Optimization of non-coplanar beam arrangements:
Expanding our delivery space

Daniel O'Connor

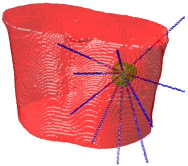
History of sparsity

Beam angle selection via group sparsity

Proximal algorithms and FISTA

Fraction-variant beam orientation optimization

Simultaneous beam angle and spot selection for IMPT



History of sparsity in signal processing

Classic problem: sparse approximation

- "Overcomplete basis": $a_1, \ldots, a_n \in \mathbb{R}^m, \quad n \gg m$
- Goal: Represent a vector $b \in \mathbb{R}^m$ as a linear combination of vectors a_i

$$b \approx \sum_i x_i a_i$$

  with most coefficients x_i equal to 0
- In matrix notation, we want

$$b \approx Ax$$

Key point: We want the vector x to be sparse

History of sparsity: Orthogonal matching pursuit

Goal:

- $F(x) = \frac{1}{2}\|Ax - b\|_2^2$ should be small
- Most components of x should be 0

Early attempt: Orthogonal matching pursuit

Strategy: Activate components of x one by one (greedy algorithm)

Repeat:

1. Select component i such that

$$\frac{\partial F}{\partial x_i} = a_i^T \underbrace{(Ax - b)}_{\text{residual}}$$

   has largest magnitude

   (In other words, select a_i that is most correlated with residual $Ax - b$)

2. Minimize $\frac{1}{2}\|Ax - b\|_2^2$ allowing only selected components to be active

Many popular beam angle selection algorithms are similar in spirit to OMP

History of sparsity: Lasso

Goal:

- $F(x) = \frac{1}{2}\|Ax - b\|_2^2$ should be small
- Most components of x should be 0

A modern approach: Lasso

Strategy: Solve optimization problem with sparsity-inducing regularizer:

$$\underset{x}{\text{minimize}} \quad \frac{1}{2}\|Ax - b\|_2^2 + \underbrace{\lambda\|x\|_1^p}_{\text{promotes sparsity}}$$

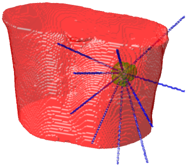
- $0 < p \leq 1$ (Lasso takes $p = 1$)
- Strategy was enabled by new optimization algorithms

Similarly, new algorithms enable the group sparsity approach to BOO

Reference: "Basis pursuit", Chen and Donoho, 1994

History of sparsity

Beam angle selection via group sparsity



Proximal algorithms and FISTA

Fraction-variant beam orientation optimization

Simultaneous beam angle and spot selection for IMPT

Beam angle selection via group sparsity

Strategy:

Solve a fluence map optimization problem with an additional **group sparsity** term that encourages most candidate beams to be inactive

$$\begin{array}{ll}\underset{x}{\text{minimize}} & \underbrace{\frac{1}{2}\|(\ell - A_0 x)_+\|_2^2}_{\text{PTV}} + \underbrace{\sum_{i=0}^{N} \frac{\alpha_i}{2}\|(A_i x - d_i)_+\|_2^2}_{\text{OARs}} + \underbrace{\sum_{b=1}^{B} w_b \|x_b\|_2^p}_{\text{group sparsity}} \\ \text{subject to} & x \geq 0\end{array}$$

- x_b is vector of beamlet intensities for bth candidate beam
- x is the concatenation of the vectors x_b
- A_0 is dose-calculation matrix for PTV
- $A_i, i = 1, \ldots, N$, is dose-calculation matrix for ith OAR

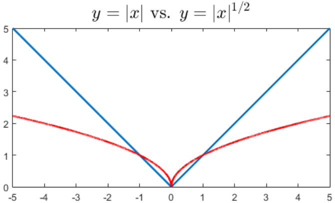
Prior work: "Beam orientation optimization for intensity modulated radiation therapy using adaptive $\ell_{2,1}$-minimization", X. Jia, C. Men, Y. Lou, and S. Jiang, 2011

Choosing the group sparsity penalty function

Previous work takes $p = 1$, solves convex problem:

- "Beam orientation optimization for intensity modulated radiation therapy using adaptive $\ell_{2,1}$-minimization", X. Jia, C. Men, Y. Lou, and S. Jiang, 2011
- "4π non-coplanar IMRT beam angle selection by convex optimization with group sparsity penalty", D. O'Connor, Y. Voronenko, D. Nguyen, W. Yin, K. Sheng, AAPM 2016

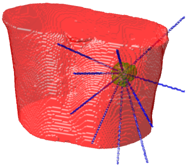
Problem: Compressed sensing informs us that $p = 1$ is only a guaranteed to be a good choice when "group restricted isometry property" is satisfied



Solution: Use a non-convex group sparsity term (we take $p = 1/2$)

History of sparsity

Beam angle selection via group sparsity



Proximal algorithms and FISTA

Fraction-variant beam orientation optimization

Simultaneous beam angle and spot selection for IMPT

Proximal algorithms

- **Proximal operator**

$$\text{prox}_{tg}(x) = \arg\min_u \quad g(u) + \frac{1}{2t}\|u - x\|_2^2$$

- **Proximal gradient method:** Minimizes

$$\underbrace{f(x)}_{\text{differentiable}} + \underbrace{g(x)}_{\text{easy prox-operator}}$$

via

$$x^{k+1} = \text{prox}_{tg}(x^k - t\nabla f(x^k))$$

- **FISTA:** accelerated proximal gradient method

$$y = x^k + \frac{k}{k+3}(x^k - x^{k-1})$$
$$x^{k+1} = \text{prox}_{tg}(y - t\nabla f(y))$$

FISTA with line search

Algorithm 1 FISTA with line search

Initialize x_0 and $t_0 > 0$, set $v_0 := x_0$, select $0 < r < 1$, $s > 1$
for $k = 1, 2, \ldots$ **do**
  $t := s\, t_{k-1}$
  repeat

$$\theta := \begin{cases} 1 & \text{if } k = 1 \\ \text{positive root of } t_{k-1}\theta^2 = t\theta_{k-1}^2(1-\theta) & \text{if } k > 1 \end{cases}$$

    $y := (1-\theta)x_{k-1} + \theta v_{k-1}$
    $x := \mathrm{prox}_{tg}(y - t\nabla f(y))$
    break if $f(x) \le f(y) + \langle \nabla f(y), x - y\rangle + \frac{1}{2t}\|x-y\|_2^2$
    $t := rt$
  $t_k := t$
  $\theta_k := \theta$
  $x_k := x$
  $v_k := x_{k-1} + \frac{1}{\theta_k}(x_k - x_{k-1})$
end for

References:

- "Fast first-order methods for composite convex optimization with line search", Goldfarb and Scheinberg, 2011
- Lieven Vandenberghe's UCLA 236C notes

Beam angle selection using FISTA

$$\underset{x}{\text{minimize}} \quad \frac{1}{2}\|(\ell - A_0 x)_+\|_2^2 + \sum_{i=0}^{N} \frac{\alpha_i}{2}\|(A_i x - d_i)_+\|_2^2 + \sum_{b=1}^{B} w_b \|x_b\|_2^{1/2} + I_{\geq 0}(x)$$

- Indicator penalty enforces nonnegativity:

$$I_{\geq 0}(x) = \begin{cases} 0 & \text{if } x \geq 0 \\ \infty & \text{otherwise} \end{cases}$$

Beam angle selection using FISTA

$$\underset{x}{\text{minimize}} \quad \underbrace{\frac{1}{2}\|(\ell - A_0 x)_+\|_2^2 + \sum_{i=0}^{N} \frac{\alpha_i}{2}\|(A_i x - d_i)_+\|_2^2}_{f(x)} + \underbrace{\sum_{b=1}^{B} w_b \|x_b\|_2^{1/2} + I_{\geq 0}(x)}_{g(x)}$$

$\uparrow$
includes both group sparsity term
and indicator function

- Indicator penalty enforces nonnegativity:

$$I_{\geq 0}(x) = \begin{cases} 0 & \text{if } x \geq 0 \\ \infty & \text{otherwise} \end{cases}$$

Beam angle selection using FISTA

$$\underset{x}{\text{minimize}} \quad \underbrace{\frac{1}{2}\|(\ell - A_0x)_+\|_2^2 + \sum_{i=0}^{N} \frac{\alpha_i}{2}\|(A_ix - d_i)_+\|_2^2}_{f(x)} + \underbrace{\sum_{b=1}^{B} w_b\|x_b\|_2^{1/2} + I_{\geq 0}(x)}_{g(x)}$$

$\uparrow$
includes both group sparsity term
and indicator function

- Indicator penalty enforces nonnegativity:

$$I_{\geq 0}(x) = \begin{cases} 0 & \text{if } x \geq 0 \\ \infty & \text{otherwise} \end{cases}$$

- Gradient of f:

$$\nabla f(x) = -A_0^T(\ell - A_0x)_+ + \sum_{i=0}^{N} \alpha_i A_i^T(A_ix - d_i)_+$$

Beam angle selection using FISTA

$$\underset{x}{\text{minimize}} \quad \underbrace{\frac{1}{2}\|(\ell - A_0 x)_+\|_2^2 + \sum_{i=0}^{N} \frac{\alpha_i}{2}\|(A_i x - d_i)_+\|_2^2}_{f(x)} + \underbrace{\sum_{b=1}^{B} w_b \|x_b\|_2^{1/2} + I_{\geq 0}(x)}_{g(x)}$$

$\uparrow$
includes both group sparsity term
and indicator function

- Indicator penalty enforces nonnegativity:

$$I_{\geq 0}(x) = \begin{cases} 0 & \text{if } x \geq 0 \\ \infty & \text{otherwise} \end{cases}$$

- Gradient of f:

$$\nabla f(x) = -A_0^T(\ell - A_0 x)_+ + \sum_{i=0}^{N} \alpha_i A_i^T (A_i x - d_i)_+$$

- Prox-operator of g:

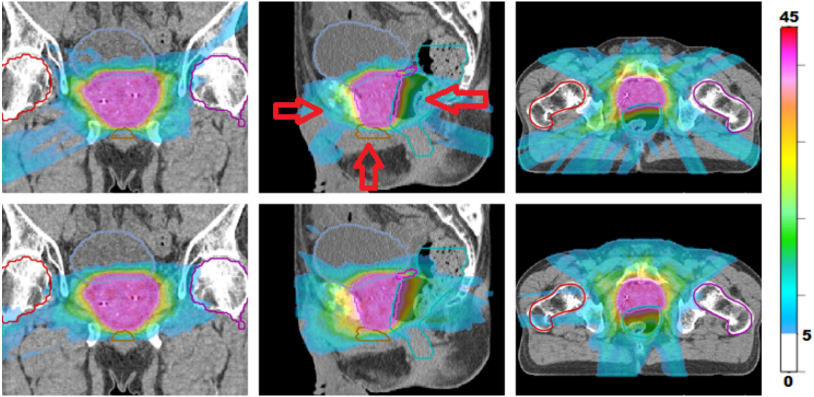
$$p_b = \text{prox}_{t w_b \|\cdot\|_2^{1/2}}(\max(x_b, 0)) = s^2 \max(x_b, 0)$$

where

$$\alpha = t w_b / \|max(x_b, 0)\|_2^{3/2}, \quad s = \begin{cases} \frac{2}{\sqrt{3}} \sin(\frac{1}{3}\arccos(\frac{3\sqrt{3}}{4}\alpha) + \frac{\pi}{2}) & \text{if } \alpha \leq \frac{2\sqrt{6}}{9} \\ 0 & \text{otherwise} \end{cases}$$

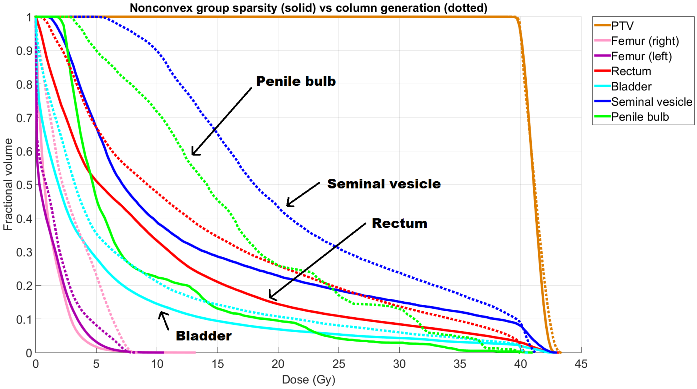
Results

Group sparsity (top row) vs. column generation (bottom row)



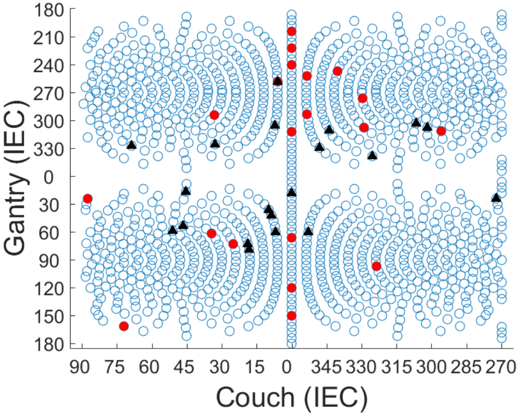
- 674 non-coplanar candidate beams
- Runtime $\approx$ 6 minutes

Results



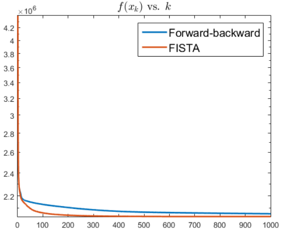
Group sparsity (solid) vs. column generation (dotted)

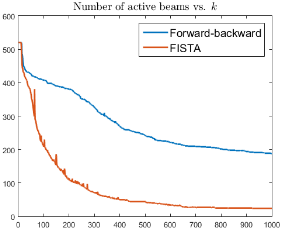
Results



Beams selected using convex (black) and non-convex (red) group sparsity terms

Proximal gradient method vs. FISTA

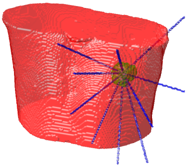




FISTA convergence rate is $O(1/k^2)$ rather than $O(1/k)$

Tricks to improve runtime:

- Prune beams as FISTA iteration progresses
- Downsample voxel grid to reduce size of matrices A_i
- Matlab trick: `A * x` $=$ `ATrans' * x`



History of sparsity

Beam angle selection via group sparsity

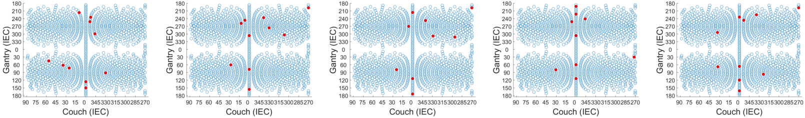
Proximal algorithms and FISTA

Fraction-variant beam orientation optimization

Simultaneous beam angle and spot selection for IMPT

Fraction-variant beam angle selection

- Standard algorithms assume same beam angles used each treatment day
- But this seems to be an unnecessary constraint
- We should simultaneously optimize beam orientations for all fractions



Potential benefits:

- Better dosimetry
- Fewer beams per fraction $\rightarrow$ faster treatment times

Optimization formulation

$$\underset{x}{\text{minimize}} \quad \underbrace{\sum_{f=1}^{F} \frac{1}{2}\|A_0 x_f - d_0/F\|_2^2}_{\text{PTV}} + \underbrace{\sum_{i=1}^{N} \frac{\alpha_i}{2}\|(\bar{A}_i x - d_i)_+\|_2^2}_{\text{OARs}} + \underbrace{\sum_{f=1}^{F}\sum_{b=1}^{B} w_b \|x_{f,b}\|_2^p}_{\text{group sparsity}}$$

$$\text{subject to} \quad x \geq 0,$$

- F is the total number of treatment fractions
- B is the number of candidate beams *per fraction*
- A_i is dose-calculation matrix for PTV ($i = 0$) or ith OAR ($i = 1, \ldots, N$)
- $x_{f,b}$ is the fluence map for the bth candidate beam on treatment day f.
-

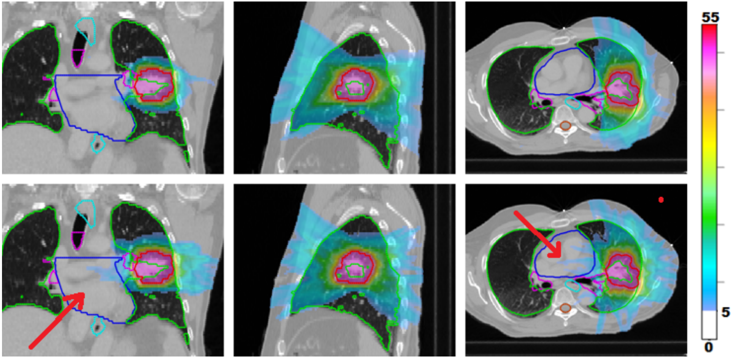
$$x_f = \begin{bmatrix} x_{f,1} \\ x_{f,2} \\ \vdots \\ x_{f,B} \end{bmatrix}, \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_F \end{bmatrix}, \qquad \bar{A}_i = \underbrace{\begin{bmatrix} A_i & A_i & \cdots & A_i \end{bmatrix}}_{F \text{ copies of } A_i}$$

- $0 < p < 1$ (we take $p = 1/2$)

Solve using accelerated proximal gradient method (FISTA)

Lung case

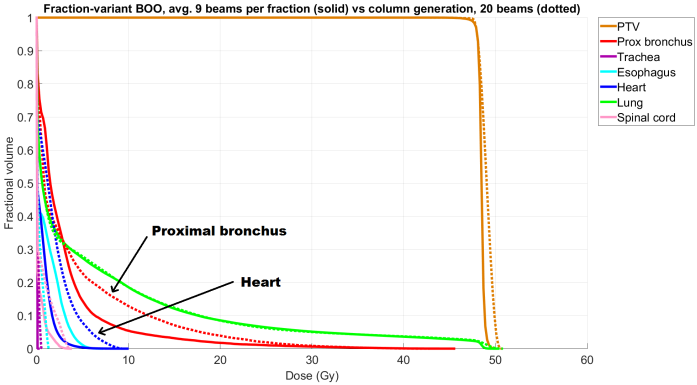
Fraction-variant plan (top row) vs. column generation plan (bottom row)



- Five fractions total, 9 beams per fraction on average
- Column generation plan used 20 beams (same beams each fraction)
- 520 candidate beams/fraction $\times$ 5 fractions $=$ 2600 candidate beams
- FISTA with line search runtime $\approx$ 65 minutes

Lung case

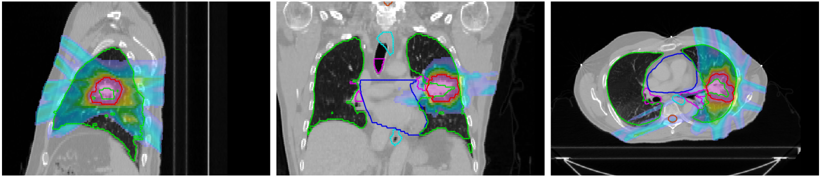
DVH for fraction-variant plan (solid) vs. column generation plan (dotted)



Fraction-variant plan: superior dosimetry, half as many beams per fraction

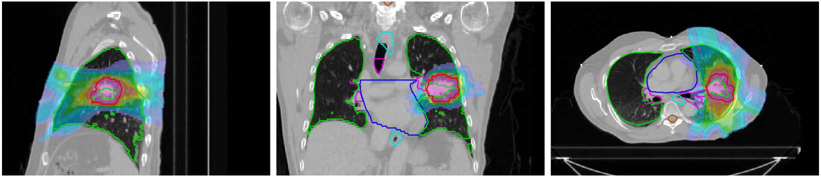
Lung case: Treatment plans for individual fractions

Fraction 1



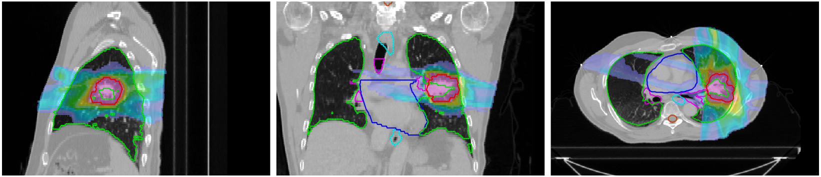
Lung case: Treatment plans for individual fractions

Fraction 2



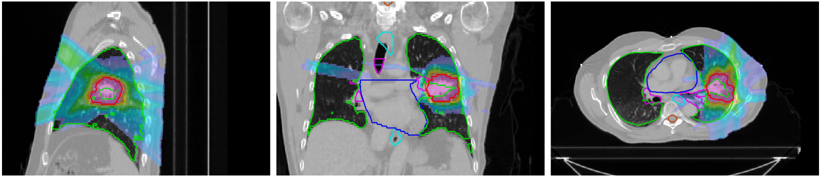
Lung case: Treatment plans for individual fractions

Fraction 3



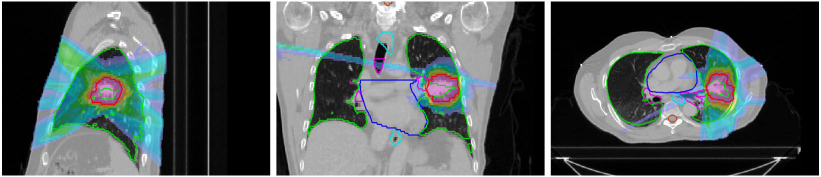
Lung case: Treatment plans for individual fractions

Fraction 4

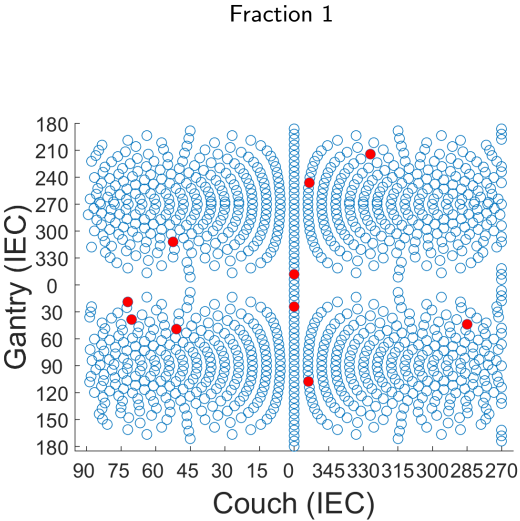


Lung case: Treatment plans for individual fractions

Fraction 5

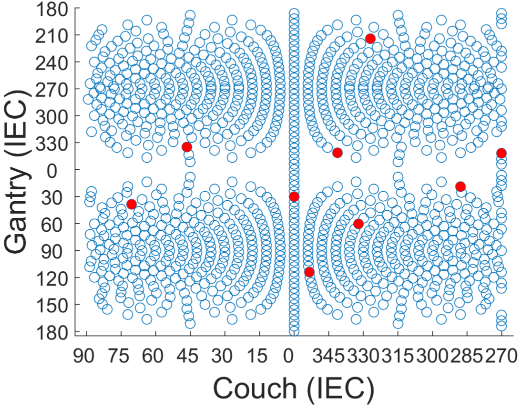


Beams selected for individual fractions

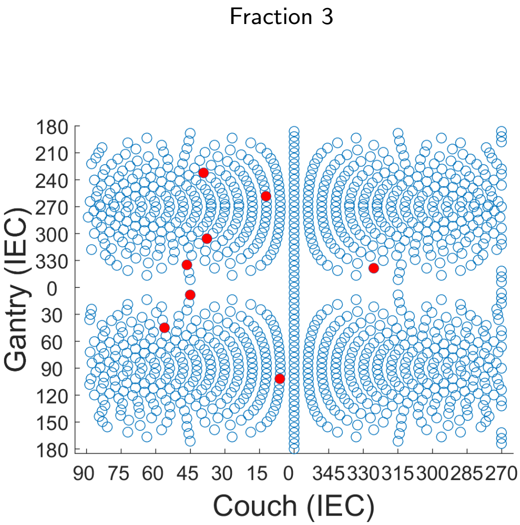


Beams selected for individual fractions

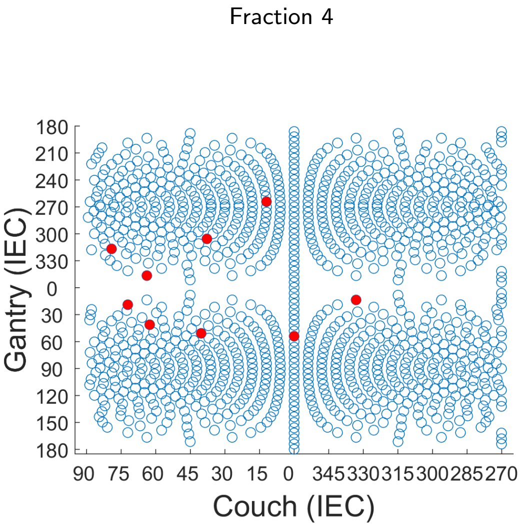
Fraction 2



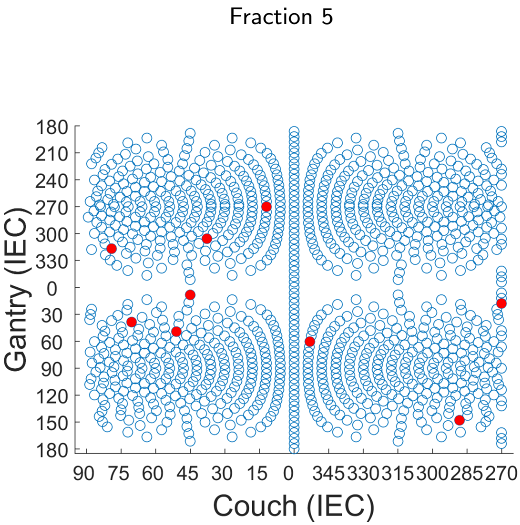
Beams selected for individual fractions



Beams selected for individual fractions

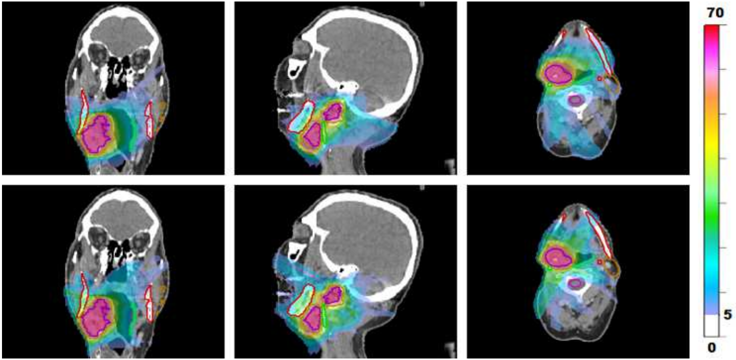


Beams selected for individual fractions



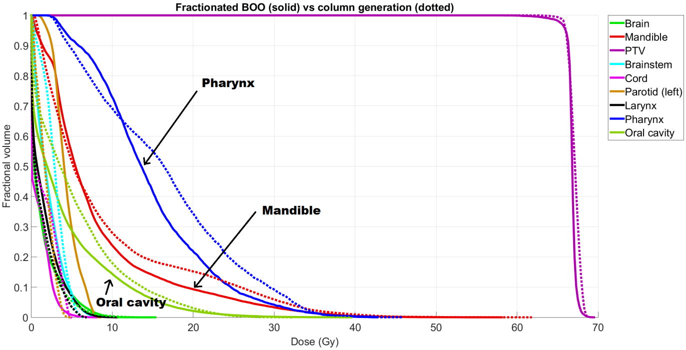
Head and neck case

Fraction-variant plan (top row) vs. column generation plan (bottom row)



- 30 fractions total, 6.5 beams per fraction on average
- Column generation plan used 20 beams (same beams each fraction)
- 811 candidate beams/fraction $\times$ 30 fractions $=$ 24,330 candidate beams
- FISTA with line search runtime $\approx$ 12 hours

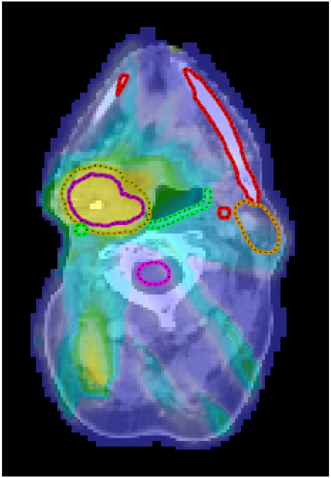
Head and neck case



Fraction-variant plan: comparable dosimetry, 1/3 as many beams per fraction

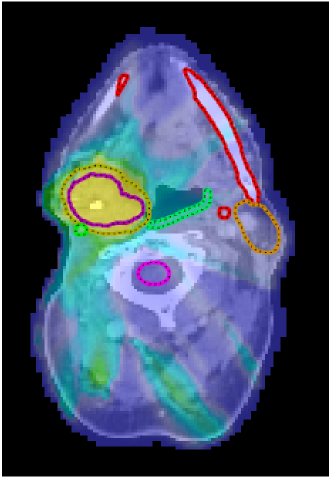
Head and neck case: Dose washes for individual fractions

Fraction 1



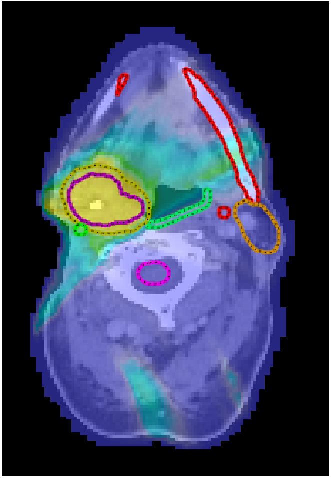
Head and neck case: Dose washes for individual fractions

Fraction 2



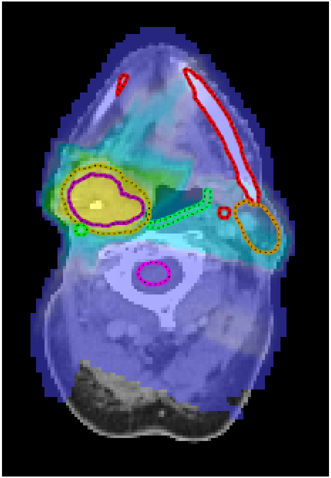
Head and neck case: Dose washes for individual fractions

Fraction 3



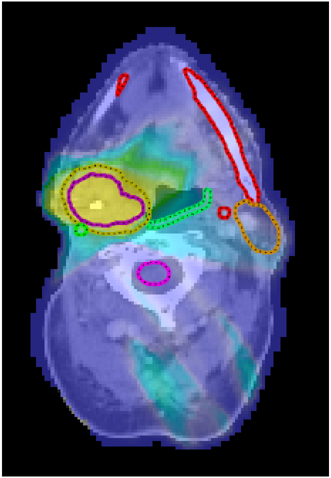
Head and neck case: Dose washes for individual fractions

Fraction 4



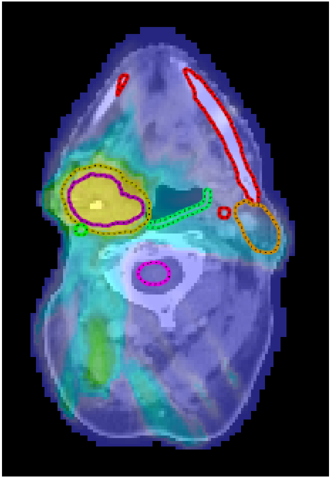
Head and neck case: Dose washes for individual fractions

Fraction 5



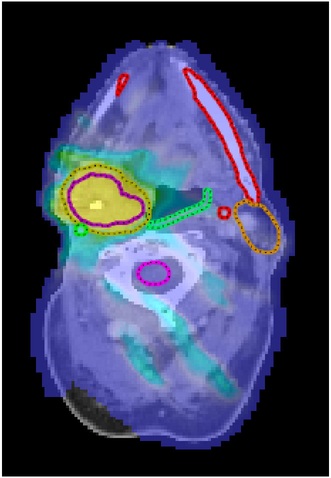
Head and neck case: Dose washes for individual fractions

Fraction 6



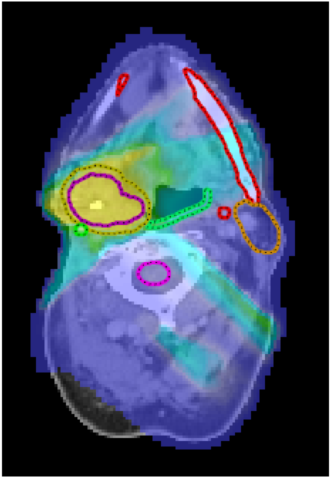
Head and neck case: Dose washes for individual fractions

Fraction 7



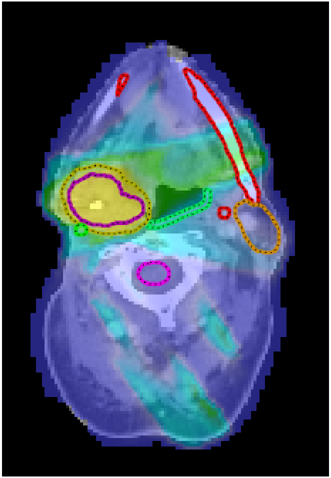
Head and neck case: Dose washes for individual fractions

Fraction 8



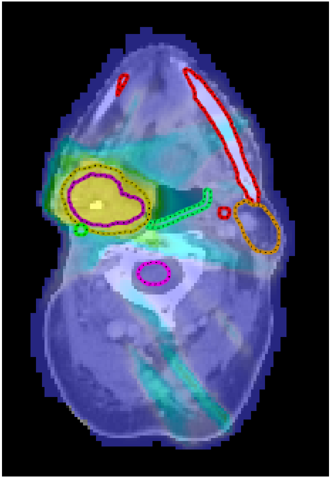
Head and neck case: Dose washes for individual fractions

Fraction 9

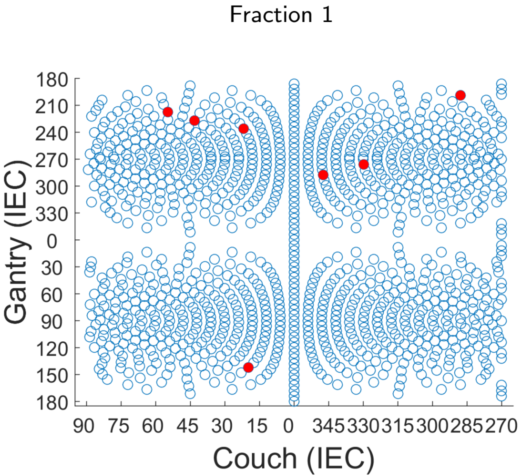


Head and neck case: Dose washes for individual fractions

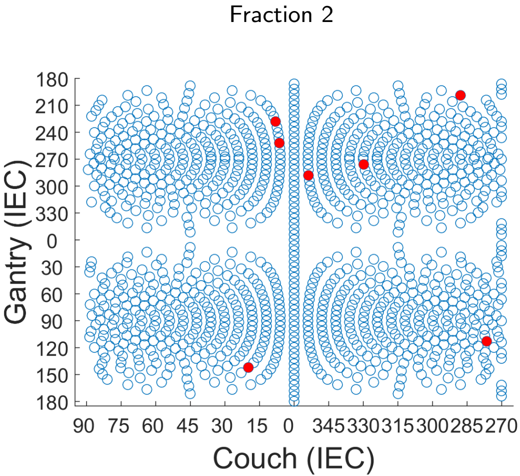
Fraction 10



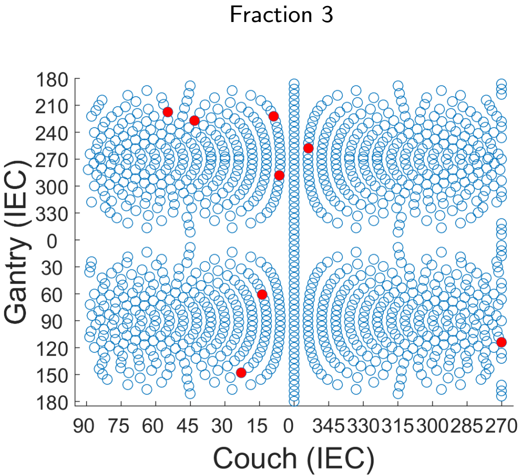
Head and neck case: Beams selected for individual fractions



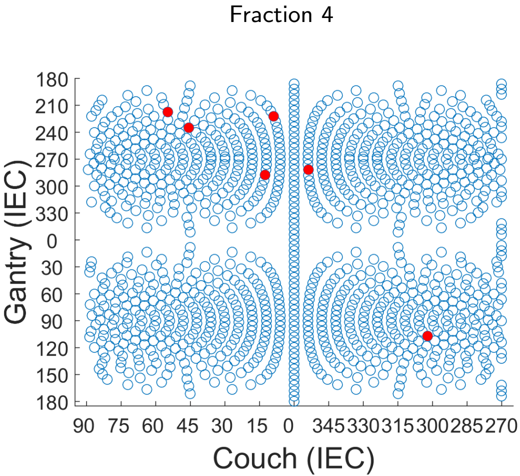
Head and neck case: Beams selected for individual fractions



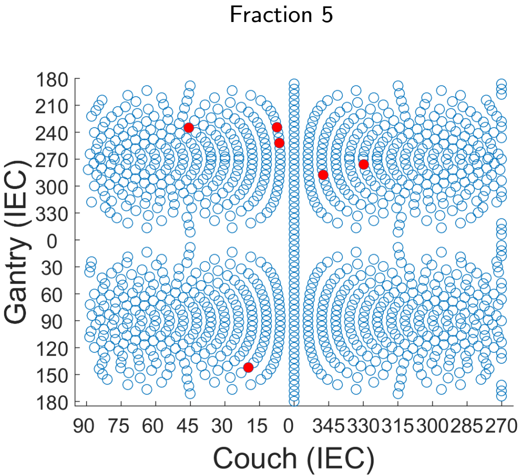
Head and neck case: Beams selected for individual fractions



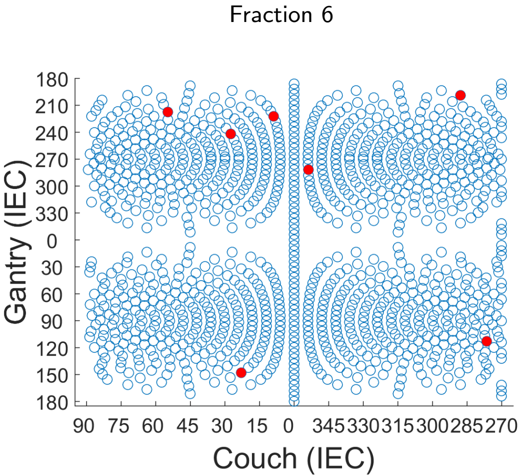
Head and neck case: Beams selected for individual fractions



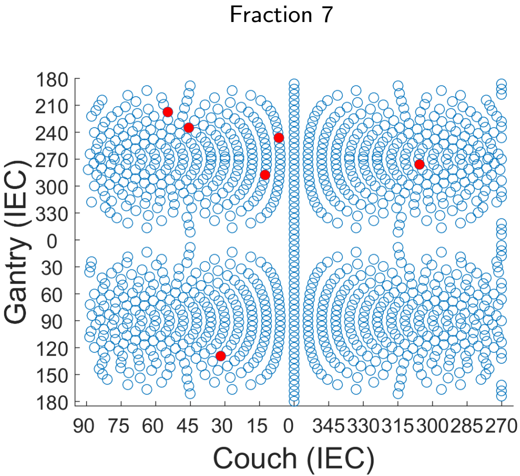
Head and neck case: Beams selected for individual fractions



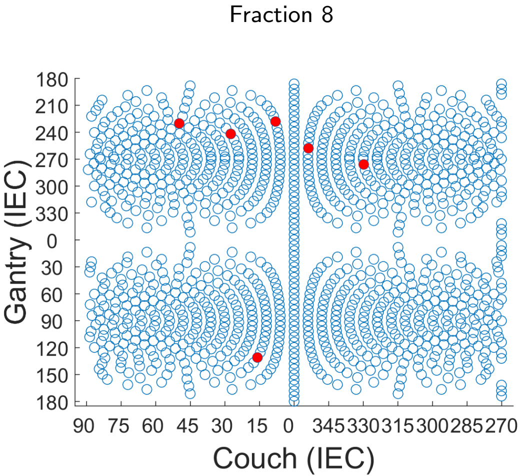
Head and neck case: Beams selected for individual fractions



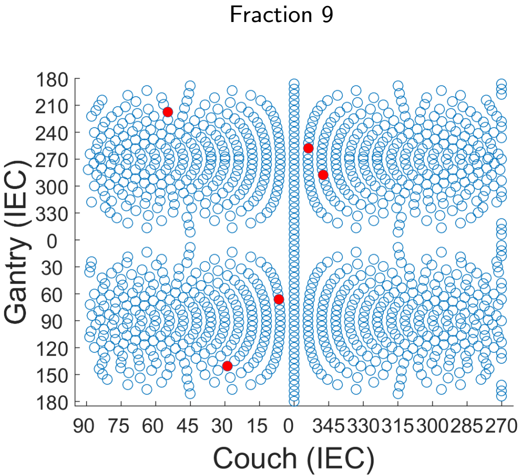
Head and neck case: Beams selected for individual fractions



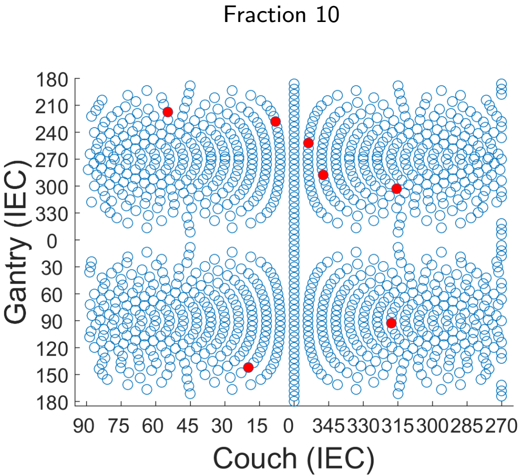
Head and neck case: Beams selected for individual fractions



Head and neck case: Beams selected for individual fractions



Head and neck case: Beams selected for individual fractions

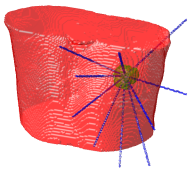


History of sparsity

Beam angle selection via group sparsity

Proximal algorithms and FISTA

Fraction-variant beam orientation optimization



Simultaneous beam angle and spot selection for IMPT

Simultaneous beam angle and spot selection for IMPT

Problem formulation:

$$\underset{x}{\text{minimize}} \quad \underbrace{\frac{1}{2}\|(\ell - A_0 x)_+\|_2^2}_{\text{PTV}} + \underbrace{\sum_{i=0}^{N} \frac{1}{2}\|(A_i x - d_i)_+\|_2^2}_{\text{OARs}} + \underbrace{\lambda \|x\|_1}_{\text{spot sparsity}} + \underbrace{\sum_{b=1}^{B} w_b \|x_b\|_2^p}_{\text{beam sparsity}}$$

$$\text{subject to} \quad x \geq 0.$$

Sparsity terms encourage x to have a hierarchical sparsity pattern:

- Sparse number of beams are active
- For each beam, sparse number of spots are active

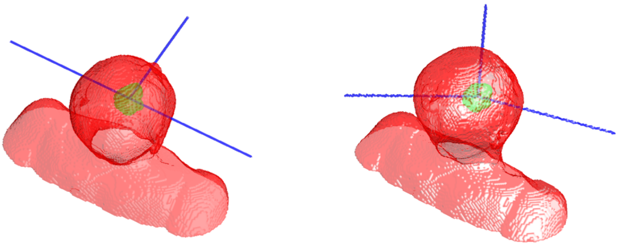
Simultaneous beam angle and spot selection for IMPT

Solution using FISTA:

$$\underset{x}{\text{minimize}} \quad \underbrace{\frac{1}{2}\|(\ell - A_0 x)_+\|_2^2 + \sum_{i=0}^{N} \frac{1}{2}\|(A_i x - d_i)_+\|_2^2}_{f(x)} + \underbrace{\lambda\|x\|_1 + \sum_{b=1}^{B} w_b \|x_b\|_2^p + I_{\geq 0}(x)}_{g(x)}$$

- Compute $\nabla f(x)$ using calculus
- Can derive explicit formula for prox-operator of g

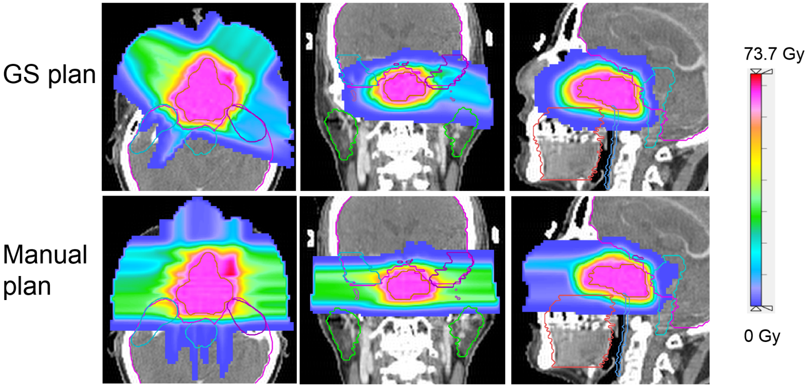
Simultaneous beam angle and spot selection for IMPT: Results



Manually selected beams (left) and group sparsity beams (right)

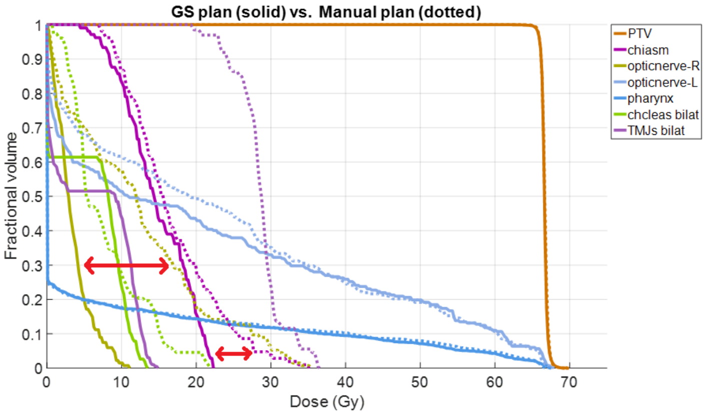
Credit: Wenbo Gu

Simultaneous beam angle and spot selection for IMPT: Results



Credit: Wenbo Gu

Simultaneous beam angle and spot selection for IMPT: Results



Credit: Wenbo Gu

Conclusions

The group sparsity approach is now efficient for non-coplanar BOO

- Improved algorithms (FISTA with line search)
- Improved hardware (more RAM)
- Potential for further improvement using GPU

Group sparsity is a useful tool in radiation treatment planning

- Non-coplanar beam orientation optimization
- Fraction-variant beam orientation optimization
- Simultaneous beam angle and spot selection for IMPT

Disclosure

Research supported by:

- NIH U19AI067769
- NIH R43CA183390
- NIH R44CA183390
- NIH R01CA188300
- DE-SC0017057
- DE-230295
- Varian Medical Systems

Thank you!

daniel.v.oconnor@gmail.com

Sheng Lab

UCLA Health

Acknowledgments

Dan Nguyen

Wenbo Gu

Yevgen Voronenko

Victoria Yu

Dan Ruan

Ke Sheng

Bonus slide: Alternative formulation

$$
\begin{aligned}
\underset{x}{\text{minimize}} \quad & \sum_{f=1}^{F} \left(\underbrace{\frac{1}{2}\|A_0 x_f - d_0/F\|_2^2 + \sum_{i=1}^{N} \frac{\alpha_i}{2}\|(A_i x_f - d_i)_+\|_2^2}_{\text{controls fractional doses to PTV and OARs}} \right) \\
& + \underbrace{\sum_{i=1}^{N} \frac{\beta_i}{2}\|\bar{A}_i x\|_2^2}_{\text{controls total dose to OARs}} + \underbrace{\gamma\|Dx\|_1^{(\mu)}}_{\text{fluence map deliverability}} + \underbrace{\sum_{f=1}^{F}\sum_{b=1}^{B} w_b \|x_{f,b}\|_2^{1/2}}_{\text{group sparsity}} \\
\text{subject to} \quad & x \geq 0.
\end{aligned}
$$

Bonus slide: References

X. Jia, C. Men, Y. Lou, and S. Jiang.
Beam orientation optimization for intensity modulated radiation therapy using adaptive $\ell_{2,1}$-minimization.
Physics in Medicine and Biology, 56(19), 2011.

D. O'Connor, Y. Voronenko, D. Nguyen, W. Yin, and K. Sheng.
4π non-coplanar IMRT beam angle selection by convex optimization with group sparsity penalty.
Medical Physics, 43(6):3895–3895, 2016.

J. Unkelbach.
Non-uniform spatiotemporal fractionation schemes in photon radiotherapy.
In *World Congress on Medical Physics and Biomedical Engineering, June 7-12, 2015, Toronto, Canada*, pages 401–404. Springer, 2015.

J. Unkelbach, M. Bussière, P. Chapman, J. Loeffler, and H. Shih.
Spatiotemporal fractionation schemes for irradiating large cerebral arteriovenous malformations.
International Journal of Radiation Oncology Biology* Physics*, 95(3):1067–1074, 2016.

J. Unkelbach and D. Papp.
The emergence of nonuniform spatiotemporal fractionation schemes within the standard BED model.
Medical Physics, 42(5):2234–2241, 2015.

J. Unkelbach, C. Zeng, and M. Engelsman.
Simultaneous optimization of dose distributions and fractionation schemes in particle radiotherapy.
Medical Physics, 40(9), 2013.

P. Dong, P. Lee, D. Ruan, T. Long, E. Romeijn, Y. Yang, D. Low, P. Kupelian, and K. Sheng.
4π non-coplanar liver SBRT: A novel delivery technique.
International Journal of Radiation Oncology Biology* Physics*, 85(5):1360–1366, 2013.

P. Dong, D. Nguyen, D. Ruan, C. King, T. Long, E. Romeijn, D. A. Low, P. Kupelian, M. Steinberg, Y. Yang, and K. Sheng.
Feasibility of prostate robotic radiation therapy on conventional C-arm linacs.
Practical Radiation Oncology, 4(4):254–260, 2014.