This document describes the TROTS data format. The data is stored in Matlab files, version 7.3. The files are fully HDF5 compliant, and can therefore be read using general HDF5 tools.

Each file contains 2 structures: `problem`, defining the optimisation problem, and `data`, containing the numerical data matrices. For further details, we refer to the publication *An interior-point implementation developed and tuned for radiation therapy treatment planning* by Breedveld, Van den Berg and Heijmen.

**The problem structure**

The problem structure is a list, where each entry defines an objective or constraint. Each entry has the following fields:

- **dataID** Reference index to the `data` entry, containing the respective numerical data.

- **Name** A name that refers to the origin of this constraint/objective. Is ignored by the solver.

- **Minimise True** for an objective means to be minimised, for a constraint this is handled as a maximum constraint. If False, vice versa.

- **Type** Indicates the hard-coded cost-function (see below).

- **Objective** For a constraint, this is the value the cost-function is constrained to. For an objective, this value is ignored.
- **Weight** Scalar to apply to the *objective*, useful to scalarise and weigh multiple objectives.

- **Parameters** Sets parameters to configure the *cost-function*, given in **Type** (see below).

- **Active** Can be **True** or **False** to enable or disable this entry.

- **IsConstraint** If **True**, this entry is a *constraint*, and an *objective* otherwise.

- **Chain** Extra information for *chain* function type (see below).

### The data structure

The data structure contains 2 entries: **matrix**, containing the numerical data, and **misc**, containing auxiliary data to configure the problem. The **matrix** structure has the following fields:

- **A** The data **matrix**. Each matrix in this structure has an equal number of columns, equal to the number of decision variables. In radiation therapy, this matrix is generally the *pencil-beam* matrix. The number of rows typically indicate the number of voxels.

- **b** Offset vector, is 0 unless you are doing something exciting.

- **c** A **scalar** for quadratic matrices, empty otherwise.

- **Type** Indicating the matrix type. When **Type=0**, this is a "normal" matrix operating in the *fluence-to-dose* domain, where the argument $d$ for the cost-functions is computed as $d = Ax + b$. When **Type=1**, this matrix operates in the *fluence* domain only (i.e. linear smoothing of the fluence map), but is treated equally as **Type=0** by the solver. Finally, **Type=2** indicates a *quadratic* matrix, meaning that when the problem is extended with auxiliary decision variables (e.g. to solve mini-max problems when minimising a pointwise maximum), the padding should be done both in rows and columns.

The **misc** structure has the following fields:

- **size** Indicates the number of decision variables of the problem, equalling the number of columns for all **A** matrices.

- **real** Original problem size. The set $\{x_i : i \in 1, \ldots, real\}$ of the decision variables identifies the part representing the actual radiation therapy problem, i.e. the problem before the introduction of auxiliary decision variables (i.e. mini-max problems). This is required to be able to retrieve the original problem, especially for warmstarting the optimisation.

- **WarmstartMatrix** List of indices to the **matrix** structure required for warmstarting the problem. In general, these matrices represent the tumour matrices.
• **WarmstartReferences** Right hand side for the least-squares warmstarting.

• **WarmstartRegularisation** Reference to a matrix structure that regulates the least-squares problem.

See the paper for detailed information on warmstarting.

**The cost-functions**

The cost-functions are referenced by type index. Except for the quadratic type, we use $d(x) = Ax + b$ as argument for all functions, and $m$ is the number of voxels/length of $d$.

1. **linear** type, computes $f(x) = \min(d(x))$ or $f(x) = \max(d(x))$

2. **quadratic** type, computes $f(x) = \frac{1}{2}x^T Ax + b^T x + c$

3. **gEUD** type, the generalised mean, or in radiation therapy better known as the *generalised equivalent uniform dose*, computed as $f(x) = \left(\frac{1}{m} \sum_{i=1}^{m} d_i(x)^a\right)^{\frac{1}{a}}$, where $a$ is given as Parameters. For $a \geq 1$, this function is convex and can be minimised, for $a \leq -1$, this function is concave and needs to be maximised.

4. **LTCP** type, the *logarithmic tumour control probability*, exclusively used for the tumour. $f(x) = \frac{1}{m} \sum_{i=1}^{m} e^{-\alpha(d_i(x) - d_p)}$ where $d_p$ is the prescribed dose, given as first in Parameters, and $\alpha$ the cell-sensitivity, given as the second in Parameters.

5. **DVH** type, the *dose-volume histogram* or *partial-volume* cost-function. In its exact form, $f(x) = \frac{1}{m} \sum_{i=1}^{m} I_{d_i(x) < d_c}(d_i(x))$ where $d_c$ is the critical dose level (given as first Parameter) and $I$ the indicator function. This simply results in the fraction of $d$ larger than $d_c$. For its smoothed approximation, $f(x) = \frac{1}{m} \sum_{i=1}^{m} \left(\frac{d_i(x)}{d_c}\right)^p$, where the parameter $p$ is the steepness, given as second in Parameters.

6. **Chain** type, describing functions of the form $a \cdot g(x) \leq x_i$, where $a$ is a scalar and $x_i$ denotes element $i$ in decision vector $x$. This formulation is required for the LRPM representation of problems, and this special function type simplifies configuration of such constraints, without requiring the addition of many separate mini-max constraints, and duplicating the data matrices because of it. The Chain field consists of 3 elements: the first is an index to the problem structure, identifying the target cost-function and respective parameters for $g(x)$. The second element sets $a$ and third element sets the index $i$. 
