Factor Grouping for Efficient Bundle Adjustment

Tin Chon Chan

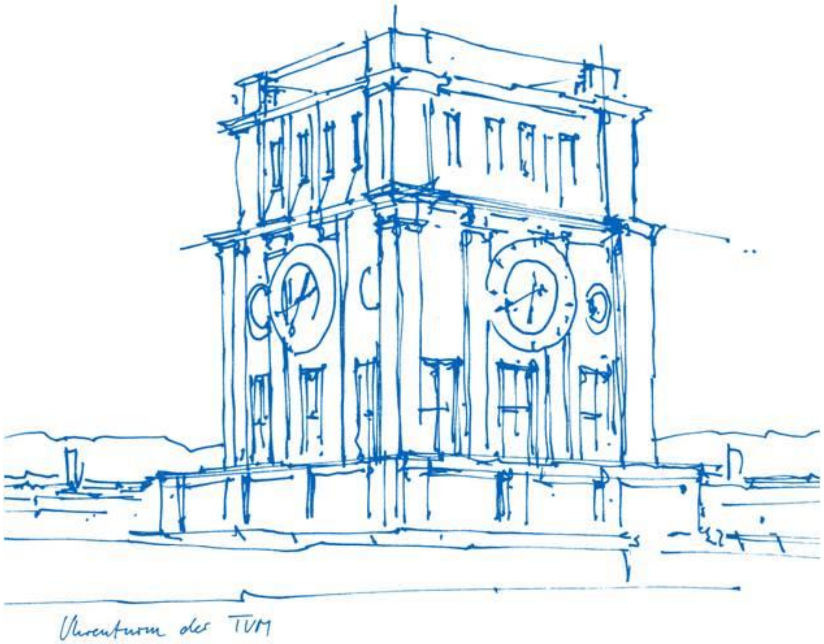
Technische Universität München

Chair of Computer Vision & AI

Master Thesis

Advisor: Nikolaus Demmel, Simon Weber

Supervisor: Prof. Dr. Daniel Cremers



Outline

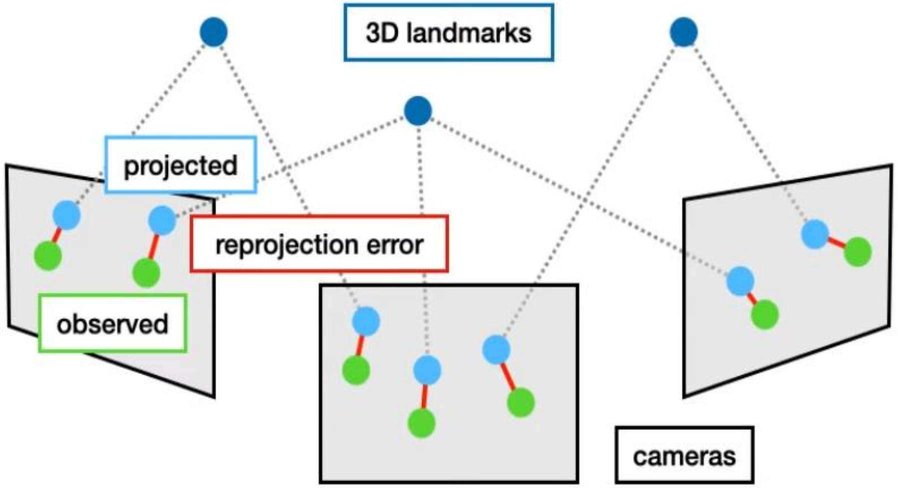
- Motivation & Background
- Schur complement trick
- Factor grouping
- Square Root Bundle Adjustment
- Experiments & Analysis

Motivation

Structure from Motion (SfM)

- 3D reconstruction from a set of unordered images looking at a scene
- Jointly estimate structure and camera parameters
- Bottleneck of SfM system

Background



residual landmark

$$r_{ij} = u_{ij} - \pi(R_i X_j + t_i; c_i)$$

pixel location

camera parameters

π : projection function

R_i: camera rotation

c_i: camera intrinsic

t_i: camera translation

$$\mathrm{E}(x_p, x_l) = \sum_{i,j} \|r_{ij}\|^2 = \|r(x_p, x_l)\|^2$$

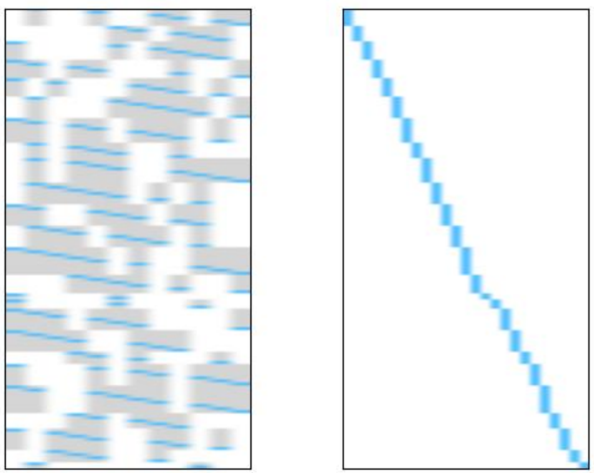
Background

Solve by Levenberg–Marquardt

$$\min_{x} \mathrm{E}(x) = \|r(x)\|^2$$

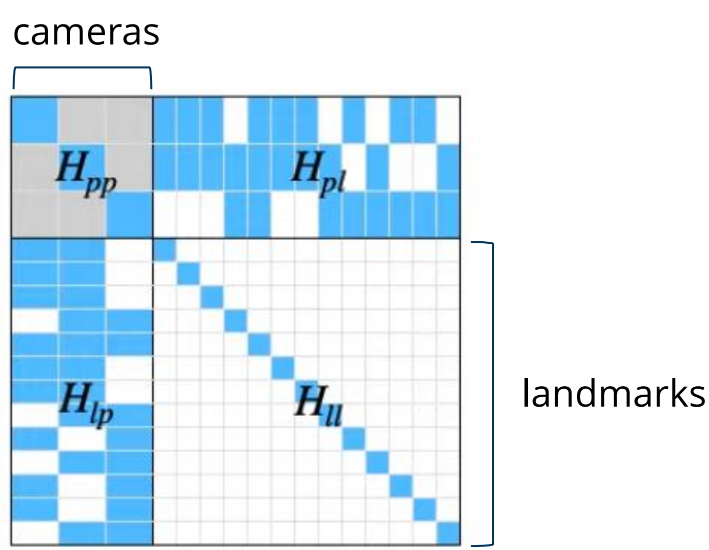
Linearize in every LM iteration

$$\min_{\Delta x} \mathrm{E}_{lin}(\Delta x) = \|r + Jx\|^2$$
$$= r^\top r + 2r^\top J\Delta x + \Delta x^\top H \Delta x$$



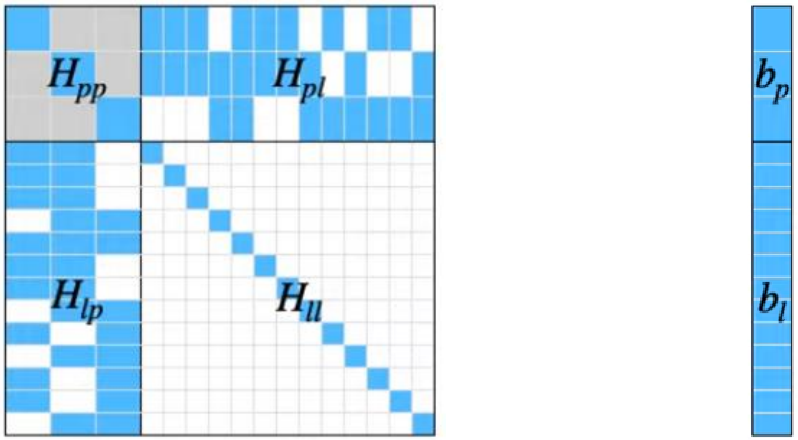
J_p J_l

Jacobian J



Hessian H

Schur complement trick



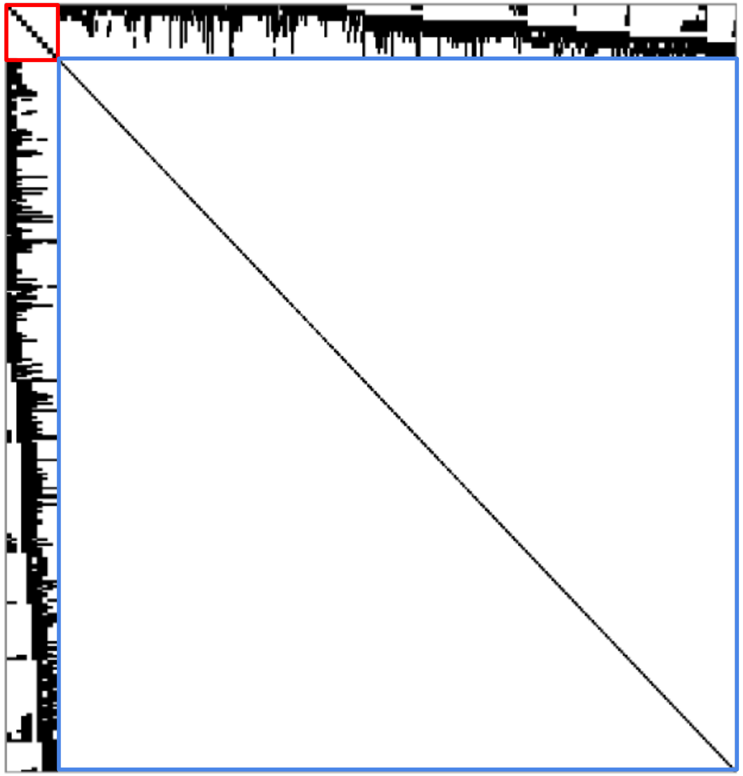
$$\begin{pmatrix} H_{pp} & H_{pl} \\ H_{lp} & H_{ll} \end{pmatrix} \begin{pmatrix} -\Delta x_p \\ -\Delta x_l \end{pmatrix} = \begin{pmatrix} b_p \\ b_l \end{pmatrix}$$

Reduced Camera System (RCS)

$$\widehat{\boldsymbol{H}}_{\boldsymbol{pp}}(-\Delta \boldsymbol{x}_{\boldsymbol{p}}) = \widetilde{\boldsymbol{b}}_{\boldsymbol{p}}$$

$$\widehat{H}_{pp} = H_{pp} - H_{pl} H_{ll}^{-1} H_{lp}$$
$$\widetilde{b}_p = b_p - H_{pl} H_{ll}^{-1} b_l$$

Schur complement trick



Hessian H

10 cameras
982 landmarks

Schur complement trick

Solve by
Preconditioned Conjugate Gradient

$$\widehat{H}_{pp}(-\Delta x_p) = \tilde{b}_p$$

$$\widehat{H}_{pp} = H_{pp} - H_{pl}H_{ll}^{-1}H_{lp}$$

Explicit SC

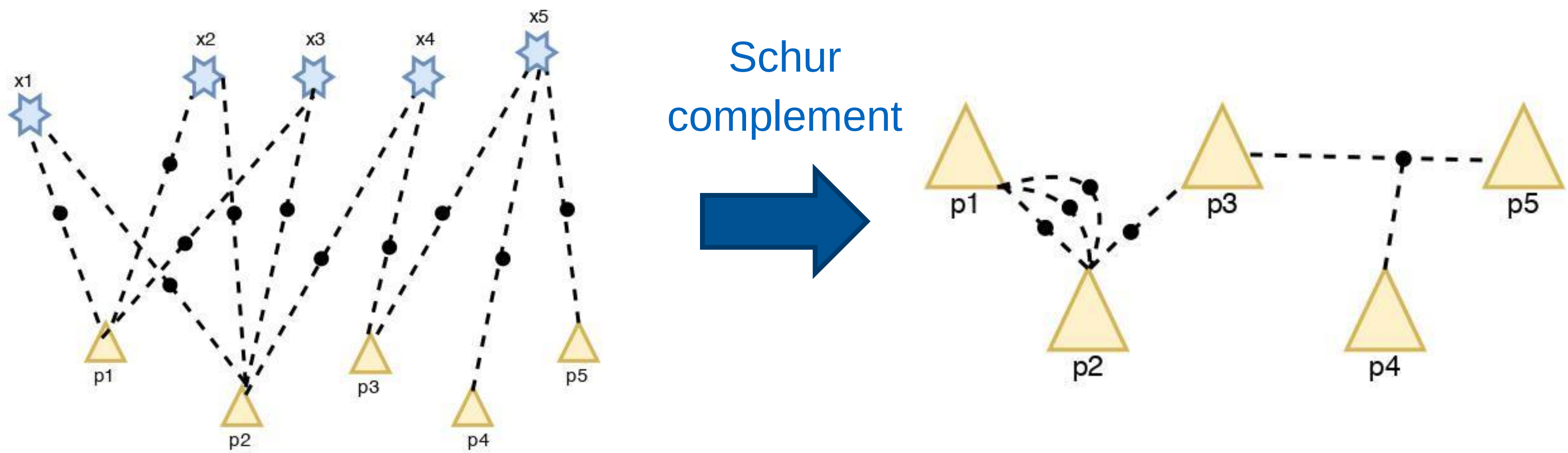
$$\widehat{H}_{pp}\, v$$

Implicit SC

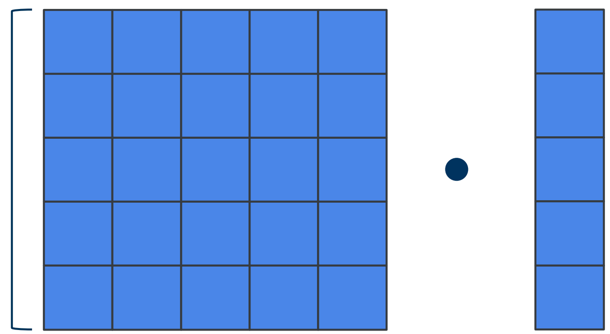
$$H_{pp}v - H_{pl}H_{ll}^{-1}H_{lp}v$$

Diagonal blocks of $\widehat{H}_{pp}$ are used as preconditioner

Factor grouping



Factor grouping



Time complexity of matrix-vector multiplication

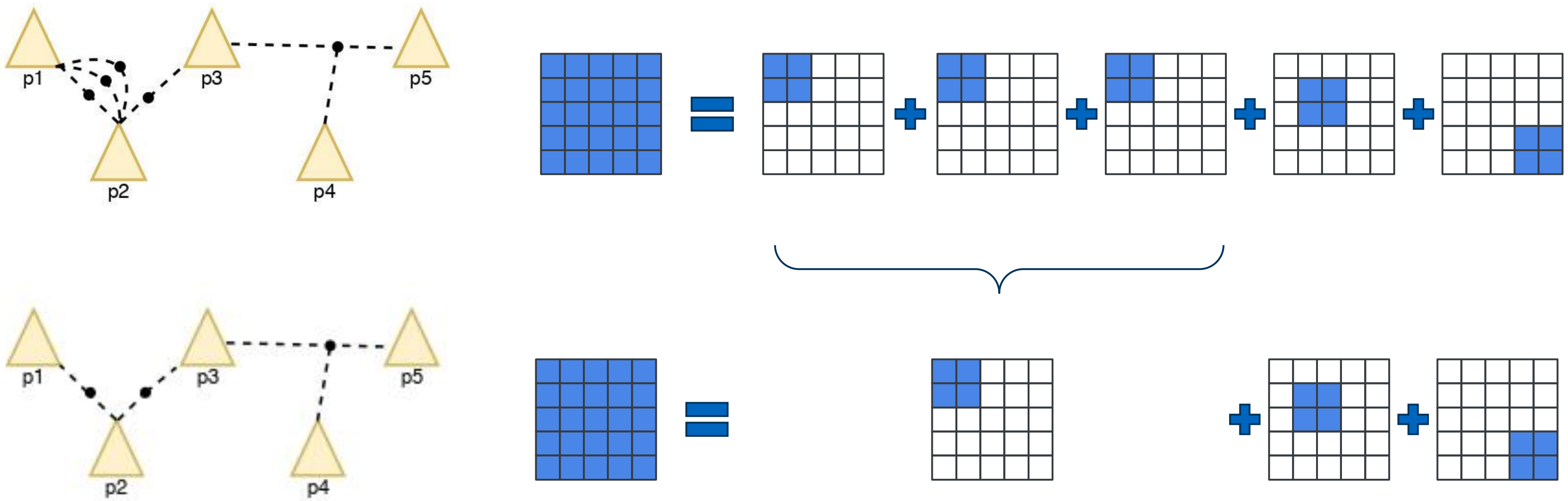
#cameras = P, #landmarks = L

Explicit: $\boldsymbol{O}(\boldsymbol{P}^2)$ Implicit: $\boldsymbol{O}(\boldsymbol{PL})$

(assume all cameras see all landmarks)

Explicit is better than Implicit only if $\boldsymbol{P} < \boldsymbol{L}$

Factor grouping



Explicit

Implicit

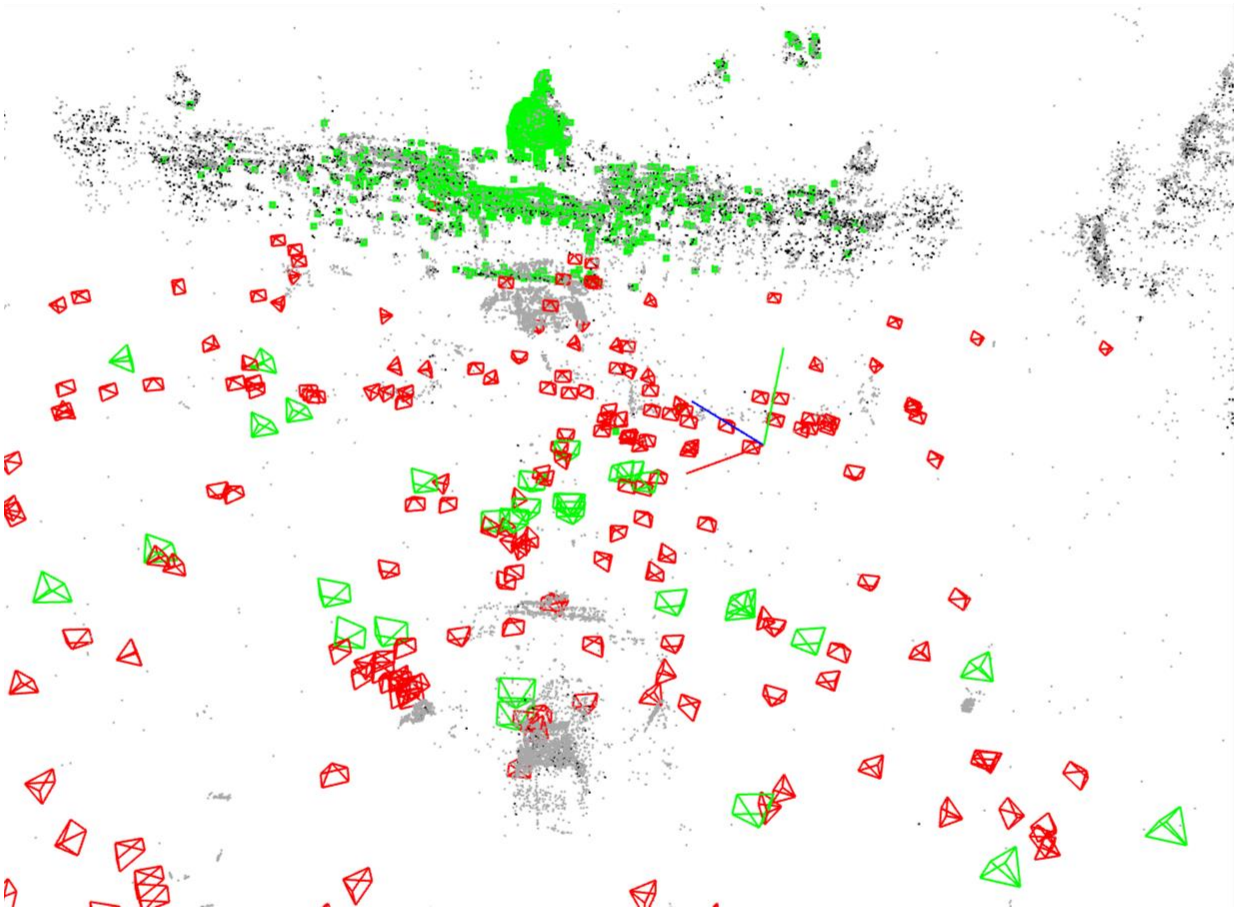
$$P < L$$
$$P = 2, L = 3$$

Explicit: $O(P^2)$ Implicit: $O(PL)$

Factor grouping

Frequent Pattern Tree

group landmarks having a common set of camera observations



Square Root Bundle Adjustment

$$\min_{\Delta x_p, \Delta x_l} \mathrm{E}_{lin}(\Delta x_p, \Delta x_l) = \left\| r + \begin{pmatrix} J_p & J_l \end{pmatrix} \begin{pmatrix} \Delta x_p \\ \Delta x_l \end{pmatrix} \right\|^2$$

Nullspace Marginalization

QR decomposition

$$J_l = \begin{pmatrix} Q_1 & Q_2 \end{pmatrix} \begin{pmatrix} R_1 \\ 0 \end{pmatrix}$$

$$\min_{\Delta x_p} \left\| Q_2^T r + Q_2^T J_p \Delta x_p \right\|^2$$

Project Q_2

Square Root Bundle Adjustment (√BA)

$$\min_{\Delta x_p, \Delta x_l} \mathrm{E}_{lin}(\Delta x_p, \Delta x_l) = \left\| r + \begin{pmatrix} J_p & J_l \end{pmatrix} \begin{pmatrix} \Delta x_p \\ \Delta x_l \end{pmatrix} \right\|^2$$

Nullspace Marginalization

Schur complement

QR decomposition

$$J_l = \begin{pmatrix} Q_1 & Q_2 \end{pmatrix} \begin{pmatrix} R_1 \\ 0 \end{pmatrix}$$

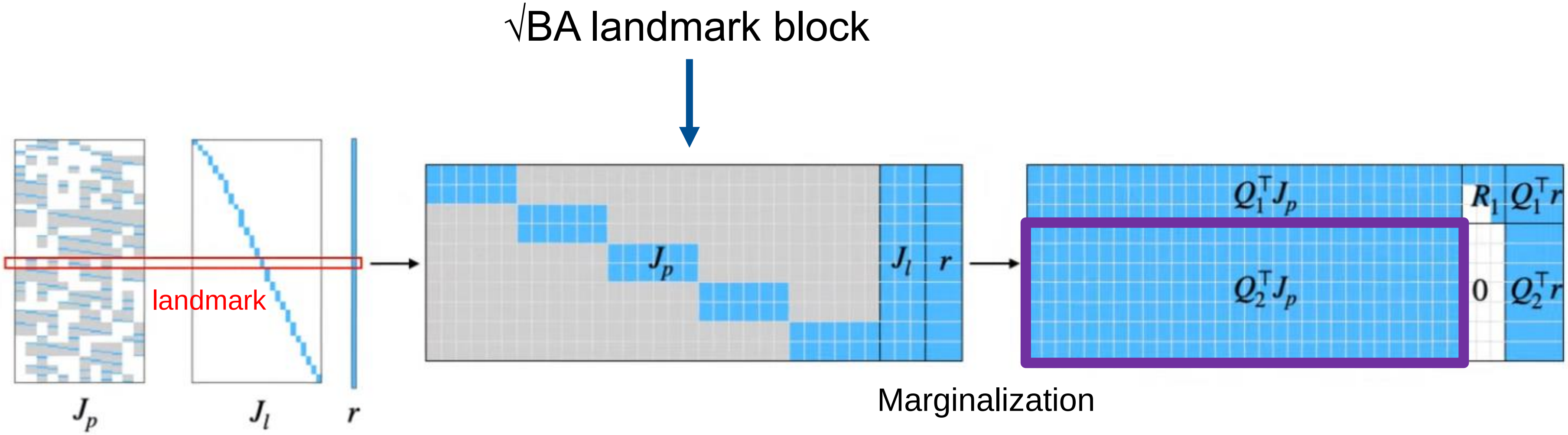
$$\min_{\Delta x_p} \left\| Q_2^T r + Q_2^T J_p \Delta x_p \right\|^2$$

Project Q_2

equivalent

$$\widehat{H}_{pp}\left(-\Delta x_p\right) = \tilde{b}_p$$

Square Root Bundle Adjustment (√BA)



Preconditioned Conjugate Gradient

√BA

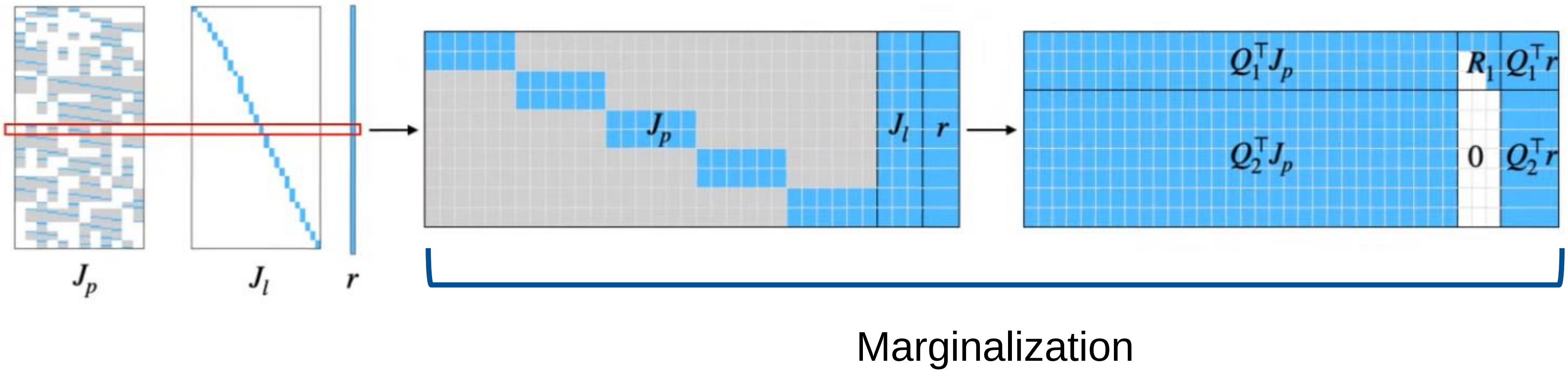
$$\left(Q_2^T J_p\right)^T \left(Q_2^T J_p\right) v$$

Schur complement

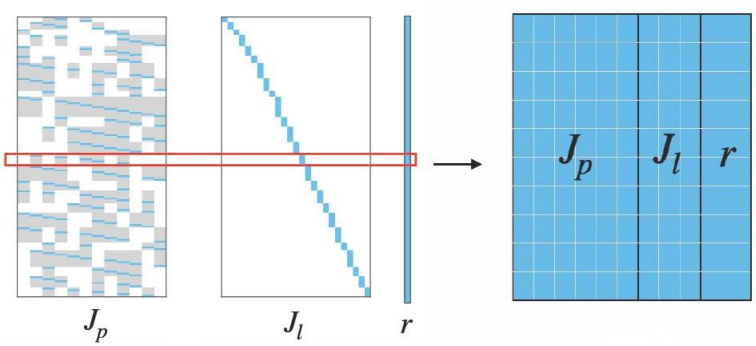
$$\hat{H}_{pp} v$$

Square Root Bundle Adjustment (√BA)

√BA



Implicit √BA



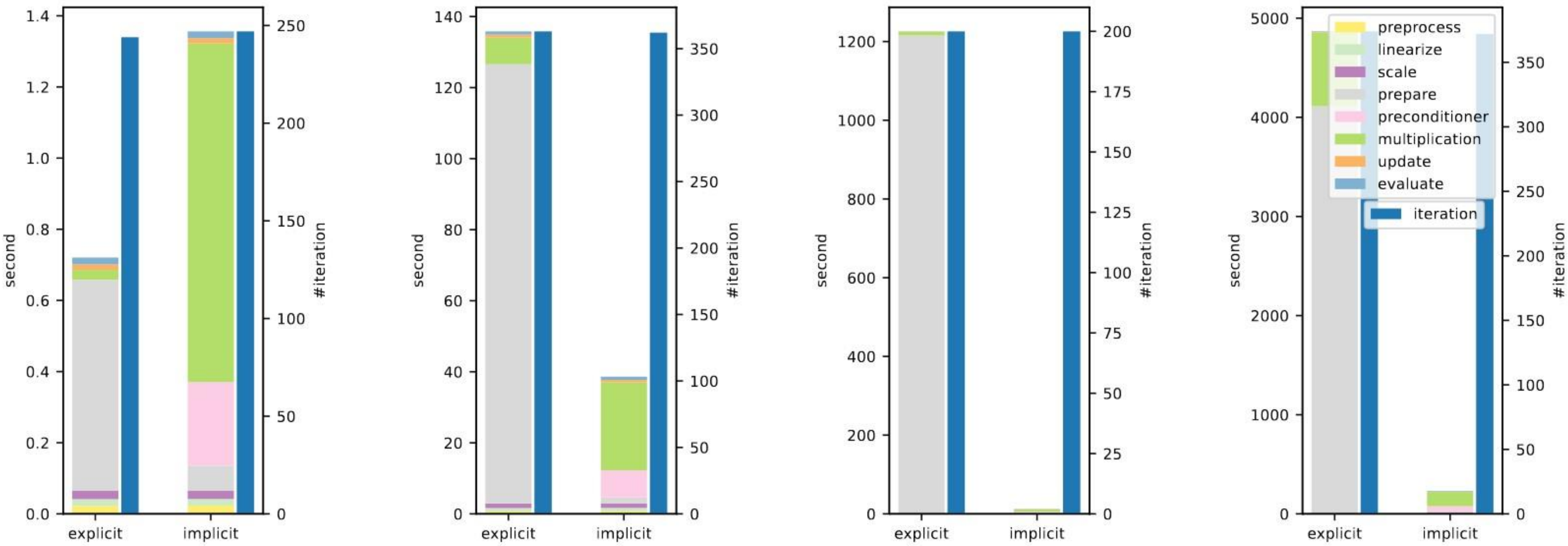
For every multiplication:

1. Copy to √BA landmark
2. Perform marginalization

Factor grouping + Implicit √BA → **Factor √BA**

Experiments & Analysis

Explicit and Implicit SC



(a) trafalgar50
(small-scale)

(b) venice1184
(medium-scale)

(c) final961
(dense)

(d) final13682
(sparse, large-scale)

Matrix-Vector Multiplication

Compute Preconditioner

Prepare RCS

Explicit and Implicit SC

Explicit SC

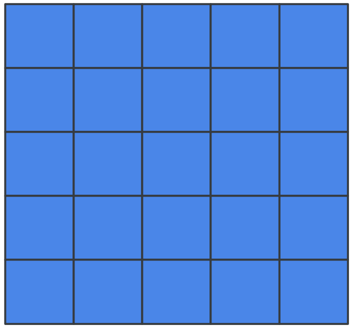
- **<u>Expensive</u>** to prepare RCS ($\widehat{H}_{pp}$)
- **<u>Cheap</u>** to compute preconditioner
- **<u>Cheap</u>** multiplication

Implicit SC

- **<u>Cheap</u>** to prepare RCS
- **<u>Expensive</u>** to compute preconditioner
- **<u>Expensive</u>** multiplication

Time complexity

	Explicit SC	Implicit SC
Prepare RCS	$\boldsymbol{O(LP^2)}$	$\boldsymbol{O(LP)}$
Preconditioner	$\boldsymbol{O(P)}$	$\boldsymbol{O(LP)}$
Multiplication	$\boldsymbol{O(P^2)}$	$\boldsymbol{O(LP)}$

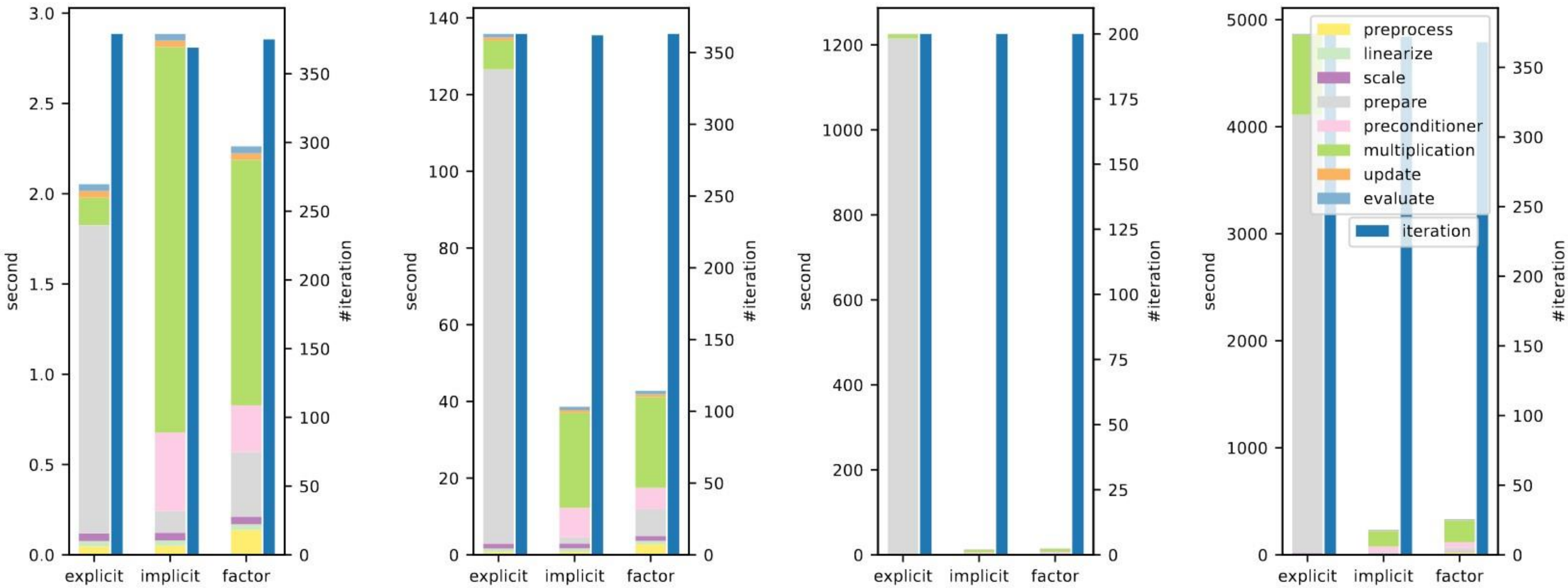


$\widehat{H}_{pp}$

#cameras = P, #landmarks = L

L>>P

Factor grouping



(a) trafalgar161 (small-scale)

(b) venice1184 (medium-scale)

(c) final961 (dense)

(d) final13682 (sparse,large-scale)

Matrix-Vector Multiplication

Compute Preconditioner

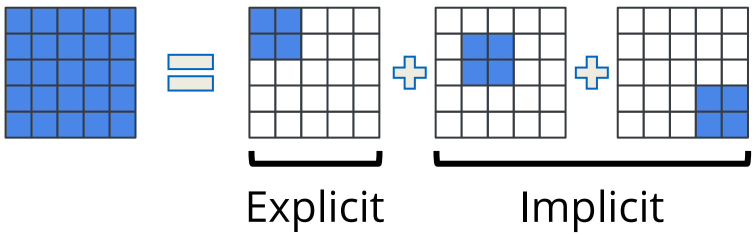
Prepare RCS

Factor grouping

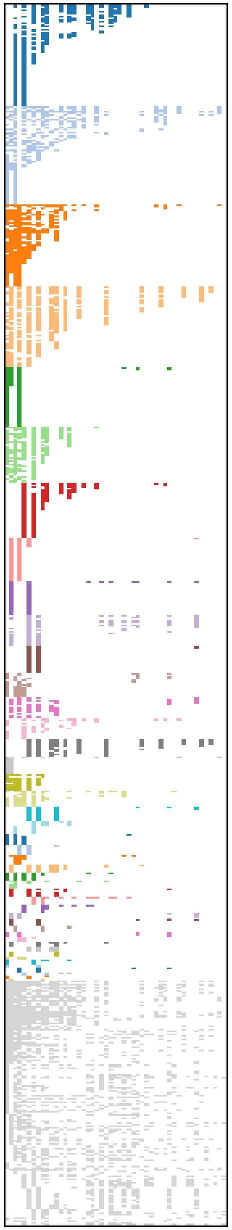
Factor SC

- **<u>Relatively cheap</u>** to prepare RCS ($\widehat{H}_{pp}$)
- **<u>Relatively cheap</u>** to compute preconditioner
- **<u>Relatively cheap</u>** multiplication

Explicit SC performed worse when given many cameras (large-scale or dense problems)



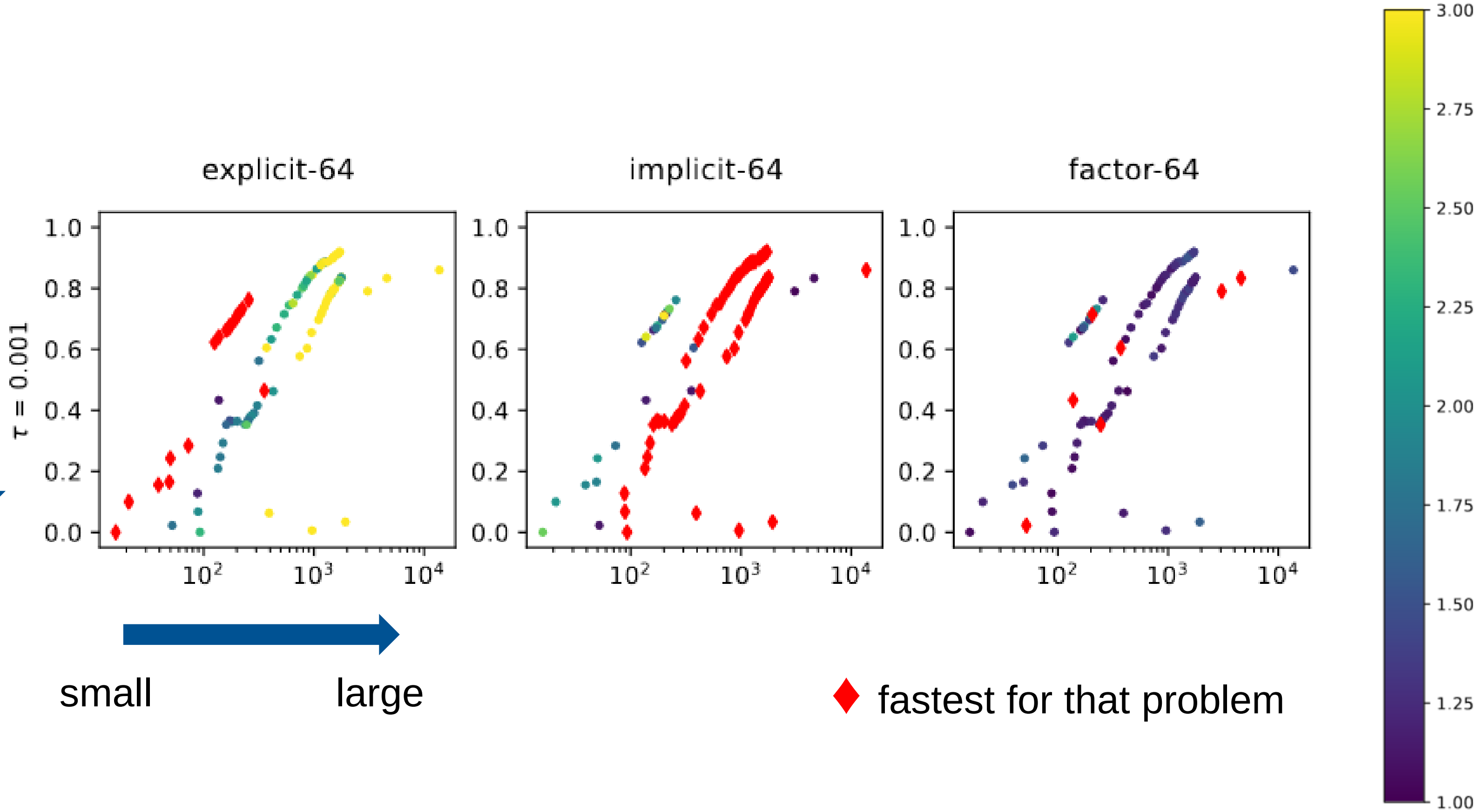
- Less cameras → Mitigate the quadratic complexity
- Improve overall multiplication runtime



Factor grouping

sparse

dense



small large

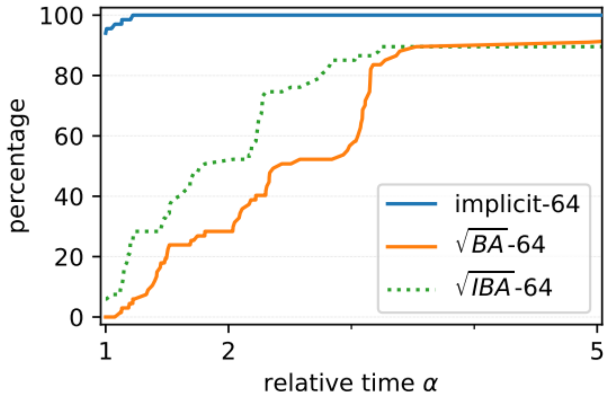
♦ fastest for that problem

Speed compared to the fastest

Square Root Bundle Adjustment (√BA)

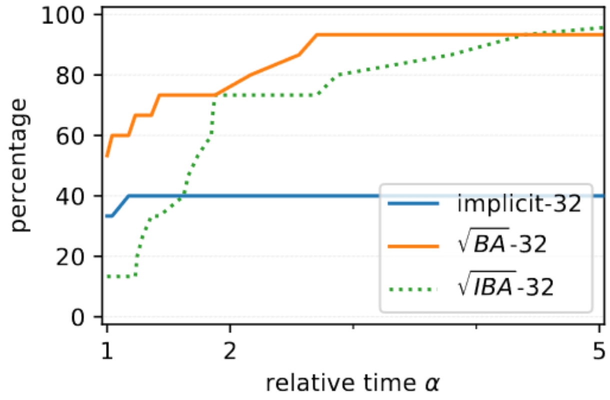
easier to solve, double precision

(Bundle Adjustment in the Large)



harder to solve, single precision

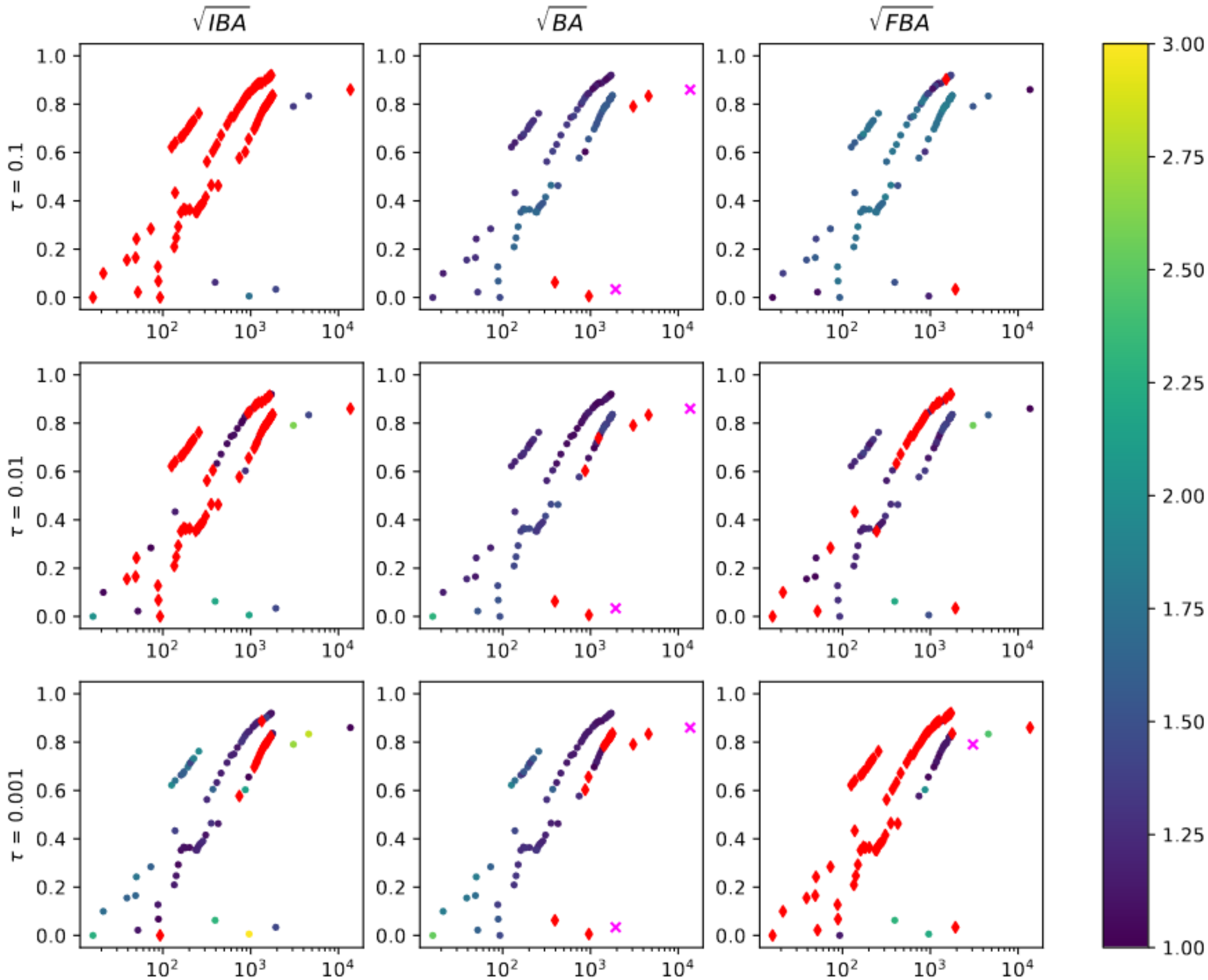
(1DSfM)



Square Root Bundle Adjustment (√BA)

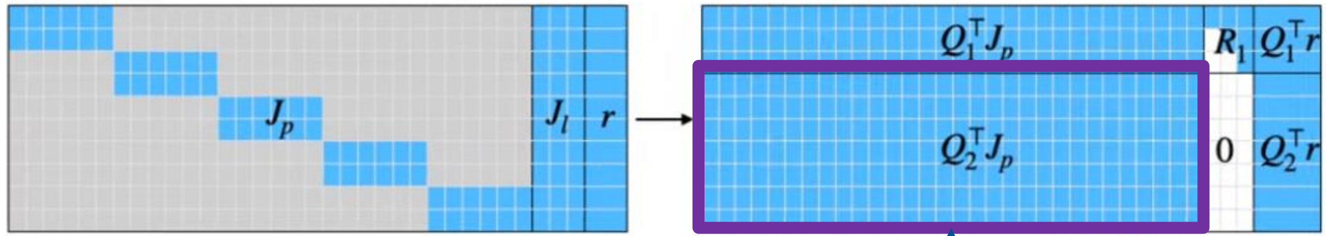
Lower accuracy

Higher accuracy



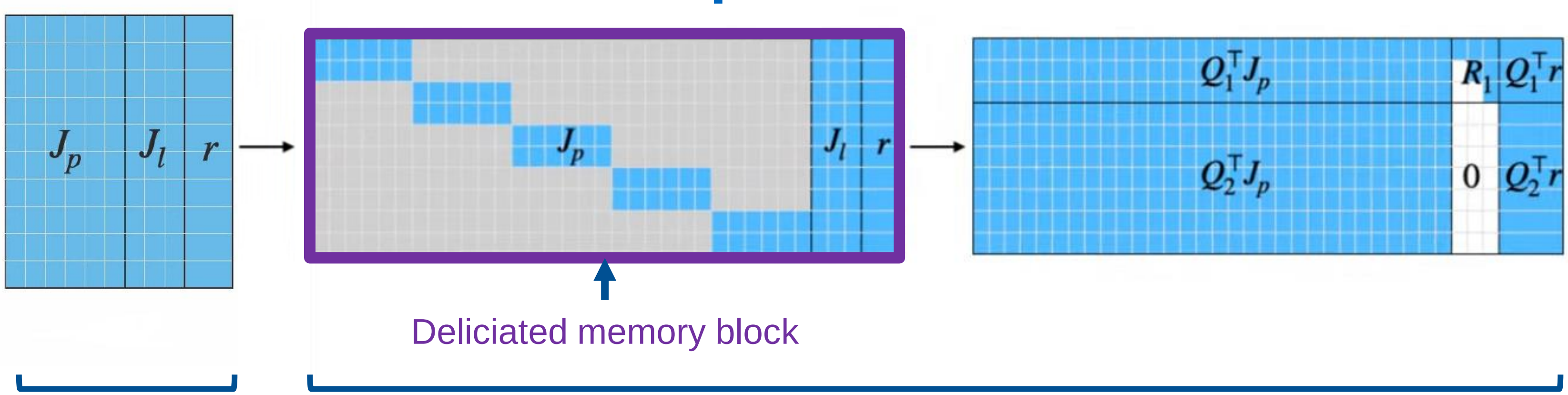
√BA and Implicit √BA

√BA



Recall from memory for every multiplication

Implicit √BA

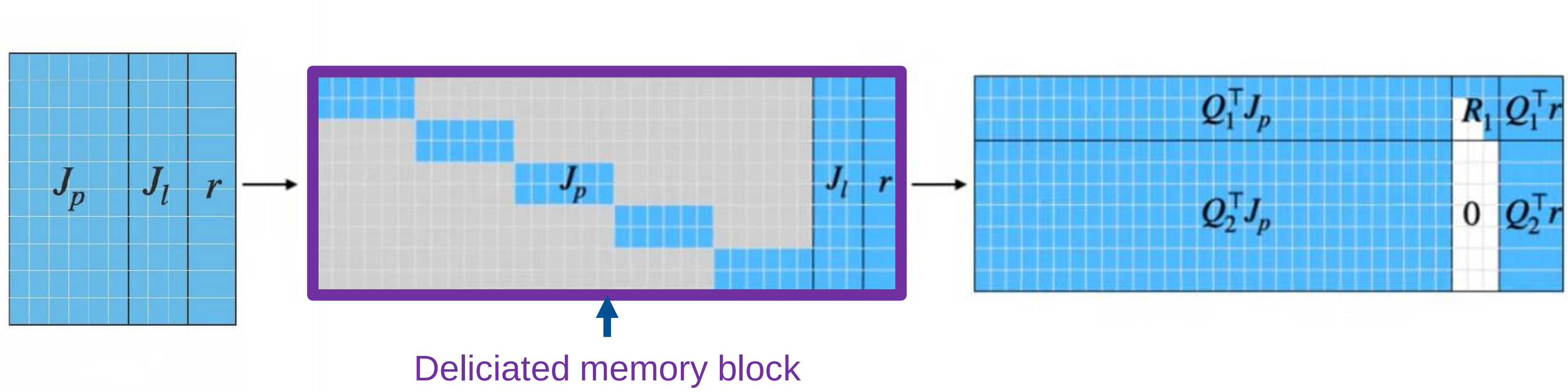


Deliciated memory block

Recall from memory for every multiplication

Recompute for every multiplication

Implicit √BA

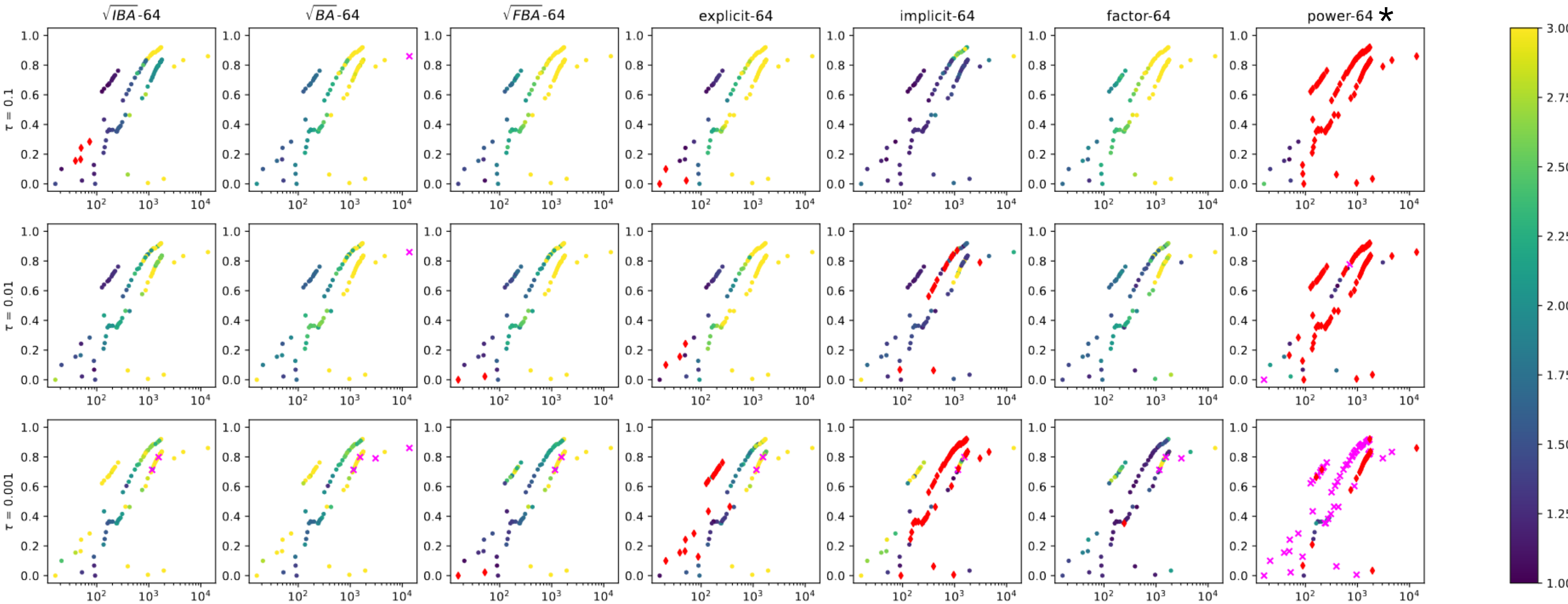


- Write result to closer memory location (linearization)
- Recall and consume less amount of memory
- Marginalize on same memory block

Better memory access pattern

(spatial and temporal locality)

Overall performance



♦ fastest for that problem

× solution didn't reach the accuracy

* Power Bundle Adjustment for Large-Scale 3D Reconstruction

Conclusion

- Implemented 5 solvers (highly parallelized and vectorized)
- Analyzed the performance of 7 solvers
- Analyzed time complexity and memory access pattern
- Recomputing intermediate result can be faster than recalling from memory

Recommendations

- Implicit SC for large-scale and dense problems
- Factor SC for small- and medium-scale problems
- Implicit √BA for solving in single precision
- Power BA for solving double precision and values speed over accuracy

Reference

- Zhou, Lei, et al. "Stochastic bundle adjustment for efficient and scalable 3d reconstruction." European Conference on Computer Vision. Springer, Cham, 2020.
- S. Agarwal, N. Snavely, S. M. Seitz, and R. Szeliski. “Bundle adjustment in the large.” In: European conference on computer vision. Springer. 2010, pp. 29–42.
- K. Wilson and N. Snavely. “Robust Global Translations with 1DSfM.” In: Proceedings of the European Conference on Computer Vision (ECCV). 2014.
- L. Carlone, P. Fernandez Alcantarilla, H.- P. Chiu, Z. Kira, and F. Dellaert. “Mining Structure Fragments for Smart Bundle Adjustment.” In: Proceedings of the British Machine Vision Conference. BMVA Press, 2014.
- J. Han, H. Cheng, D. Xin, and X. Yan. “Frequent pattern mining: current status and future directions.” In: Data mining and knowledge discovery 15.1 (2007), pp. 55–86.
- N. Demmel, C. Sommer, D. Cremers, and V. Usenko. “Square Root Bundle Adjustment for Large-Scale Reconstruction.” In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2021.
- S. Weber, N. Demmel, T. C. Chan, and D. Cremers. Power Bundle Adjustment for Large-Scale 3D Reconstruction. Manuscript submitted for publication. 2022.