad (5)$$

where D_V are the dose influence matrices for V.
Since Ω_V and $\mathbb{E}[D]$ can be precomputed during scenario dose calculation, we can now build objective functions depending on the total variance $x^T\Omega_V x$ or the expected dose $\mathbb{E}[D]x$ without requiring storage of scenarios during optimization. As $\mathbb{E}[D]$ and D have similar structure and Ω is a symmetric square matrix with number of rows/columns equal to the number of beamlets, optimization performance exhibits minimal overhead and is independent to the number of error scenarios used in the precomputation of Ω_V's and $\mathbb{E}[D]x$.

2.3 Multicriteria formulation and optimization

In general, an MCO problem is of the form

$$\text{"min"}_{x\in\mathscr{X}} \left[f_1(x), \ldots, f_k(x)\right] \quad (6)$$

where f_1 to f_k are objective functions for the treatment planning problem which should ideally all be minimized simultaneously by x. Hard constraints that must all hold (e.g. non-negativity of x) define the set $\mathscr{X}$. It is typically true that there exists no $x \in \mathscr{X}$ such that all objective function attain their individual minimum. Interesting solutions to (6) have the property that if one of the individual objectives should be further improved, at least one other objectives would necessarily have to degrade. This is often referred to as "efficiency" or "Pareto-efficiency" (after Vilfredo Pareto, an Italian economist). We label the set of all such efficient treatment plans the *Pareto front* (in some publications, this set is also referred to as the *Pareto surface*).
We formulate the MCO problem in this work by defining the following objectives for the planning structures. We distinguish between risk structures for which we want to analyze the trade-offs against all other objectives, and those for which we just want the dose to be within certain bounds.

- for a target structure $\mathscr{T}$, we define
  1. an objective $f^{lsq}_{\mathscr{T}}(\mathbb{E}[d_{\mathscr{T}}(x)])$ as defined above to minimize the squared deviation of expected dose from a reference dose (normalized by $\frac{1}{|\mathscr{T}|}$), and
  2. the total variance objective $f^{\Omega}_{\mathscr{T}}(x) := \frac{1}{|\mathscr{T}|}x^T\Omega_{\mathscr{T}}x$
- for risk structures $\mathscr{R}$ which we want to include in a MCO trade-off analysis, we define
  1. the 2-EUD as objective: $f^{EUD}_{\mathscr{R}}(\mathbb{E}[d(x)]) = \frac{1}{|\mathscr{R}|}\mathbb{E}[d]^T_{\mathscr{R}}\mathbb{E}[d]_{\mathscr{R}}$, and
  2. the total variance objective $f^{\Omega}_{\mathscr{R}}(x) := \frac{1}{|\mathscr{R}|}x^T\Omega_{\mathscr{R}}x$

Aside from the physical requirement $x \geq 0$, $\mathscr{X}$ also contains hard constraints on the over- and underdosing of target structures. We found that the constraints $g^{UD}_{\mathscr{T}}(\mathbb{E}[d_{\mathscr{T}}(x)]) := \frac{1}{|\mathscr{T}|}(d^{ref}_{\mathscr{T}} - \mathbb{E}[d_{\mathscr{T}}(x)])_+ \leq 1$ and $g^{OD}_{\mathscr{T}}(\mathbb{E}[d_{\mathscr{T}}(x)]) := \frac{1}{|\mathscr{T}|}(\mathbb{E}[d_{\mathscr{T}}(x)] - d^{ref}_{\mathscr{T}})_+ \leq 0.5$ work well for each involved target T. Here, $(\cdot)_+ := max(0,\cdot)$. Finally, for each risk structure not included in the MCO analysis, we add an overdose constraint $g^{OD}_{\mathscr{R}}(\mathbb{E}[d_{\mathscr{R}}(x)]) \leq 2$ with d^{ref} at half of the main target prescription dose.
Since all our objectives and constraints are convex functions, the Pareto front will be the surface of a k-dimensional convex body and we may directly apply well established methods *sandwiching methods* [5–8] to approximate this set. These methods iteratively produce efficient treatment plans and stop when the approximation quality is good enough. Details can be found in the references given above.

2.4 Implementation and Software

Computation of $\mathbb{E}[D]$ and Ω as well as the robustness analyses were done in matRad [9]. We use an in-house sandwiching method implemented by the Fraunhofer ITWM to approximate the Pareto fronts. Objectives are scaled to the [0,1] interval to prevent numerical distortions due to differing value domains. We set the termination criterion to stop the approximation when the ε-indicator [5–8] for the scaled objectives is below 0.05. For each case, we state the number of computed solutions to reach the required quality. This number corresponds to the number of plan optimizations. We use Ipopt [10] and KNITRO [11] to do the numerical optimization of the resulting scalar optimizations.

3 Results

3.1 TG-119 C-Shape

For a first low-dimensional demonstration, we used the TG119 phantom (directly available in matRad), a single proton field was set to cover the C-shaped Target and expose the distal Core structure to range uncertainties.
Setup uncertainties of $\pm 2.25\,\text{mm}$ and range uncertainties of $\pm 3.5\,\% \pm 1\,\text{mm}$ were assumed as standard deviation to sample 10 error scenarios from the respective multivariate normal distribution. These 10 scenarios were used to estimate $\mathbb{E}[D]$ and Ω before optimization.
Objectives and constraints were chosen according to section 2.3, resulting in a 5D MCO problem requiring 40 solutions for an acceptable Pareto surface approximation. For all solutions, robustness analysis was performed with DVHs and Standard-deviation-volume-histograms (SDVHs). SDVHs relate dose standard deviation to percentage volume for a given VOI to quantify the VOI-specific robustness. We illustrate the observed trade-off within the Core OAR in Figure 1.

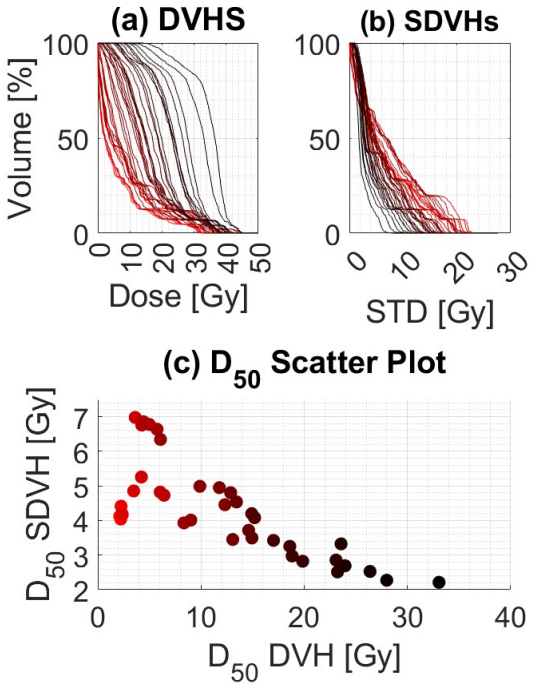

Figure 1: DVHs (a) and SDVHs (b) for all 40 solutions in the Core. (c) Scatter plot relates median dose and standard deviation (D_{50}). Similar colors correspond to the same solution in (a-c).

Figure 1 shows that the exposition of the Core to range uncertainties creates a direct trade-off between the Core robustness (total variance objective) and the dosage (EUD objective). With MCO, this trade-off can now be directly explored with Pareto surface navigation.

3.2 Liver case

To demonstrate our approach on a high-dimensional patient case, we planned a two-field proton plan for the Liver case available in matRad. One of the fields was chosen to expose the heart to range uncertainties. Similar uncertainty assumptions as in section 3.1 were applied using the CTV as target and liver and heart as OAR. A larger set of 100 scenarios was sampled to construct $\mathbb{E}[D]$ and Ω to demonstrate that our approach works with commonly unsuitably large numbers of scenarios. For optimization, again the protocol defined in section 2.3 was applied, resulting in a 7D MCO problem requiring 60 Pareto-optimal solutions. For all solutions, robustness analysis was again performed reloading the 100 scenarios initially used to construct $\mathbb{E}[D]$ and Ω.
Figure 2 illustrates the 7D MCO solution with navigation interface allowing trading robustness and dosimetric obejctives.

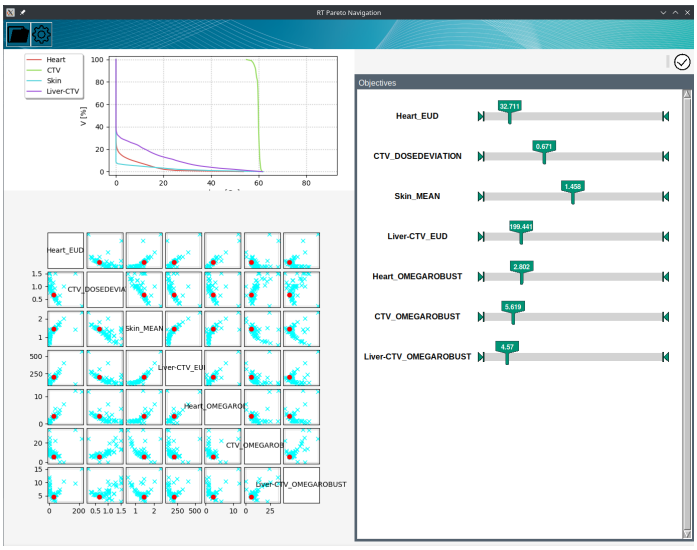

Figure 2: Pareto navigation for the liver case. Each slider on the right corresponds to an objective function. To improve one objective, the slider has to be pulled to the left. The bottom left is a scatterplot matrix of all pairs of objectives, showing the 2D projection of the Pareto front (in turquoise) as well as the currently navigated point (red). All elements then update in real time.

Figure 3 shows two exemplary solutions illustrating how heart robustness can be traded against CTV robustness and coverage; improved heart robustness will downregulate the proton beam exposing the heart to range uncertainties.
Figure 4 adds DVHs for the CTV and SDVHs for the Heart to the analysis and illustrates the correlation of target coverage with heart robustness, indicating a clear trade-off between the two that can be explored by Pareto navigation.

4 Discussion

One of the main clinical benefits of the presented methods is difficult to convey on paper: robustness can now be easily included in the interactive planning process. A common stage in modern multicriteria radiotherapy treatment planning is the interactive *Pareto Navigation* [12, 13] as implemented in Varian's Eclipse Treatment System and in RaySearch Laboratories' RayStation. This decision support conveys the shape

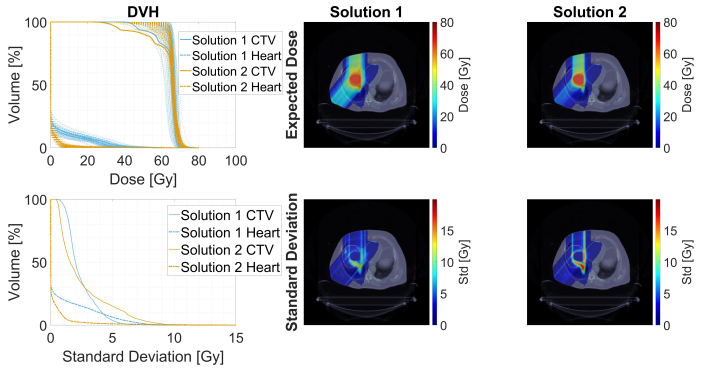

Figure 3: On the left, DVHs and SDVH for the CTV and Heart structures. Histograms are reported for all the 100 error scenarios nad for two solutions on the Pareto-front. On the right, expected dose and SD distributions for the two solutions are compared.

of a high-dimensional Pareto front and its implied trade-offs between the objectives to the decision-maker. It lets the planner explore the chances and limitations of the treatment by interactively manipulating the plan quality.
By adding the total variance objectives to the multicriteria formulation, the planner can interactively trade off the expected plan quality against the robustness on a structure level. We hope that this interactive analysis will improve the planner's and treating physician's confidence to apply robustness to their treatment plans.

5 Conclusion

We successfully introduced robustness as an objective into MCO for proton therapy. The approach limits computational overhead by precomputing expected dose and total variance influence before optimization, avoiding the need for explicit scenario evaluation. Trade-offs between robustness and dose metrics could be effectively calculated and visualized using Pareto surface approximation techniques.

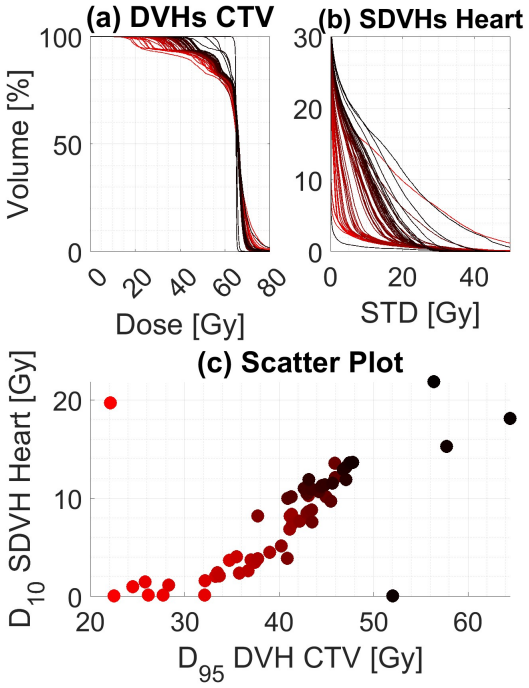

Figure 4: (a) CTV's DVHs and (b) Heart's SDVHs for all the obtained solution. (c) Scatter plot of the CTV's DVHs D_{95} and the Heart's SDVHs D_{10} points.

References

[1] W. Chen, J. Unkelbach, A. Trofimov, et al. "Including robustness in multi-criteria optimization for intensity-modulated proton therapy." *Physics in Medicine & Biology* 57.3 (2012), p. 591. DOI: 10.1088/0031-9155/57/3/591.

[2] J. Unkelbach, M. Alber, M. Bangert, et al. "Robust radiotherapy planning." *Physics in Medicine & Biology* 63.22 (2018), 22TR02. DOI: 10.1088/1361-6560/aae659.

[3] A. Fredriksson. "A characterization of robust radiation therapy treatment planning methods—from expected value to worst case optimization." *Medical Physics* 39.8 (2012), pp. 5169–5181. DOI: https://doi.org/10.1118/1.4737113.

[4] M. Bangert, P. Hennig, and U. Oelfke. "Analytical probabilistic modeling for radiation therapy treatment planning." *Physics in Medicine & Biology* 58.16 (2013), p. 5401. DOI: 10.1088/0031-9155/58/16/5401.

[5] K. Klamroth, J. Tind, and M. M. Wiecek. "Unbiased approximation in multicriteria optimization." *Mathematical Methods of Operations Research* 56 (2003), pp. 413–437. DOI: 10.1007/s001860200217.

[6] J. I. S. Hernandez. "Multi-objective optimization in Mixed Integer Problems with application to the Beam Selection Optimization Problem in IMRT." PhD thesis. Technical University of Kaiserslautern, 2011.

[7] R. Bokrantz and A. Forsgren. "An algorithm for approximating convex pareto surfaces based on dual techniques." *INFORMS Journal on Computing* 25 (2013), pp. 377–393. DOI: 10.1287/ijoc.1120.0508.

[8] I. Lammel. "Approximation of Pareto Surfaces in Multi-Criteria Optimization." PhD thesis. RPTU Kaiserslautern, 2023.

[9] H.-P. Wieser, E. Cisternas, N. Wahl, et al. "Development of the Open-Source Dose Calculation and Optimization Toolkit matRad." *Medical Physics* 44.6 (2017), pp. 2556–2568. DOI: 10.1002/mp.12251.

[10] A. Wächter and L. T. Biegler. "On the Implementation of a Primal-Dual Interior Point Filter Line Search Algorithm for Large-Scale Nonlinear Programming." *Mathematical Programming* 106 (2006), pp. 25–57. DOI: 10.1007/s10107-004-0559-y.

[11] R. H. Byrd, J. Nocedal, and R. Waltz. "KNITRO: An integrated package for nonlinear optimization." *Large-Scale Nonlinear Optimization*. Springer, 2006, pp. 35–59. DOI: 10.1007/0-387-30065-1_4.

[12] M. Monz, K.-H. Küfer, T. R. Bortfeld, et al. "Pareto navigation—algorithmic foundation of interactive multi-criteria IMRT planning." *Physics in Medicine and Biology* 53 (2008), pp. 985–998. DOI: 10.1088/0031-9155/53/4/011.

[13] C. Collicott, E. Bonacker, I. Lammel, et al. "Interactive navigation of multiple convex patches." *Journal of Multi-Criteria Decision Analysis* 28.5-6 (2021), pp. 311–321. DOI: 10.1002/mcda.1768.