Basis Selection by GRASPG-Based Machine Learning

Chaofan Shi

Institute of Modern Physics,

Fudan University

12 June, 2025

Contents

Contents

For High-Precision Atomic Calculations

- Configuration State Function Generator(CSFG) – **GRASPG** package[1-3]

```
1s ( 2)  7s ( 1)  7p-( 1)
            1/2       1/2
                        0-
```

- {7s,7p-,7p,7d-,7d}
  - **200 CSFGs →1 200 CSFs**
- {10s,10p-,10p,10d-,10d}
  - **200 CSFGs →7 844CSFs**

Calculate the angular coefficients between **200 CSFGs**.

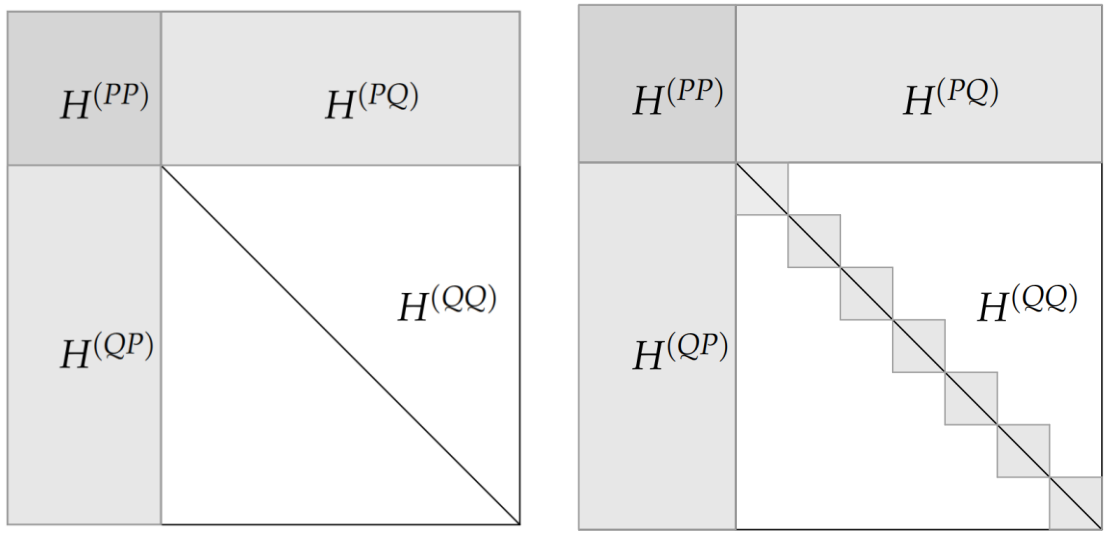
$$H_{ij} = \langle \Phi(\gamma_i J) \| \mathcal{H}_{\mathrm{DC}} \| \Phi(\gamma_j J) \rangle = \sum_{ab} t^{ij}_{ab} I(a,b) + \sum_{abcd;k} v^{ij}_{abcd;k} R^k(ab,cd)$$

- The **spin-angular integration is independent** of the principal quantum numbers of the orbitals.
- Reduce both the **CPU times** and the **memory load** by introducing CSFGs.

[1] Atoms 11, 12 (2023) [2] Comp. Phys. Comm. 283，108562 (2023) [3] Comp. Phys. Comm. 312，109604 (2025)

➢ CSFGs based applications

- Improvement on zero-first order perturbation.[4]



- Rmixaccumulate - *a priori condensation* techniques

Higher-Precision ⟷ Approximating complete basis spaces

Table . The number of CSFs increases with atomic complexity (represented by atomic number Z)

Z	Atom	CSFGs (billion)
5	B I	9.5
6	C I	36.2
7	N I	101.5
8	O I	163.2
9	F I	209.5

[4] Phys. Rev. A 111, 042805 (2025)

Machine Learning in Atomic calculations

- Machine learning in *Quantum Chemistry*[5]

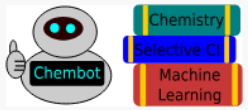
Chembot: A Machine Learning Approach to Selective Configuration Interaction

Sergio D. Pineda Flores*

Cite This: *J. Chem. Theory Comput.* 2021, 17, 4028−4038 | Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: We introduce a machine learning-based approach to selective configuration interaction, dubbed Chembot, that utilizes many novel choices for its model design and training. These choices include the use of a support vector machine to select important configurations, the use of the charge density matrix and configuration energy as features, and heuristics to improve the quality of training data. We test Chembot's ability to obtain near full configuration interaction quality energies and find that it definitively outperforms its purely Stochastic cousin Monte Carlo configuration interaction by requiring fewer iterations to converge, fewer determinants in the variational space, and fewer important configurations to achieve the same energy. In addition, Chembot at times requires fewer determinants in its variational space than the heat-bath configuration interaction method to achieve the same energy. We demystify Chembot's innards and then showcase our claims on the set of small but challenging systems: the hydrogen ring (H_4), stretched methylene (H_2C), and stretched water (H_2O).



Machine learning approach to Selective CI[5]

Deep-Learning Approach for the Atomic Configuration Interaction Problem on Large Basis Sets

Pavlo Bilous[1,2,*], Adriana Pálffy[3,2], and Florian Marquardt[1,4]

Show more

Phys. Rev. Lett. **131**, 133002 – **Published 27 September, 2023**

DOI: https://doi.org/10.1103/PhysRevLett.131.133002

A neural network approach to running high-precision atomic computations

Pavlo Bilous,[1,*] Charles Cheung,[2] and Marianna Safronova[2]

[1] *Max Planck Institute for the Science of Light, Staudtstr. 2, 91058 Erlangen, Germany*
[2] *Department of Physics and Astronomy, University of Delaware, Delaware 19716, USA*

(Dated: August 2, 2024)

Machine Learning in Atomic calculations[6,7]

- Deep Learning & GRASP[6]
- Neutral Network & pCI[7]

[5] J. Chem. Theory Comput. 17, 4028 (2021). [6] Phys. Rev. Lett.131.133002 (2023) [7] Phys. Rev. A 110, 042818 (2024)

Contents

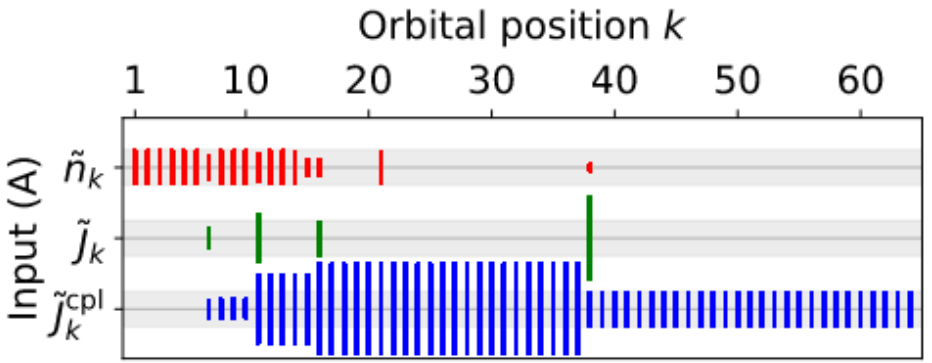
Characteristics of CSF(G)

- CSF(G)s based on N orbitals are characterized by $3N$ quantum numbers.
- Just the 3 lines of CSF(G)s, for orbital k

```
2s ( 1)   2p-( 1)   6s ( 1)   7s ( 1)
   1/2        1/2        1/2        1/2
                         1          1/2        0-
```

n_k
j_k
j_k^{cpl}

- $3N$ parameters after **normalization** would be fed into Neutral Network.



Color representation of the exemplary CSFG[6]

[6] Phys. Rev. Lett.131.133002 (2023)

How does NN work on Basis Selection

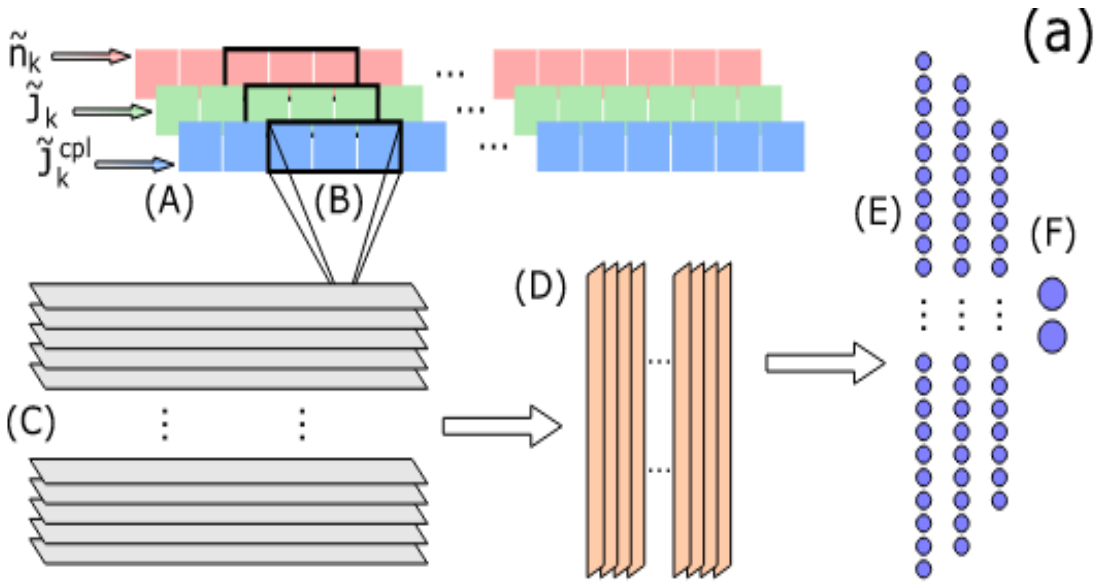
A. CI on **a small Part of the large CSF(G) set.**

$$\Psi(\gamma PJM) = \sum_{r=1}^{N_{CSFG}} c_r \Phi(\gamma_r PJM)$$

CSF(G) **important or not**

B. **Train NN** based on the **limited** CI result.

C. **Apply** trained NN model to **large set.**



NN architecture used in the Algorithms[6]

- CSFG compared to CSF.

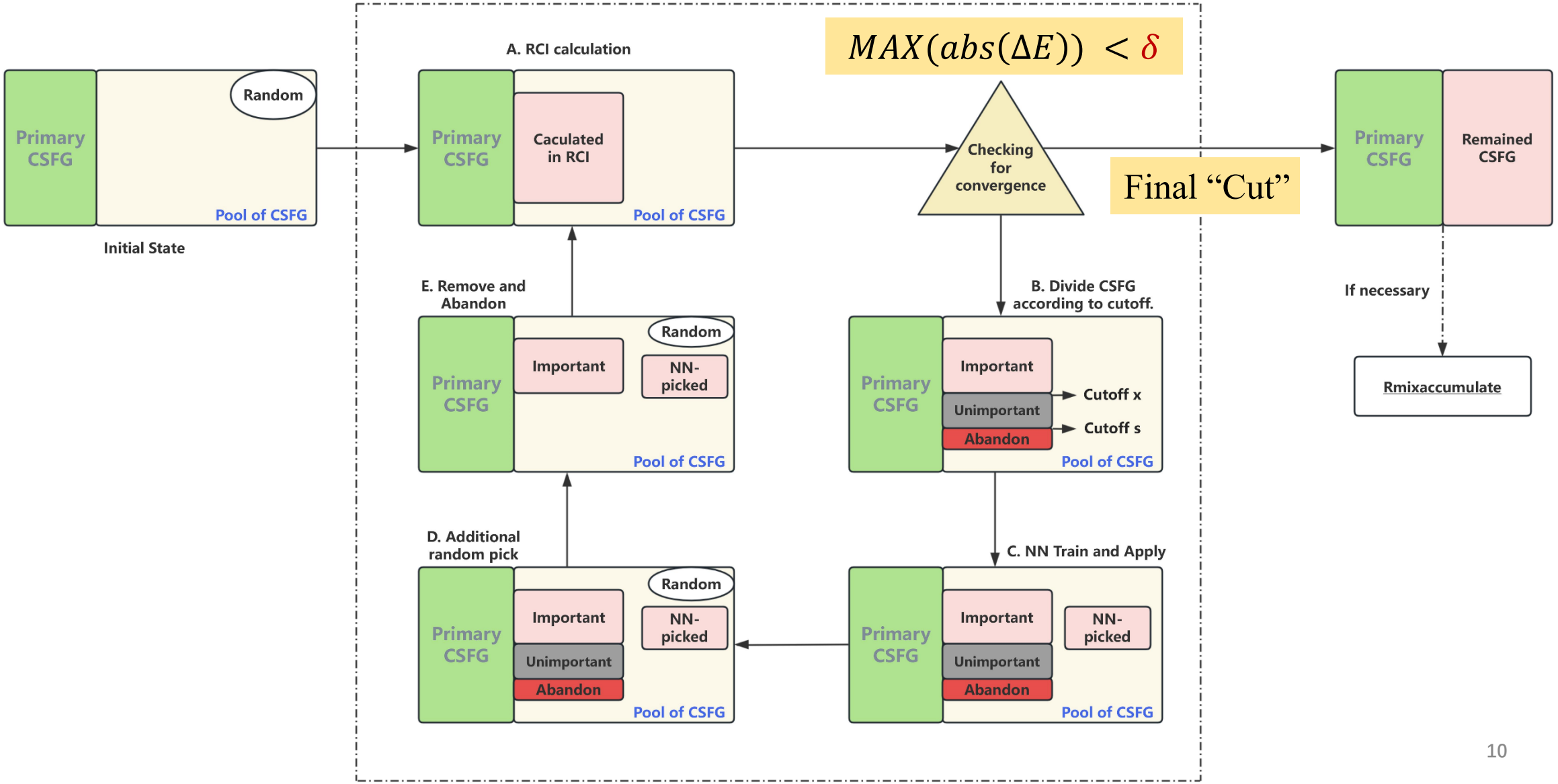
I. CSFGs **Pre-classify** CSFs.

Less Training Volume, time & **Accelerate** Convergence

II. **Efficient** CI & **Low** memory load & **Larger** Scale

[6] Phys. Rev. Lett.131.133002 (2023)

Program Flowchart



program flowchart of GRASPG-Based NN Algorithms

Key points in NN&GRASPG

- **Weight** of CSFG:

$$\tilde{c} = \sqrt{\sum_{i=1}^{n} \omega_i}, \omega_i = \frac{\sum_{j=1}^{k} {c_{ij}}^2}{k}$$

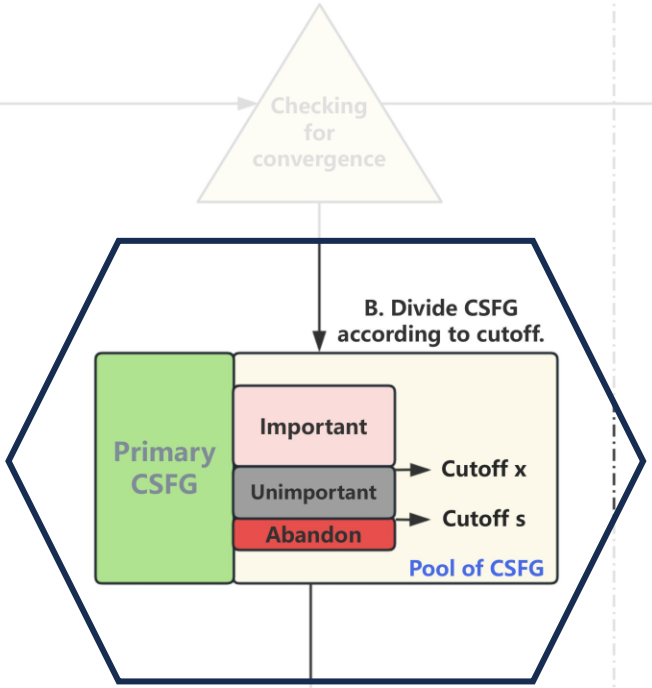
- Cutoff x:

$$\sum \tilde{c}^2 = \rho, \rho = 1 - 10^{\sigma}$$

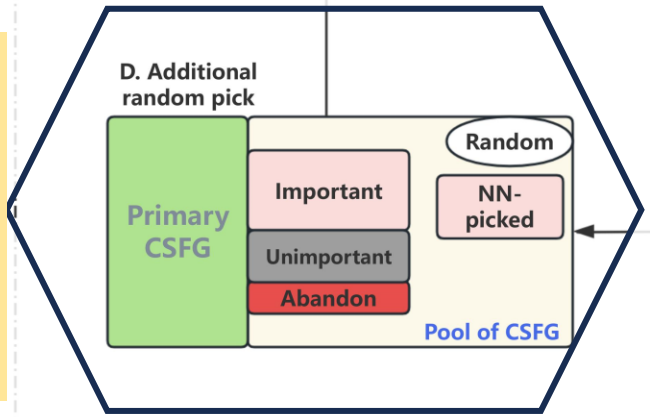
e.g. $\sigma: -6.0 \sim -9.0$

- Abandon Cutoff s:

e.g. $s = x * 10^{-8}$



Key points in NN&GRASPG



- Unimportant part, Back into the pool.
- Abandon part, Remove from pool

- Additionally pick CSFG to **avoid losing potentially important terms** ignored by NN

Benchmarking Case

➢ 5 lowest levels of Ni^{12+} $[Ar]\ 3s^2\ 3p^4$

- Multi Reference： $3s(2,*)3p(4,*) \rightarrow \{3s, 3p, 3d\}$
- Active Space： MR $\rightarrow \{7s, 7p, 7d, 7f, 7g, 7h, 7i\}(n < 8, l < 7)$
- Primary set： **VV** Full set: VV **+ CV +CC**

block	Primary Set		Full Set	
	CSFG	CSF	CSFG	CSF
even1	676	4 558	129 793	920 490
even2	1 672	12 260	566 686	4 340 535
even3	2 537	18 009	1 130 776	8 752 928
total	4 885	34 827	1 827 255	**14 013 953**

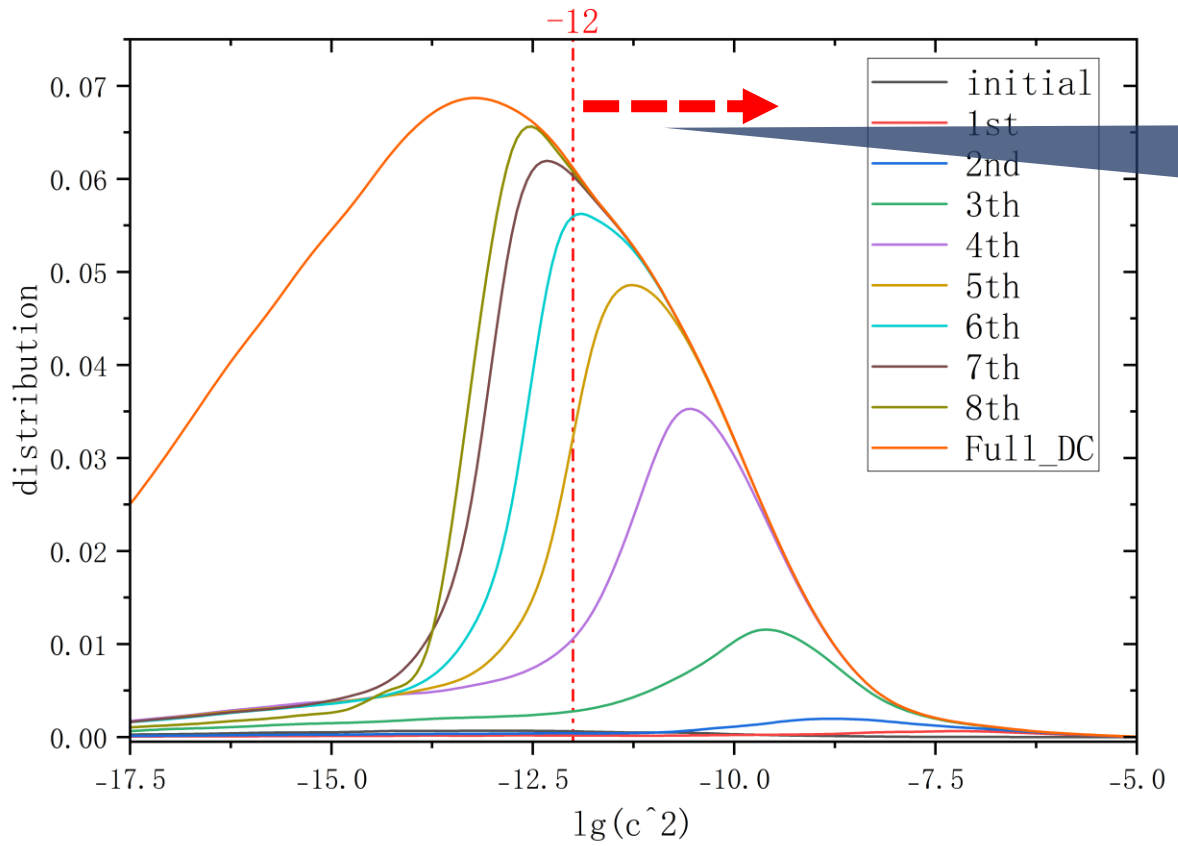
Accuracy and Convergence

$MAX(abs(\Delta E)) < 1cm^{-1}$

No.	Term	$E^{iter} - E^{FCI}(cm^{-1})$								
		0	1	2	3	4	5	6	7	8
0	$3s^2 3p^4\ ^3P_2$	82681.32	4845.63	1297.53	194.81	26.36	5.93	1.71	0.46	**0.24**
1	$3s^2 3p^4\ ^3P_1$	79707.79	3971.72	1305.35	205.85	31.56	6.72	1.95	0.83	**0.55**
2	$3s^2 3p^4\ ^3P_0$	81282.71	34038.06	16707.51	2215.64	145.27	8.67	1.62	0.57	**0.46**
3	$3s^2 3p^4\ ^1D_2$	83778.76	7758.74	1512.42	223.18	33.34	7.62	2.11	0.68	**0.35**
4	$3s^2 3p^4\ ^1S_0$	83151.12	32993.49	14339.40	1607.21	110.64	6.74	2.28	0.79	**0.66**
CSF number		557 156	507 852	1 238 956	3 185 860	5 530 609	6 772 329	7 912 725	8 907 649	**9 114 461**

- Energies gradually converge to the **directly calculated values(E^{FCI})**

➢ **CSFG distribution** for **Full set** & each iteration (even 3).



CSFGs distribution for Full set & each iteration

Total weight of CSFGs on the right side
=
99.99999% of Full set

- **"Important CSFGs"** is progressively included in the CI calculation via **NN**

Compared to GRASP-Based NN

➢ Iteration number

GRASPG	GRASP
8	11
8	13
8	13

No.	Term	$E(cm^{-1})$		
		FCI	NN&GRASPG	NN&GRASP
1	$3s^2 3p^4\ ^3P_1$	19953.68	19955.10	19954.25
2	$3s^2 3p^4\ ^3P_0$	20367.62	20367.07	20368.38
3	$3s^2 3p^4\ ^1D_2$	47632.06	47633.80	47633.00
4	$3s^2 3p^4\ ^1S_0$	98651.09	98650.83	98650.79
CSF number		14 013 953	5 527 293	**6 760 986**

➢ Time performance

NN time(s)	GRASPG	GRASP	ratio
even1	38	195	5.16
even2	135	832	6.16
even3	270	1 699	6.31
average			**5.88**

CI time(s)	GRASPG	GRASP	ratio
even1	10	30	2.94
even2	17	159	9.57
even3	50	441	8.76
average			**7.09**

Total time(s)	GRASPG	GRASP	ratio
even1	442	4 098	9.27
even2	1 541	28 952	18.79
even3	3 608	90 801	25.17
total	5 591(1.55h)	123 851(34.40h)	**22.15**

Introduce "Abandon Train"

A. When "**Abandon part**" is over certain percentage (e.g. 5%) of **current CI**

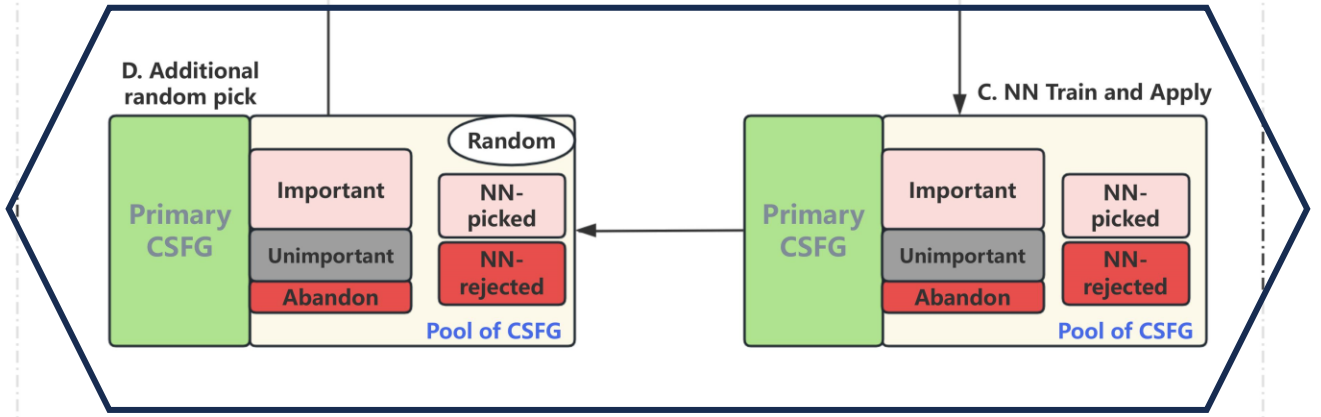
B. NN train on the such **CSFGs**

C. Predict CSFGs that **should** be abandoned

- Abandon & Abandon train
  - Reduce the **size of pool**.
  - Avoid **multiple selections** for CSFGs that contribute too little.

➢ **Time Saving**

➢ **Accelerate** Convergence



Introduce "Abandon Train"

- The **Same Result**

No.	Term	$E(cm^{-1})$	
		No_Abandon_train	Abandon_train
1	$3s^2 3p^4\ ^3P_1$	19955.10	19955.40
2	$3s^2 3p^4\ ^3P_0$	20367.07	20367.02
3	$3s^2 3p^4\ ^1D_2$	47633.80	47633.76
4	$3s^2 3p^4\ ^1S_0$	98650.83	98650.51
CSF number		**5 527 293**	**5 526 042**

- **Less** Iteration

No_Abandon_train	Abandon_train
8	7
8	7
8	8

- **Time** performance

total time(s)	No_Abandon_train	Abandon_train
even1	442	**394**
even2	1 541	**1 226**
even3	3 608	**3 492**
total	5 591	**5 112**

Contents

More MR & Larger AS

➢ Expand configuration space as large as possible

- Multi Reference： $3s(2,*)3p(4,*) \rightarrow \{4s, 4p, 4d, 4f\}$
- Active Space： MR $\rightarrow \{13s, 13p, 13d, 13f, 13g, 12h, 12i, 12k, 12l, 12m\}$
- Primary set： **VV** Full set: VV **+ CV +CC**

➢ nrconfgenerate_csfg

block	Primary Set		Full Set	
	CSFG	CSF	CSFG	CSF
even1	4 820	38 788	6 902 942	169 831 039
even2	12 724	109 090	26 039 297	732 302 209
even3	18 436	166 866	45 851 067	1 362 297 565
total	35 980	314744	78 793 306	**2 264 430 813**

GRASPG-Based NN

➢ Remained CSF(G) after GRASPG-Based NN

$MAX(abs(\Delta E)) < 1cm^{-1}$

block	Origin		NN&GRASPG	
	CSFG	CSF	CSFG	CSF
even1	6 902 942	169 831 039	696 513	19 457 315
even2	26 039 297	732 302 209	1 949 727	56 996 493
even3	45 851 067	1 362 297 565	4 620 951	152 472 848
Total	78 793 306	**2 264 430 813**	7 267 191	**228 926 656**

➢ Original CSFs: **2 264 430 813**
➢ Remained CSF(G)s: **228 926 656 (7 267 191)**
➢ Ratio ≈ 9.9 (311.6)

➢ Far **beyond** computational resources

➢ **Calculable** Scale
- Rmixaccumulate & ZF
- Multi - Server

Result

$\{29(spdfg), 28(hi\ldots)\}$

6 electrons: $\{20(spdfghikl)\}$
All electrons: $\{18(spdfg)\}$

No.	Term	Exp(cm^{-1}) [9,10]	NN&GRASPG		CI + all-order[8]		16-electron pure CI[8]	
			$E(cm^{-1})$	diff	$E(cm^{-1})$	diff	$E(cm^{-1})$	diff
1	$3s^2 3p^4\ {}^3P_1$	19542	19535	-7	19547	5	19550	8
2	$3s^2 3p^4\ {}^3P_0$	20060	**20063**	**3**	20086	26	20081	21
3	$3s^2 3p^4\ {}^1D_2$	47033	47051	18	47063	30	47051	18
4	$3s^2 3p^4\ {}^1S_0$	97836	**97881**	**45**	97894	58	97771	-65

➢ **Lower** basis space expansion.

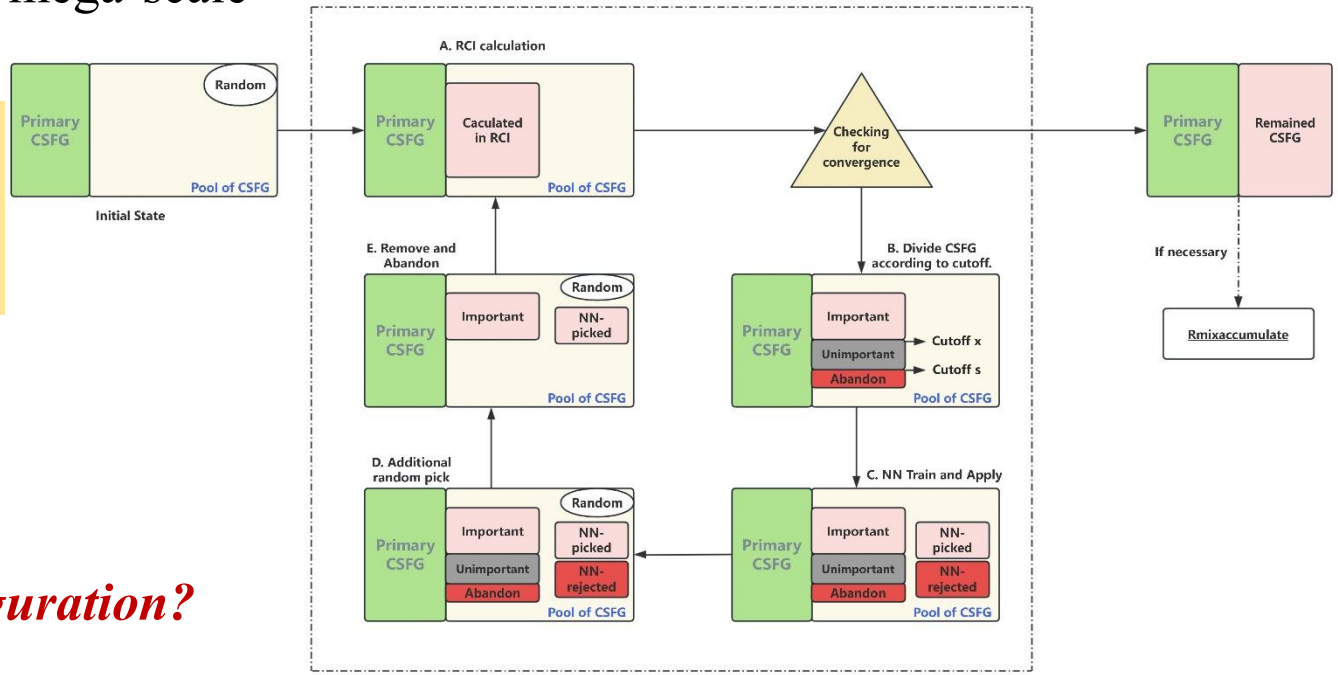
[8] arXiv:2502.05386 [9] Phys. Rev. Research 6, 013030 (2024) [10] NIST Atomic Spectra Database

Contents

NN approach in Atomic calculation

- ***GRASPG-Based NN***: An approach to handling mega-scale

- Original CSFs: **2 264 430 813**
- Condensed CSF(G)s: **228 926 656 (7 267 191)**
- Ratio ≈ 9.9 (311.6)



- NN approach based on ***(non-)Relativistic Configuration?***
- **Further Pre-classification** of CSFs

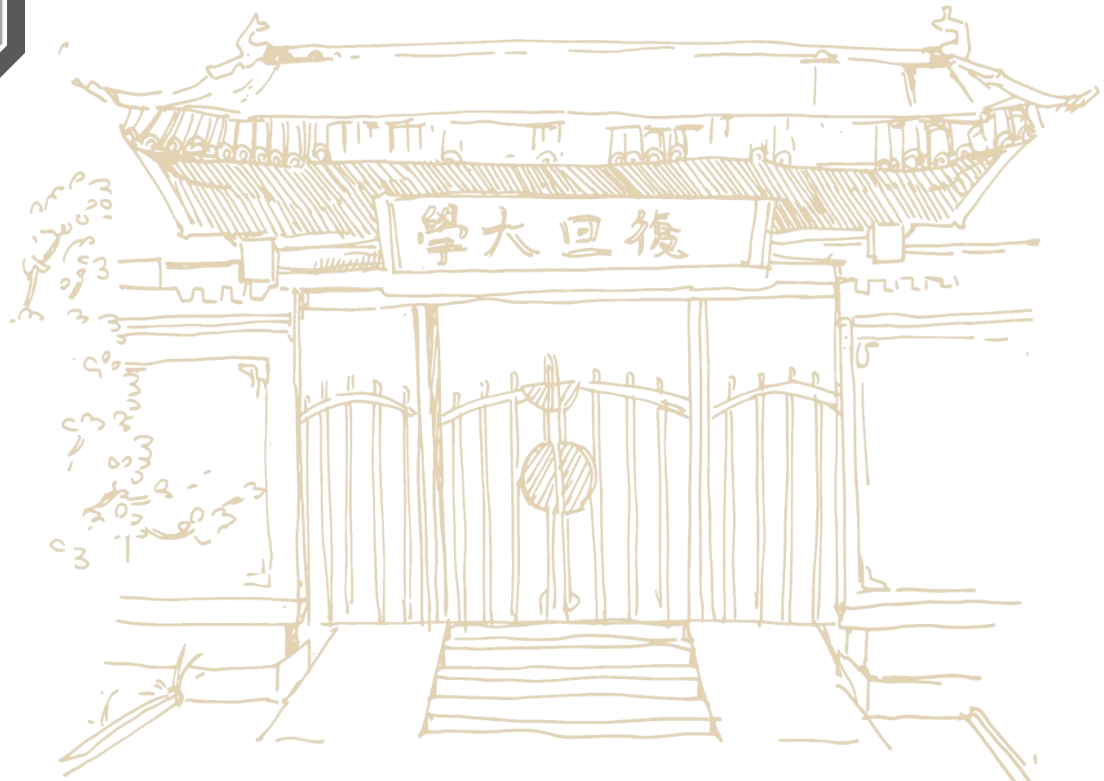
1 non-Relativistic Conf → 9 Relativistic Conf → 29 CSFGs → … CSFs

1s(2)2s(2)2p(6)3s(2)3p(2)3d(2)

Acknowledgements

➢ ***Team & Partner***

- Prof. Chongyang Chen, Prof. Ran Si
- Prof. Kai Wang
- Prof. Pavlo Bilous
- Dr. Yanting Li, Dr. Changxian Song
- Sijie Wu, Xiaohan Zhang, Shaowei Tian.



Thanks for your attention!