An Integrated Alternating Direction Method of Multipliers for Treatment Planning Optimization

Masoud Zarepisheh, Yinyu Ye, Lei Xing

Stanford University

Department of Radiation Oncology

Introduction

ADMM: Alternating Direction Method of Multipliers

- has been around since 1970s
- has received a lot of attention recently

Distributed optimization and statistical learning via the alternating direction method of multipliers
S Boyd, N Parikh, E Chu, B Peleato, J Eckstein - Foundations and Trends® in Machine Learning, 2011
Cited by 2071 - Related articles - All 31 versions

ADMM Problem

Least Squares Problem:

$$Min \, \left| |Dx - p| \right|$$
$$x \geq 0$$

Augmented Lagrangian:

$$Min \, \left| |Dx - p| \right| + y(x - z) + \frac{\rho}{2} \left| |x - z| \right|$$
$$z \geq 0$$

ADMM

$$\begin{cases} x^{k+1} = argmin_x \left| |Dx - p| \right| + y^k(x - z^k) + \frac{\rho}{2} \left| |x - z^k| \right| \\ z^{k+1} = argmin_{z \geq 0} \left| |Dx^{k+1} - p| \right| + y^k(x^{k+1} - z) + \frac{\rho}{2} \left| |x^{k+1} - z| \right| \\ y^{k+1} = y^k + \rho(x^{k+1} - z^{k+1}) \end{cases}$$

ADMM well-known issue:

- it could be really slow if used with an improper penalty parameter

Accelerating ADMM

$$x^{k+1} = argmin_x ||Dx - p|| + y^k(x - z^k) + \frac{\rho}{2}||x - z^k||$$

$$z^{k+1} = argmin_{z\geq 0} \left|\left|Dx^{k+1} - p\right|\right| + y^k(x^{k+1} - z) + \frac{\rho}{2}\left|\left|x^{k+1} - z\right|\right|$$

Gauss -Seidel

$$y^{k+1} = y^k + \rho(x^{k+1} - z^{k+1})$$

Fixed Stepsize Gradient Ascent

Small ρ → Slow Gradient Ascent

- BB Method Acceleration
  - Barzilai , J. M. Borwein, **IMA**, (1988)

Large ρ → Slow Gauss Seidel

- Line Search Acceleration
  - Dax, **Linear Alg. Applics**. (1990)

The Integrated ADMM

❑ADMM + Barzilai-Borwein+ Line Search

ADMM-BB-LS :

Gauss-Seidel:

$$\begin{cases} x^{k+1} = (D^T D + \frac{\rho}{2} I)^{-1} (D^T p - \frac{y^k}{2} + \frac{\rho}{2} z^k) \\ z^{k+1} = max(0, x^{k+1} + \frac{y^k}{\rho}) \end{cases}$$

Line Search:

$$\begin{cases} u^k = x^{k+1} - x^k, \; w^k = z^{k+1} - z^k \\ \alpha^k = -\frac{(Du^k)^t (Dx^k - p) + (u^k - w^k)^t (y^k + \rho(x^k - z^k))}{(Du^k)^t Du^k + \rho(u^k - w^k)^t (u^k - w^k)} \\ (x^{k+1}, z^{k+1}) = (x^k, z^k) + \alpha^k (x^{k+1} - x^k, z^{k+1} - z^k) \\ z^{k+1} = max(0, z^{k+1}) \end{cases}$$

Lagrange Variables Update Using BB:

$$\begin{cases} \triangle y^k = y^k - y^{k-1}, \quad \triangle d^k = (x^{k+1} - z^{k+1}) - (x^k - z^k) \\ \beta^k = -\frac{<\triangle d^k, \triangle y^k>}{<\triangle d^k, \triangle d^k>}, \text{ if } (\beta^k < 0 \text{ or } \beta^k = NAN) : \beta^k = \rho \\ y^{k+1} = y^k + \beta^k (x^{k+1} - z^{k+1}) \end{cases}$$

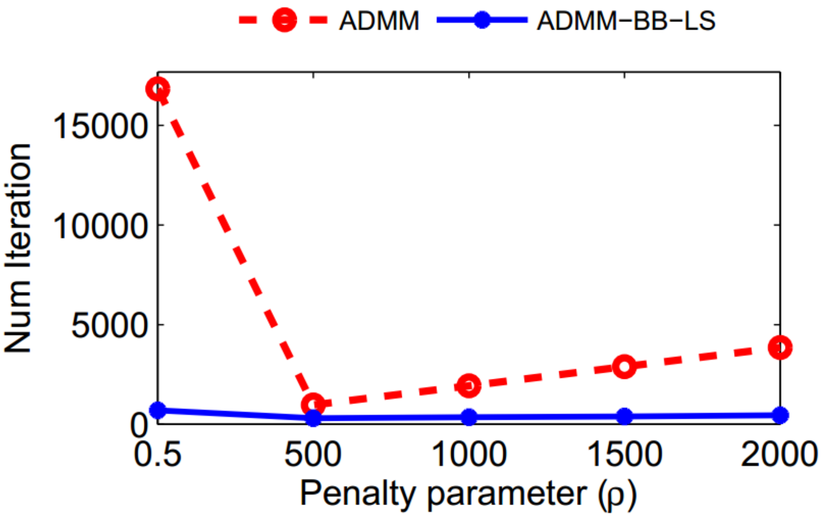
Results

❑ Integrated ADMM VS Regular ADMM

Prostate Case:

~430,000 Voxels

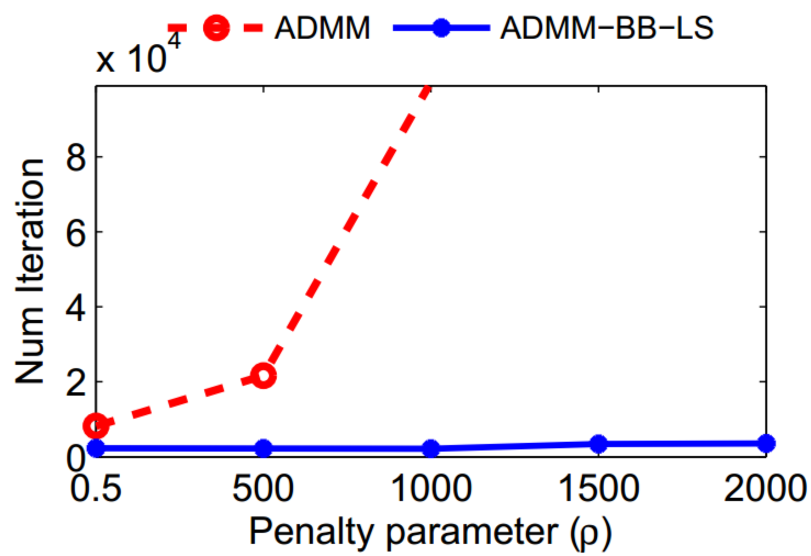
~9,000 Beamlets



H&N Case:

~365,000 Voxels

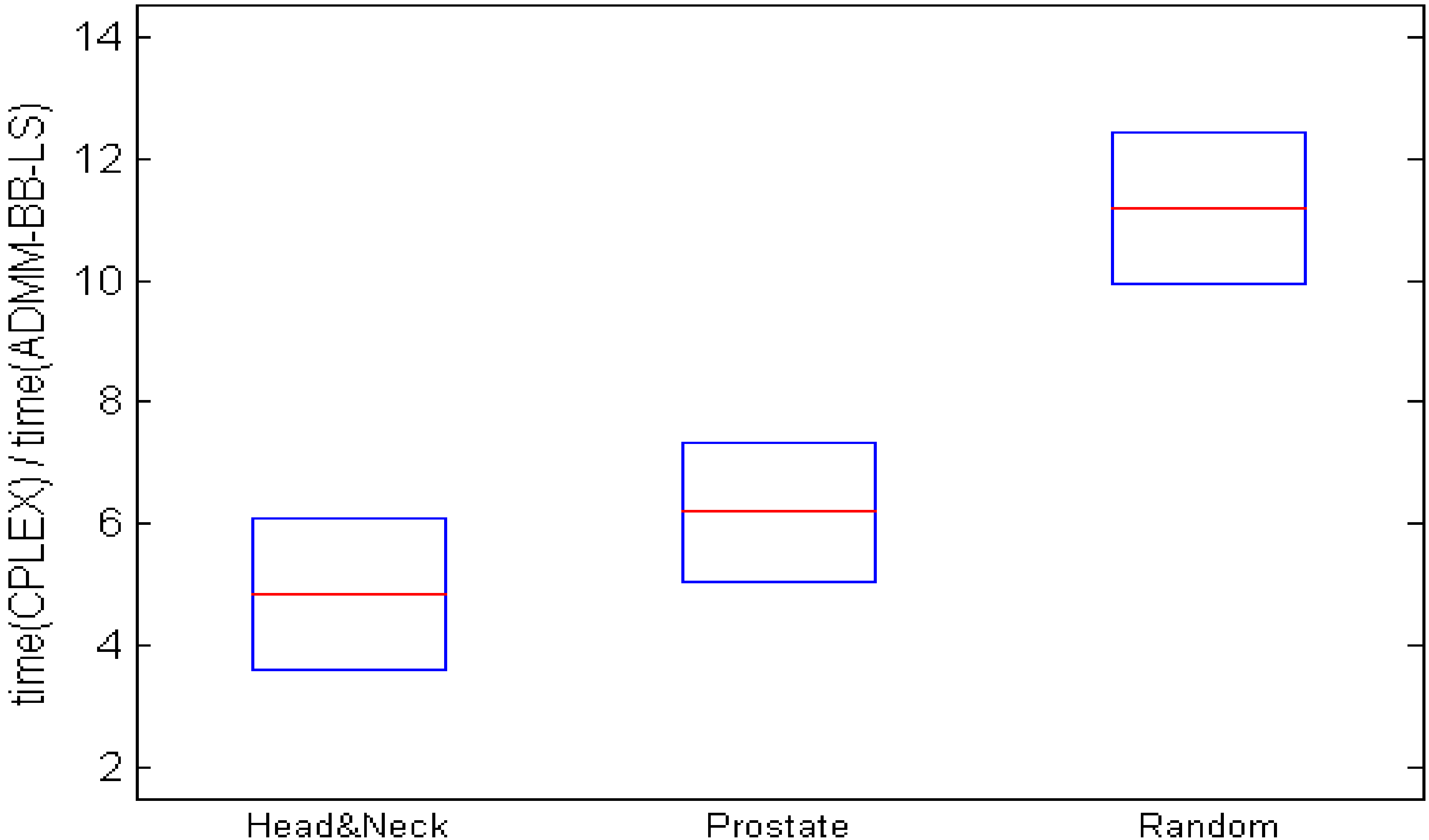
~20,000 Beamlets



Results

❑ Integrated ADMM VS Commercial CPLEX

Relative Error ≤ 5E-5



Summary

- We introduced a new version of ADMM which is much faster and less parameter independent

- Even a serial implementation of ADMM beats the commercial software CPLEX

- Distributed version of the integrated ADMM could be used for treatment planning in distributed settings like cloud