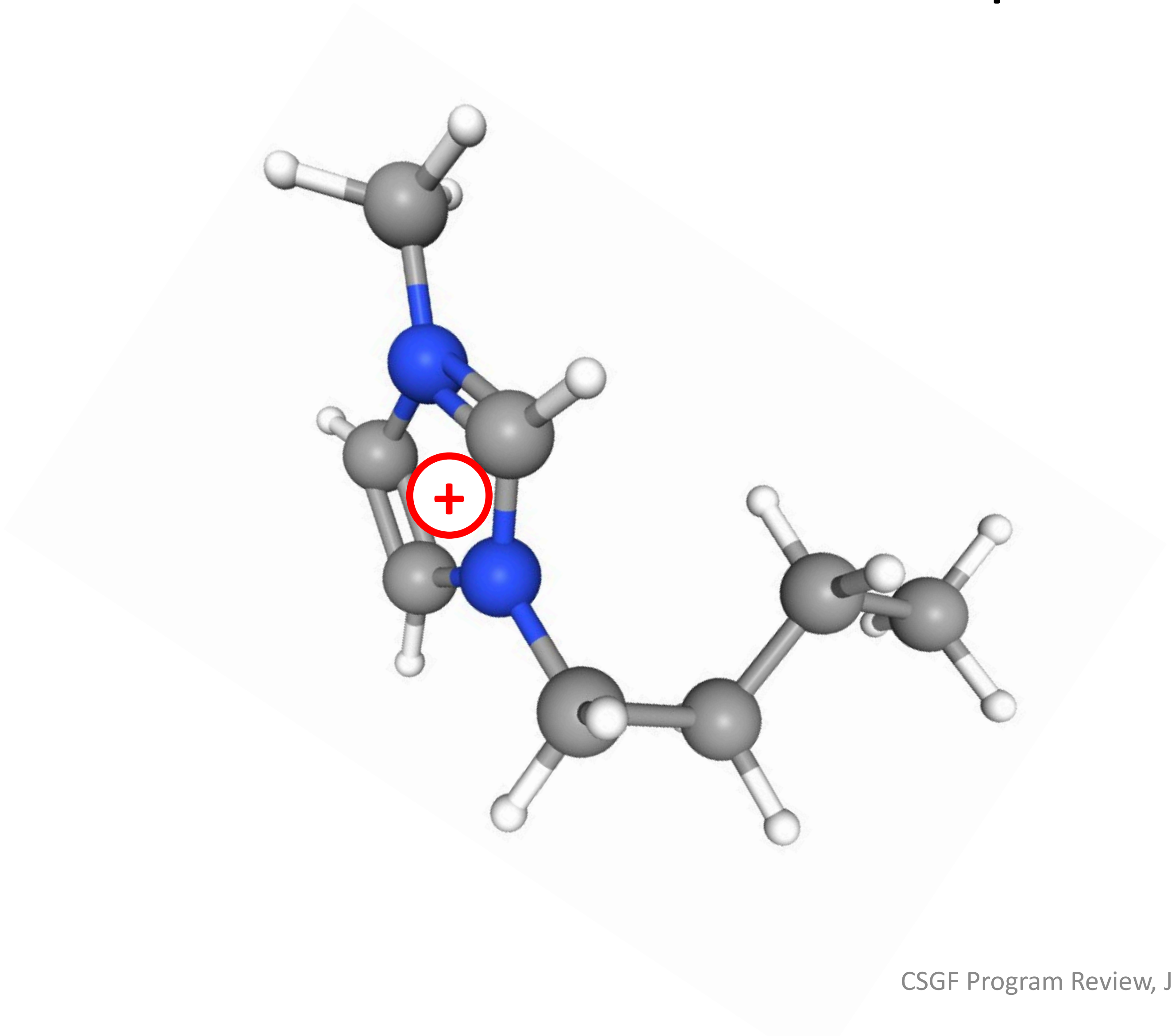
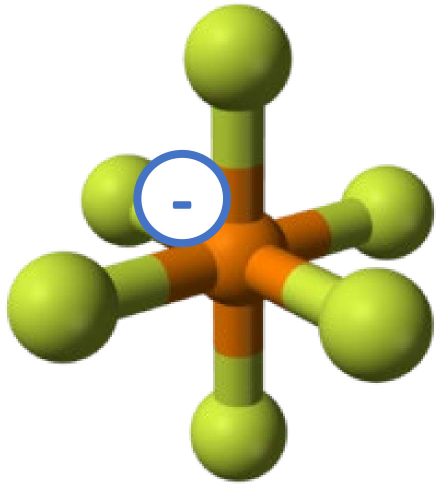
What Do Ionic Liquids Have To Do With Linear Algebra?

Devin Matthews

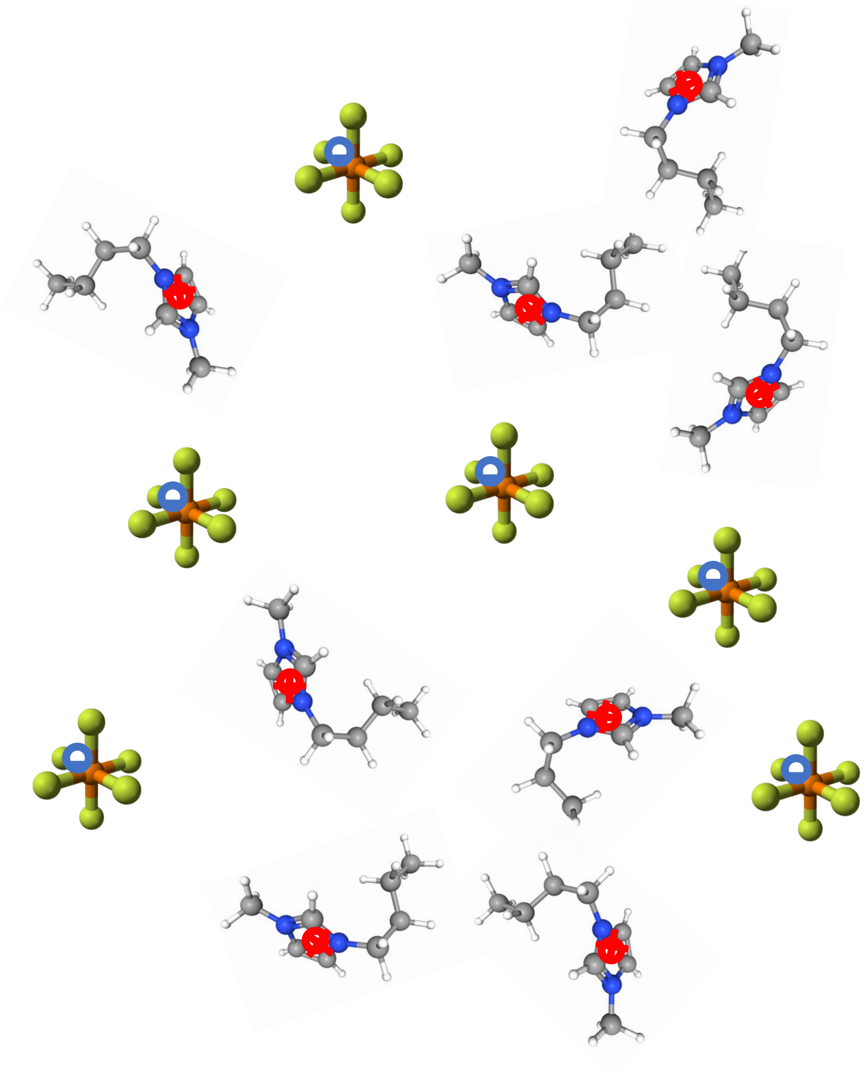
Southern Methodist University

Motivation: Ionic Liquids



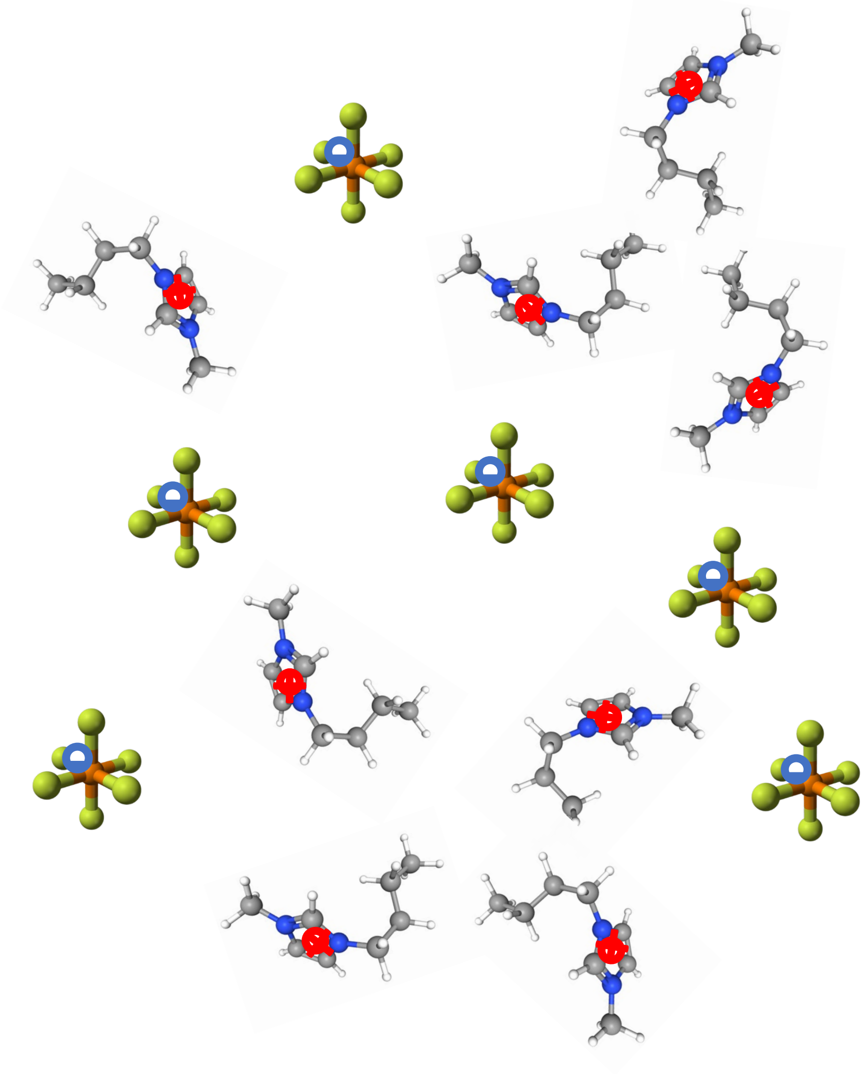


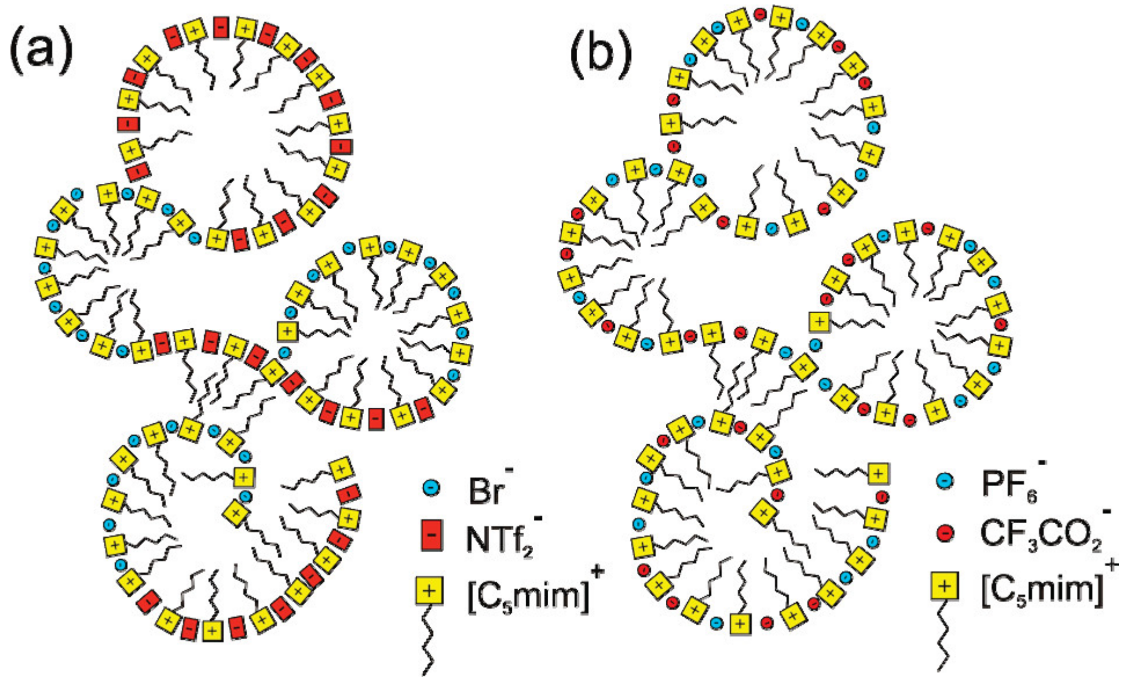
Motivation: Ionic Liquids



Both long-range and short-range *dynamical* order

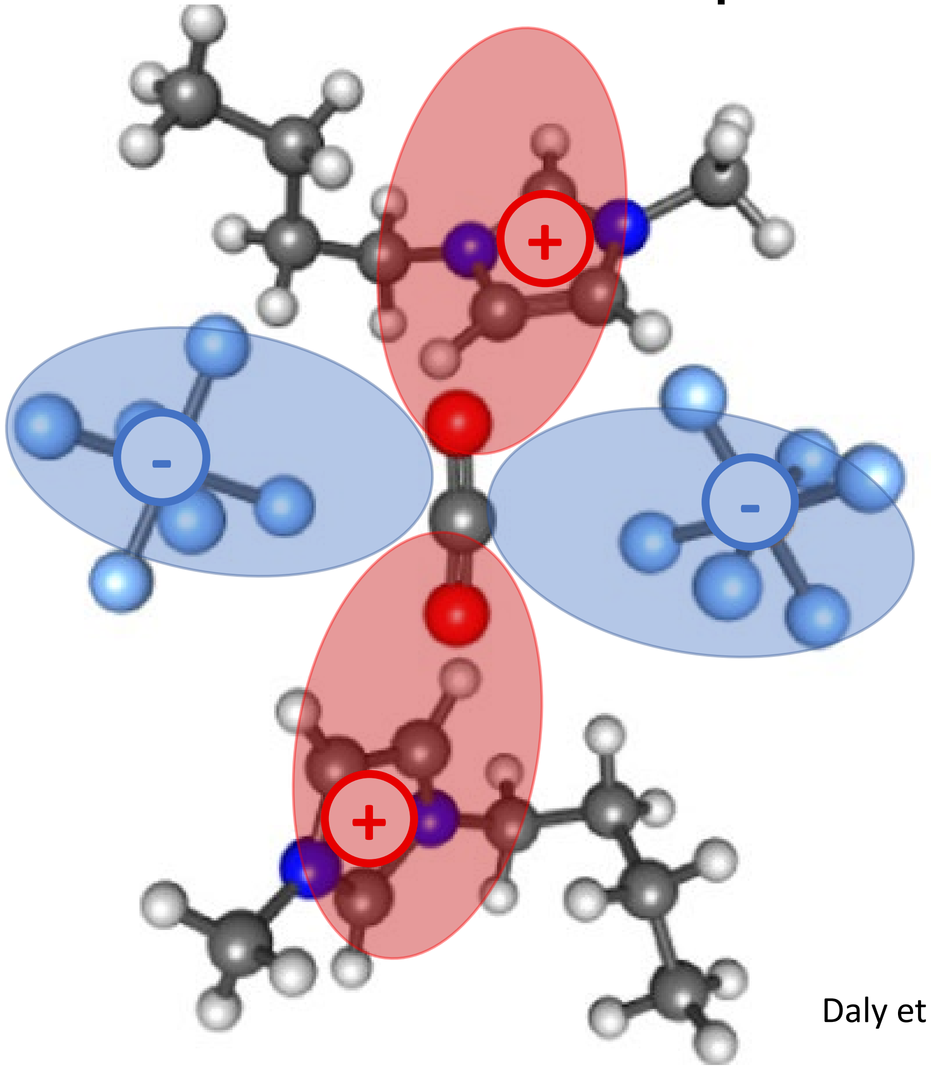
Motivation: Ionic Liquids





Xiao et al. *J. Phys. Chem. B* 2008, **112**, 42, 13316–13325

Motivation: Ionic Liquids

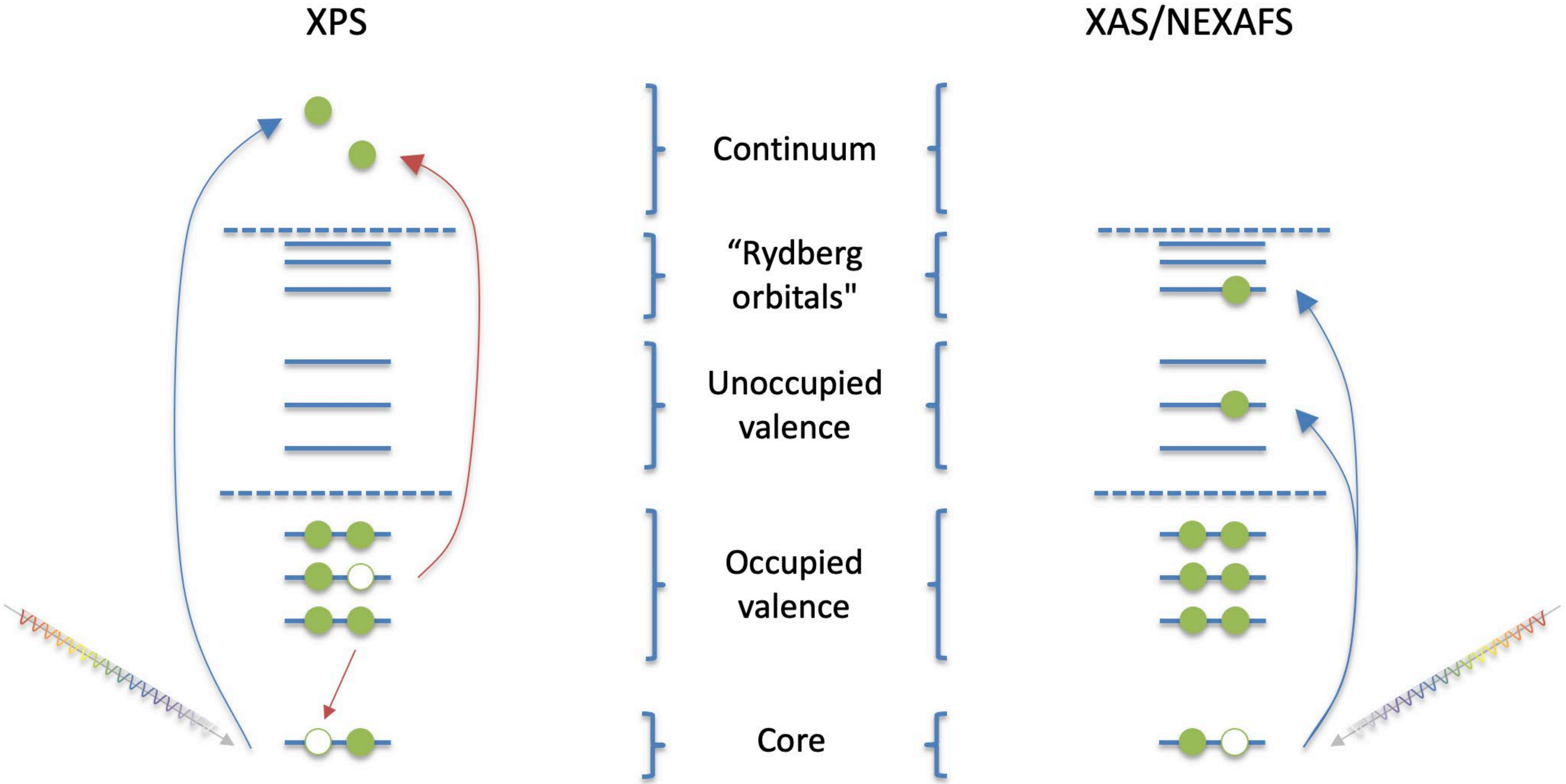


Electric fields in the range of 0.01-0.02 au ≈ 5-10 V/nm

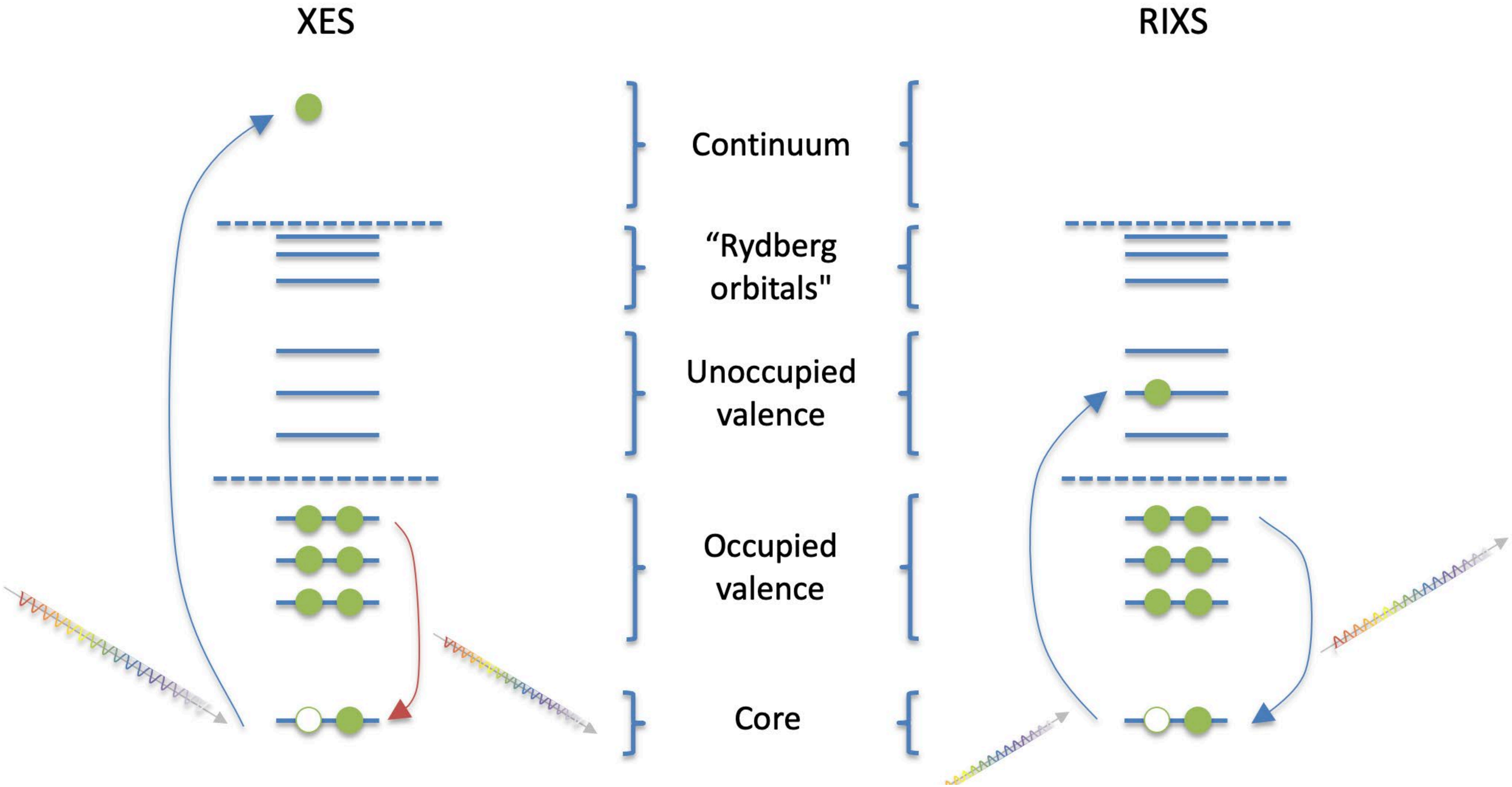
Strong orientation and solvent caging effects

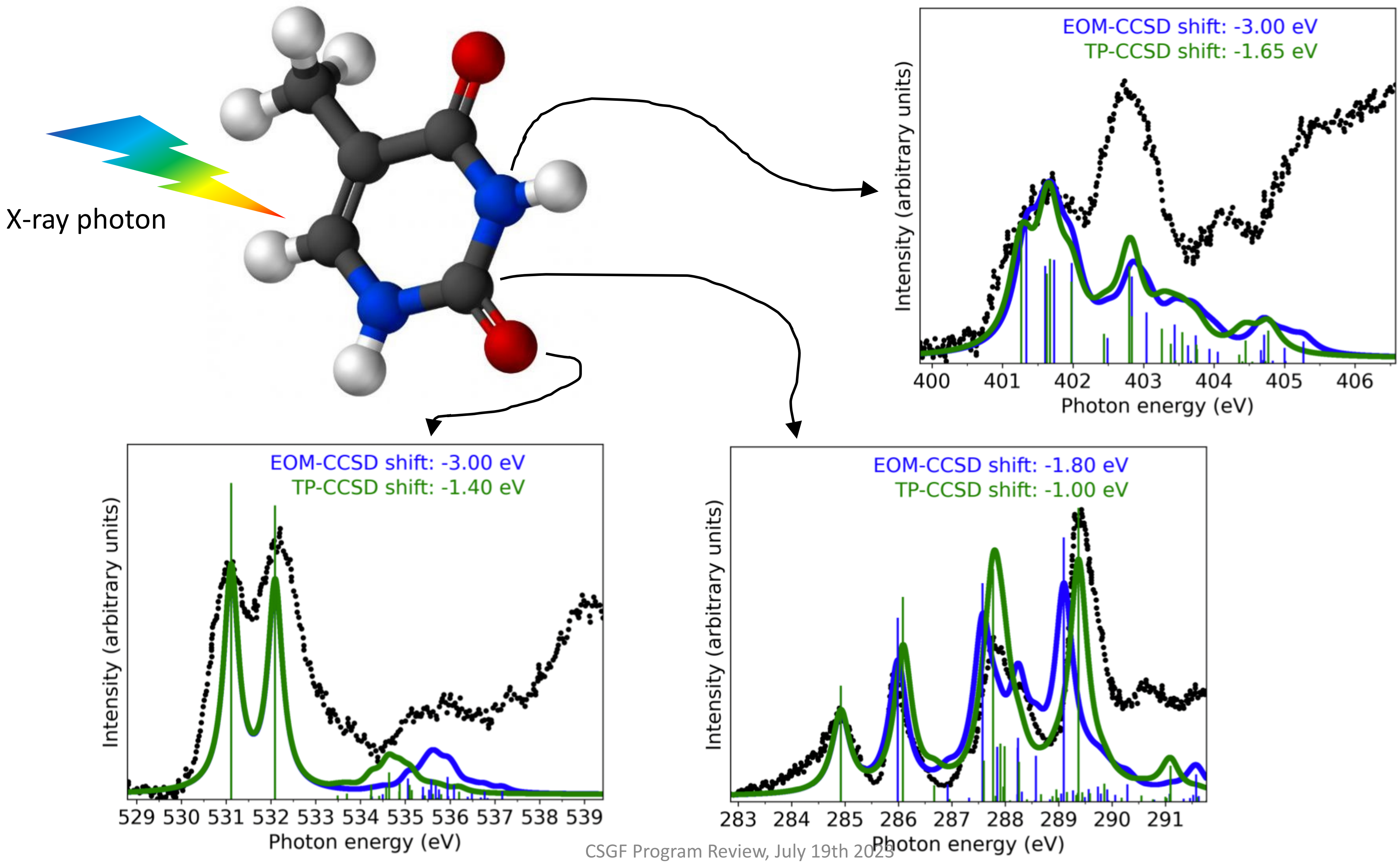
Daly et al., *J. Phys. Chem. B* 2016, **120**, 49

Soft X-ray Spectroscopy



Soft X-ray Spectroscopy

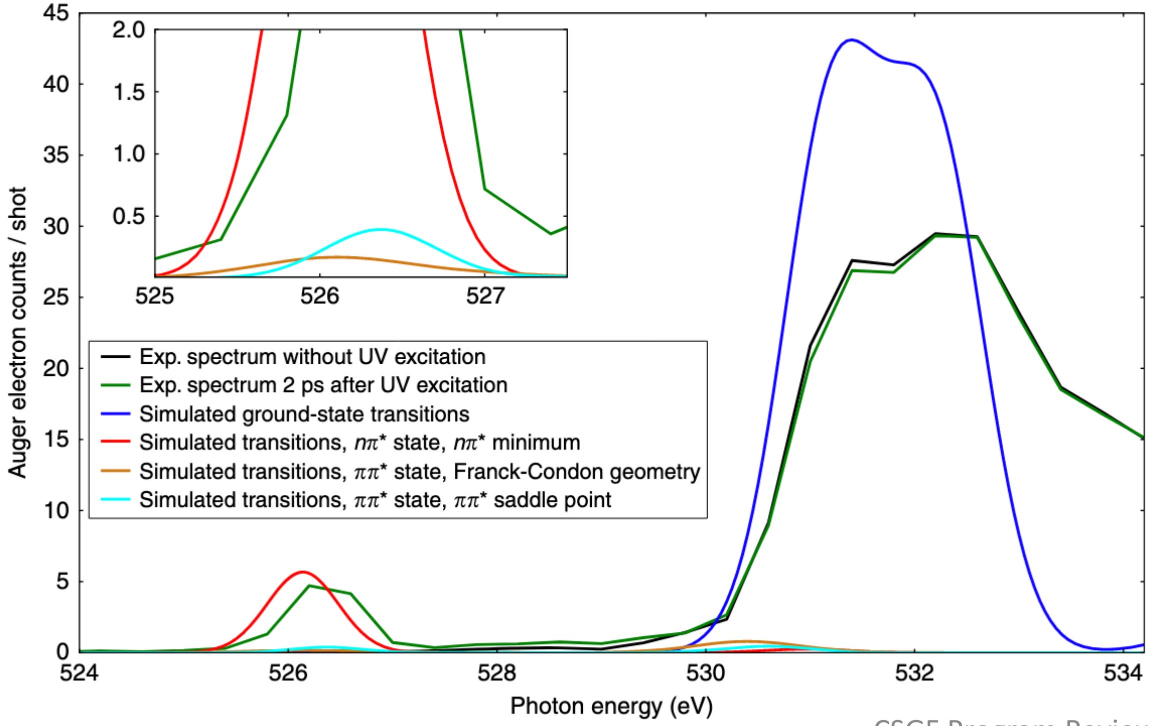



X-ray photon
EOM-CCSD shift: -3.00 eV
TP-CCSD shift: -1.65 eV
Intensity (arbitrary units)
400 401 402 403 404 405 406
Photon energy (eV)
EOM-CCSD shift: -3.00 eV
TP-CCSD shift: -1.40 eV
Intensity (arbitrary units)
529 530 531 532 533 534 535 536 537 538 539
Photon energy (eV)
EOM-CCSD shift: -1.80 eV
TP-CCSD shift: -1.00 eV
Intensity (arbitrary units)
283 284 285 286 287 288 289 290 291
Photon energy (eV)

Ultrafast X-Ray Spectroscopy

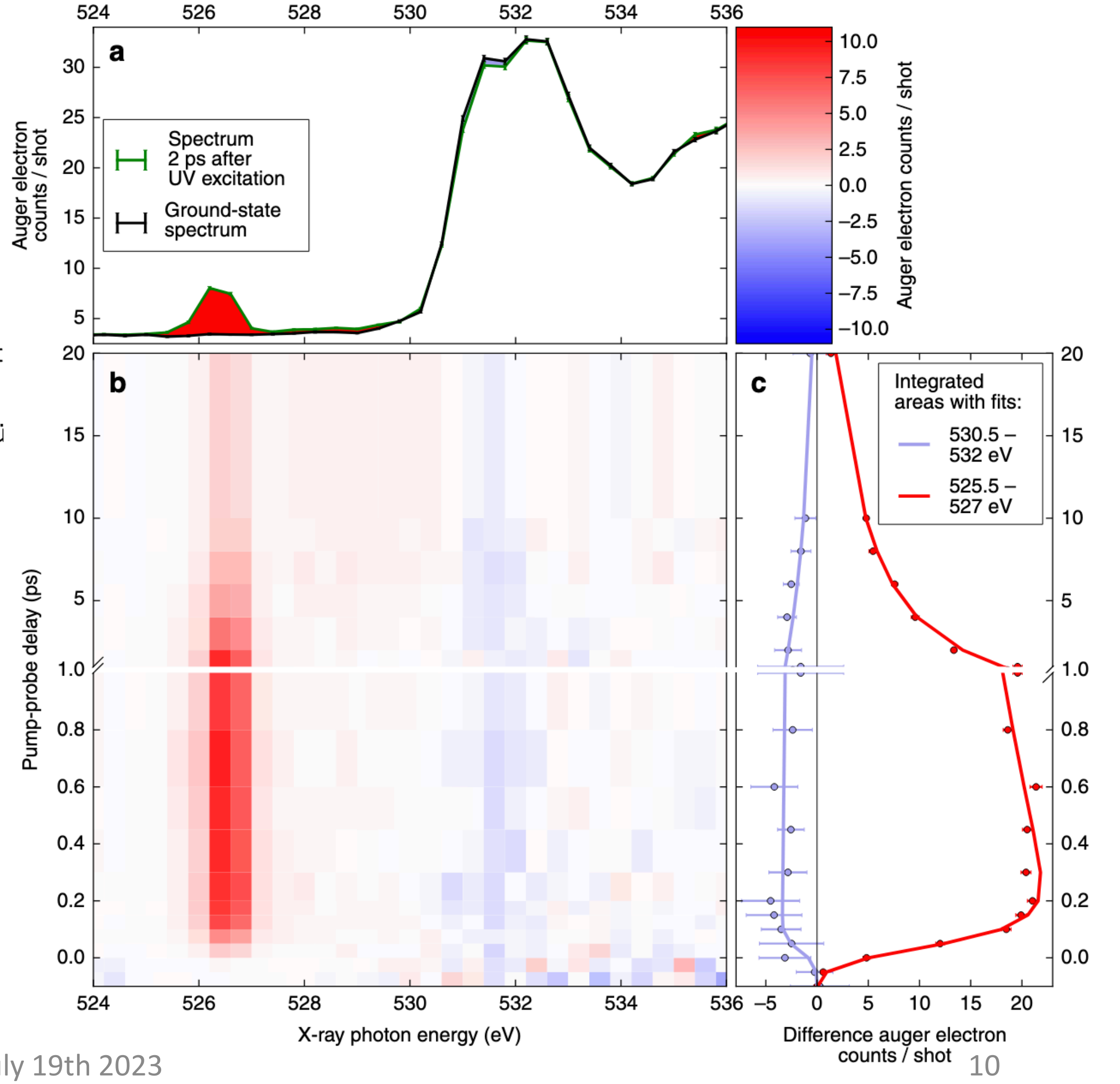
DOI: 10.1038/s41467-017-00069-7 OPEN

Probing ultrafast $\pi\pi^*/n\pi^*$ internal conversion in organic chromophores via K-edge resonant absorption



dt

Gü



Computational Methodology

- **Orbital relaxation** is a dominant effect: how to explicitly or implicitly handle this?
- How to recover **electron correlation**?
- What is the optimal **basis set**?
- How to study the effect of the **environment**?

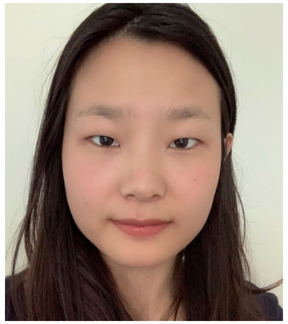

Catherine Wright


Duc Anh Lai

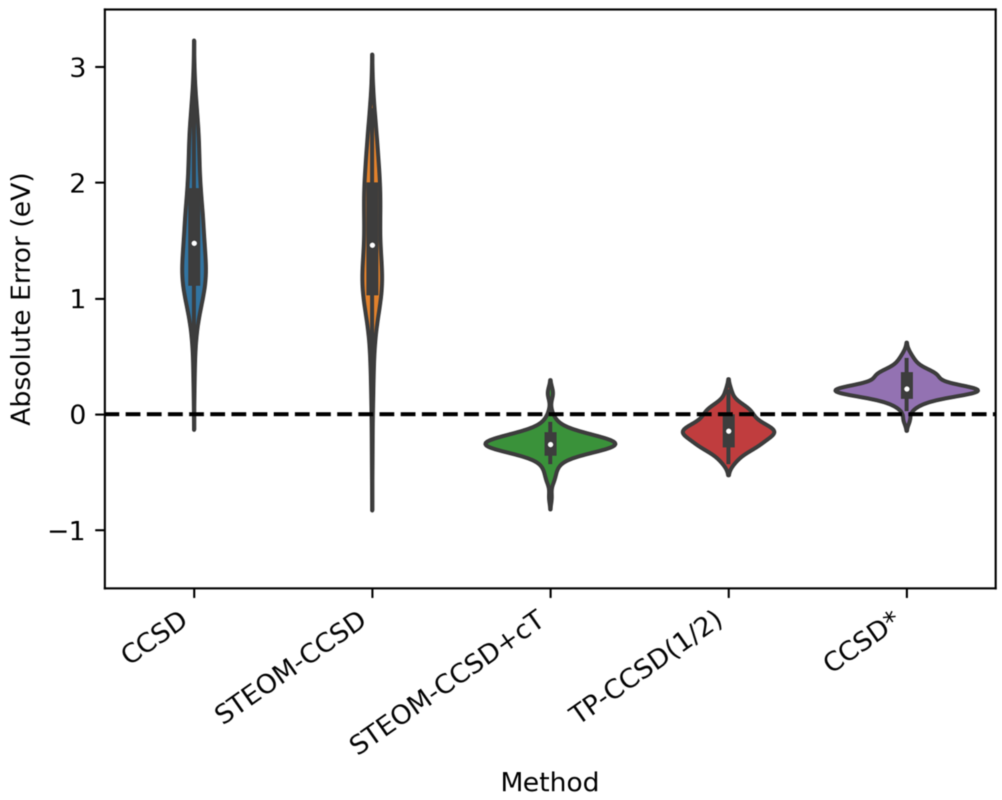

Dr. Avdhoot Datar

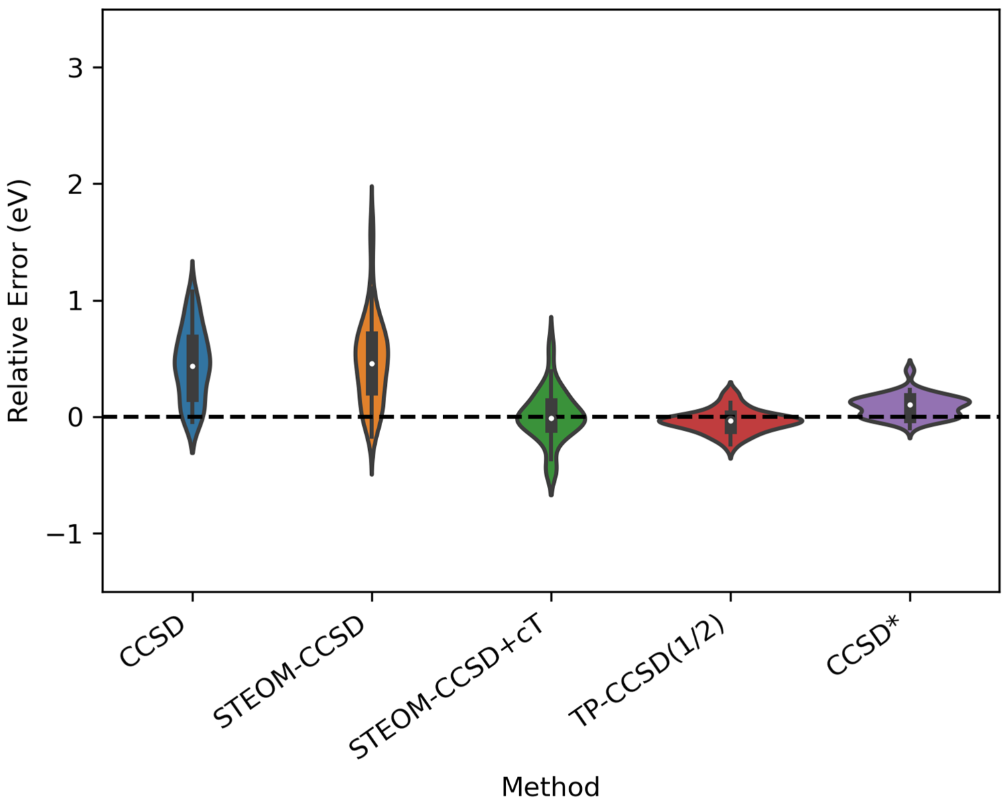

Dr. Megan Simons


Dr. Alexis Delgado

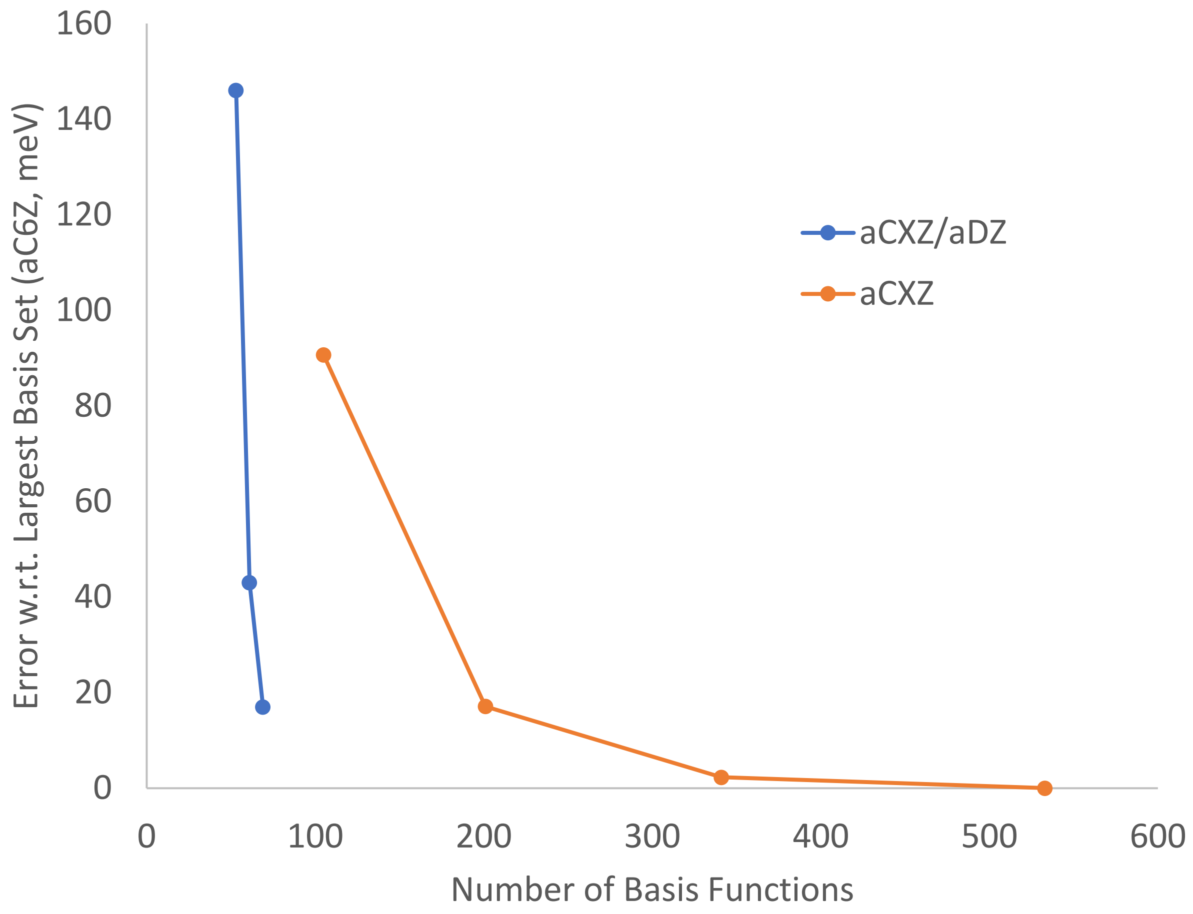

Tingting Zhao

Computational Methodology

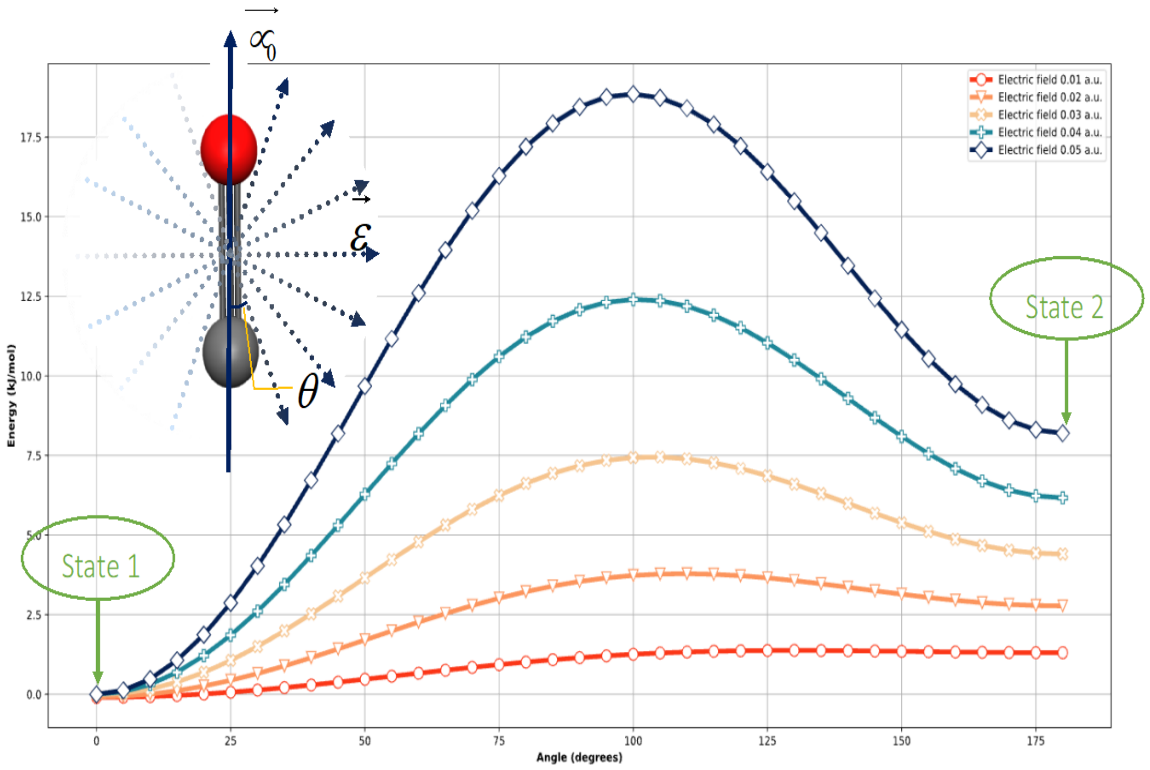


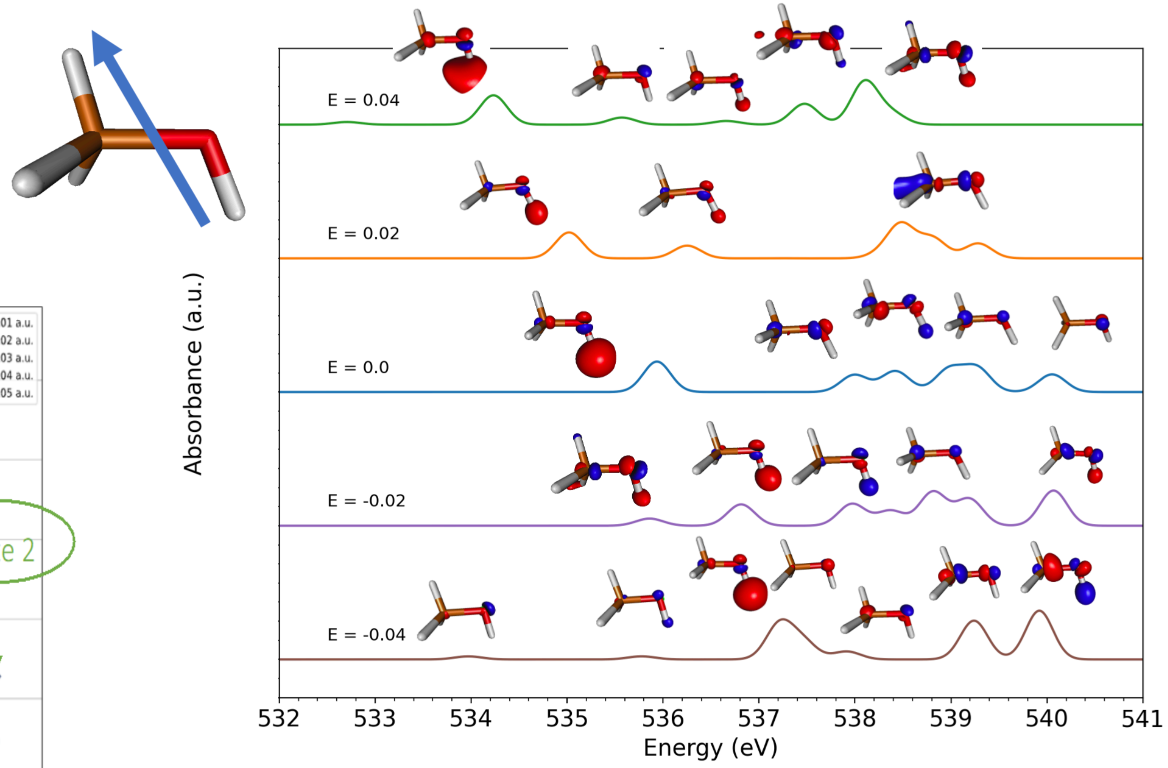


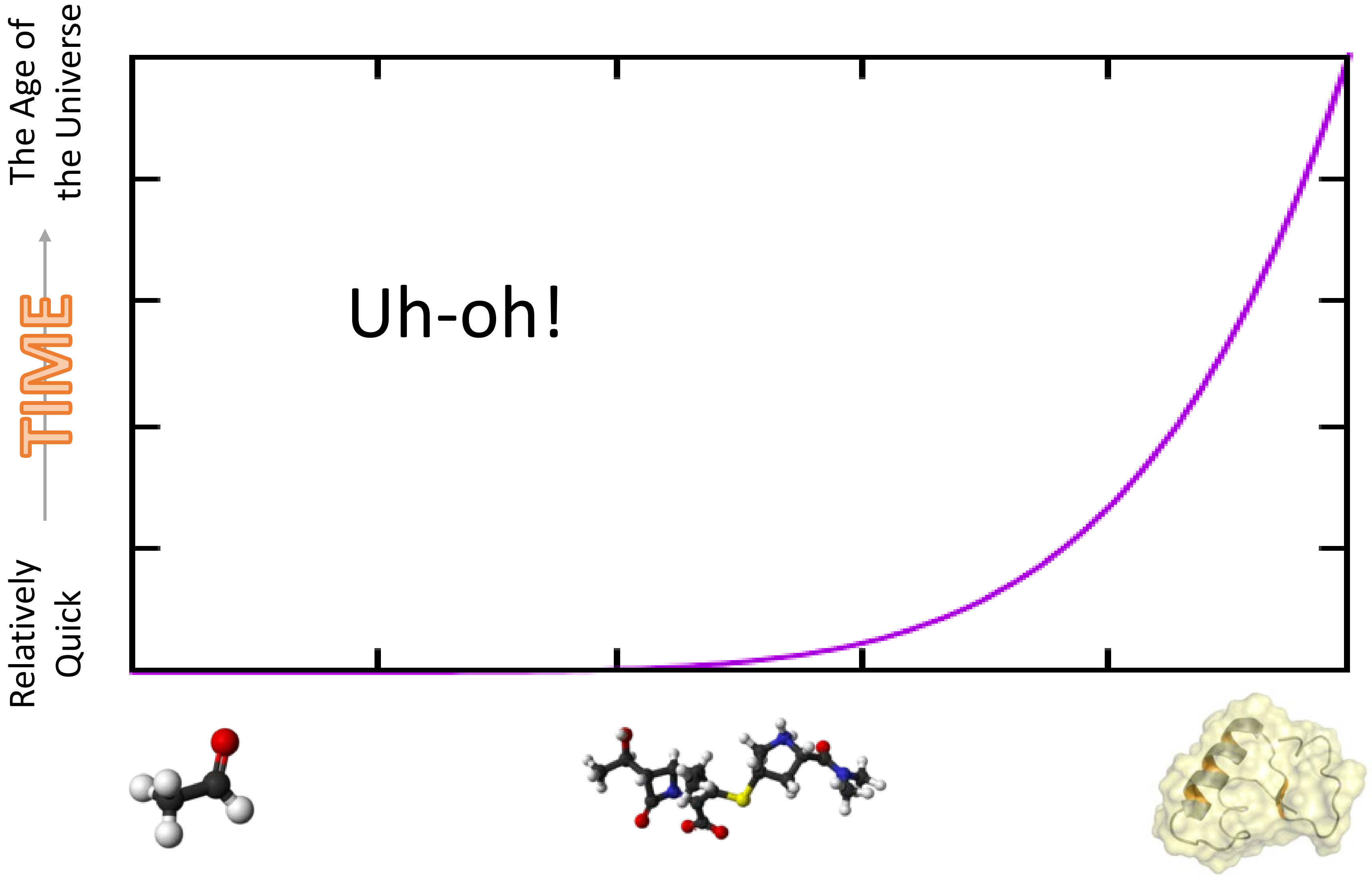
Computational Methodology



Computational Methodology





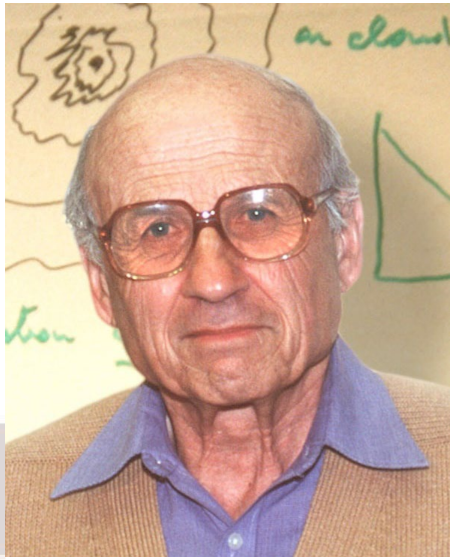

The Age of
the Universe
TIME
Relatively
Quick
Uh-oh!

Tensor Factorization

Density Functional and Density Matrix Method Scaling Linearly with the Number of Atoms

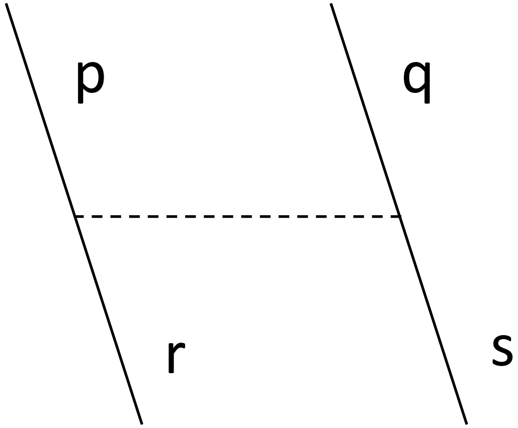
W. Kohn
Phys. Rev. Lett. **76**, 3168 – Published 22 April 1996

Article | References | Citing Articles (694) | PDF | Export Citation

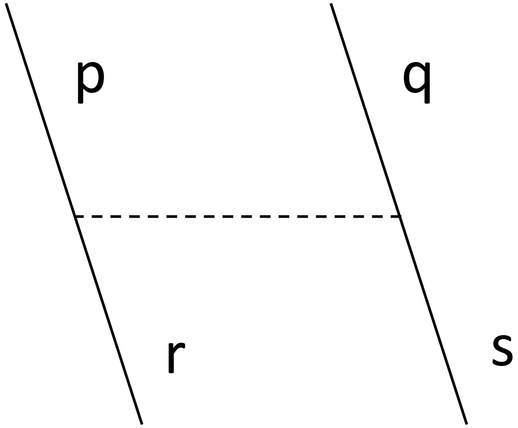


I first discuss a widely applicable physical principle which explains why $O(N)$ methods can exist. I call this principle the *nearsightedness* of equilibrium systems consisting of *many* quantum mechanical particles moving in an external potential $v(r)$.

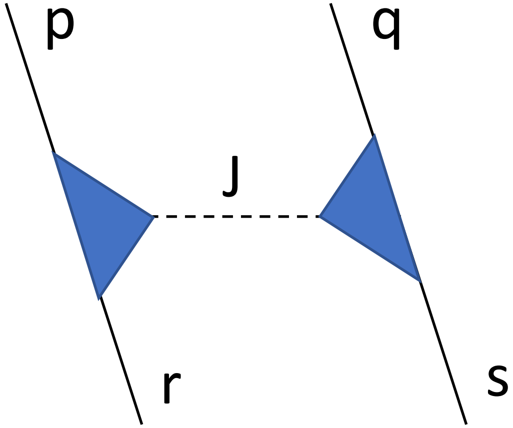
Tensor Factorization



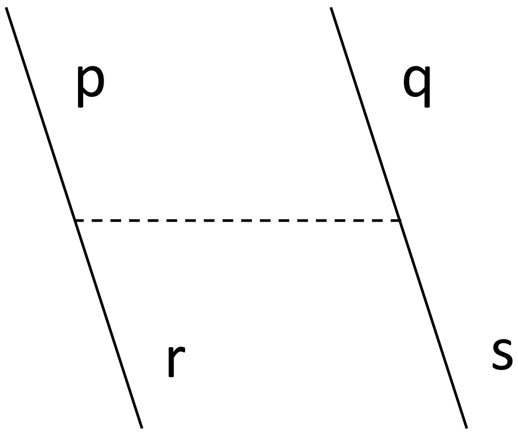
Tensor Factorization



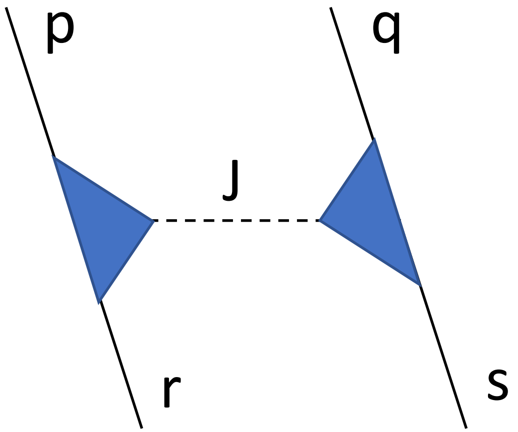
$$\langle pq|rs\rangle \approx \sum_{J} B_{pr}^{J} B_{qs}^{J}$$



Tensor Factorization



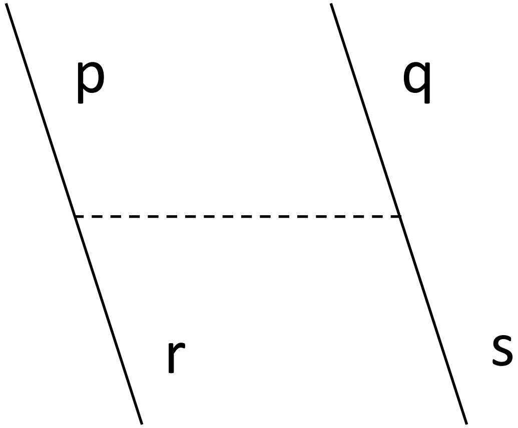
$$\langle pq|rs\rangle \approx \sum_J B^J_{pr} B^J_{qs}$$



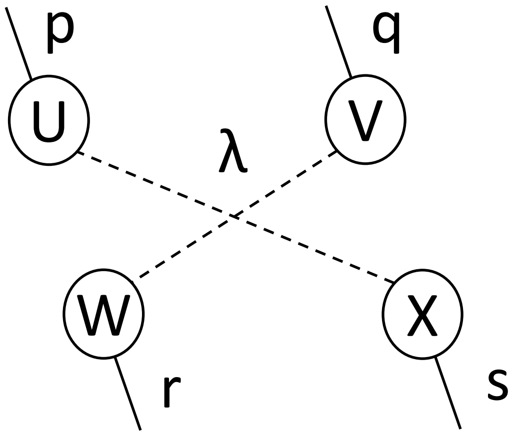
= Density Fitting, Cholesky Decomposition, RI, etc.

Reduced storage, but **NOT** reduced cost!

Tensor Factorization



$$\langle pq|rs\rangle \approx \sum_{\lambda} U_p^{\lambda} V_q^{\lambda} W_r^{\lambda} X_s^{\lambda}$$

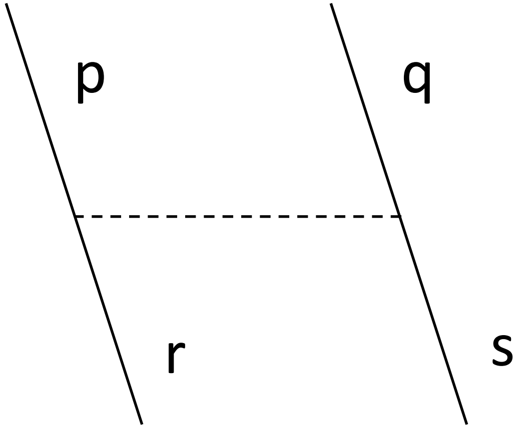


Tensor decomposition in post-Hartree–Fock methods. I. Two-electron integrals and MP2

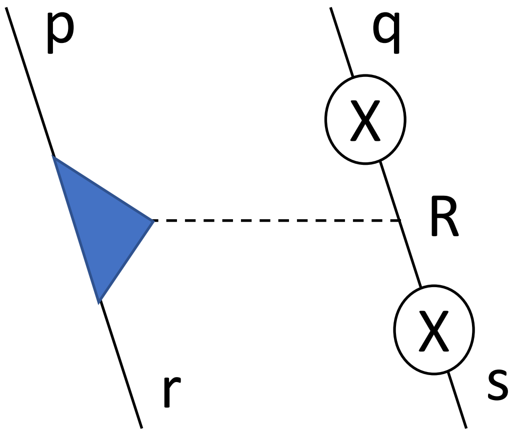
Cite as: J. Chem. Phys. **134**, 054118 (2011); https://doi.org/10.1063/1.3514201
Submitted: 02 August 2010 . Accepted: 19 October 2010 . Published Online: 07 February 2011

Udo Benedikt, Alexander A. Auer, Mike Espig, and Wolfgang Hackbusch

Tensor Factorization



$$\langle pq|rs\rangle \approx \sum_J (B\gamma)_{pr}^R X_q^R X_s^R$$



Communication: Acceleration of coupled cluster singles and doubles via orbital-weighted least-squares tensor hypercontraction

J. Chem. Phys. **140**, 181102 (2014); https://doi.org/10.1063/1.4876016

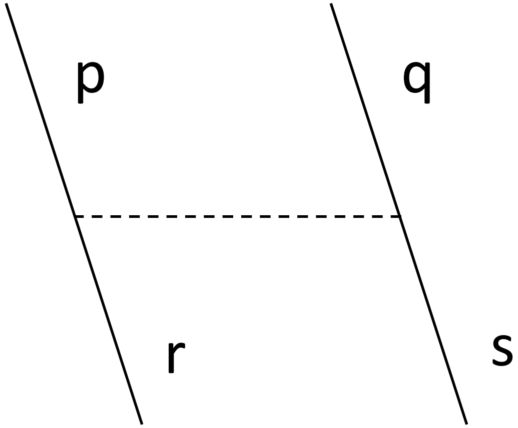
Robert M. Parrish[1], C. David Sherrill[1, a)], Edward G. Hohenstein[2], Sara I. L. Kokkila[2], *and* Todd J. Ma

Robust Approximation of Tensor Networks: Application to Grid-Free Tensor Factorization of the Coulomb Interaction

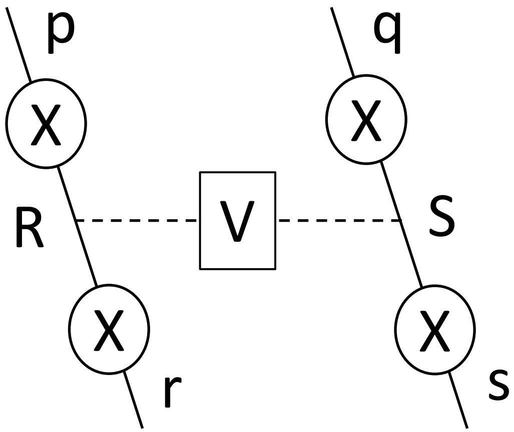
Karl Pierce, Varun Rishi, and Edward F. Valeev*

Cite this: *J. Chem. Theory Comput.* 2021, 17, 4, 2217–2230
Publication Date: March 29, 2021
https://doi.org/10.1021/acs.jctc.0c01310

Tensor Factorization



$$\langle pq|rs\rangle \approx \sum_{RS} X_p^R X_r^R V_{RS} X_q^S X_s^S$$

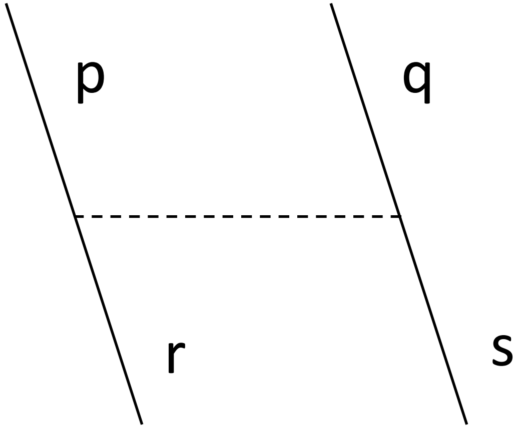


Tensor hypercontraction density fitting. I. Quartic scaling second- and third-order Møller-Plesset perturbation theory

J. Chem. Phys. **137**, 044103 (2012); https://doi.org/10.1063/1.4732310

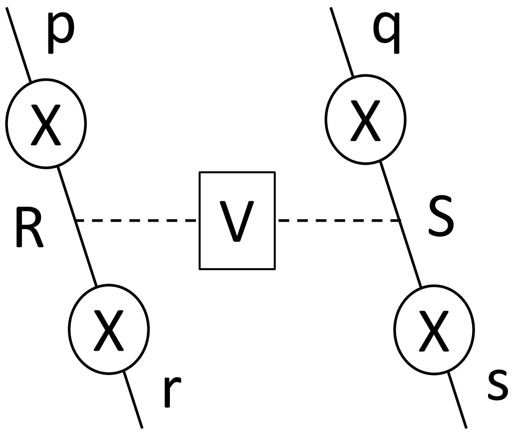
Edward G. Hohenstein[1,2], Robert M. Parrish[3], *and* Todd J. Martínez[1,2]

Tensor Factorization



A Grid!

$$\langle pq|rs\rangle \approx \sum_{RS} X_p^R X_r^R V_{RS} X_q^S X_s^S$$



Tensor hypercontraction. II. Least-squares renormalization

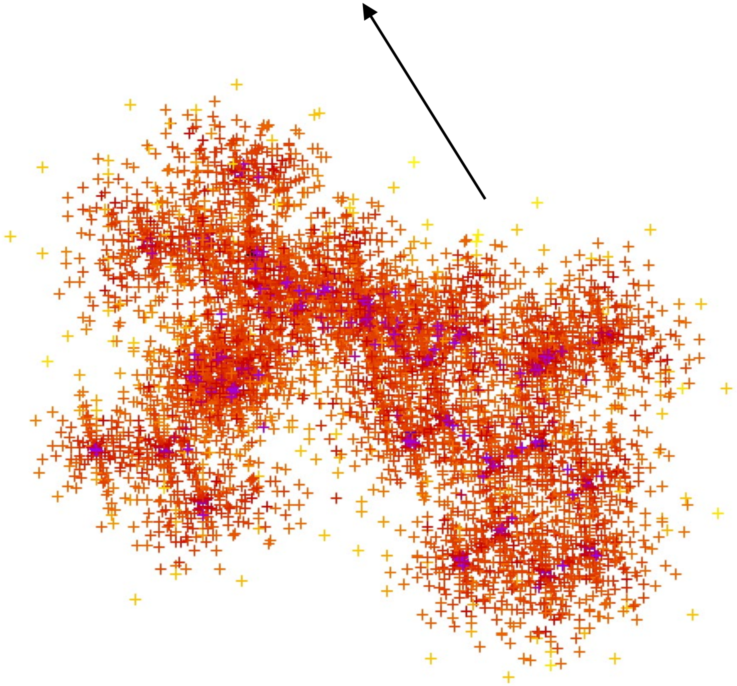
J. Chem. Phys. **137**, 224106 (2012); https://doi.org/10.1063/1.4768233

Robert M. Parrish[1], Edward G. Hohenstein[2,3], Todd J. Martínez[2,3, a)], *and* C. David Sherrill[1,4, b)]

Tensor Factorization

$$\langle pq|rs\rangle \approx \sum_{RS} X_p^R X_r^R V_{RS} X_q^S X_s^S$$

$$X_p^R = \psi_p(x_R)$$



$$\langle pq|rs\rangle = \iint \psi_p(r_1)\psi_r(r_1)\frac{1}{|r_1 - r_2|}\psi_q(r_2)\psi_s(r_2)dr_1 dr_2$$

$$\approx \sum_{R\neq S} X_p^R X_r^R \frac{w_R w_S}{|r_R - r_S|} X_q^S X_s^S + \text{“diagonal term”}$$

$$\approx \sum_{RS} X_p^R X_r^R V_{RS} X_q^S X_s^S$$

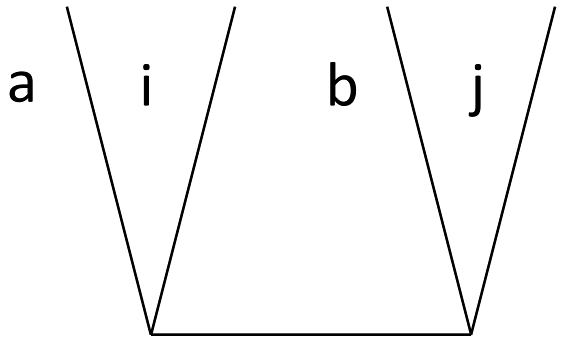
Quite similar to DFT grids:

$$E_{xc} = \int f(\rho(r), \tau(r), \dots)dr$$

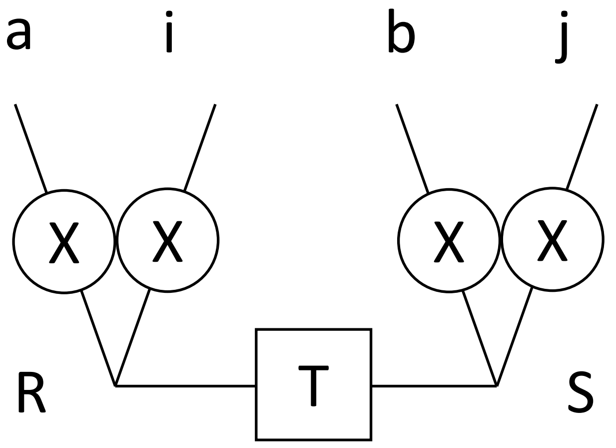
$$\approx \textstyle\sum_R f(\rho(x_R), \tau(x_R), \dots)w_R$$

$$\rho(x_R) = \textstyle\sum_{\mu\nu} X_\mu^R X_\nu^R P_{\mu\nu}$$

Tensor Factorization



$$t_{ij}^{ab} \approx \sum_{RS} X_a^R X_i^R T_{RS} X_b^S X_j^S$$



Communication: Tensor hypercontraction. III. Least-squares tensor hypercontraction for the determination of correlated wavefunctions

Edward G. Hohenstein,[1,2] Robert M. Parrish,[3] C. David Sherrill,[3] and Todd J. Martínez[1,2]
[1]*Department of Chemistry and the PULSE Institute, Stanford University, Stanford, California 94305, USA*
[2]*SLAC National Accelerator Laboratory, Menlo Park, California 94025, USA*
[3]*Center for Computational Molecular Science and Technology, School of Chemistry and Biochemistry, and School of Computational Science and Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0400, USA*

(Received 8 October 2012; accepted 5 November 2012; published online 11 December 2012)

Systematically Improvable Tensor Hypercontraction: Interpolative Separable Density-Fitting for Molecules Applied to Exact Exchange, Second- and Third-Order Møller–Plesset Perturbation Theory

Joonho Lee,[*,†,‡] Lin Lin,[§,∥] and Martin Head-Gordon[*,†,‡]
[†]Department of Chemistry, University of California, Berkeley, California 94720, United States
[‡] Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States
[§]Department of Mathematics, University of California, Berkeley, California 94720, United States
[∥]Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States

A critical analysis of least-squares tensor hypercontraction applied to MP3

Cite as: J. Chem. Phys. 154, 134102 (2021); doi: 10.1063/5.0038764
Submitted: 25 November 2020 • Accepted: 17 March 2021 •
Published Online: 1 April 2021

View Online

Devin A. Matthews[a)]

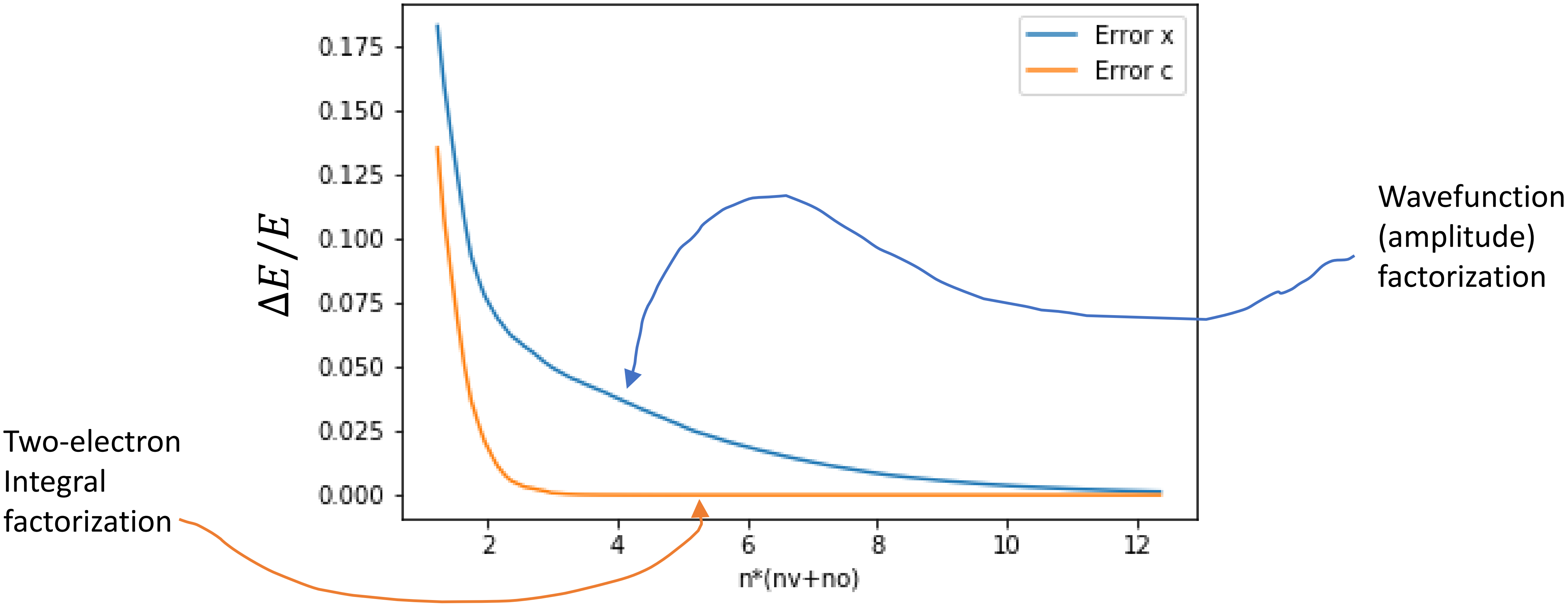
Rank-reduced coupled-cluster. III. Tensor hypercontraction of the doubles amplitudes

Cite as: J. Chem. Phys. 156, 054102 (2022); doi: 10.1063/5.0077770
Submitted: 6 November 2021 • Accepted: 6 January 2022 •
Published Online: 1 February 2022

View Online Export C

Edward G. Hohenstein,[1,2,a)] B. Scott Fales,[1,2] Robert M. Parrish,[3] and Todd J. Martínez[1,2]

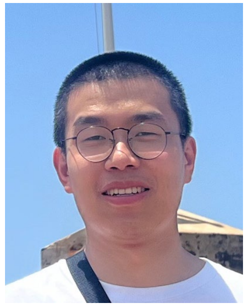
Tensor Factorization



Which Basis Functions?


Sara Beth Becker


Dr. James Thorpe


Chao Yin

Option 1:

$$T_2 = U\Sigma V^T$$

$$X_a^R X_i^R = Y_{ai}^R = QR$$

$$f = |U\Sigma V^T Q|_F^2 / |\Sigma|_F^2$$

Option 2:

$$\Delta E_{THC}(y) = |E_{THC}(Y \cup y) - E_{THC}(Y)|$$

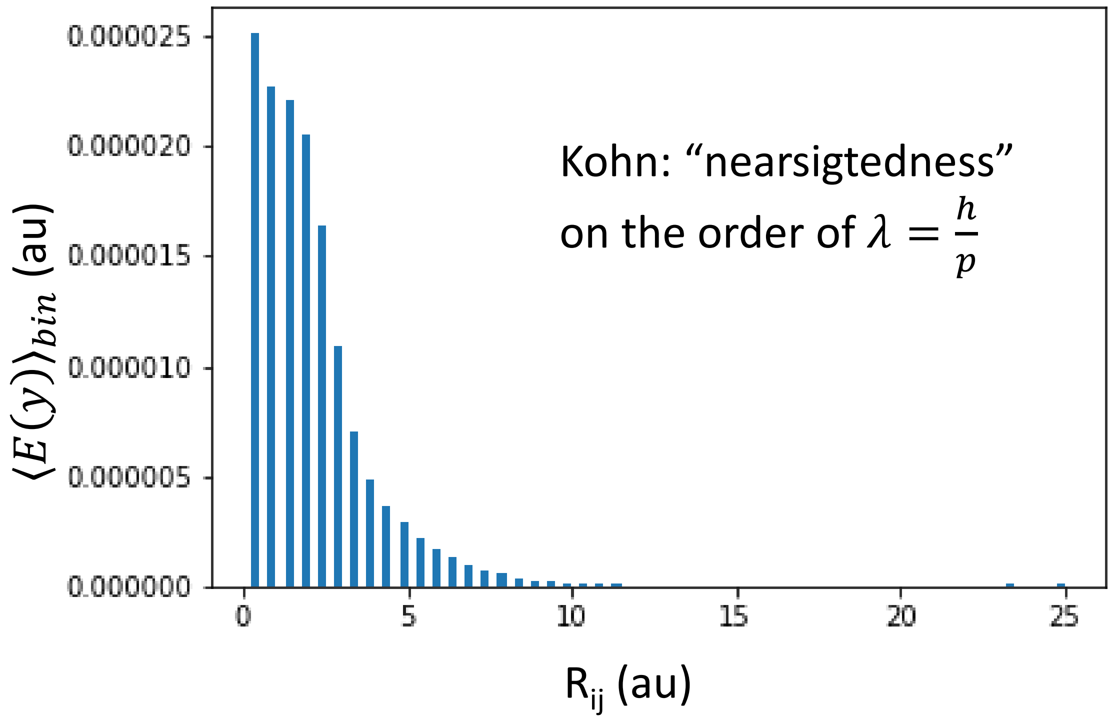
Which Basis Functions?

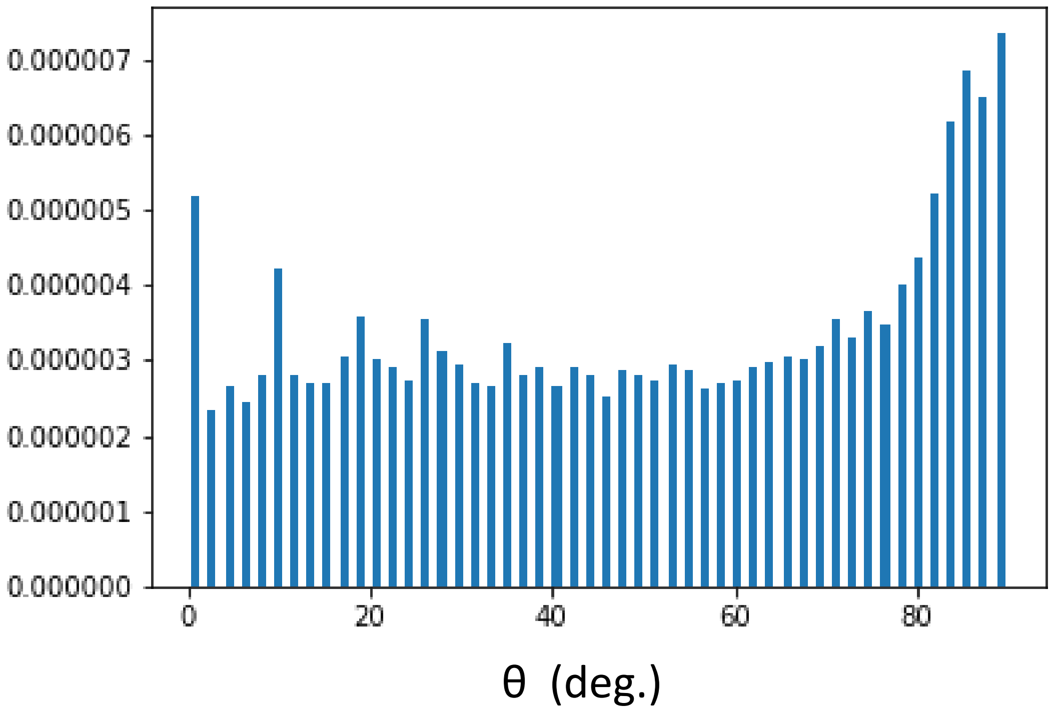
$$E_{THC} = \mathrm{Tr}[\tilde{V} T_{THC}] = \mathrm{Tr}[\tilde{V} Y S^{-1} Y^T T Y S^{-1} Y^T]$$

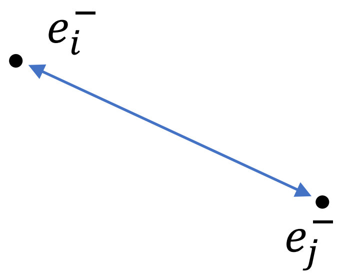
$$\Delta E_{THC}(y) = 2\mu^{-1} \mathrm{Tr}[\tilde{V} B B^T T Y S^{-1} Y^T] + \mu^{-2} \mathrm{Tr}[\tilde{V} B B^T T B B^T]$$

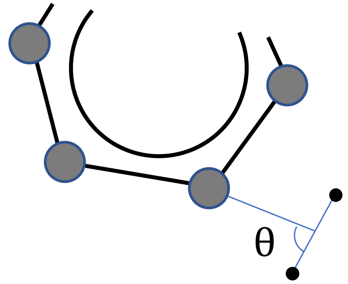
$$B = (I - Y S^{-1} Y^T) y$$
$$\mu = y^T (I - Y S^{-1} Y^T) y$$

Which Basis Functions?

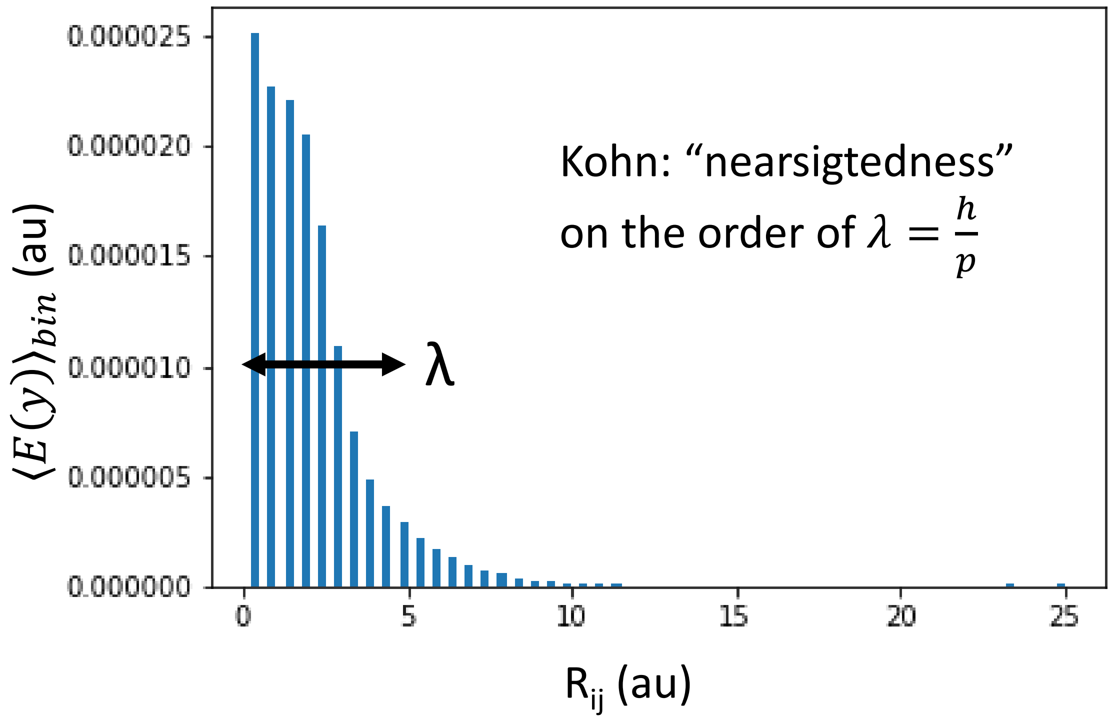






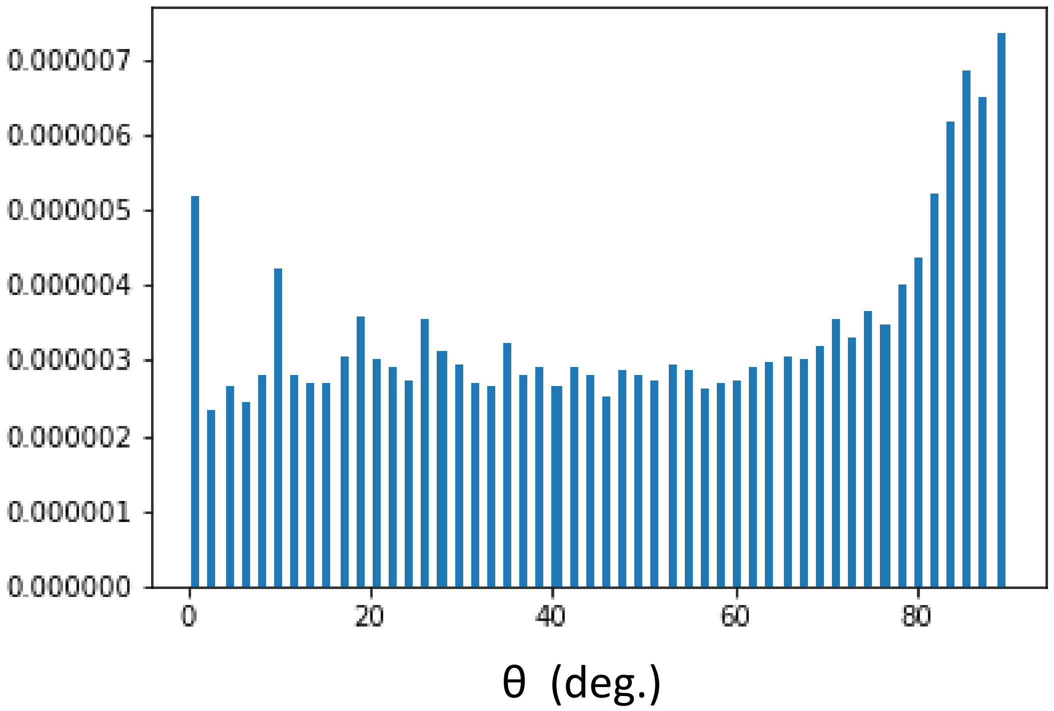


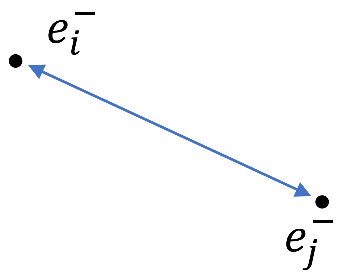
Which Basis Functions?

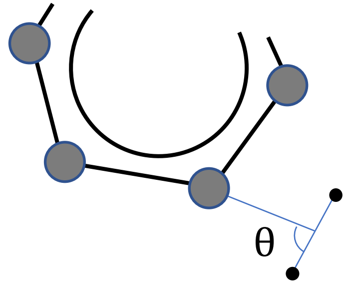


Kohn: "nearsigtedness"

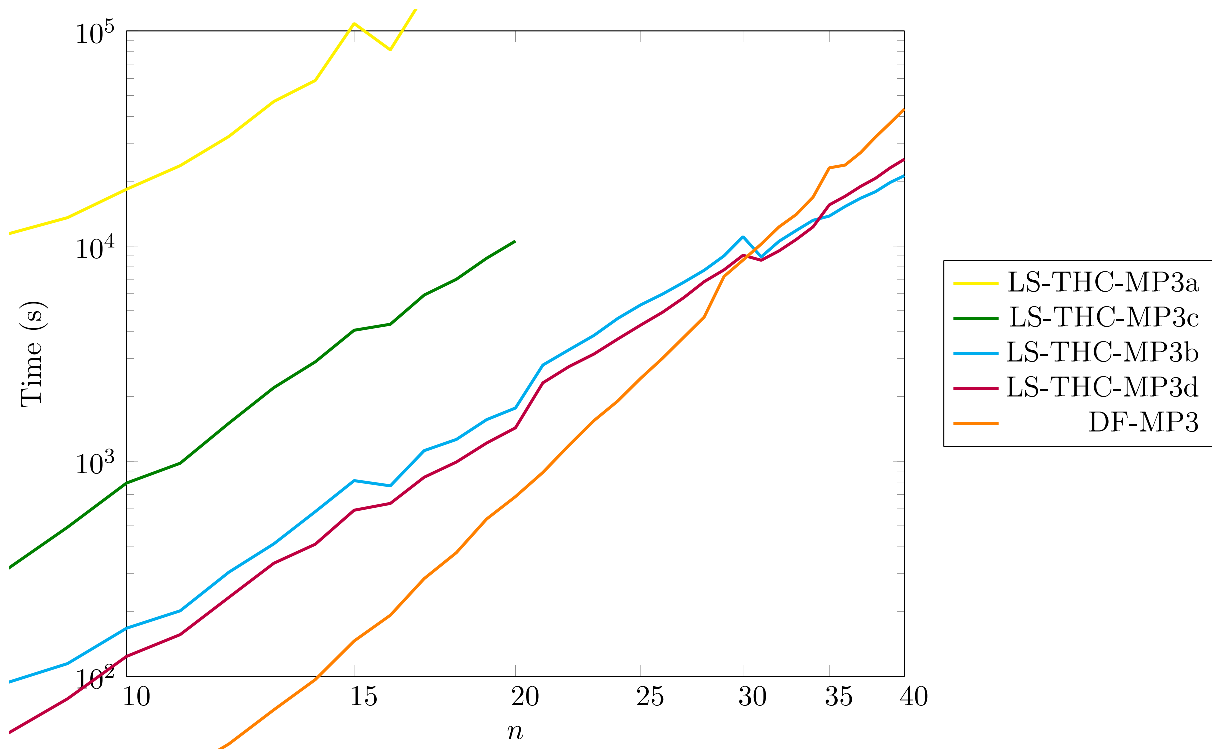
on the order of $\lambda = \frac{h}{p}$



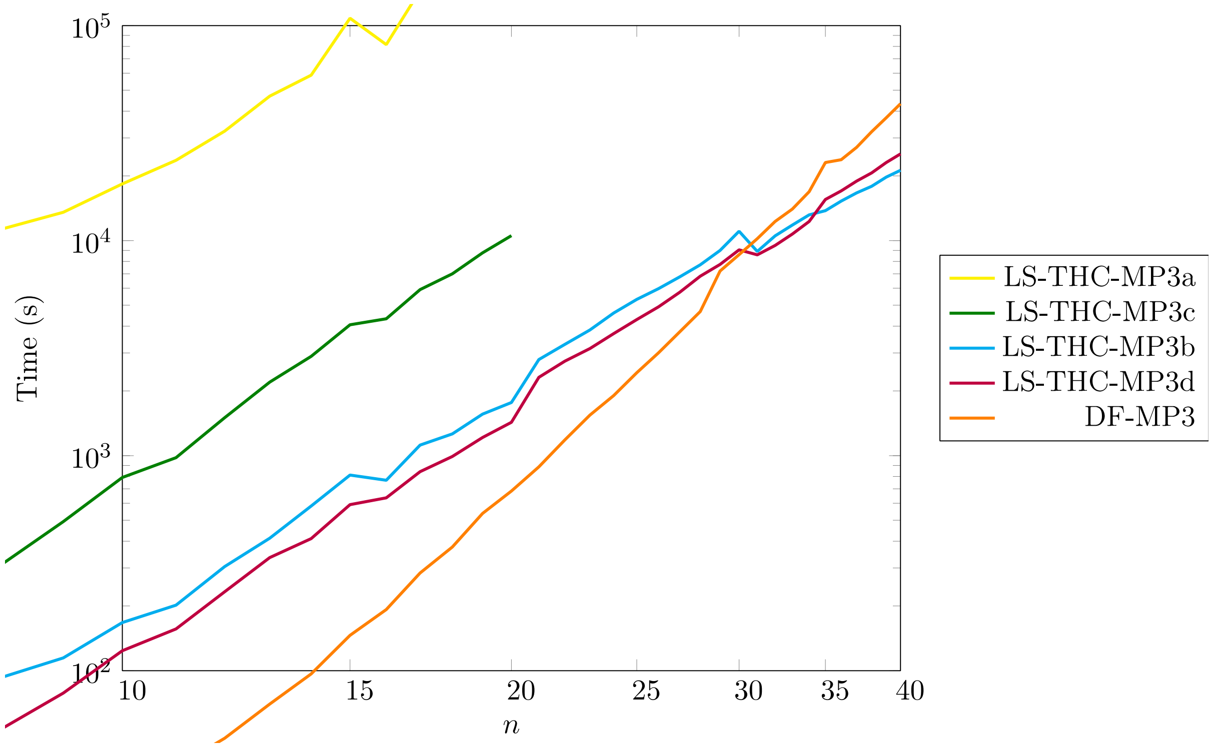


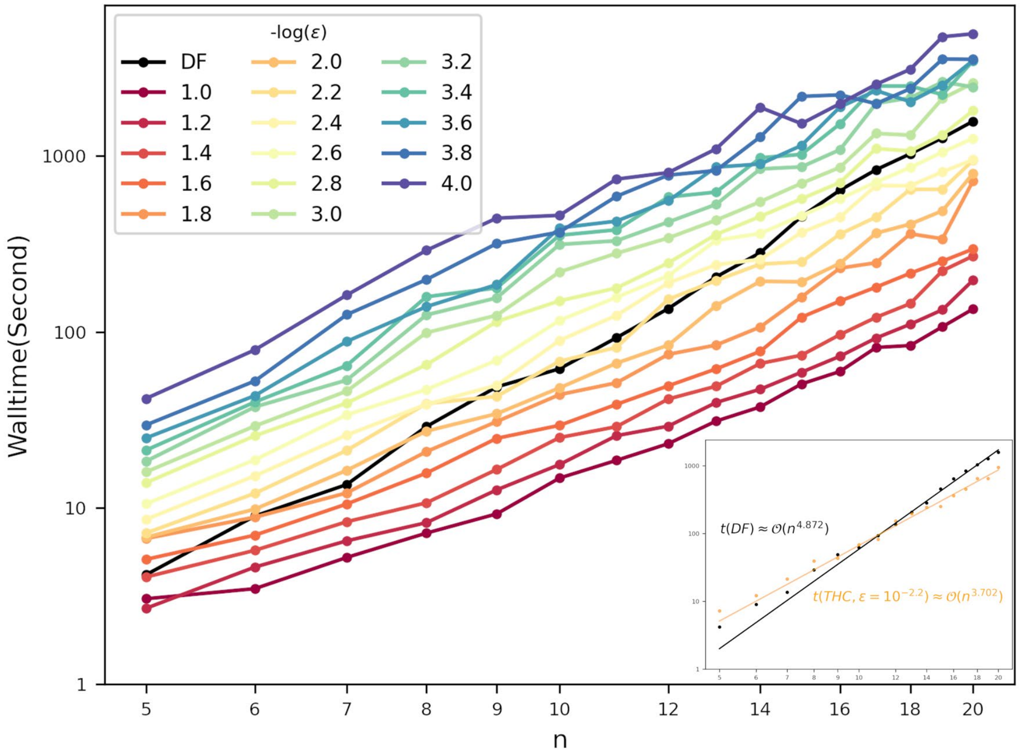


HPC Implementation

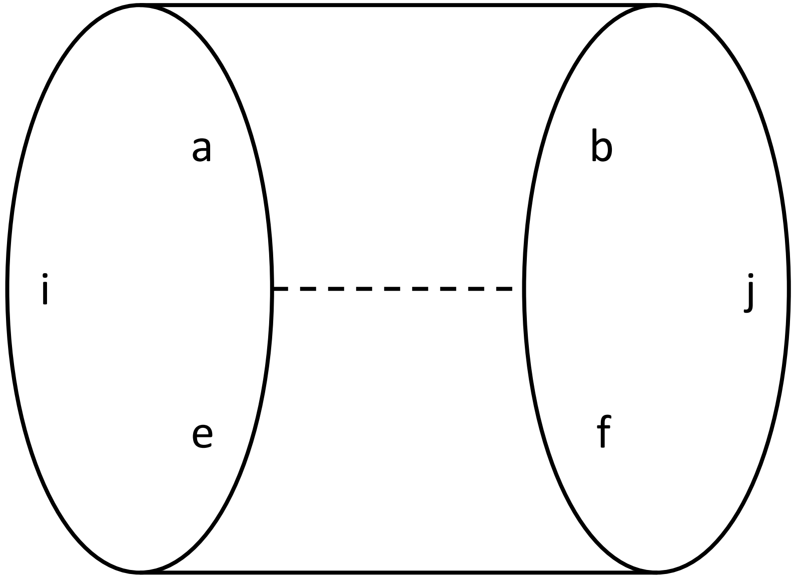


HPC Implementation

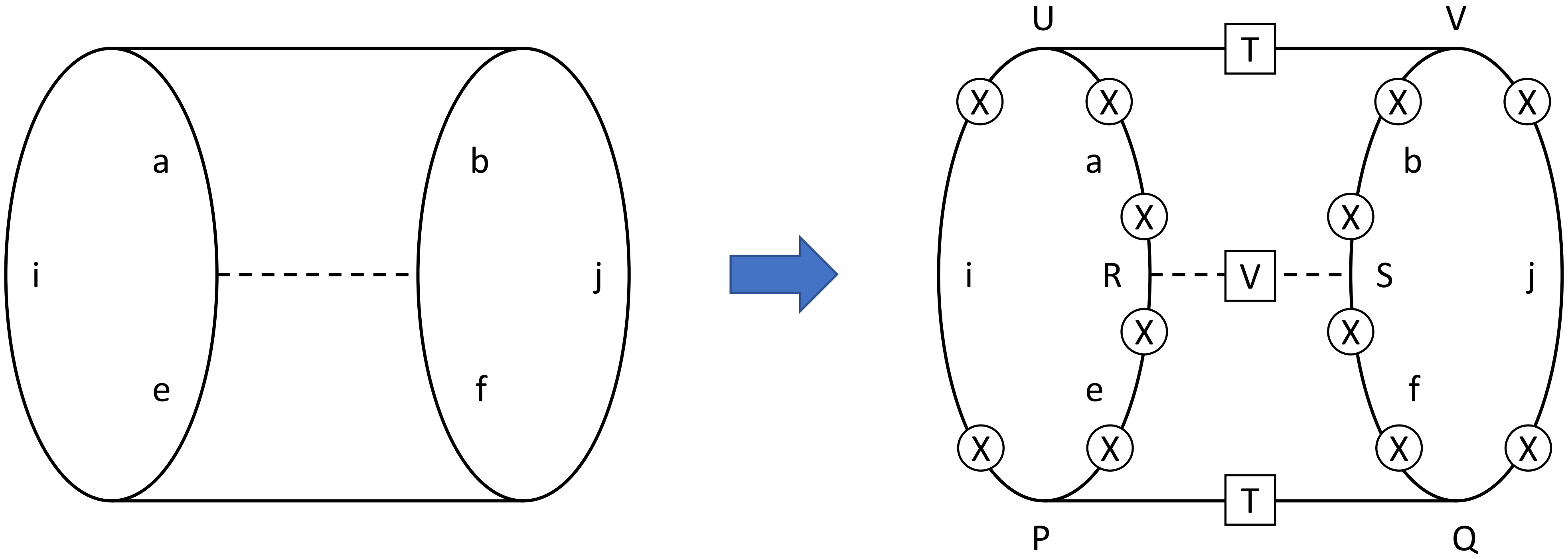




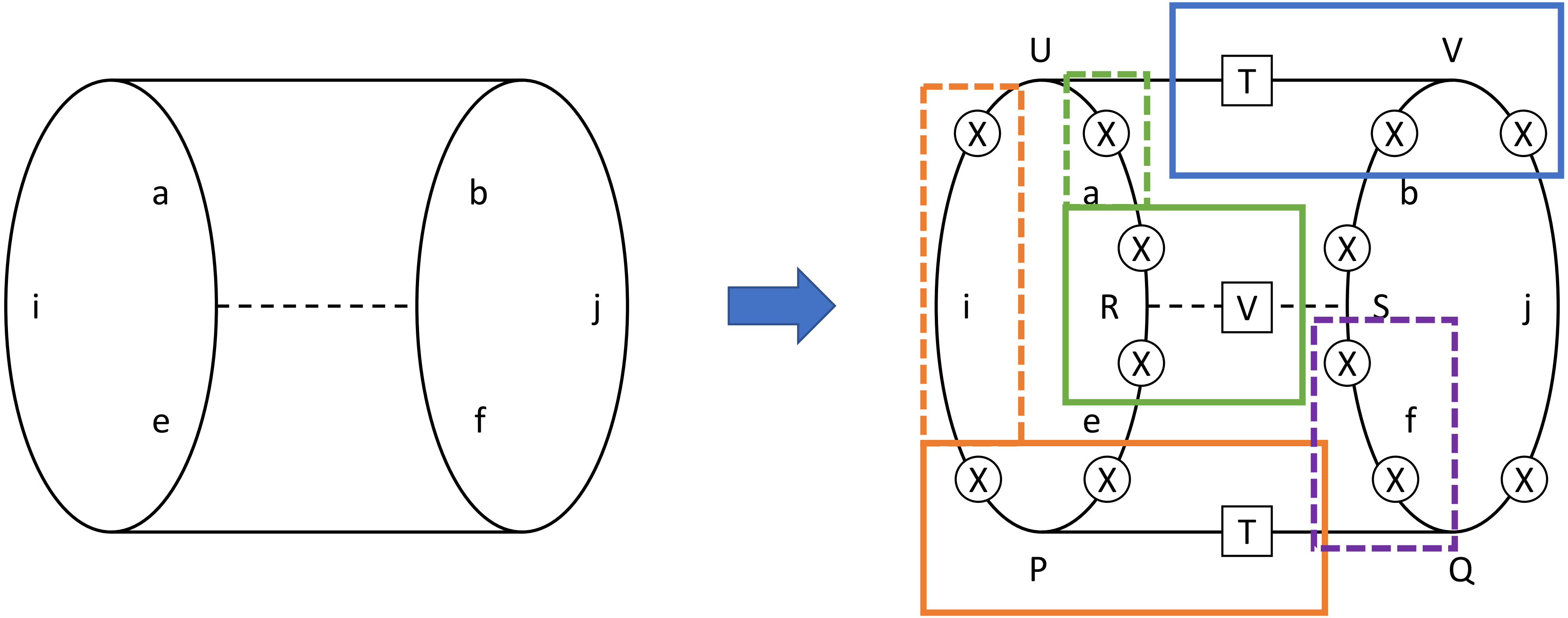
HPC Implementation



HPC Implementation

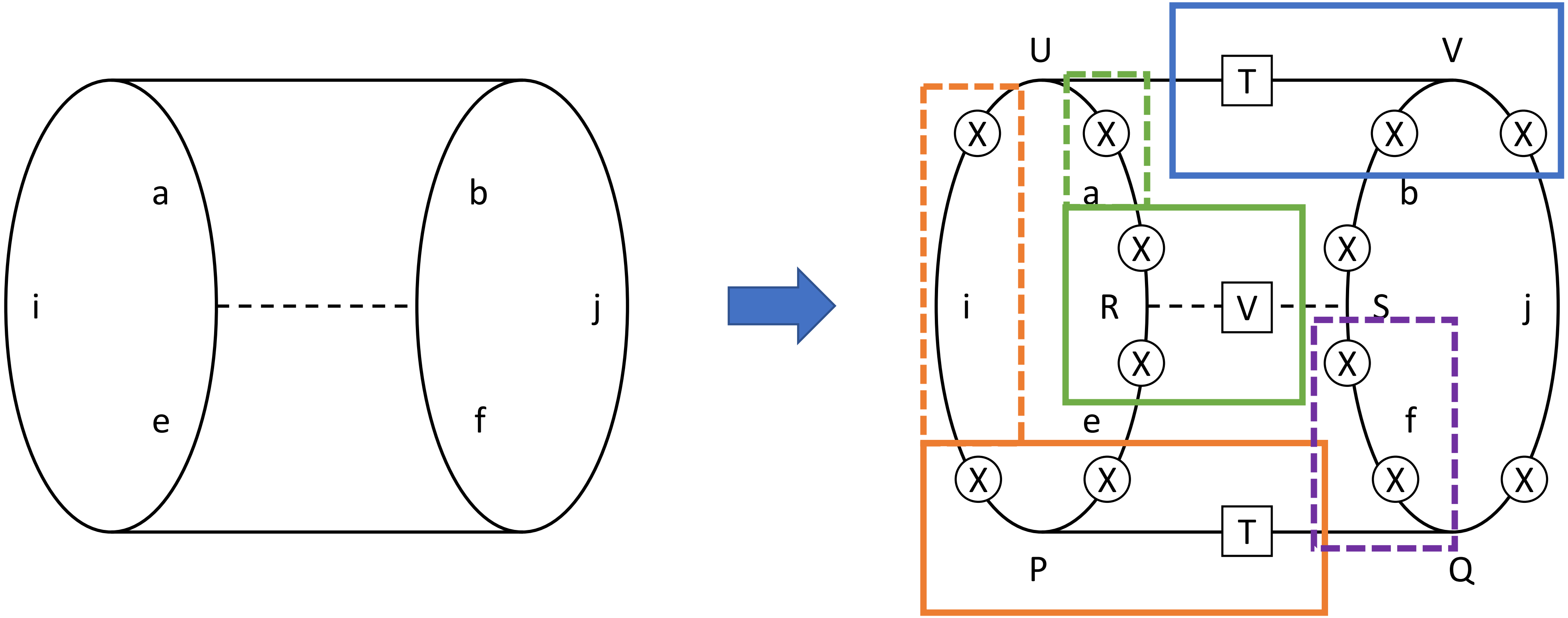


HPC Implementation



$$E_{PPC} = 2\sum_{Ubj}\left(\sum_{S}\left(\sum_{Q}\left(\sum_{e}\left(\sum_{a}\left(\sum_{R}\tilde{\tilde{X}}_a^R\tilde{\tilde{X}}_e^R\tilde{\tilde{V}}_{RS}\right)\tilde{X}_a^U\right)\left(\sum_{i}\left(\sum_{P}\tilde{X}_e^P\tilde{X}_i^P T_{PQ}^{[1]}\right)\tilde{X}_i^U\right)\right)\left(\sum_{f}\tilde{\tilde{X}}_f^S\tilde{X}_f^Q\right)\tilde{X}_j^Q\right)\tilde{\tilde{X}}_b^S\right)\left(\sum_{V}\tilde{X}_b^V\tilde{X}_j^V T_{UV}^{[1]}\right)$$

HPC Implementation



$$E_{PPC} = 2\sum_{Ubj}\left(\sum_{S}\left(\sum_{Q}\left(\sum_{e}\left(\sum_{a}\left(\sum_{R}\tilde{\tilde{X}}_a^R\tilde{\tilde{X}}_e^R\tilde{\tilde{V}}_{RS}\right)\tilde{X}_a^U\right)\left(\sum_{i}\left(\sum_{P}\tilde{X}_e^P\tilde{X}_i^P T_{PQ}^{[1]}\right)\tilde{X}_i^U\right)\right)\left(\sum_{f}\tilde{\tilde{X}}_f^S\tilde{X}_f^Q\right)\tilde{X}_j^Q\right)\tilde{\tilde{X}}_b^S\right)\left(\sum_{V}\tilde{X}_b^V\tilde{X}_j^V T_{UV}^{[1]}\right)$$

? ? BLAS ?

$$\alpha = \sum_{ijk} A_{ik} B_{kj}?$$

$$C_{ij} = \sum_{k} A_{ik} B_{kj} x_j?$$

$$C_{ij} = \sum_{k} A_{ik} x_{j+k}$$

$$D_{ij} = C_{ij} \sum_{k} A_{ik} B_{kj}?$$

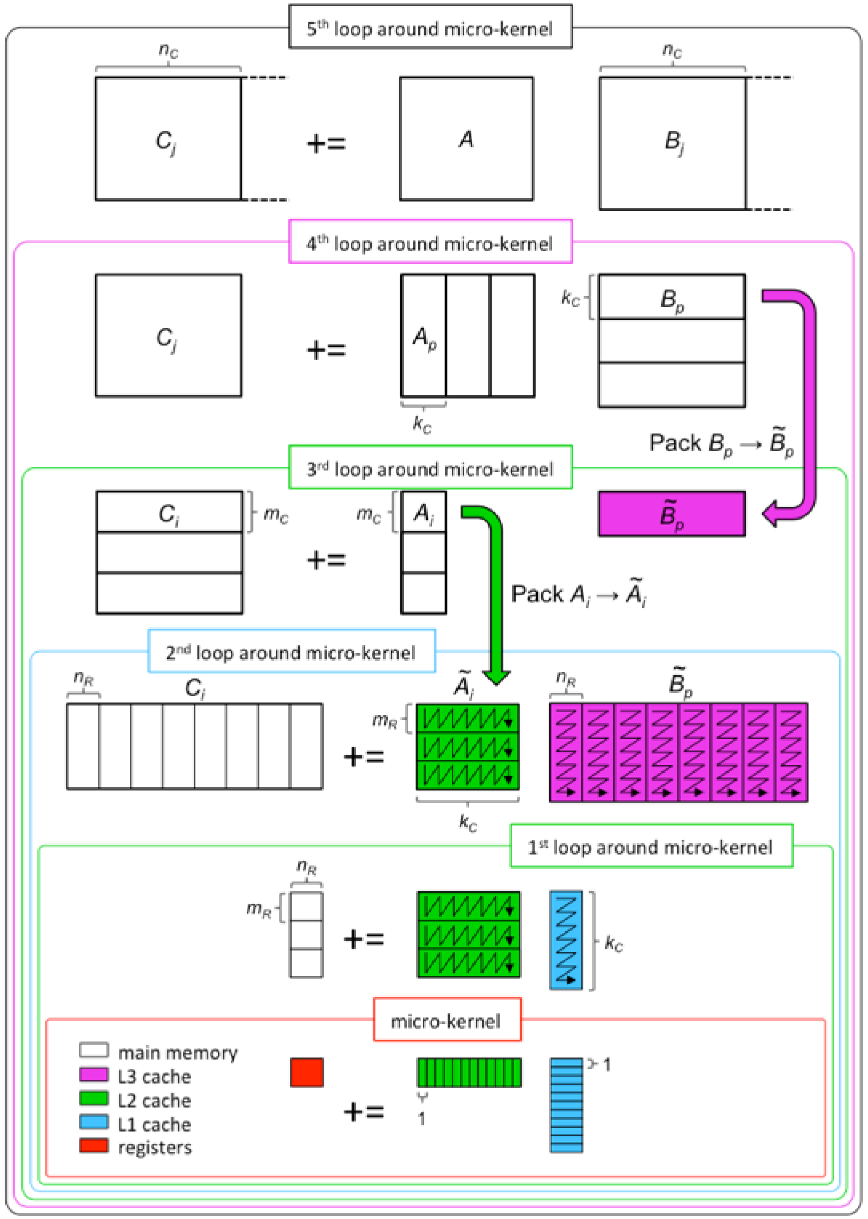
$$z_i = \sum_{j} \exp(x_i^2 + y_j^2 - 2 \sum_{k} A_{ik} B_{kj})$$

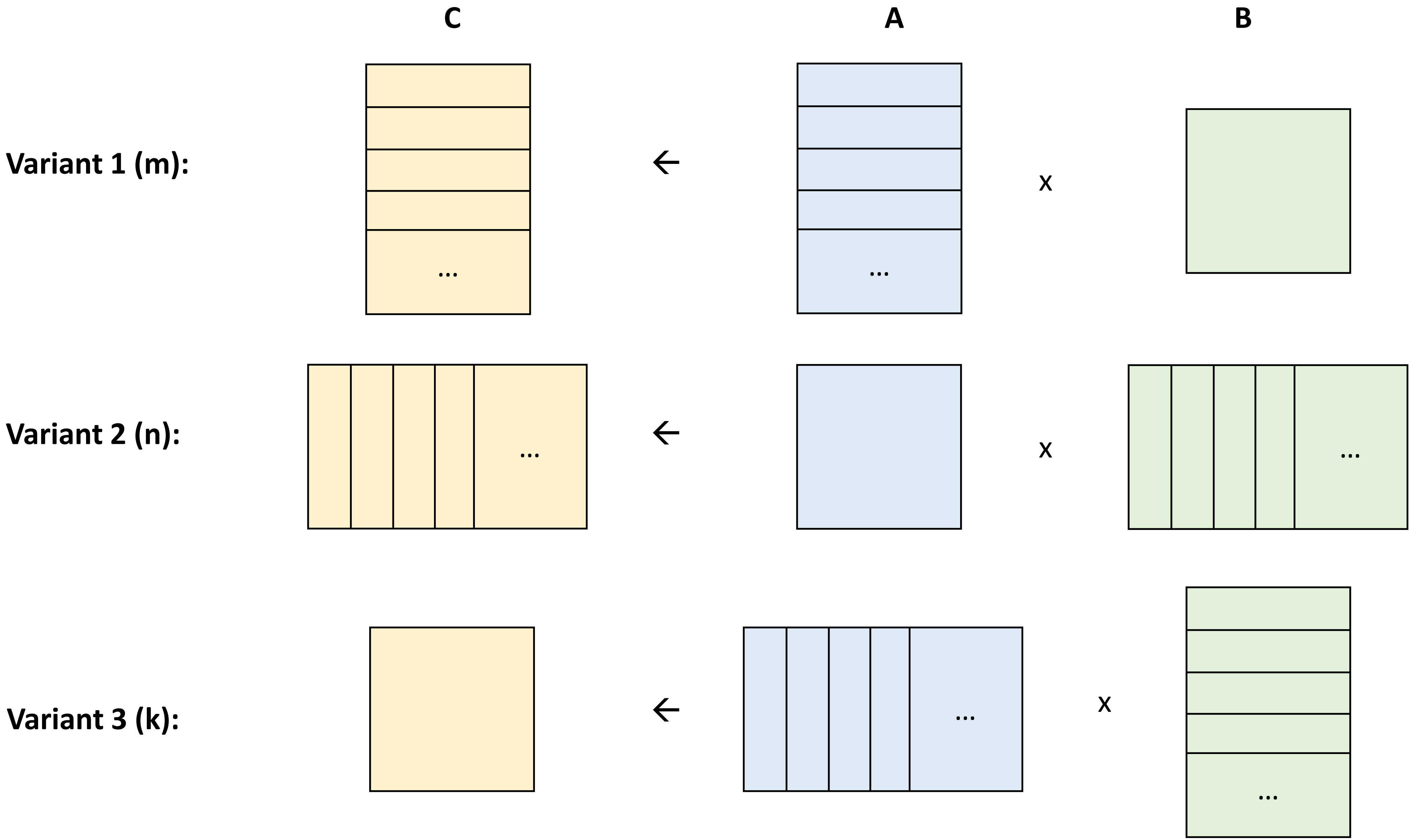
$$\alpha = \sum_{ijk} C_{ij} A_{ik} B_{kj} A_{jk} B_{ki}?$$

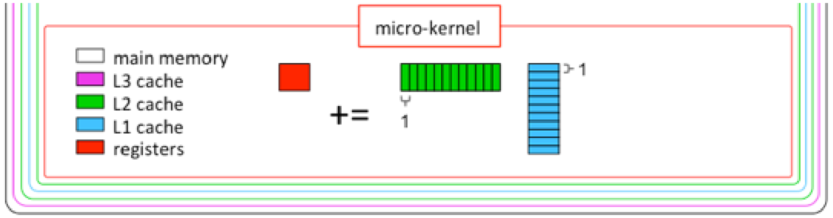
How ~~BLAS~~ BLIS Works

BLAS-Like
Library
Instantiation
Software

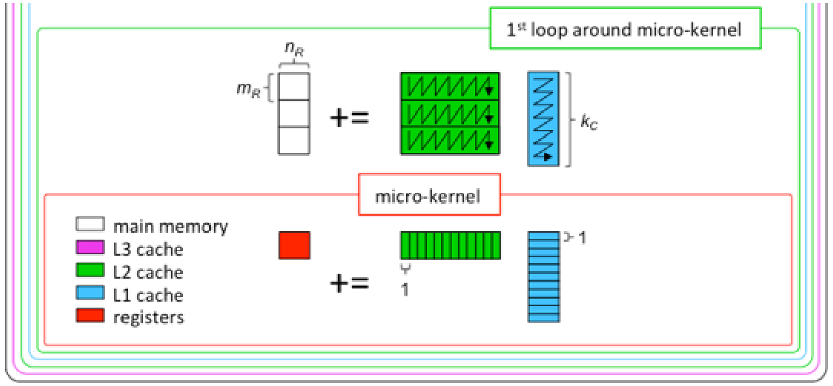
Field Van Zee


5th loop around micro-kernel
n_C
C_j
+=
A
B_j
4th loop around micro-kernel
C_j
+=
A_p
k_C
B_p
Pack $B_p \rightarrow \tilde{B}_p$
3rd loop around micro-kernel
C_i
m_C
A_i
$\tilde{B}_p$
Pack $A_i \rightarrow \tilde{A}_i$
2nd loop around micro-kernel
n_R
C_i
$\tilde{A}_i$
m_R
$\tilde{B}_p$
1st loop around micro-kernel
micro-kernel
main memory
L3 cache
L2 cache
L1 cache
registers
1

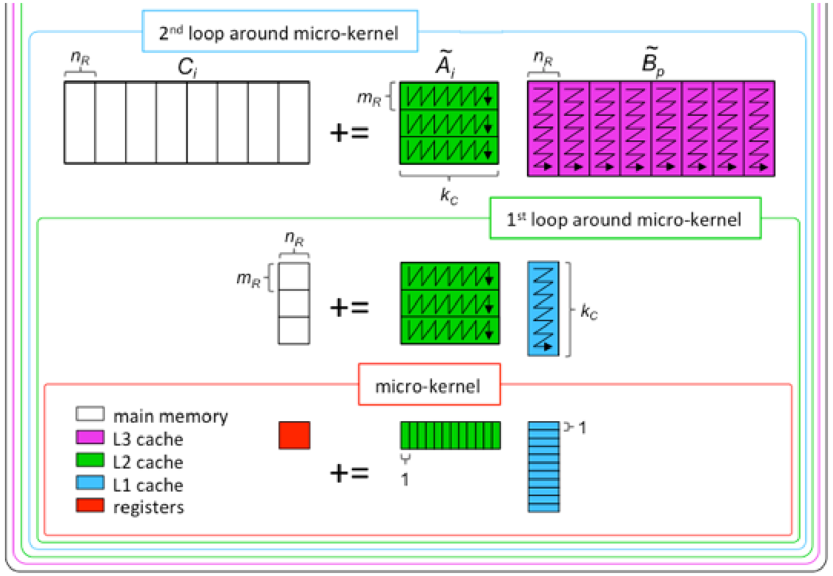

C
A
B
Variant 1 (m):
x
...
Variant 2 (n):
Variant 3 (k):



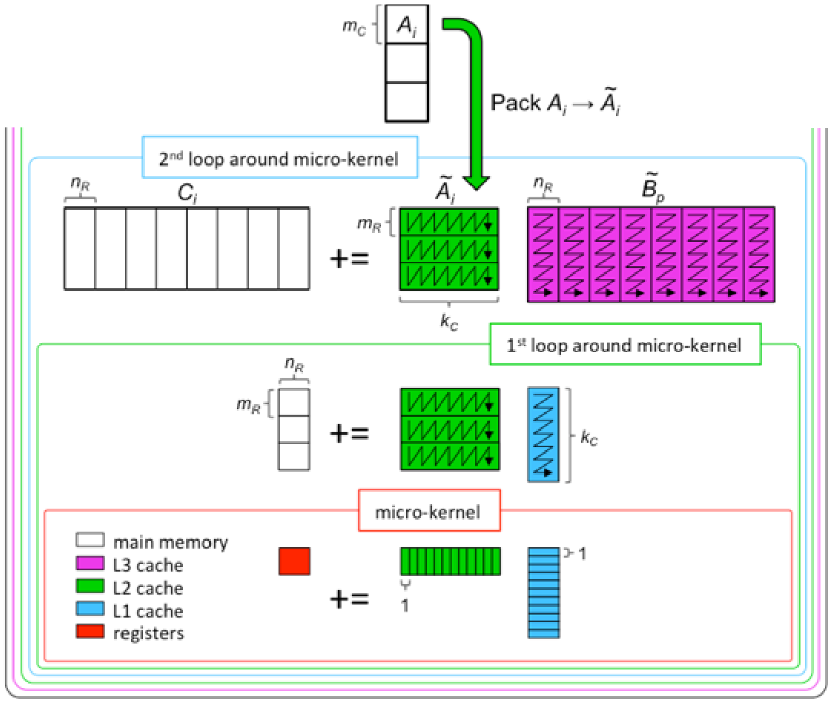
Variant	Name	Blocking factor(s):	Reuse data from:
3 (k)	Micro-kernel	1	Registers



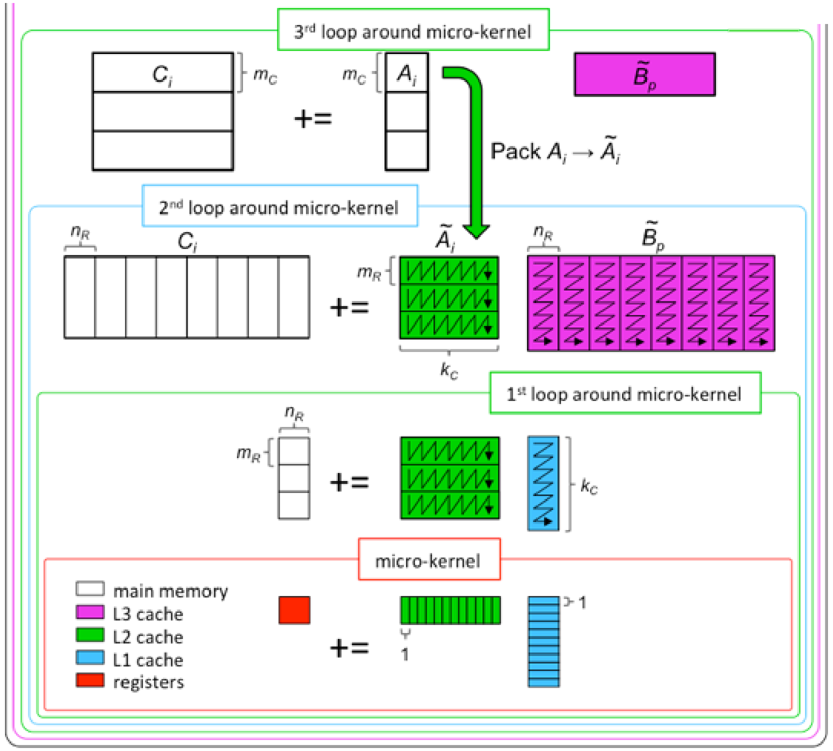
Variant	Name	Blocking factor(s):	Reuse data from:
1 (m)	IR loop	MR	L1 cache
3 (k)	Micro-kernel	1	Registers



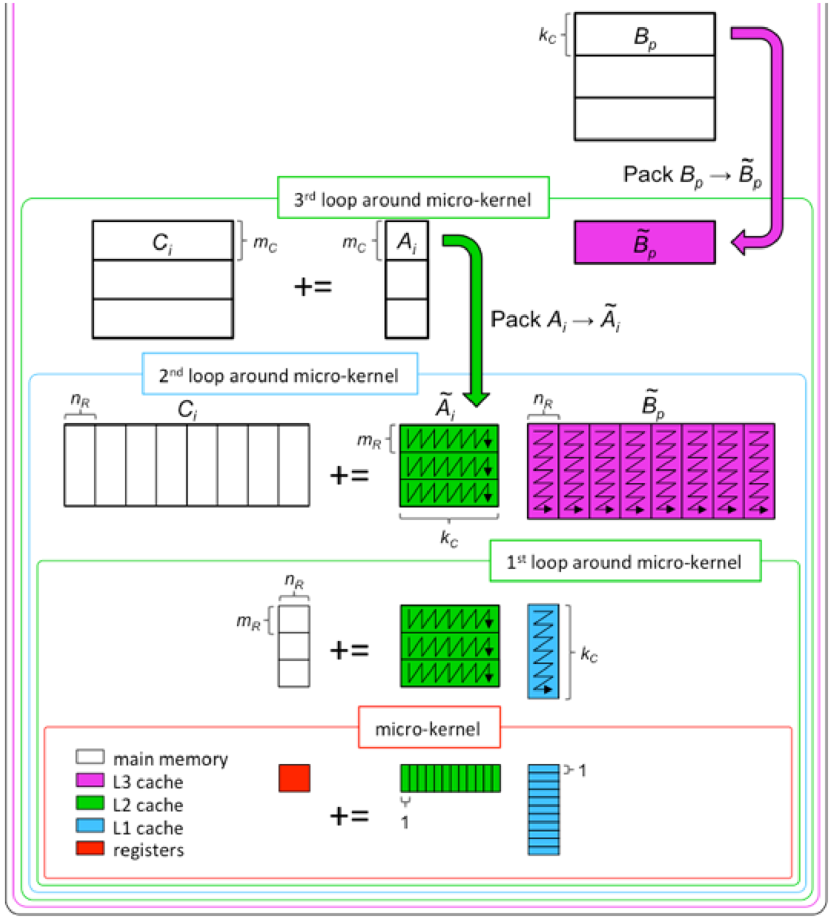
Variant	Name	Blocking factor(s):	Reuse data from:
2 (n)	JR loop	NR	L2 cache
1 (m)	IR loop	MR	L1 cache
3 (k)	Micro-kernel	1	Registers



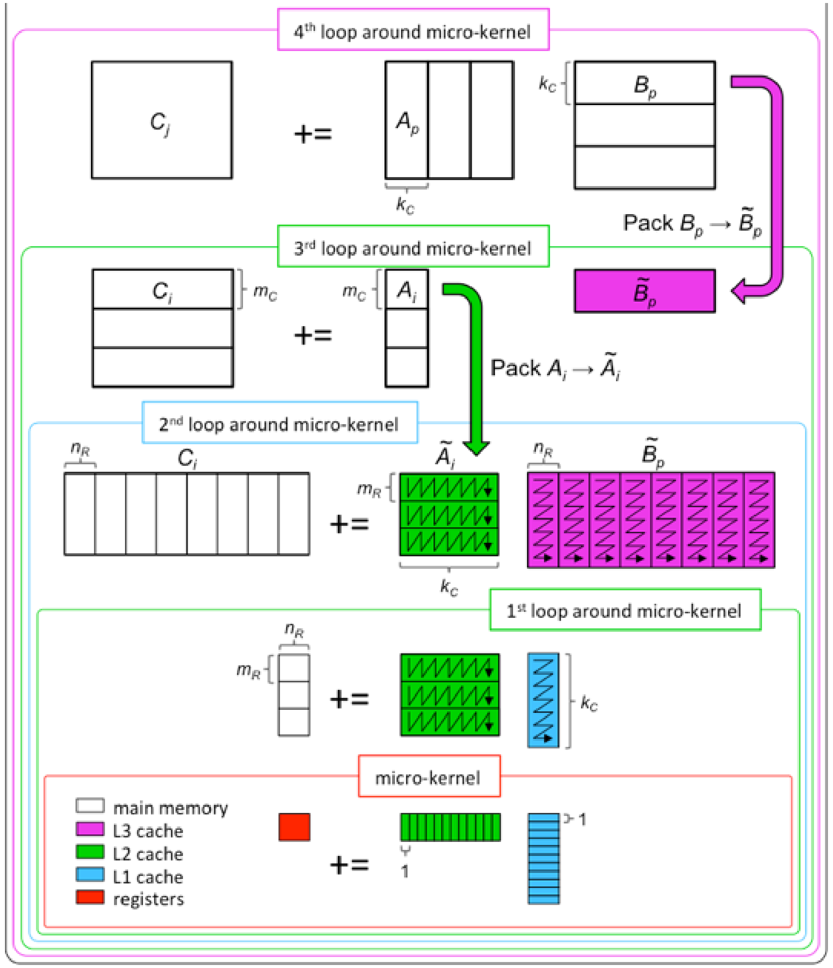
Variant	Name	Blocking factor(s):	Reuse data from:
	Pack A	MR, KC	
2 (n)	JR loop	NR	L2 cache
1 (m)	IR loop	MR	L1 cache
3 (k)	Micro-kernel	1	Registers



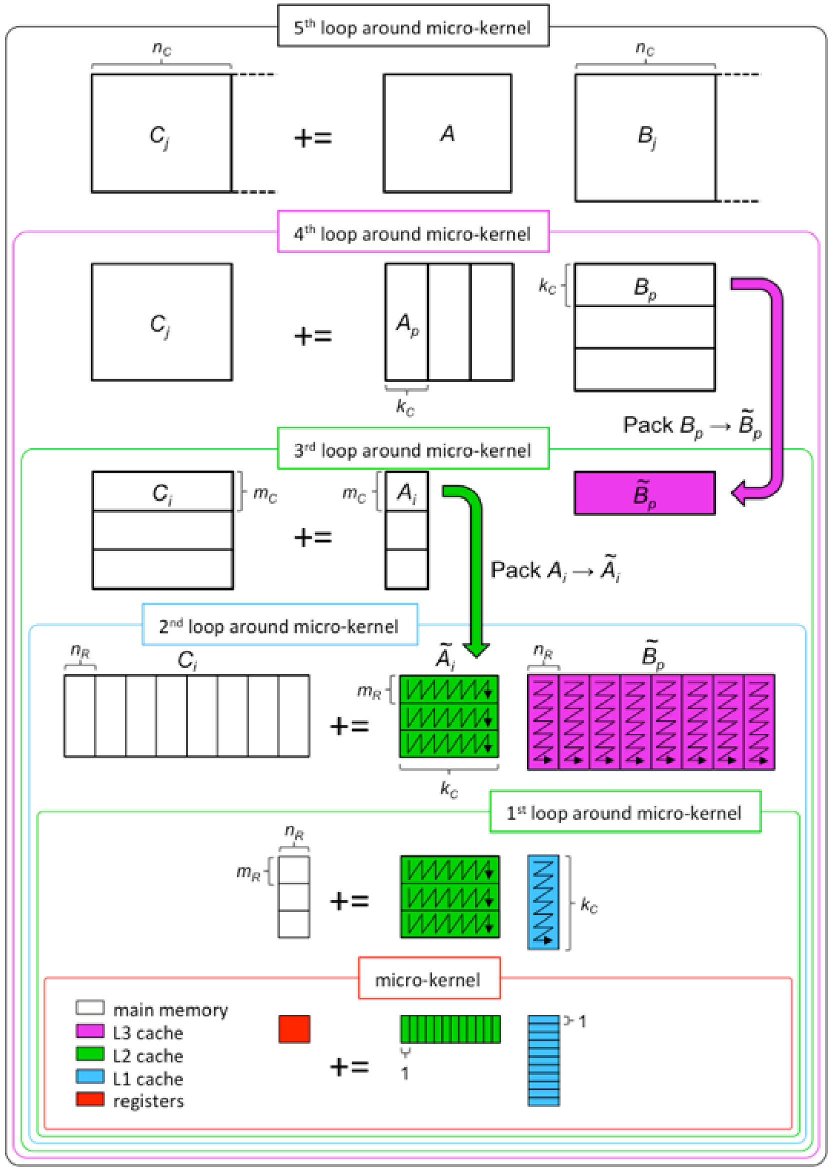
Variant	Name	Blocking factor(s):	Reuse data from:
1 (m)	IC loop	IC	L3 cache
	Pack A	MR, KC	
2 (n)	JR loop	NR	L2 cache
1 (m)	IR loop	MR	L1 cache
3 (k)	Micro-kernel	1	Registers



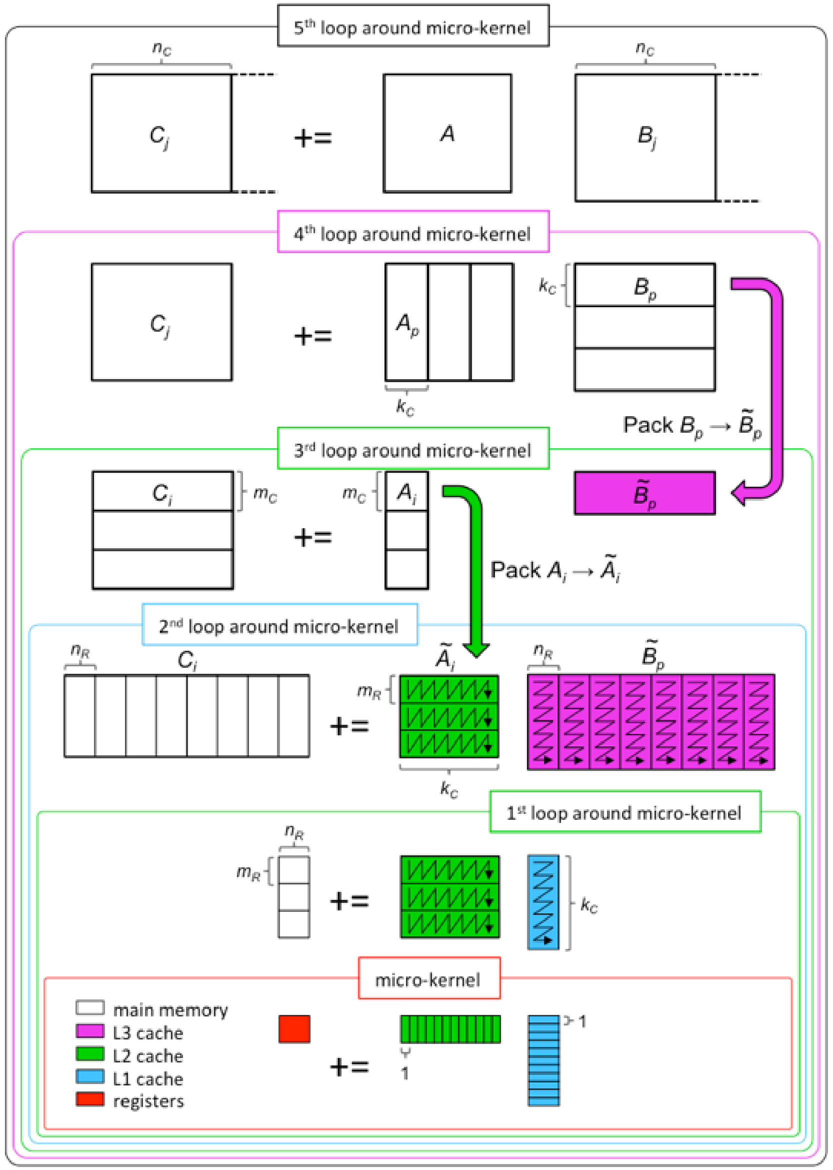
Variant	Name	Blocking factor(s):	Reuse data from:
	Pack B	NR, KC	
1 (m)	IC loop	IC	L3 cache
	Pack A	MR, KC	
2 (n)	JR loop	NR	L2 cache
1 (m)	IR loop	MR	L1 cache
3 (k)	Micro-kernel	1	Registers

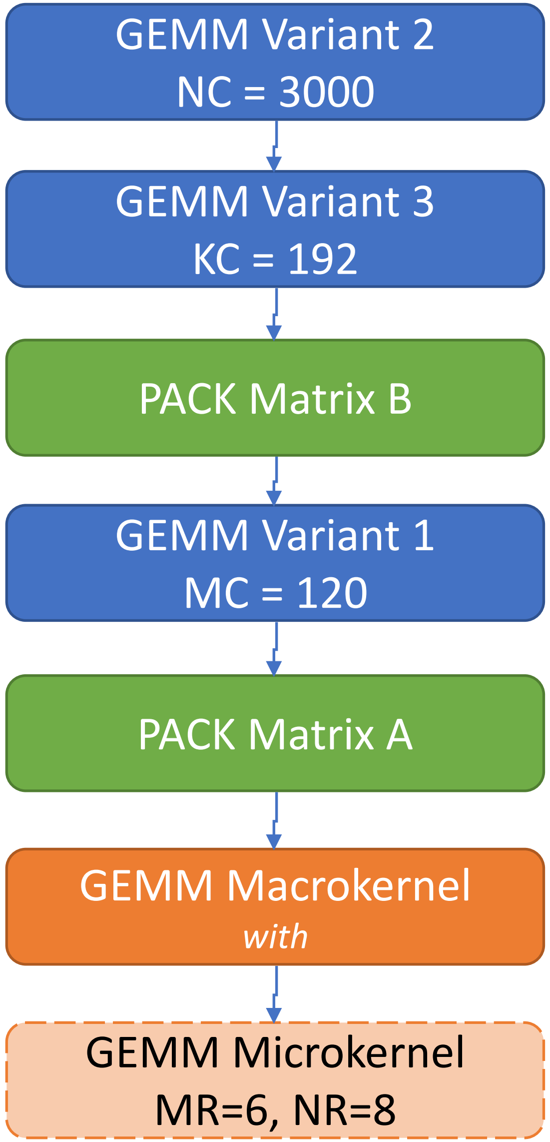


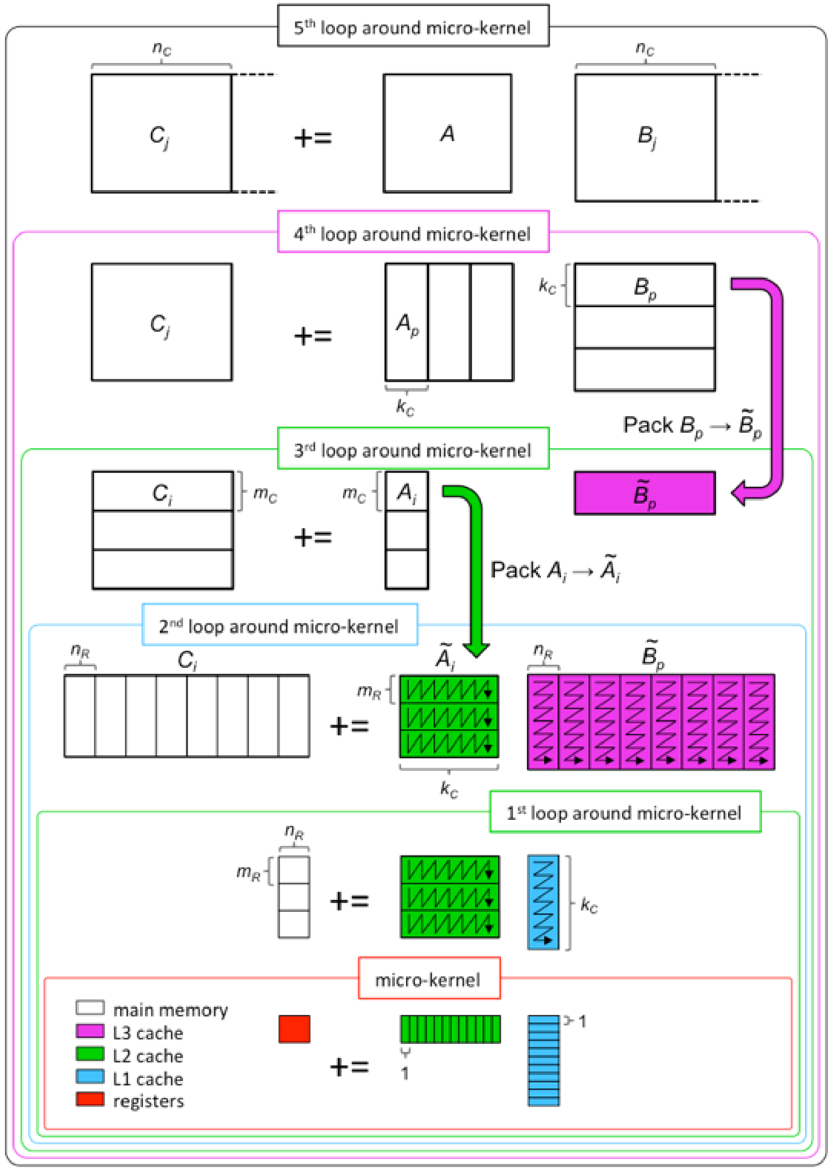
Variant	Name	Blocking factor(s):	Reuse data from:
3 (k)	PC loop	KC	
	Pack B	NR, KC	
1 (m)	IC loop	IC	L3 cache
	Pack A	MR, KC	
2 (n)	JR loop	NR	L2 cache
1 (m)	IR loop	MR	L1 cache
3 (k)	Micro-kernel	1	Registers



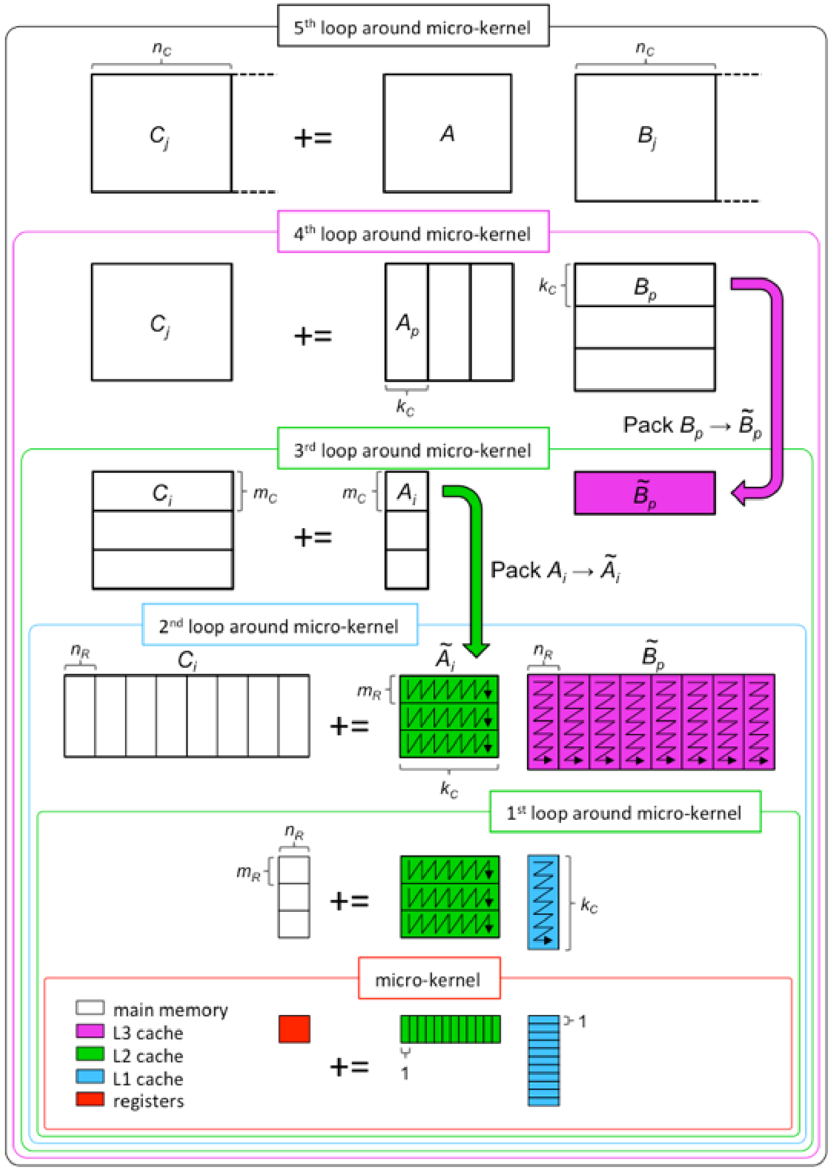
Variant	Name	Blocking factor(s):	Reuse data from:
2 (n)	JC loop	NC	
3 (k)	PC loop	KC	
	Pack B	NR, KC	
1 (m)	IC loop	IC	L3 cache
	Pack A	MR, KC	
2 (n)	JR loop	NR	L2 cache
1 (m)	IR loop	MR	L1 cache
3 (k)	Micro-kernel	1 (KR)	Registers

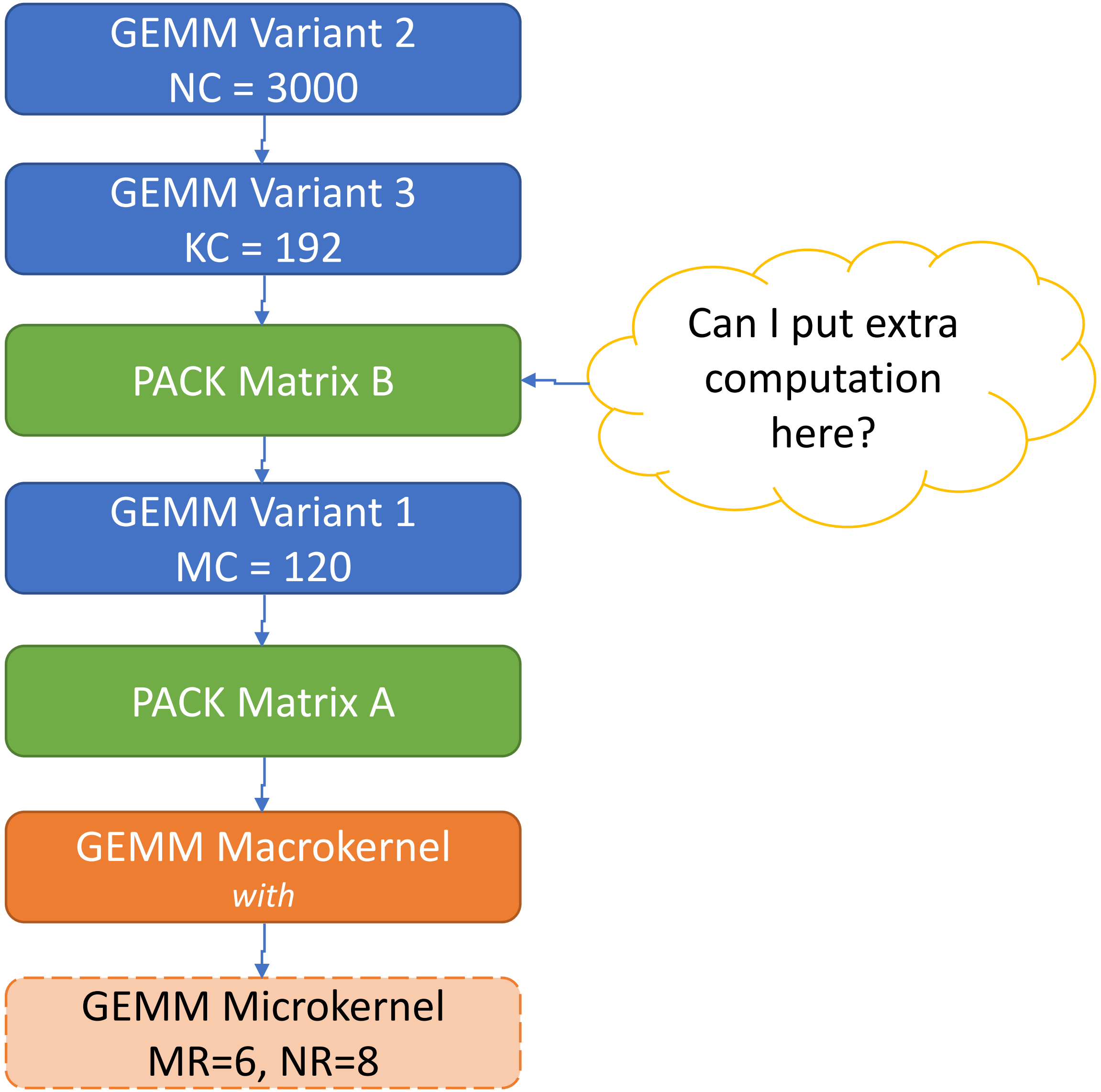

5th loop around micro-kernel
4th loop around micro-kernel
Pack $B_p \rightarrow \tilde{B}_p$
3rd loop around micro-kernel
Pack $A_i \rightarrow \tilde{A}_i$
2nd loop around micro-kernel
1st loop around micro-kernel
micro-kernel
main memory
L3 cache
L2 cache
L1 cache
registers

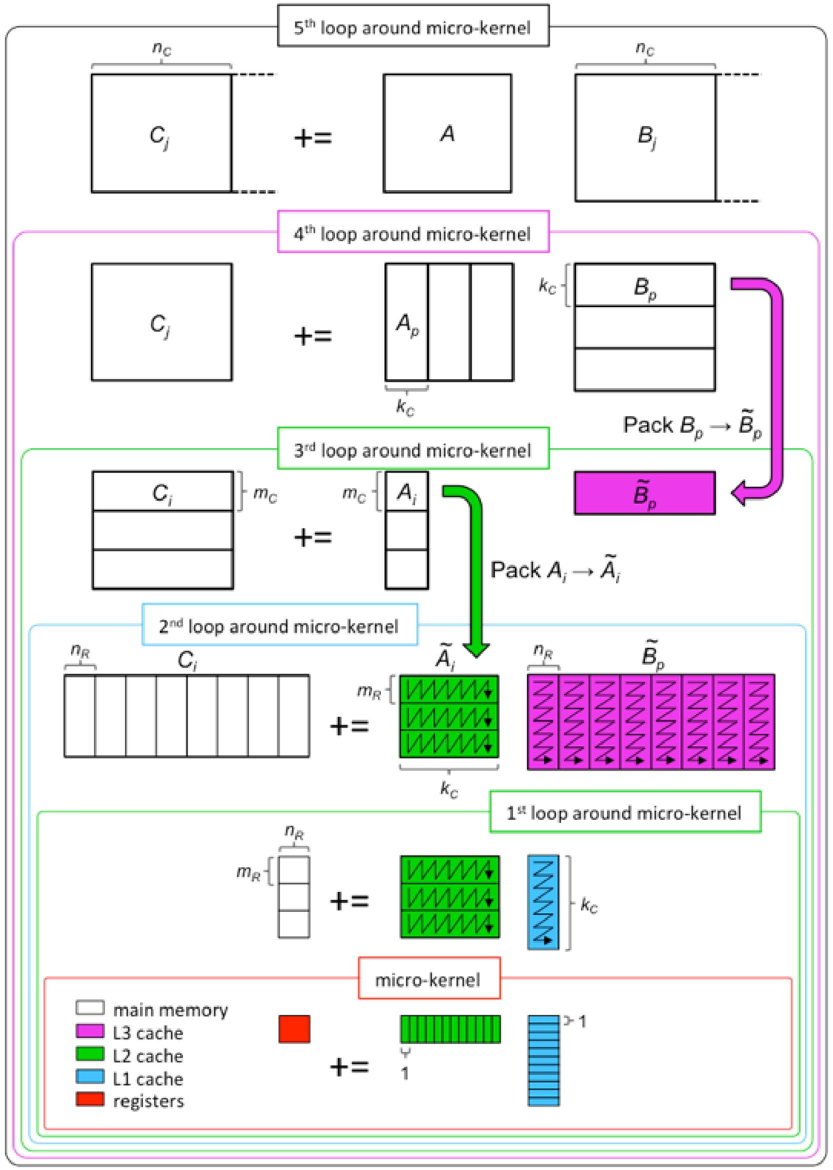

GEMM Variant 2
NC = 3000
GEMM Variant 3
KC = 192
PACK Matrix B
GEMM Variant 1
MC = 120
PACK Matrix A
GEMM Macrokernel
with
GEMM Microkernel
MR=6, NR=8

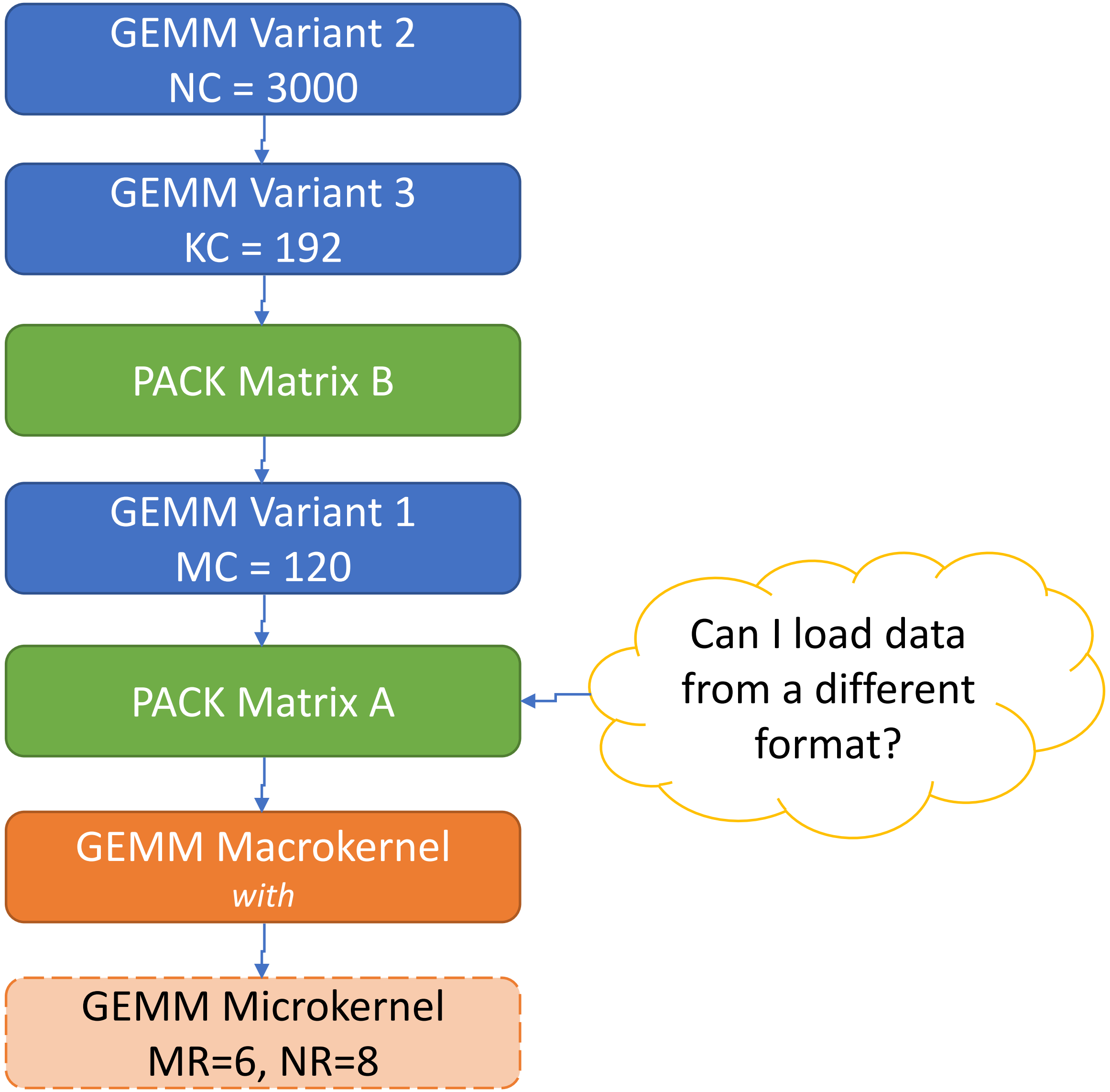


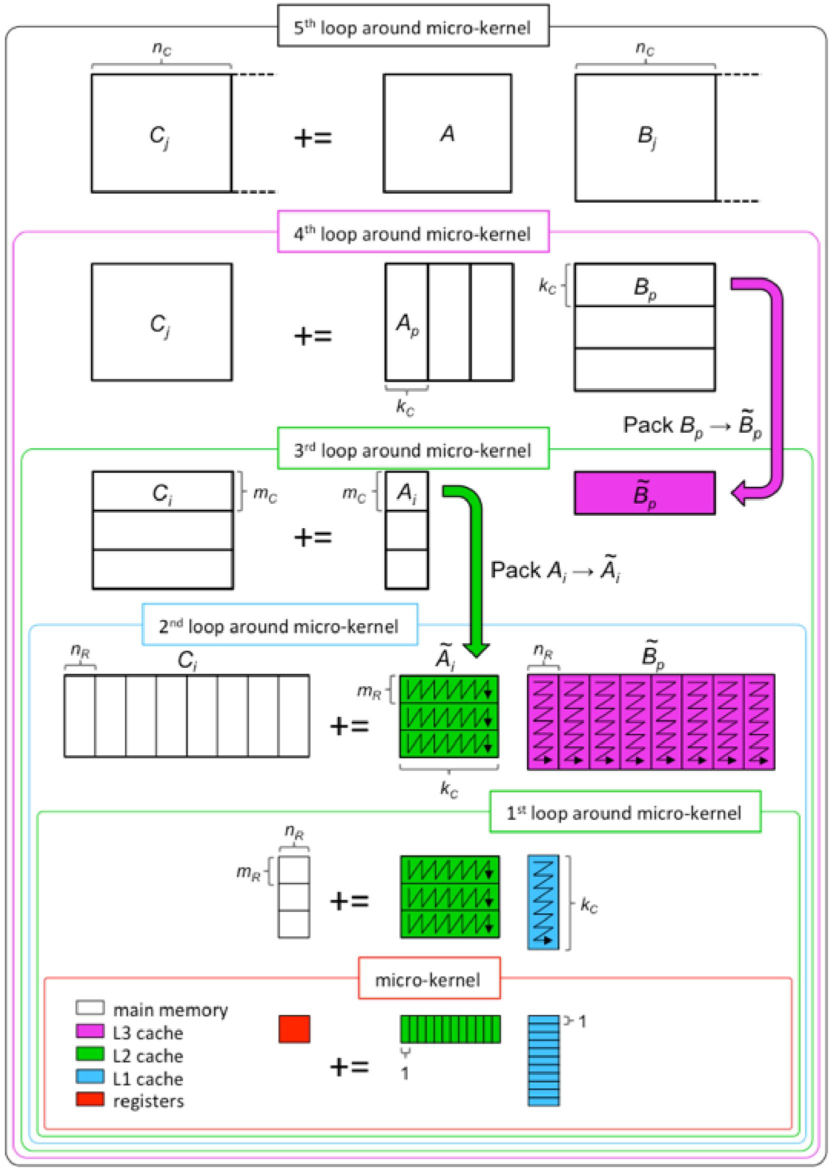
But why limit ourselves?

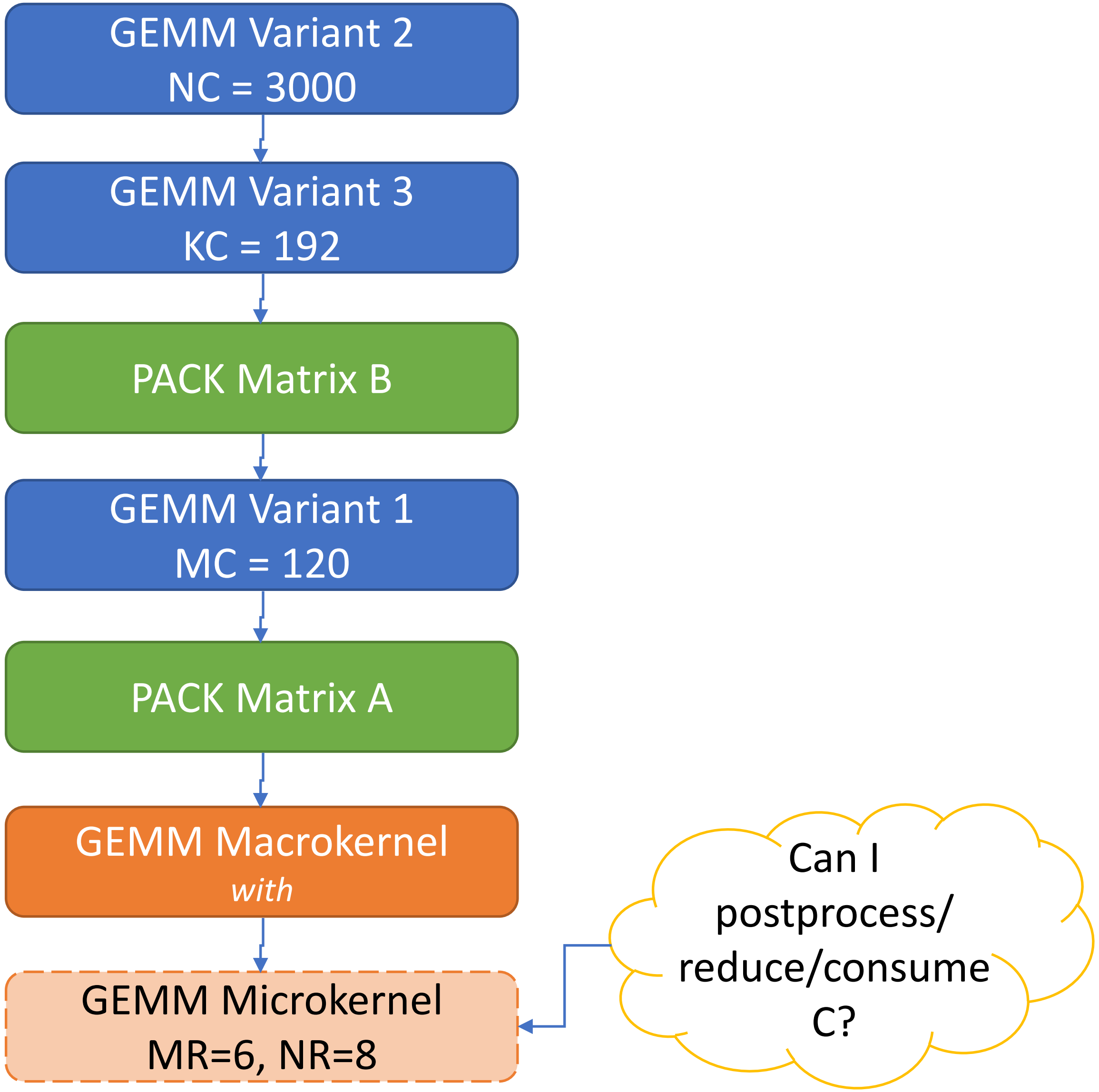

5th loop around micro-kernel
4th loop around micro-kernel
Pack Bp → B̃p
3rd loop around micro-kernel
Pack Ai → Ãi
2nd loop around micro-kernel
1st loop around micro-kernel
micro-kernel
main memory
L3 cache
L2 cache
L1 cache
registers

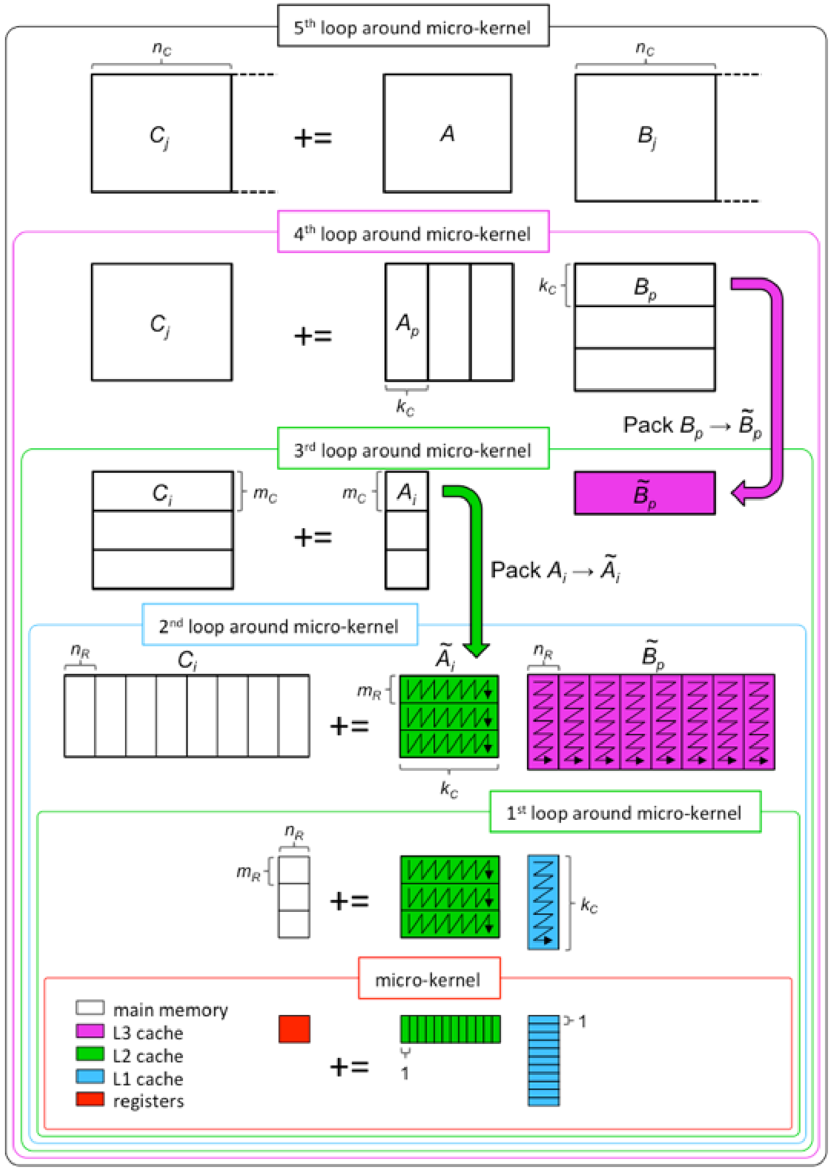

GEMM Variant 2
NC = 3000
GEMM Variant 3
KC = 192
PACK Matrix B
Can I put extra computation here?
GEMM Variant 1
MC = 120
PACK Matrix A
GEMM Macrokernel
with
GEMM Microkernel
MR=6, NR=8

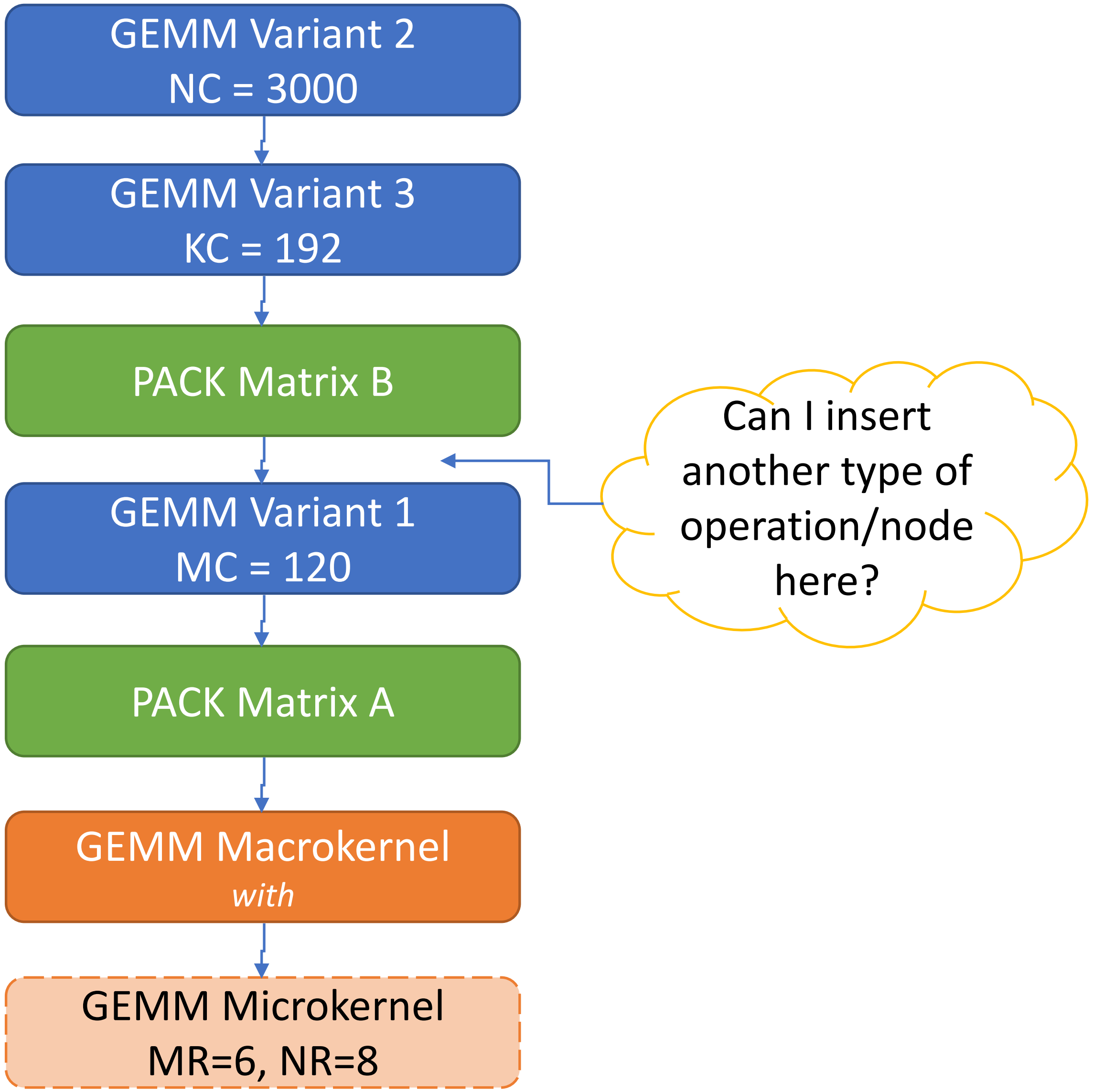

5th loop around micro-kernel
4th loop around micro-kernel
Pack $B_p \rightarrow \tilde{B}_p$
3rd loop around micro-kernel
Pack $A_i \rightarrow \tilde{A}_i$
2nd loop around micro-kernel
1st loop around micro-kernel
micro-kernel
main memory
L3 cache
L2 cache
L1 cache
registers

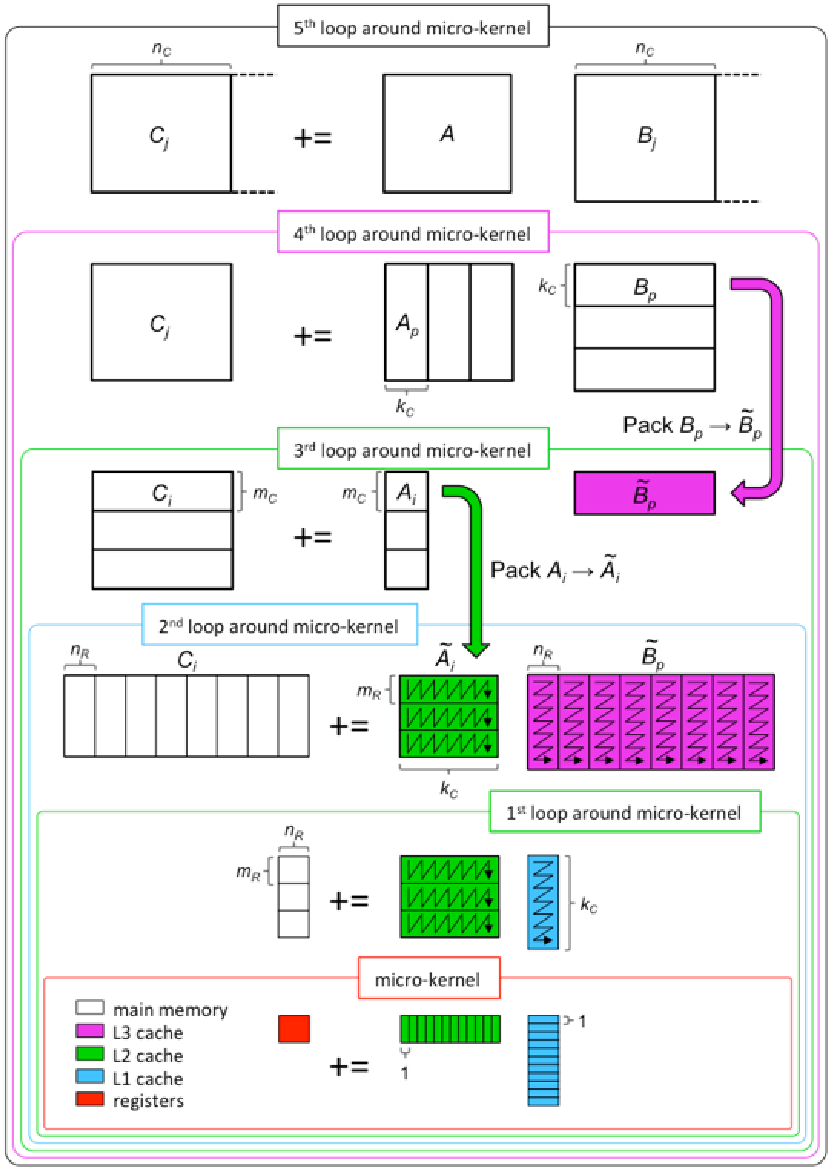

GEMM Variant 2
NC = 3000
GEMM Variant 3
KC = 192
PACK Matrix B
GEMM Variant 1
MC = 120
PACK Matrix A
GEMM Macrokernel
with
GEMM Microkernel
MR=6, NR=8
Can I load data from a different format?

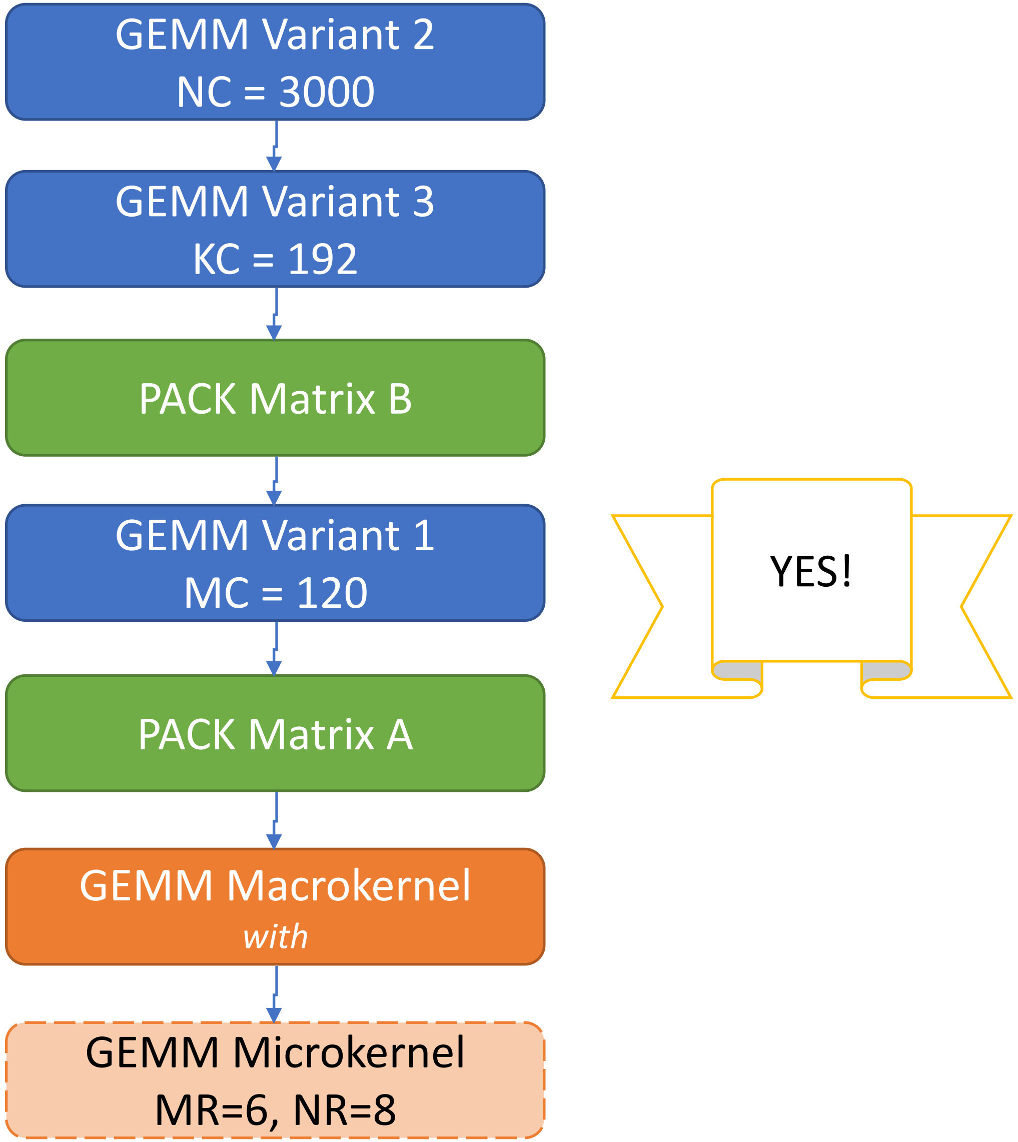

5th loop around micro-kernel
4th loop around micro-kernel
3rd loop around micro-kernel
2nd loop around micro-kernel
1st loop around micro-kernel
micro-kernel
Pack $B_p \rightarrow \tilde{B}_p$
Pack $A_i \rightarrow \tilde{A}_i$
main memory
L3 cache
L2 cache
L1 cache
registers

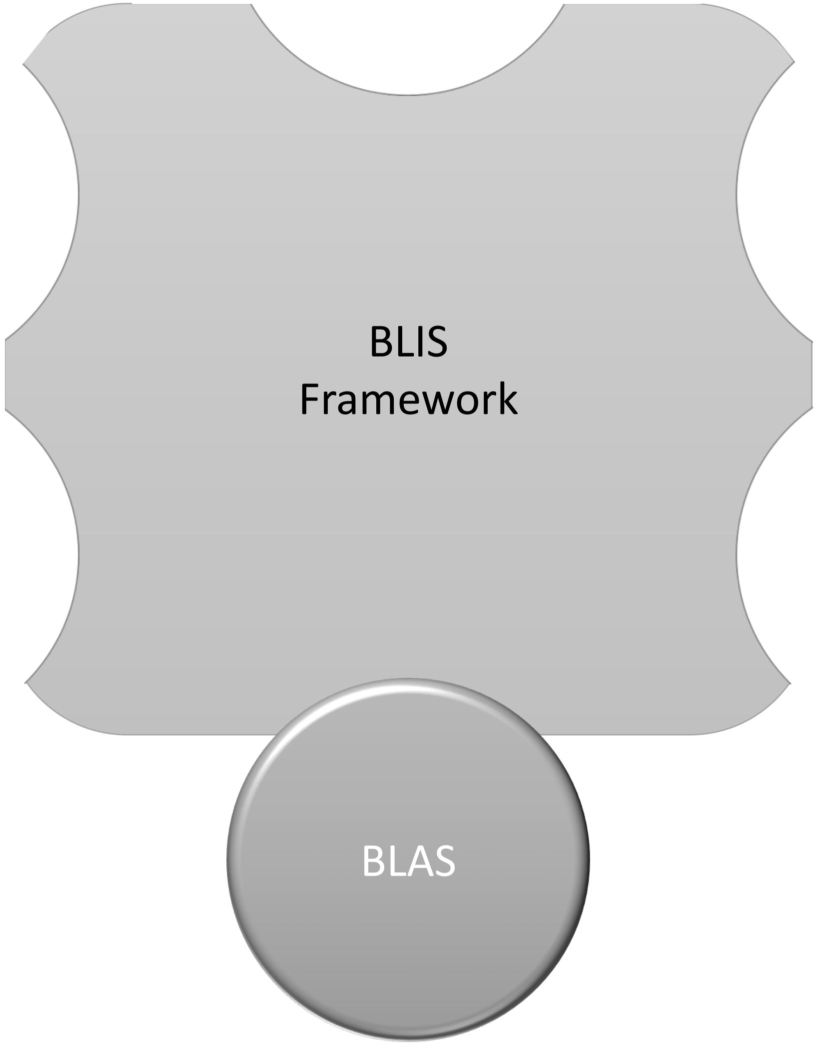

GEMM Variant 2
NC = 3000
GEMM Variant 3
KC = 192
PACK Matrix B
GEMM Variant 1
MC = 120
PACK Matrix A
GEMM Macrokernel
with
GEMM Microkernel
MR=6, NR=8
Can I postprocess/ reduce/consume C?

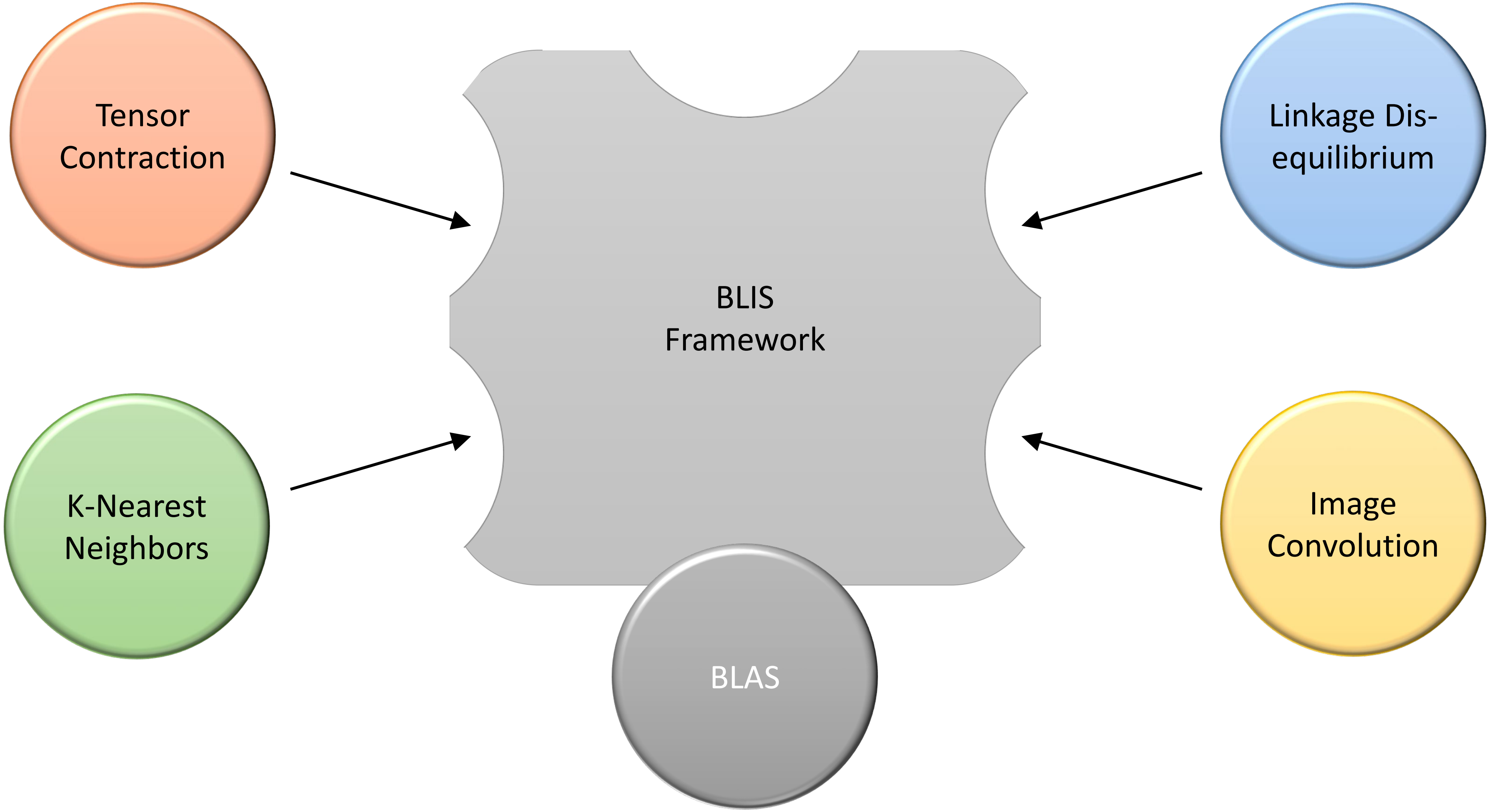

5th loop around micro-kernel
4th loop around micro-kernel
Pack $B_p \rightarrow \tilde{B}_p$
3rd loop around micro-kernel
Pack $A_i \rightarrow \tilde{A}_i$
2nd loop around micro-kernel
1st loop around micro-kernel
micro-kernel
main memory
L3 cache
L2 cache
L1 cache
registers

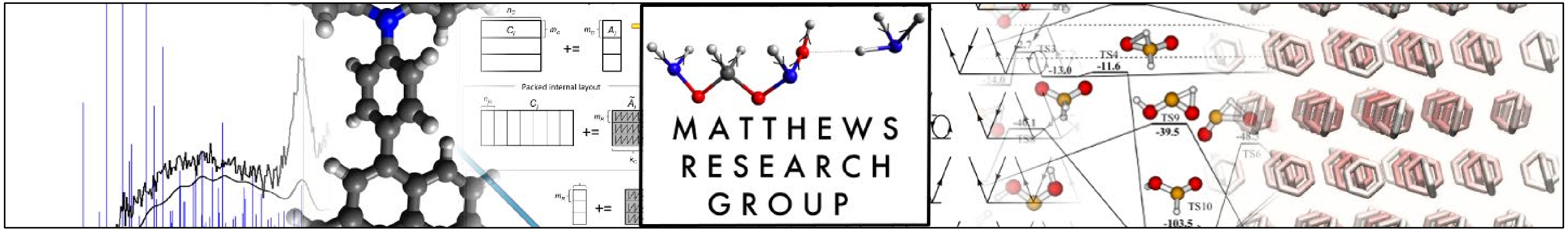

GEMM Variant 2
NC = 3000
GEMM Variant 3
KC = 192
PACK Matrix B
GEMM Variant 1
MC = 120
PACK Matrix A
GEMM Macrokernel
with
GEMM Microkernel
MR=6, NR=8
Can I insert another type of operation/node here?


5th loop around micro-kernel
4th loop around micro-kernel
3rd loop around micro-kernel
2nd loop around micro-kernel
1st loop around micro-kernel
micro-kernel
Pack $B_p \rightarrow \tilde{B}_p$
Pack $A_i \rightarrow \tilde{A}_i$
main memory
L3 cache
L2 cache
L1 cache
registers

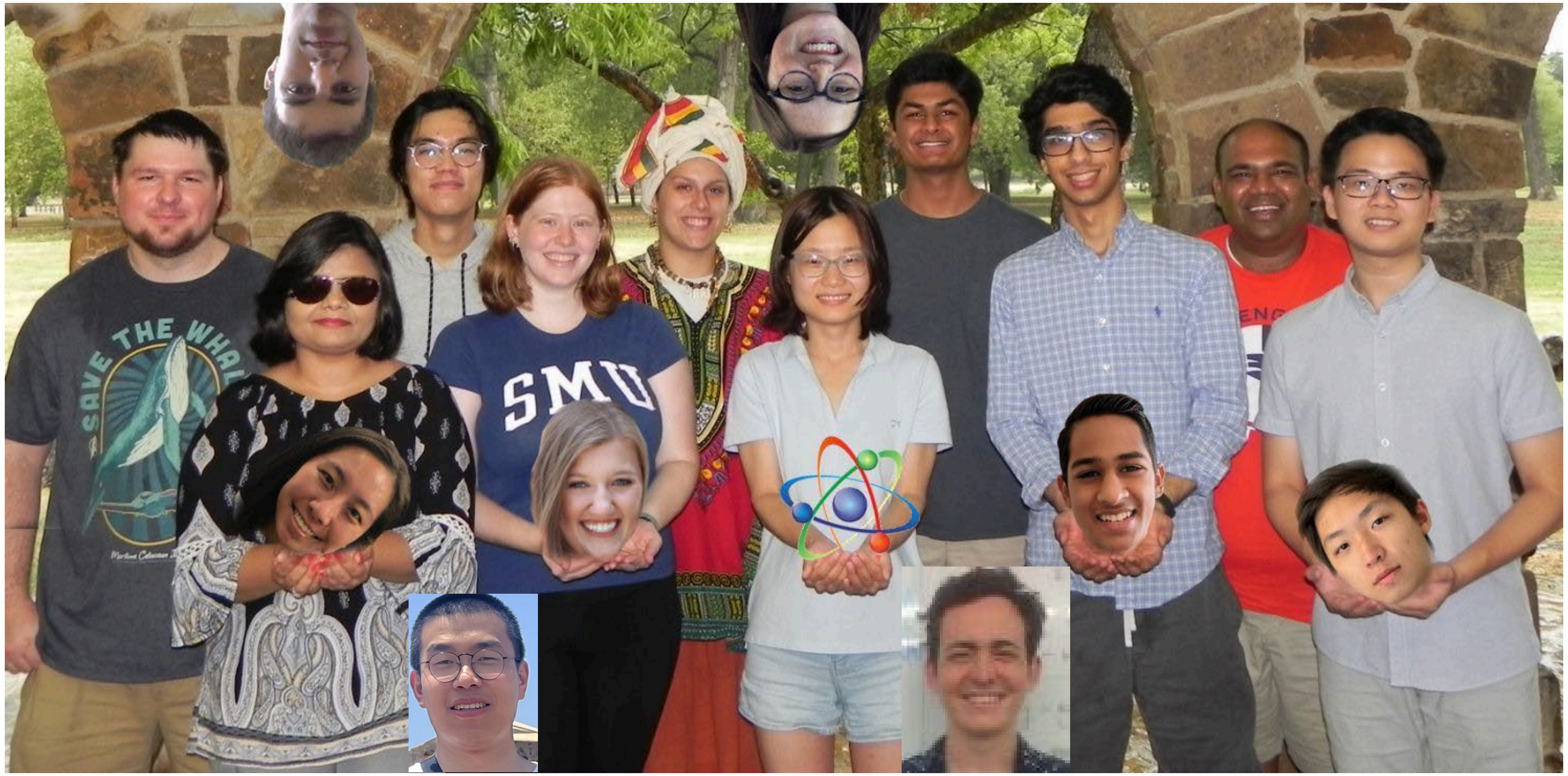

GEMM Variant 2
NC = 3000
GEMM Variant 3
KC = 192
PACK Matrix B
GEMM Variant 1
MC = 120
PACK Matrix A
GEMM Macrokernel
with
GEMM Microkernel
MR=6, NR=8
YES!


BLIS
Framework
BLAS


Tensor Contraction
Linkage Dis-equilibrium
BLIS Framework
K-Nearest Neighbors
Image Convolution
BLAS



people.smu.edu/dmatthews



CAREER CHE-2143725
OAC-2003931
CBET-2117574

THE Welch FOUNDATION

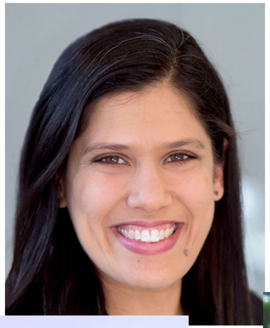
N-2072-20210327

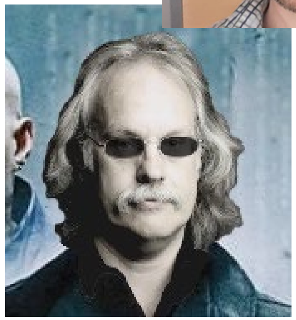
ECRP DE-SC0022893

BLIS is a **Community Project**

115+ contributors

Bryan
Marker





I’M HIRING POSTDOCS