Unleashing the Matsubara technique:

Storing and manipulating many-body response functions with exponential convergence

Markus Wallerberger*

in collaboration with:

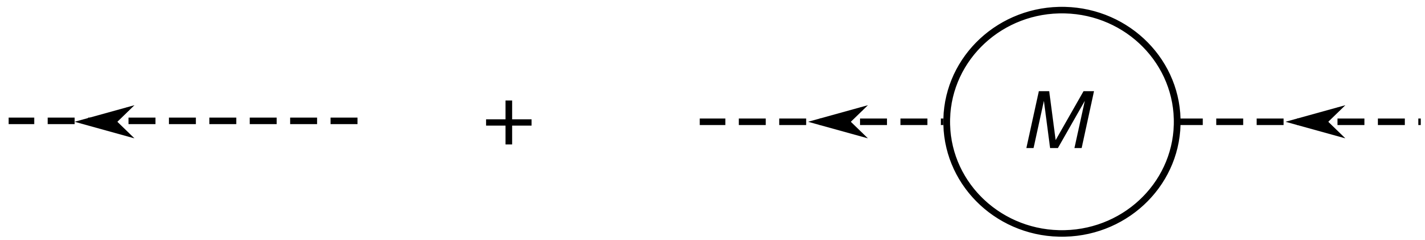
- Hiroshi Shinaoka (Saitama University, Japan)
- Jia Li, Emanuel Gull (University of Michigan)
- Anna Kauch (TU Wien)

VeCoCo P30997

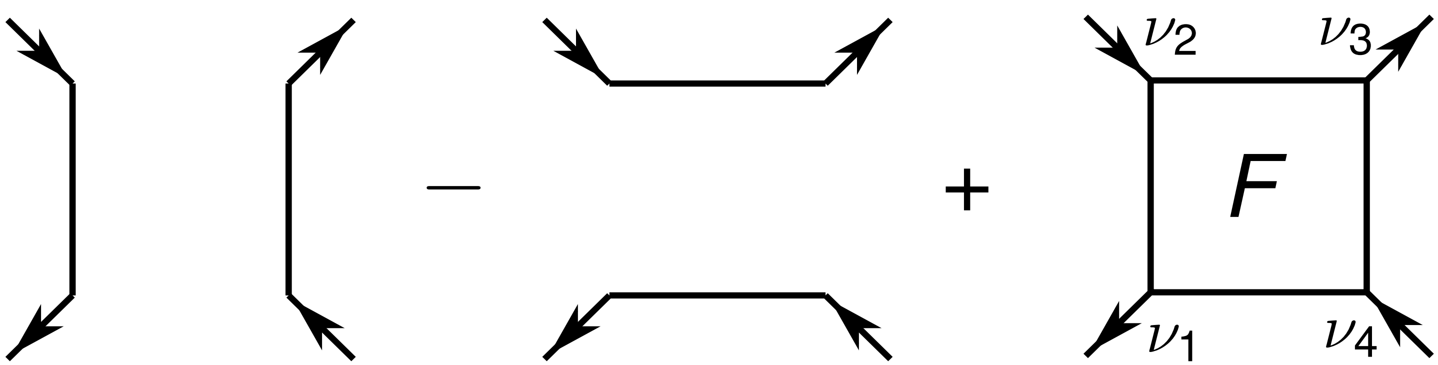
** markus.wallerberger@tuwien.ac.at*

Response functions

$G_{ab}(\omega)$



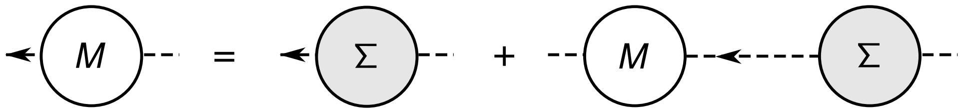
$G_{abcd}(\omega, \omega', \Omega)$



One-particle response

- Spectral function
- (Inverse) Photoemission

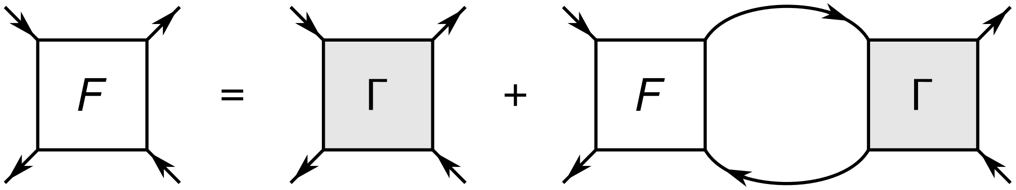
- SCF: self-energy Σ
- Approximations:
  Hartree–Fock, DMFT, ...
- Dyson equation



Two-particle response

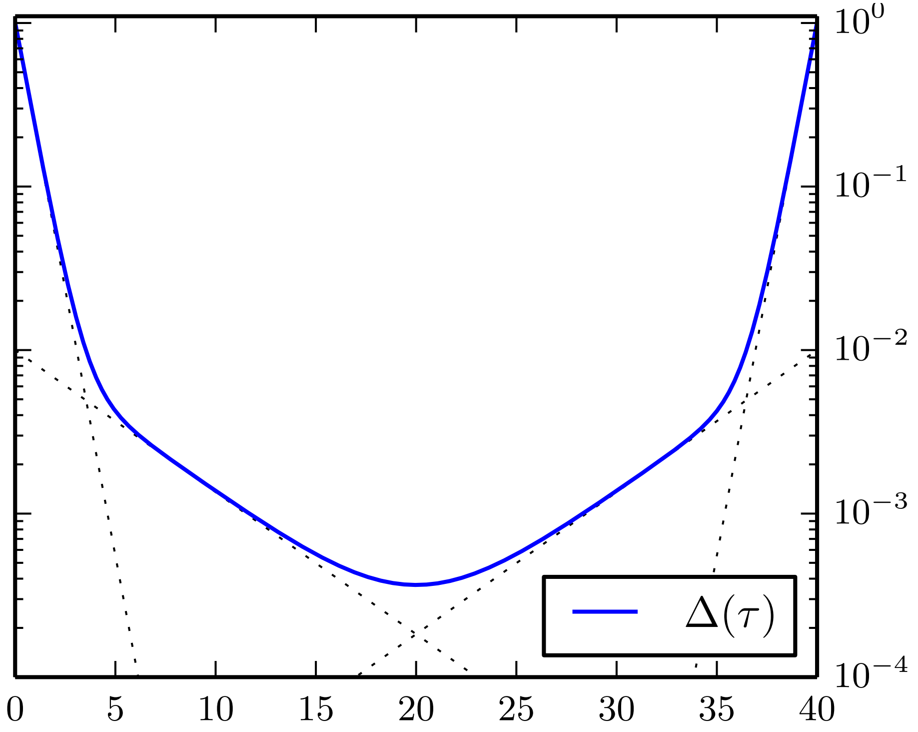
- Susceptibility
- RIXS, NMR, ...

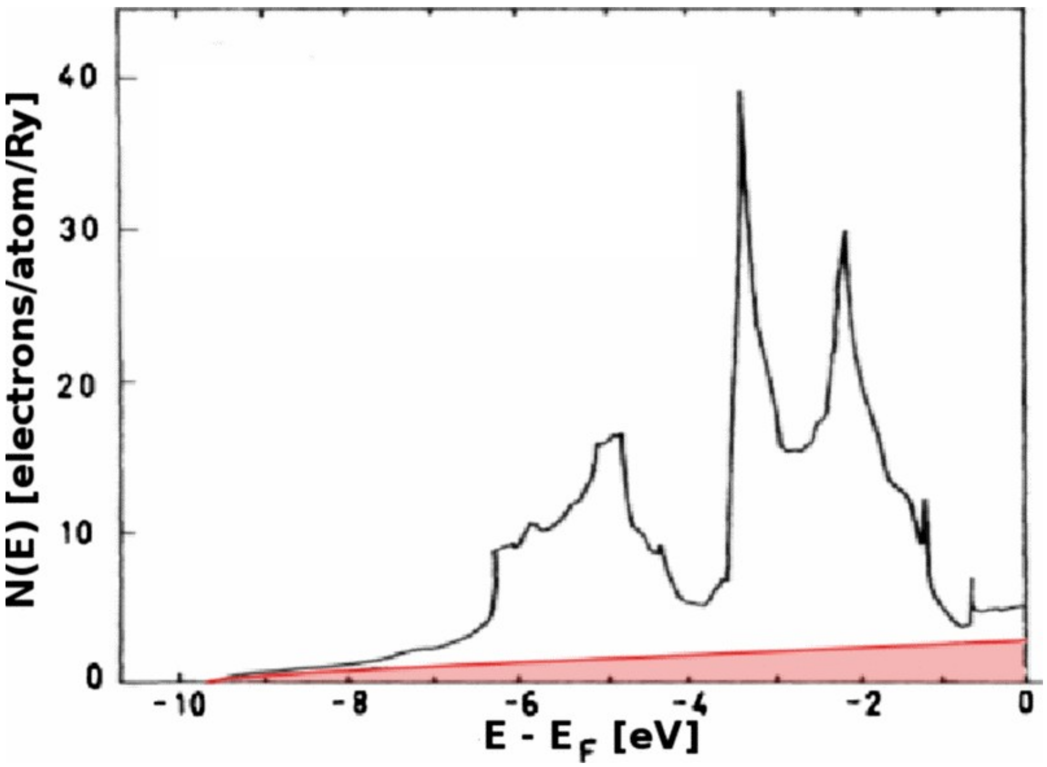
- SCF: irreducible vertex Γ
- Approximations:
  RPA, DΓA, ...
- Bethe–Salpeter equation



Imaginary time **Laplace kernel** **Real time/freq.**

$$G(\tau) = \int_{-W}^{W} \mathrm{d}\omega\, K(\tau,\omega) \times \rho(\omega)$$

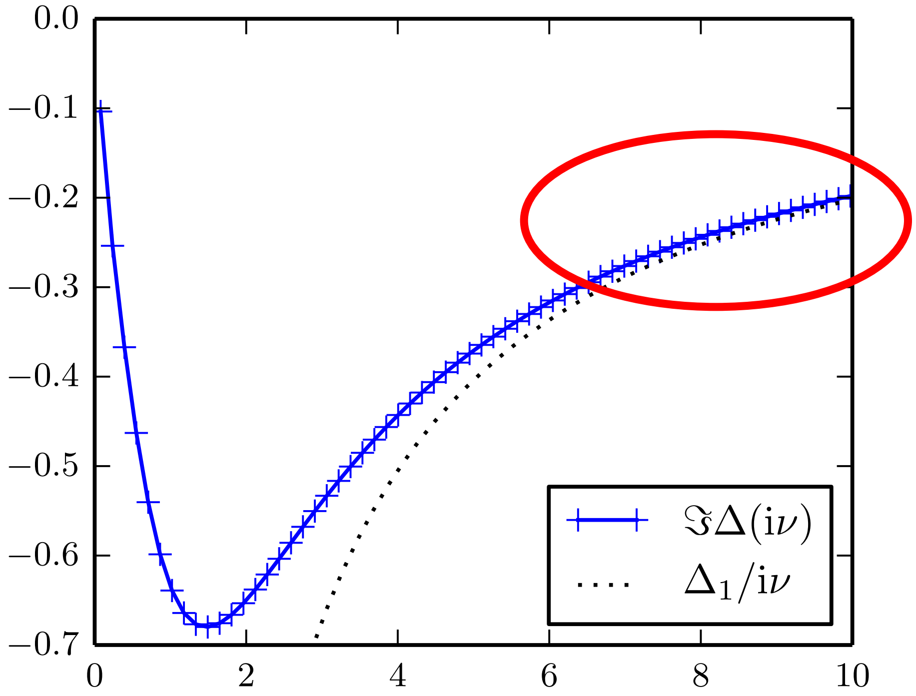




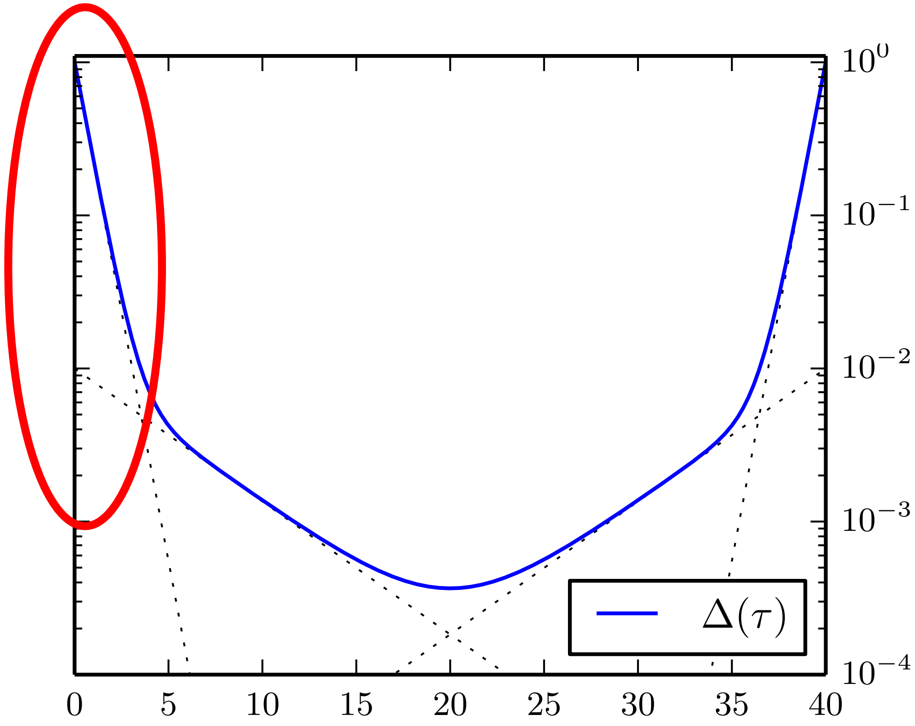
T?

Representation?

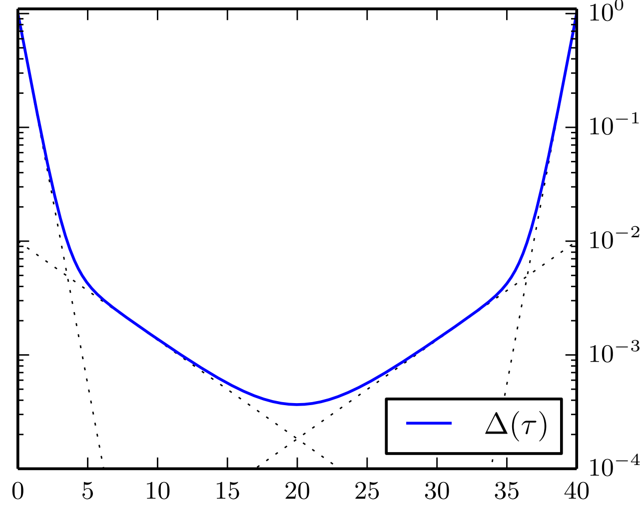
Imaginary frequency



Imaginary time



Representations



- bandwidth W, error ε, temp. $T = 1/\beta$

- Dense grid with "tails":

$$G(\tau) \approx G(i\tfrac{\beta}{N}) + G^{(K)}_{\text{model}}(\tau) \qquad N \sim \beta W \epsilon^{-K}$$

- Splines/orthogonal polynomials: [1,2]

$$G(\tau) \approx \textstyle\sum_{l=0}^{L-1} P_l(\tau) G_l \qquad L \sim \sqrt{\beta W} \log(\epsilon^{-1})$$

[1] *L. Boehnke et al., Phys. Rev. B 84, 075145 (2011)*
[2] *E. Gull et al., Phys. Rev. B 98, 075127 (2018)*

But It Works For Me!(TM)

- Does it?
  → low-energy models, moderate T, mostly 1-particle

- Otherwise:
  - W: Quantum chemistry, molecules? → ~300,000 K
  - β: superconductivity? → ~1 K^{-1}
  - ε: reliable real-frequency spectra? → 10^{-4} ~ 10^{-16}
  - N: for two-particle quantities? → N^3
  - L: Dyson equation? → L^3 (naive), L^2 [1] or back to N

[1] *X. Dong et al., J. Chem. Phys. 152, 134107 (2020)*

Task/Outline

We need representation:

- **controlled:** given error bound
- **compact (1-p)** : few coefficients, scaling $\sim \beta W$
- **fast (1-p)**: diagrammatic equations
- **compact + controlled (2-p)**: structure
- **fast (2-p)**: convolution

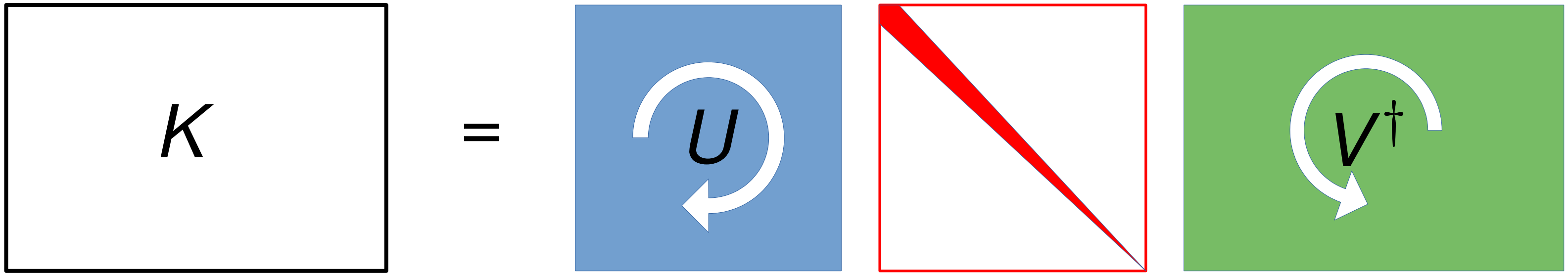
Task/Outline

We need representation:

- **controlled:** given error bound
- **compact (1-p)** : few coefficients, scaling $\sim \beta W$
- **fast (1-p)**: diagrammatic equations
- **compact + controlled (2-p)**: structure
- **fast (2-p)**: convolution

Singular Value Decomposition

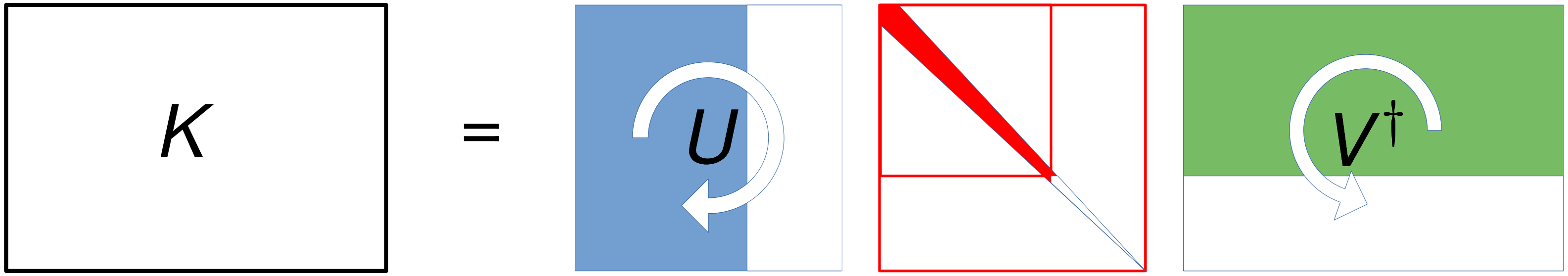
- For every matrix: $K = U\ S\ V^\dagger = \sum_{l=0}^{L-1} s_l \vec{u}_l \vec{v}_l^\dagger$
- Singular values: $s_0 \geq s_1 \geq \ldots \geq s_{L-1} \geq 0$
- Singular “bases”: $u_l^\dagger u_m = v_l^\dagger v_m = \delta_{lm}$



-

Singular Value Decomposition

- For every matrix: $K = U\ S\ V^\dagger = \sum_{l=0}^{L-1} s_l \vec{u}_l \vec{v}_l^\dagger$
- Singular values: $s_0 \geq s_1 \geq \ldots \geq s_{L-1} \geq 0$
- Singular "bases": $u_l^\dagger u_m = v_l^\dagger v_m = \delta_{lm}$

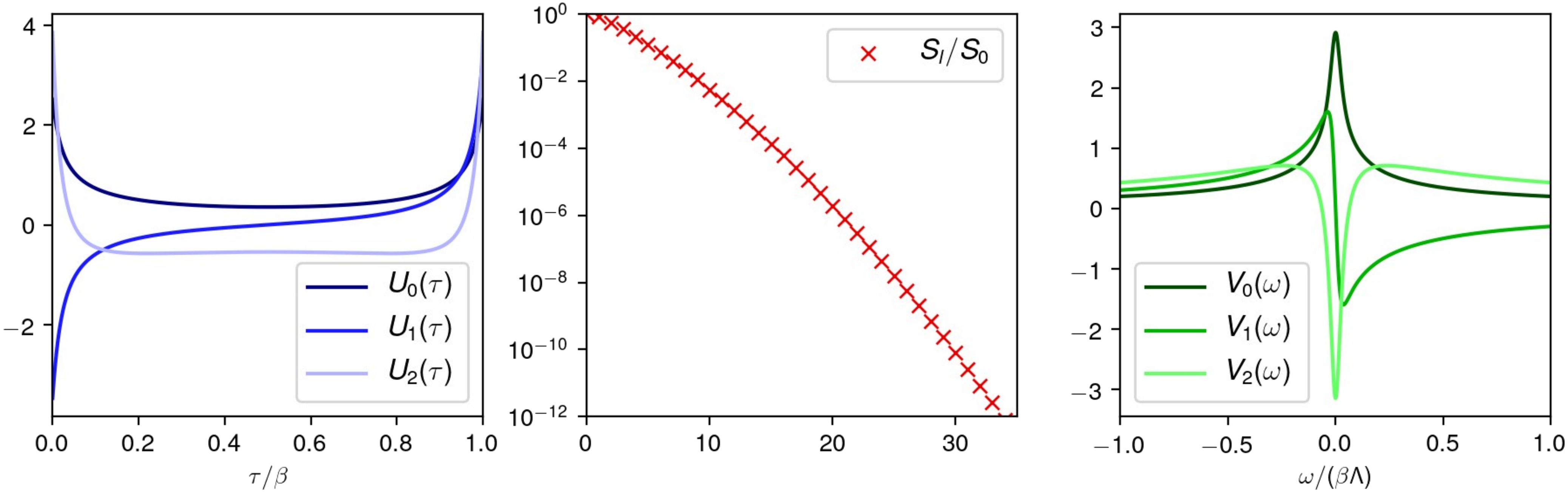


- "Best" low-rank approximation: truncate *s*

Singular Value Expansion

$$G(\tau) = \int_{-W}^{W} \mathrm{d}\omega\, K(\tau,\omega)\rho(\omega)$$

$$K(\tau,\omega) = \sum_{l=0}^{\infty} \quad U_l(\tau) \quad \times \quad S_l \quad \times \quad V_l(\omega)$$



[1] *J. Otsuki et al., Phys. Rev. E 95, 061302 (2017); H. Shinaoka et al., Phys. Rev. B 96, 035147 (2017)*

- Insert SVD:

$$G(\mathrm{i}\nu) = \int_{-W}^{W} \mathrm{d}\omega\, K(\mathrm{i}\nu, \omega)\rho(\omega)$$

$$= \sum_{l=0}^{\infty} U_l(\mathrm{i}\nu) S_l \int_{-W}^{W} \mathrm{d}\omega\, V_l(\omega)\rho(\omega)$$

- Exponential decay: $S_l \sim \exp(-\alpha \frac{l}{\log(\beta W)})$
- Saturation: $\rho_l \sim \mathrm{const}$
- **IR basis expansion**

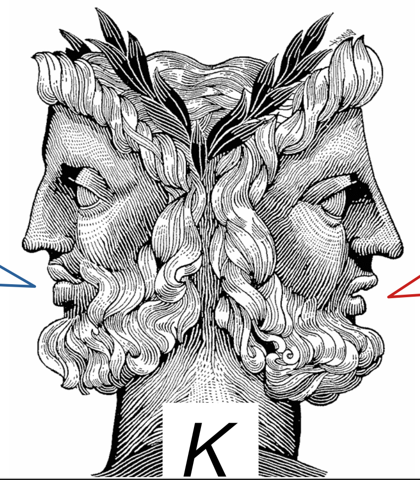
$$G(\mathrm{i}\nu) = \sum_{l=0}^{L} G_l U_l(\mathrm{i}\nu) + \epsilon_L \quad \text{with} \quad \epsilon_L \sim S_L$$

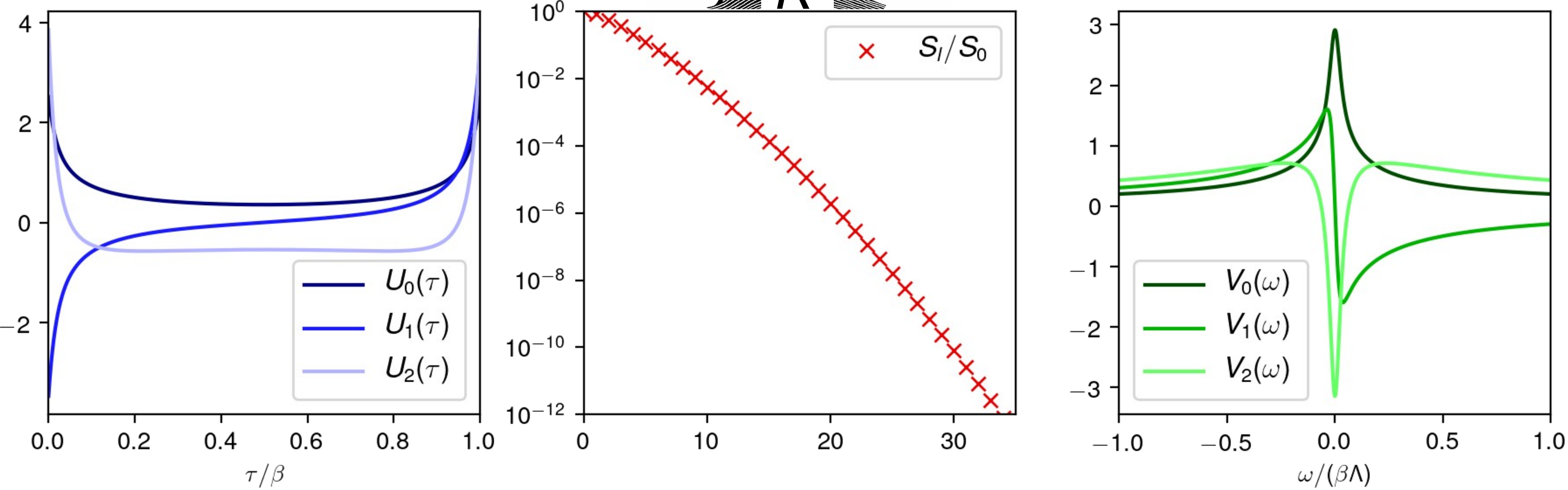
[1] *J. Otsuki, H.S., et al., Phys. Rev. E 95, 061302 (2017)*
[2] *H.S., et al., Phys. Rev. B 96, 035147 (2017)*
[3] *M.W., H.S., ..., in preperation*

No Free Lunch theorem

compression
(on the imaginary axis)

loss of information
(from the real axis)





$$G(\mathrm{i}\nu) = \sum_{l=0}^{L} G_l U_l(\mathrm{i}\nu) + \epsilon_L$$

$$L \sim \log(\beta W) \log \epsilon^{-1}$$

$$\epsilon \sim S_L / S_0$$

Task/Outline

We need representation:

- **controlled:** $\epsilon \sim S_L / S_0$
- **compact (1-p) :** $L \sim \log(\beta W) \log \epsilon^{-1}$
- **fast (1-p)**: diagrammatic equations
- **compact + controlled (2-p)**: structure
- **fast (2-p)**: convolution

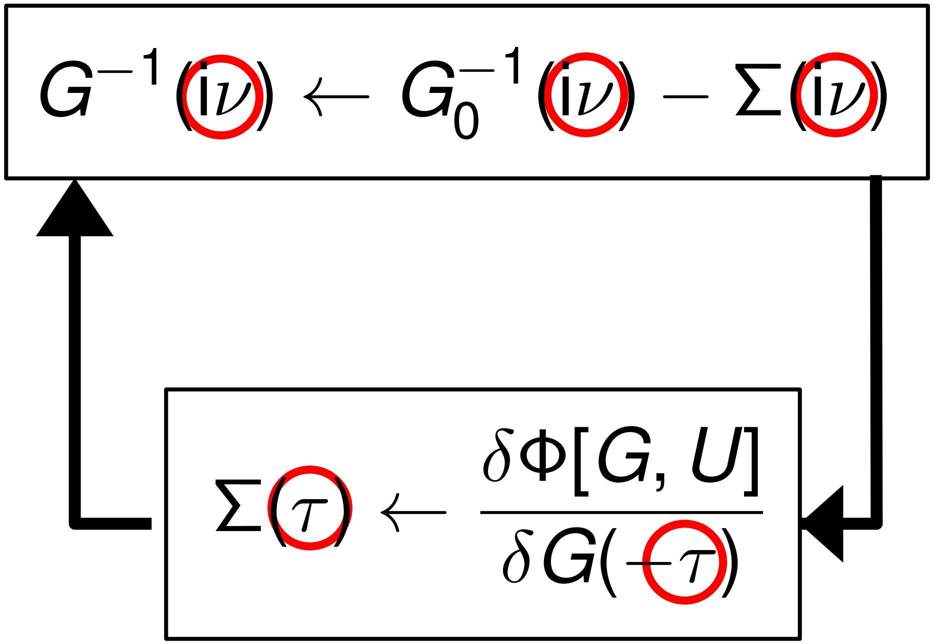
Task/Outline

We need representation:

- **controlled:** $\epsilon \sim S_L / S_0$
- **compact (1-p) :** $L \sim \log(\beta W) \log \epsilon^{-1}$
- **fast (1-p)**: diagrammatic equations
- **compact + controlled (2-p)**: structure
- **fast (2-p)**: convolution

Self-consistent field (SCF) equations

- "Dressing propagators"
  → frequency

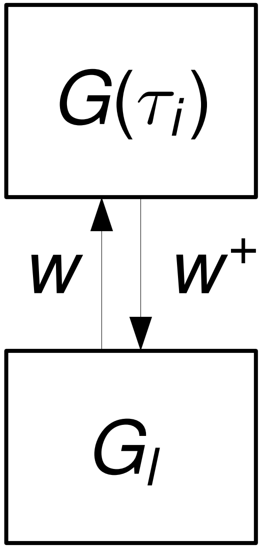


- Summing diagrams
  → *U* → time

- Costly (L^3) in IR basis

Sparse sampling in time [1,2]

- Chebyshev polynomials
- Gauss/Clenshaw–Curtis quadrature:
  - Sampling points $\{\tau_i\}$ = zeros of $T_l(\tau)$



$$G_l = \int_0^\beta \mathrm{d}\tau\, G(\tau)\, T_l(\tau) = \sum_{i=0}^{L-1} w_{li} G(\tau_i) + \epsilon$$

- Inverse: evaluation

$$G(\tau_i) = \sum_{l=0}^{\infty} U_l(\tau_i) G_l = \sum_{l=0}^{L-1} w_{il}^{+} G_l + \epsilon$$

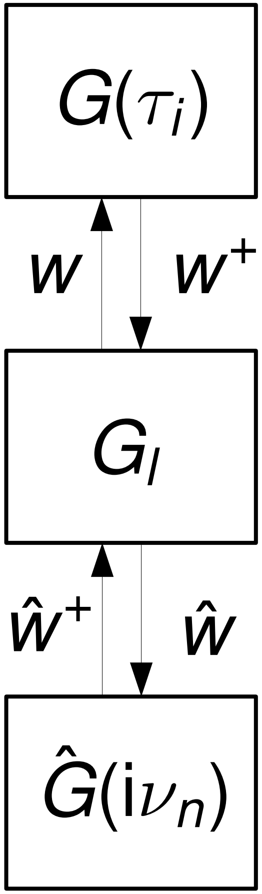
[1] *Jia Li, M. W., N. Chikano, C.N. Yeh, E. Gull, H. Shinaoka, Phys. Rev. B 101, 035144 (2020) – preprint: arXiv:1908.07575*
[2] independent work*: M. Kaltak and G. Kresse, Phys. Rev. B 101, 205145 (2020) – preprint: arXiv:1909.01740*

Sparse sampling in frequency [1,2]

- Fourier transform: polynomial in $1/\mathrm{i}\nu$
  - Sampling frequencies $\{ \nu_n \} \leftrightarrow T_L(1/\mathrm{i}\nu)$

$$G_l = \frac{1}{\beta} \sum_{\nu} \hat{G}(\mathrm{i}\nu) \hat{T}_l(\mathrm{i}\nu) = \sum_{i=0}^{L-1} \hat{w}_{ln} \hat{G}(\mathrm{i}\nu_n) + \epsilon$$

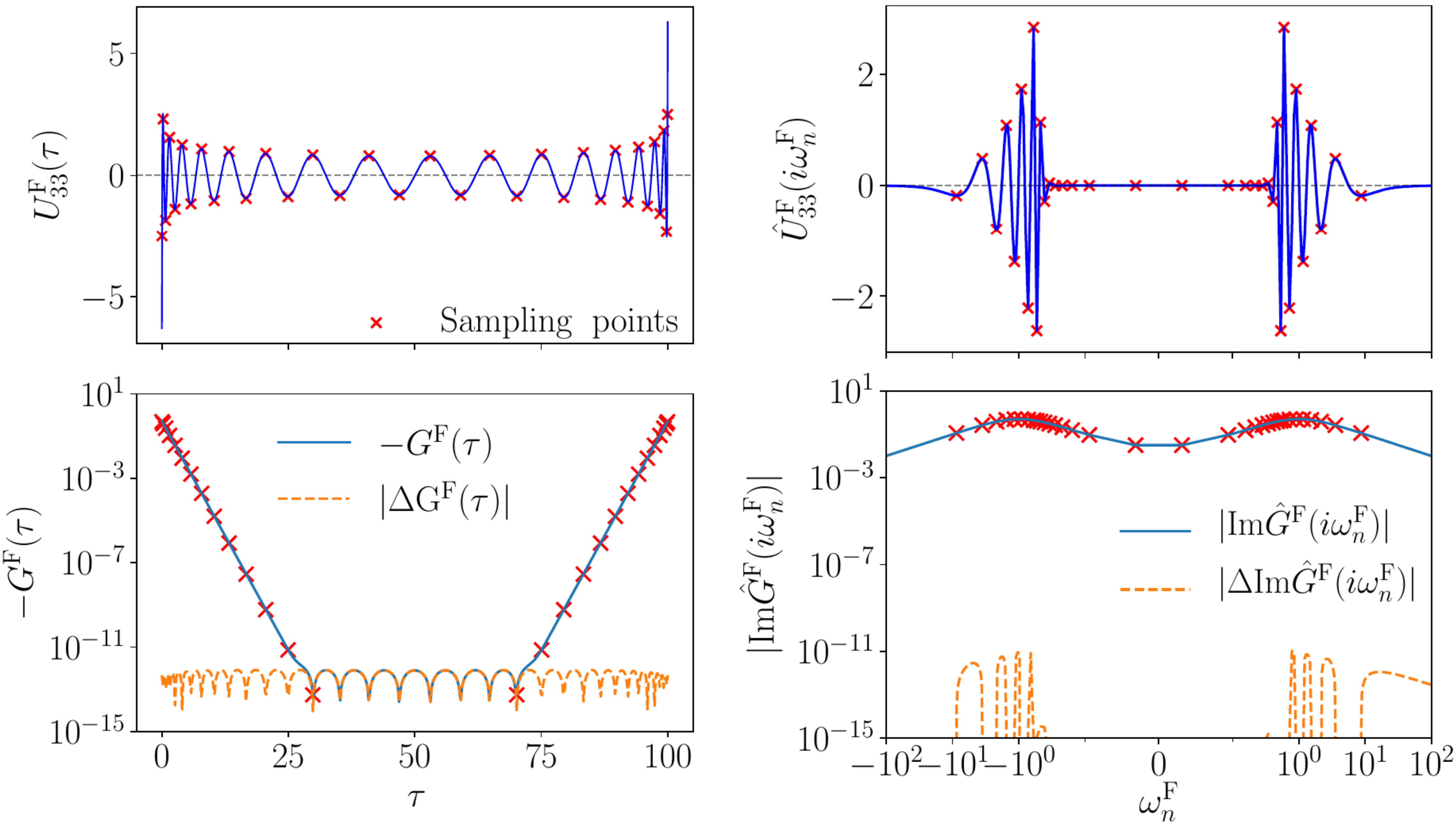
$$G(\mathrm{i}\nu_n) = \sum_{l=0}^{\infty} \hat{U}_l(\mathrm{i}\nu_n) G_l = \sum_{l=0}^{L-1} \hat{w}^+_{nl} G_l + \epsilon$$



- Basis of *K* behaves similar to polynomials[3]
  → Also works for IR basis functions![4]

[1] *Jia Li, M. W., N. Chikano, C.N. Yeh, E. Gull, H. Shinaoka, Phys. Rev. B 101, 035144 (2020) – preprint: arXiv:1908.07575*
[2] independent work: *M. Kaltak and G. Kresse, Phys. Rev. B 101, 205145 (2020) – preprint: arXiv:1909.01740*
[3] Karlin, Total positivity 1968; [4] MW, et al., in preparation

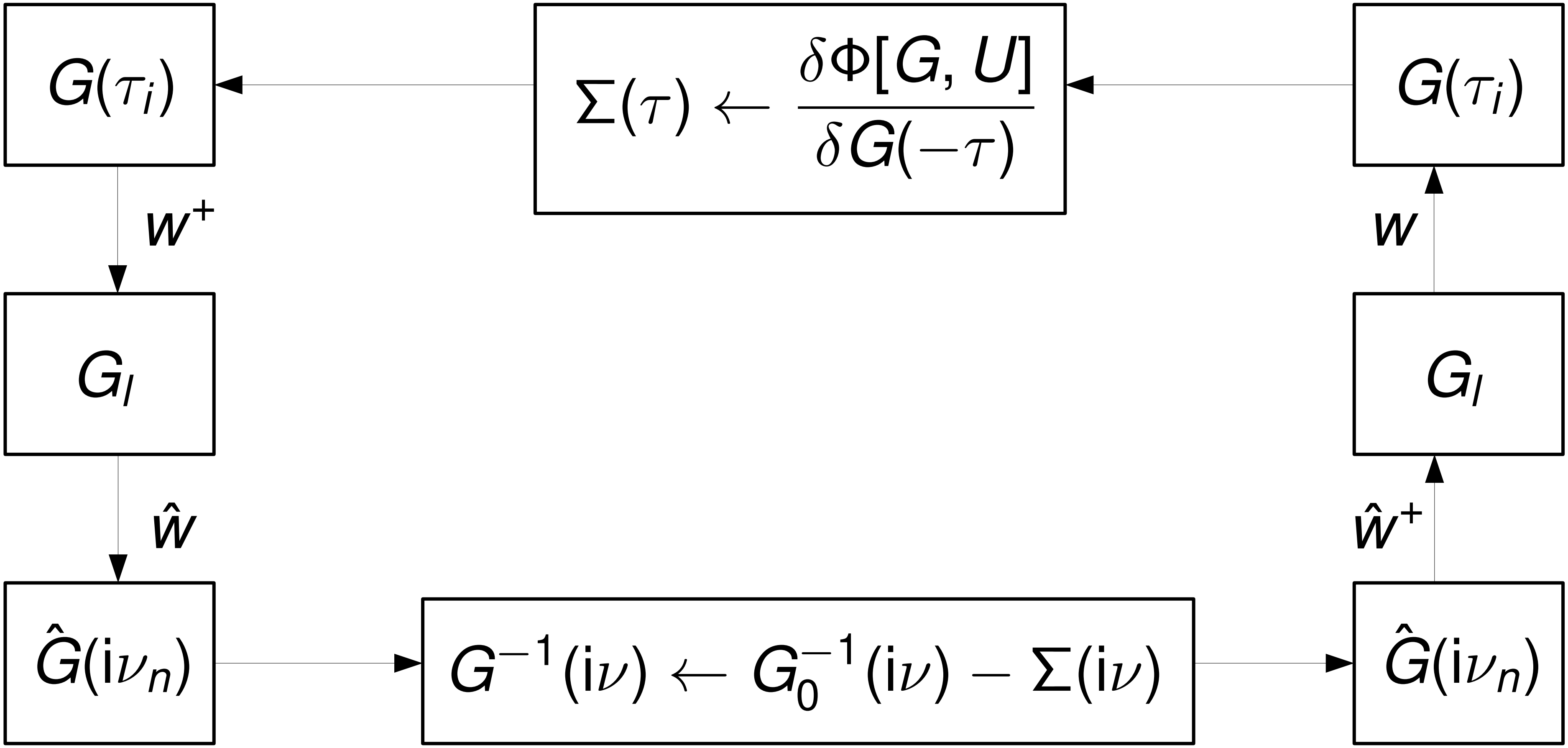
Sparse sampling



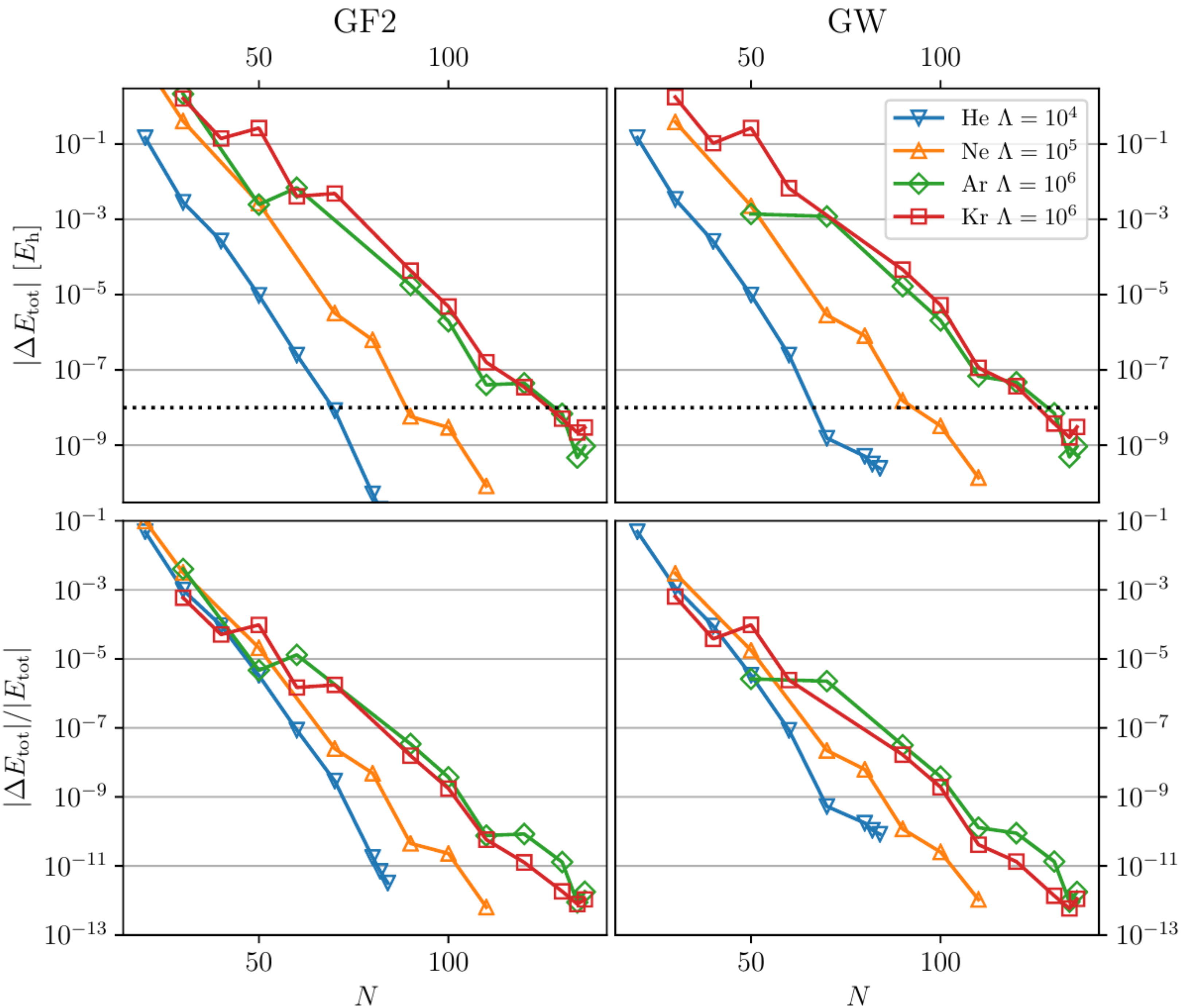
[1] *Jia Li, M. W., N. Chikano, C.N. Yeh, E. Gull, H. Shinaoka, Phys. Rev. B 101, 035144 (2020)*

$$\boxed{\Sigma(\tau) \leftarrow \frac{\delta\Phi[G, U]}{\delta G(-\tau)}}$$

$$\boxed{G^{-1}(\mathrm{i}\nu) \leftarrow G_0^{-1}(\mathrm{i}\nu) - \Sigma(\mathrm{i}\nu)}$$


$G(\tau_i)$
$\Sigma(\tau) \leftarrow \frac{\delta \Phi[G, U]}{\delta G(-\tau)}$
$G(\tau_i)$
w^+
w
G_l
G_l
$\hat{w}$
$\hat{w}^+$
$\hat{G}(\mathrm{i}\nu_n)$
$G^{-1}(\mathrm{i}\nu) \leftarrow G_0^{-1}(\mathrm{i}\nu) - \Sigma(\mathrm{i}\nu)$
$\hat{G}(\mathrm{i}\nu_n)$

Benchmark: total E of noble gases



Jia Li, M. W., N. Chikano, C.N. Yeh, E. Gull, H. Shinaoka
Phys. Rev. B 101, 035144 (2020)

Task/Outline

We need representation:

- **controlled:** $\epsilon \sim S_L / S_0$
- **compact (1-p) :** $L \sim \log(\beta W \epsilon^{-1})$
- **fast (1-p)**: $\sim \mathcal{O}(L), \quad \sim \mathcal{O}(L^2)$
- **compact + controlled (2-p)**: structure
- **fast (2-p)**: convolution

Task/Outline

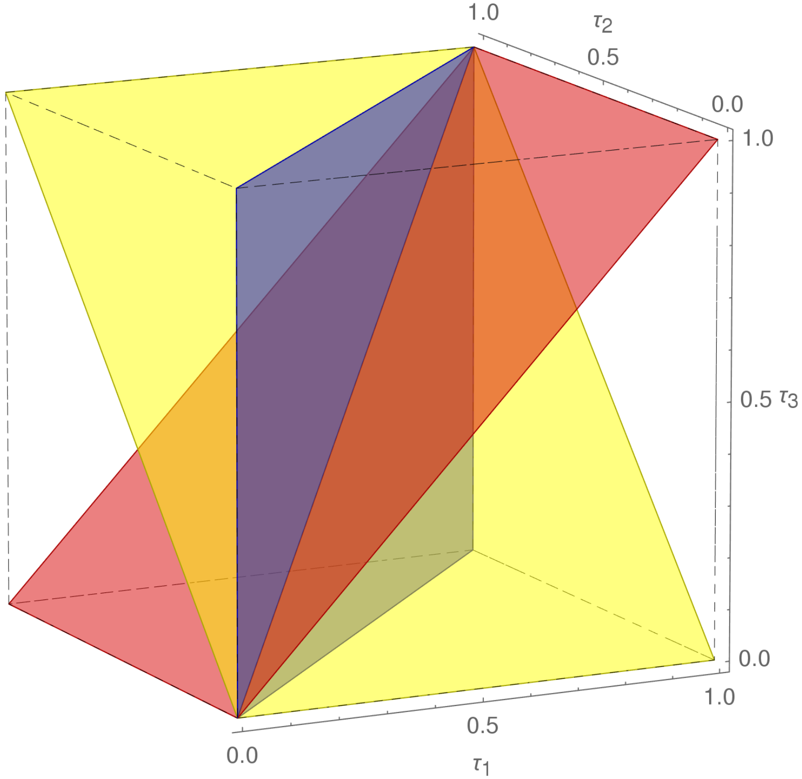
We need representation:

- **controlled:** $\epsilon \sim S_L / S_0$
- **compact (1-p) :** $L \sim \log(\beta W \epsilon^{-1})$
- **fast (1-p):** $\sim \mathcal{O}(L), \quad \sim \mathcal{O}(L^2)$
- **compact + controlled (2-p):** structure
- **fast (2-p):** convolution

The problem with 2 particles

- First try: $$G(\tau_1, \tau_2, \tau_3) = \sum_{ll'm} U_l(\tau_1) U_{l'}(\tau_2) U_m(\tau_3) G_{ll'm}$$

- Discontinuity planes
- → product basis not compact!



[1] image: *M. W., Ph. D. thesis, 2016.*

- two-particle response function[1]

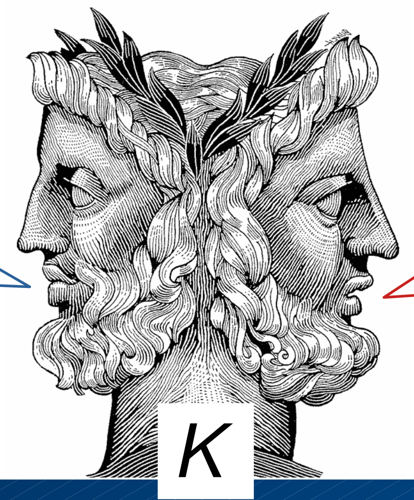
$$G(\mathrm{i}\vec{\nu}) = \sum_{r=1}^{12} T_r(\mathrm{i}\vec{\nu}; \mathrm{i}\nu, \mathrm{i}\nu', \mathrm{i}\omega) \int \mathrm{d}^3\omega K(\mathrm{i}\nu, \omega_1) K(\mathrm{i}\nu', \omega_2) K(\mathrm{i}\omega, \omega_3) \rho(\omega_1, \omega_2, \omega_3)$$

frequency translation
(→ **overcomplete**)

3x IR basis expansion
(→ **compact**)

- vertices like propagators, except Hartree–Fock term
→ augmented kernels

compression
(on the imaginary axis)



loss of information
(from the real axis)

[1] *H. S., D. Geffroy, M. W., ..., SciPost Phys. 8, 012 (2020)*

Overcomplete 2-particle basis

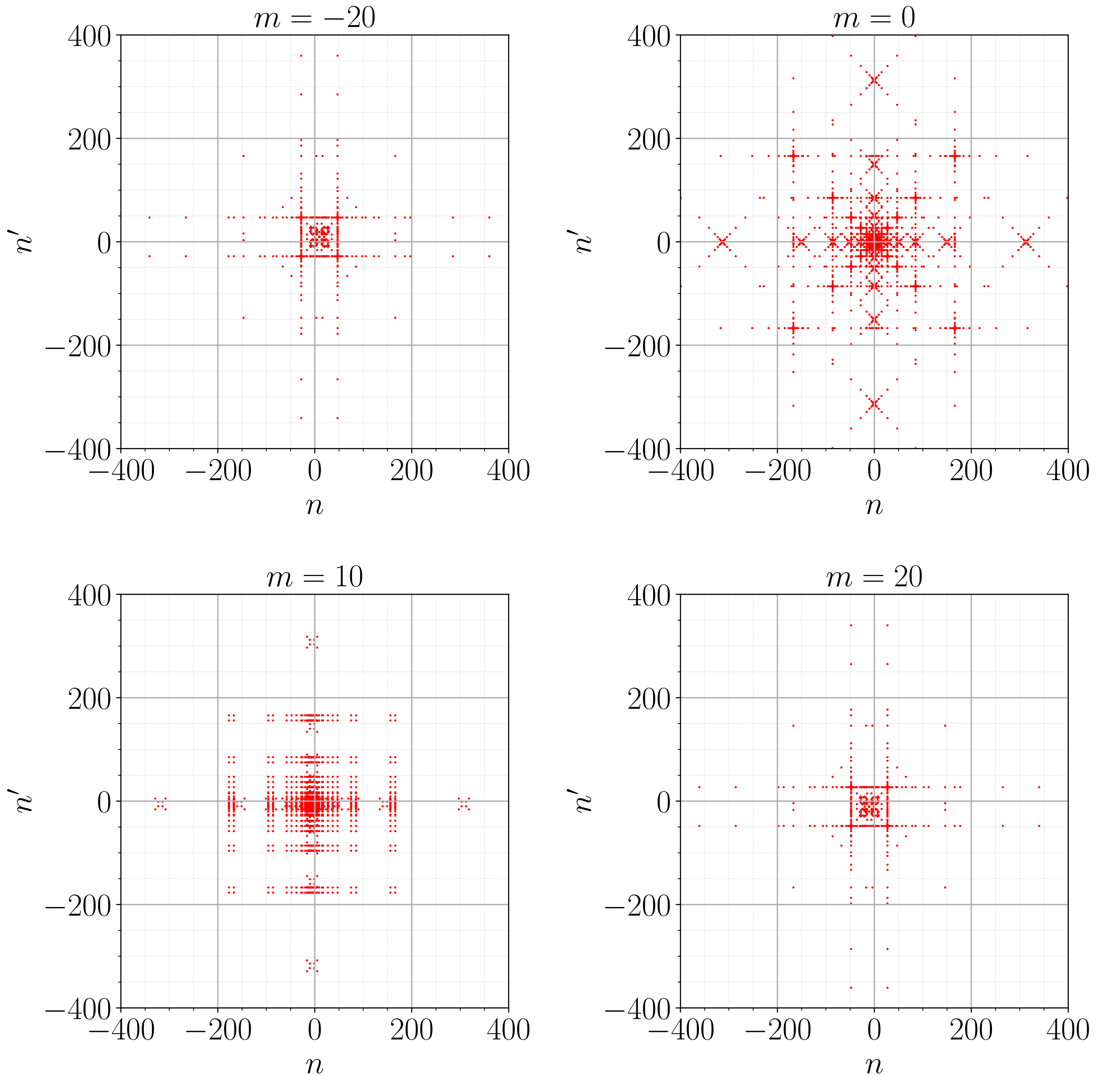
- Overcomplete basis:

$$\hat{G}(\mathrm{i}\vec{\nu}) = \underbrace{\sum_{r=1}^{12} T_r(\mathrm{i}\vec{\nu}; \mathrm{i}\nu, \mathrm{i}\nu', \mathrm{i}\omega) \sum_{ll'm} \hat{U}_l(\mathrm{i}\nu)\hat{U}_{l'}(\mathrm{i}\nu')\hat{U}_m(\mathrm{i}\omega)}_{E_{r,ll'm}(i\vec{\nu})} G_{r,ll'm}$$

- Sparse sampling on sampling frequencies:

$$G = \arg\min_G \sum_{\vec{\nu}} \left|\hat{G}(\vec{\nu}) - E(\vec{\nu})G\right|^2 + (\text{regularization})$$

[1] *H. S., D. Geffroy, M. W., ..., SciPost Phys. 8, 012 (2020)*



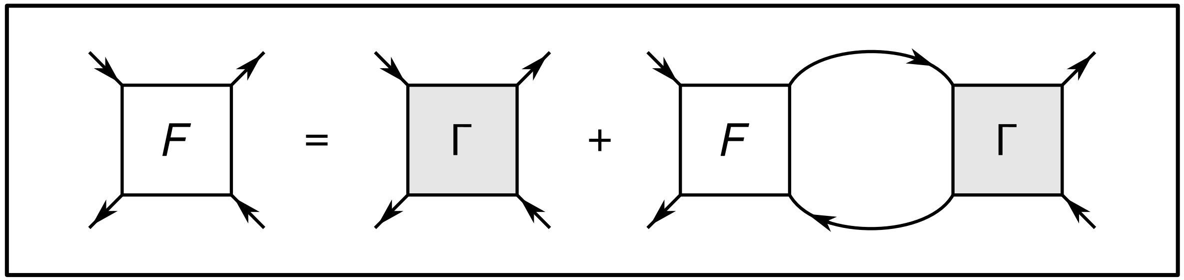
[1] *H. S., D. Geffroy, M. W., ..., SciPost Phys. 8, 012 (2020)*

Task/Outline

We need representation:

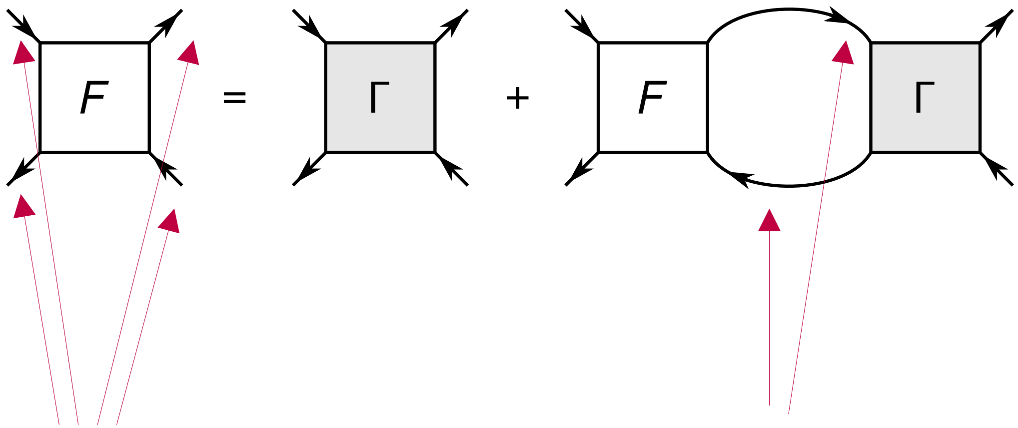
- **controlled:** $\epsilon \sim S_L / S_0$
- **compact (1-p) :** $L \sim \log(\beta W \epsilon^{-1})$
- **fast (1-p):** $\sim \mathcal{O}(L), \quad \sim \mathcal{O}(L^2)$
- **compact + controlled (2-p):** structure
- **fast (2-p):** convolution

Two-particle SCF equations



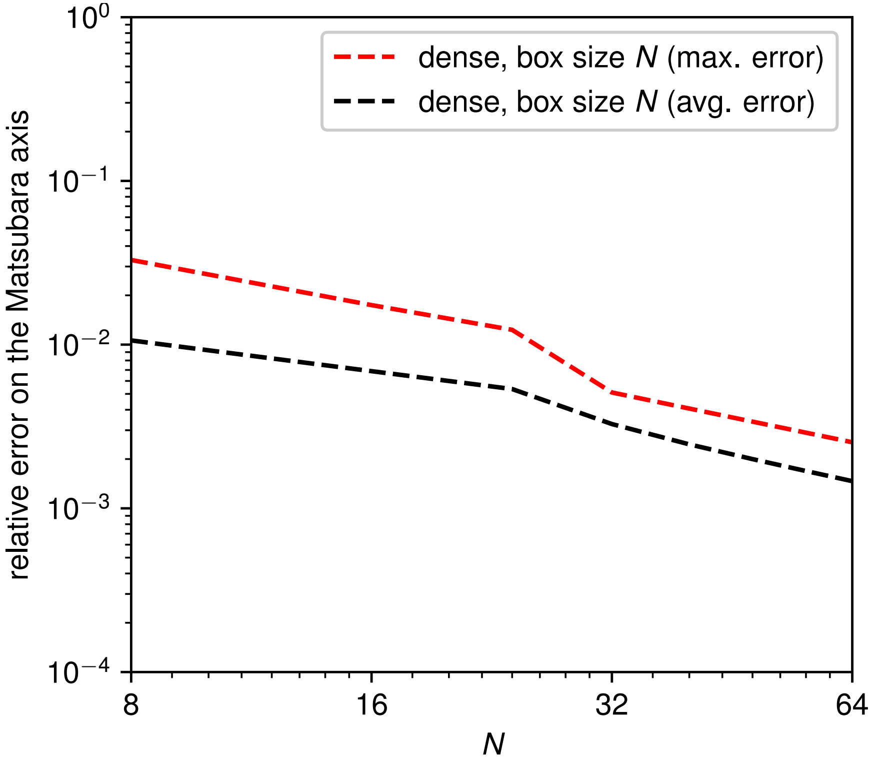
$$\Gamma \leftarrow \frac{\delta^2 \Phi[G, U]}{\delta G \delta G}$$

Bethe–Salpeter equation



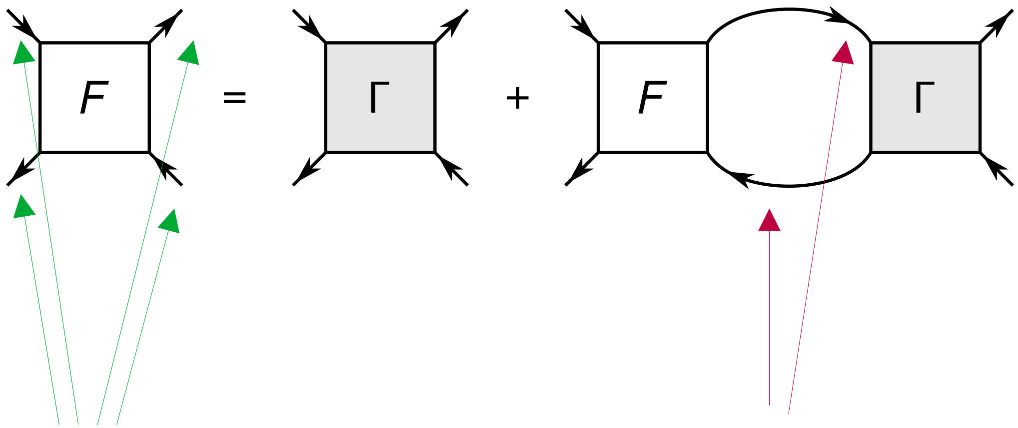
storage?
$\mathcal{O}(N_{\mathrm{orb}}^4 N^3)$

convolution?
$\mathcal{O}(N_{\mathrm{orb}}^6 N^4)$



$N \sim \epsilon^{-1} \beta W$ (accuracy * bandwidth / temperature)

Bethe–Salpeter equation



storage!

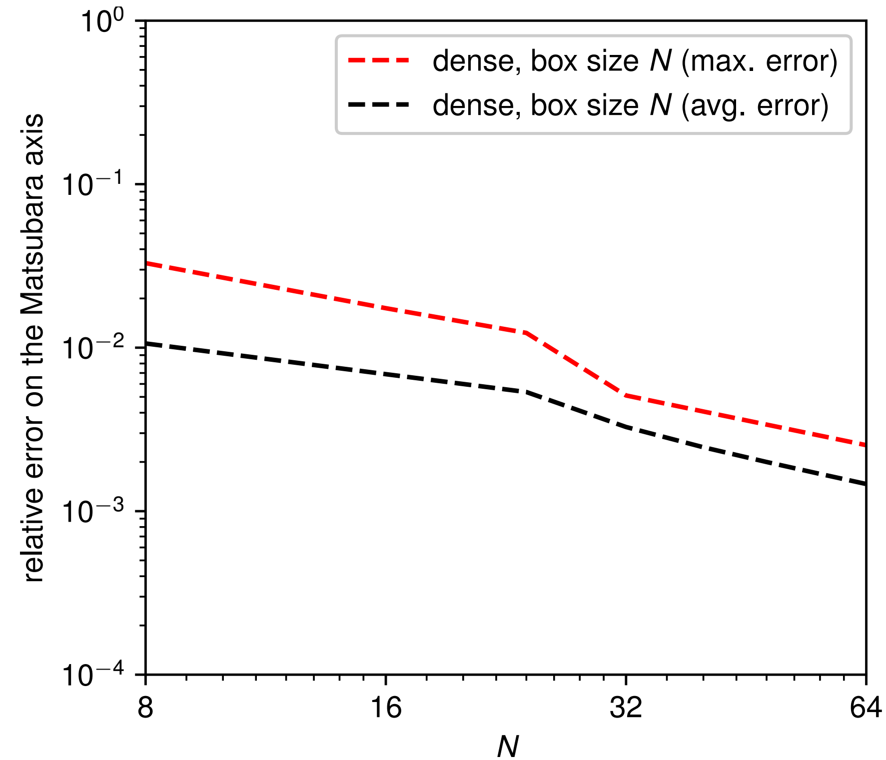
$\mathcal{O}(N_{\mathrm{orb}}^4 L^3)$

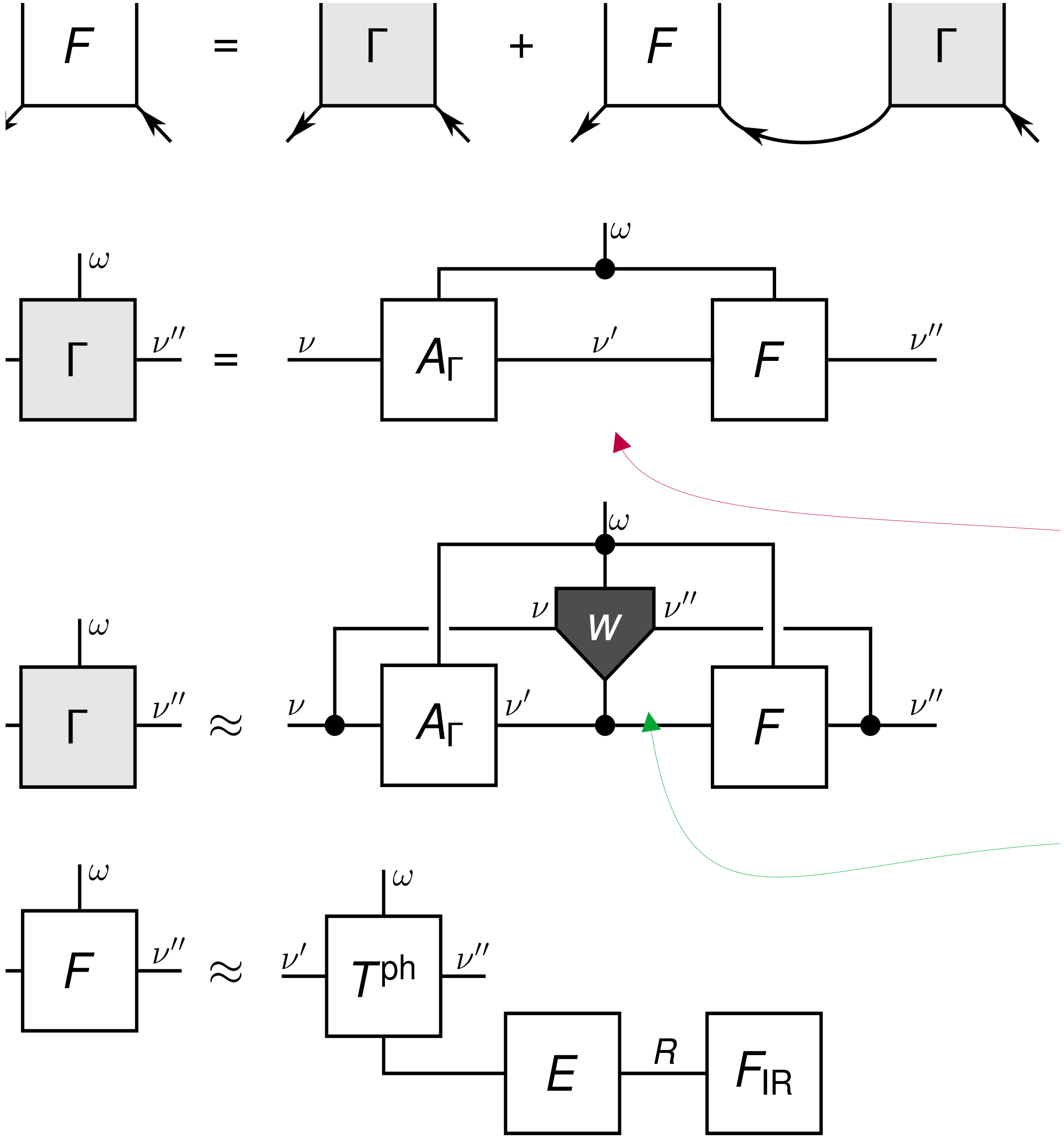
$L \sim \log(\beta W) \log \epsilon^{-1}$

convolution?

$\mathcal{O}(N_{\mathrm{orb}}^6 N^4)$

$N \sim \epsilon^{-1} \beta W$



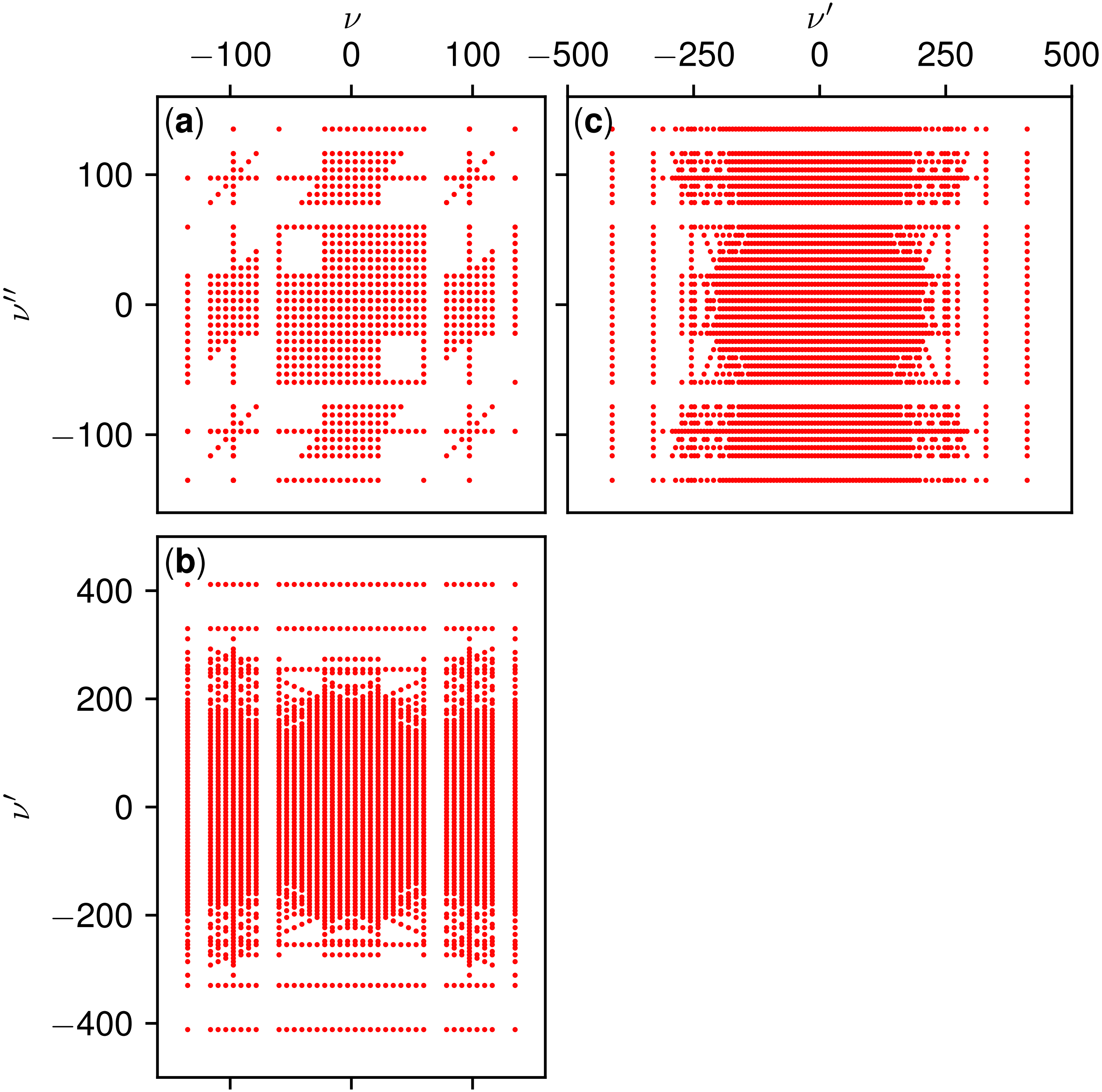


dense summation

$\sim N^4$ $\epsilon \sim N^{-1}$

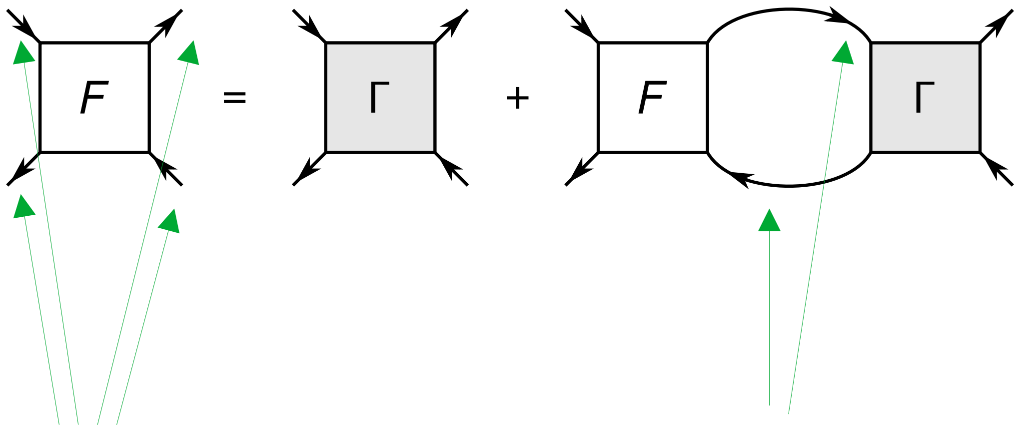
sparse summation

$\sim L^4$ $\epsilon \sim e^{-\alpha L}$



ν
−100
0
100
ν′
−500
−250
0
250
500
(a)
(c)
100
0
−100
ν′′
(b)
400
200
0
−200
−400
ν′

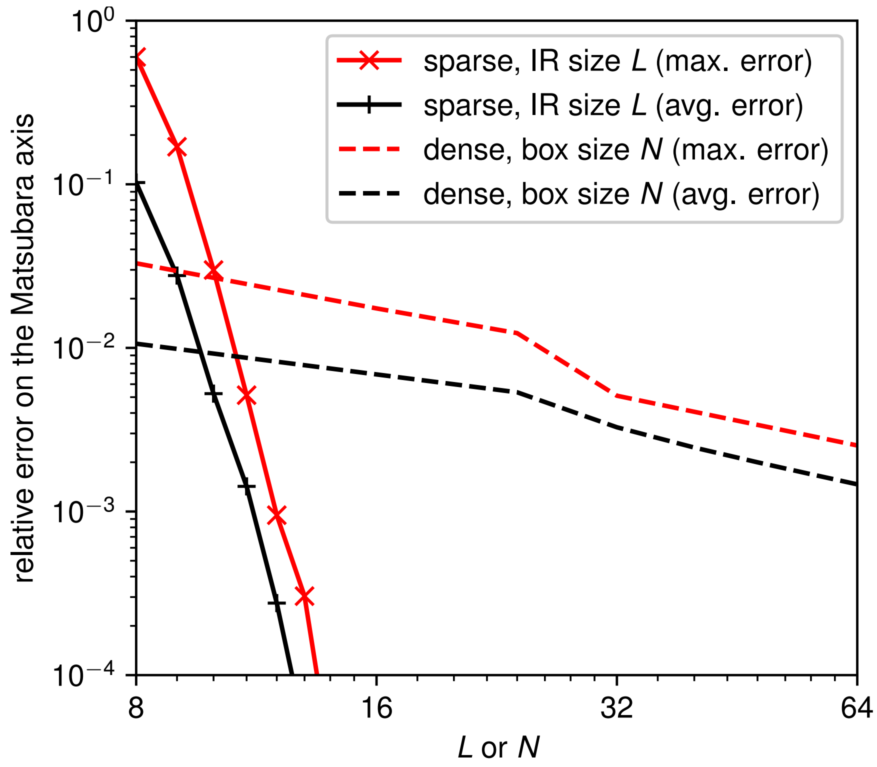
Bethe–Salpeter equation



storage!
$\mathcal{O}(N_{\text{orb}}^4 L^3)$

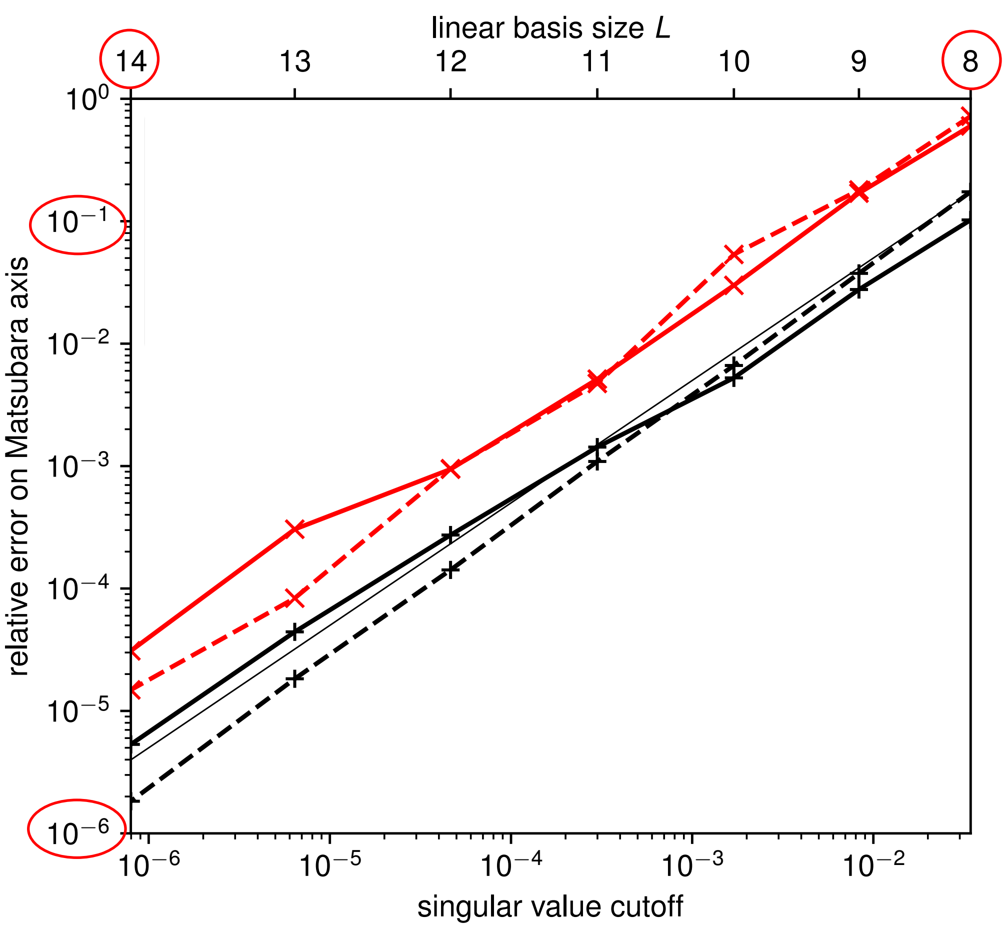
convolution!
$\mathcal{O}(N_{\text{orb}}^6 L^4)$

$L \sim \log(\beta W) \log \epsilon^{-1}$



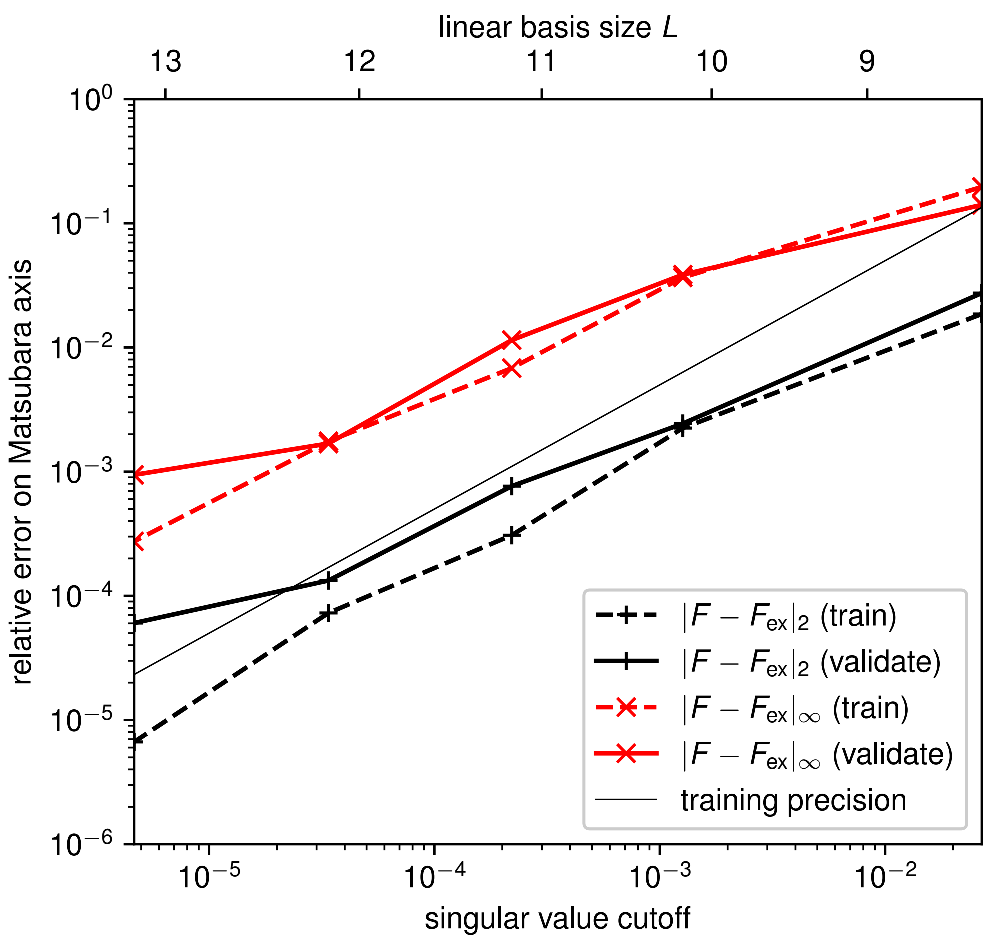
Numerical benchmarks

Hubbard atom



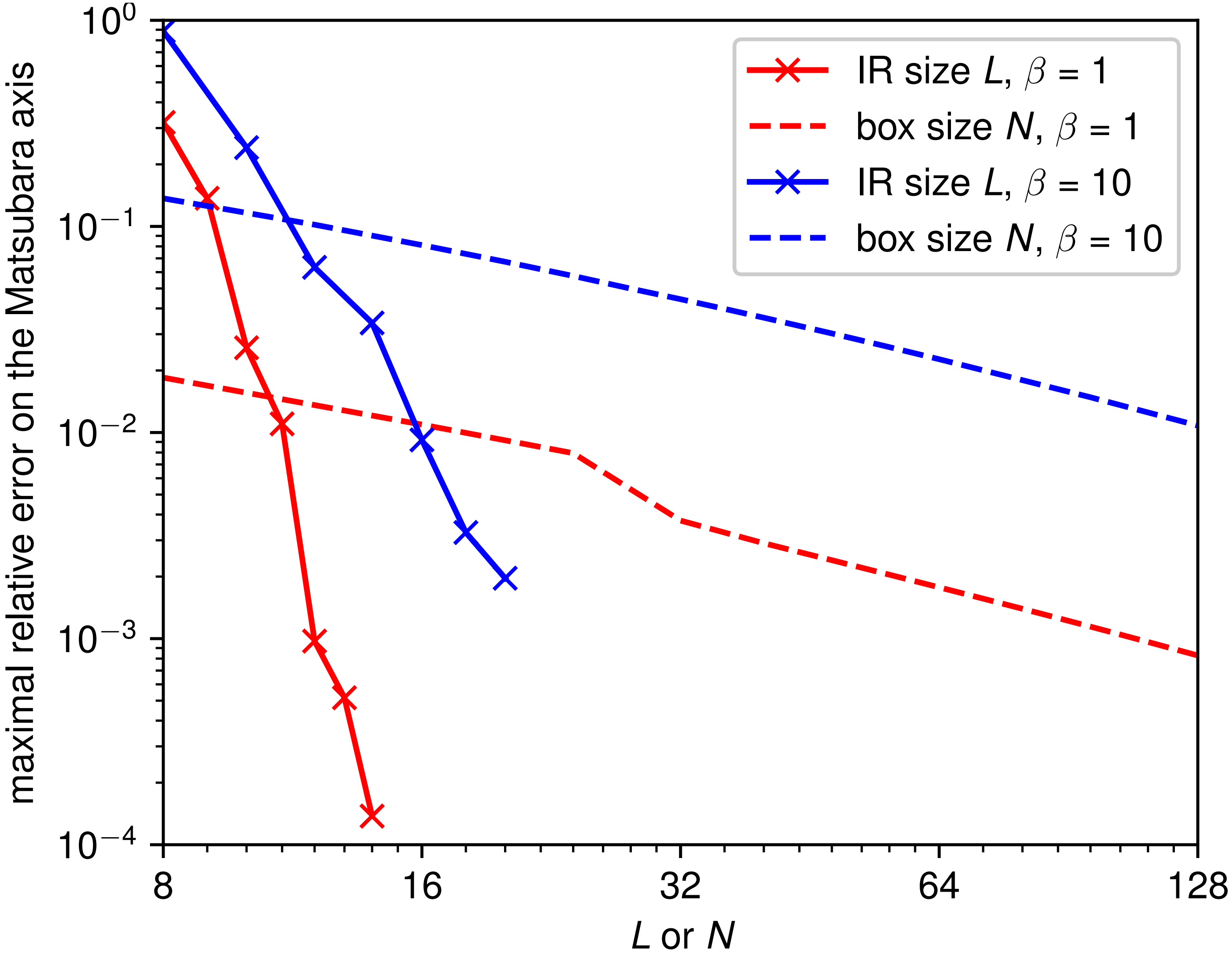
$U = 2.3, \beta = 1.5, \mu = \frac{U}{2}$

Weak coupling



$\Gamma \approx U, ||U|| = 0.3, \beta = 1.55$

Single impurity Anderson model



$\Lambda \approx U$

$U = 1.59\Delta$

$W = 5$

Task/Outline

We need representation:

- **controlled:** $\epsilon \sim S_L / S_0$
- **compact (1-p)** : $L \sim \log(\beta W) \log \epsilon^{-1}$
- **fast (1-p)**: $\sim \mathcal{O}(L), \quad \sim \mathcal{O}(L^2)$
- **compact + controlled (2-p)**: structure
- **fast (2-p)**: convolution

Conclusions and Outlook

- Compression of propagators and equations
- Cost is polynomial in $L \sim \log(\beta W) \log \epsilon^{-1}$
- Controllable, exponentially decaying error
- Accurate multi-orbital computations
- Paradigm shift: brute force HPC → sparse modelling
- Outlook: Tensor network model, Parquet equations