

Nevanlinna Analytical Continuation

Emanuel Gull



ERG 2022 Berlin

Tue, July 26, 2022

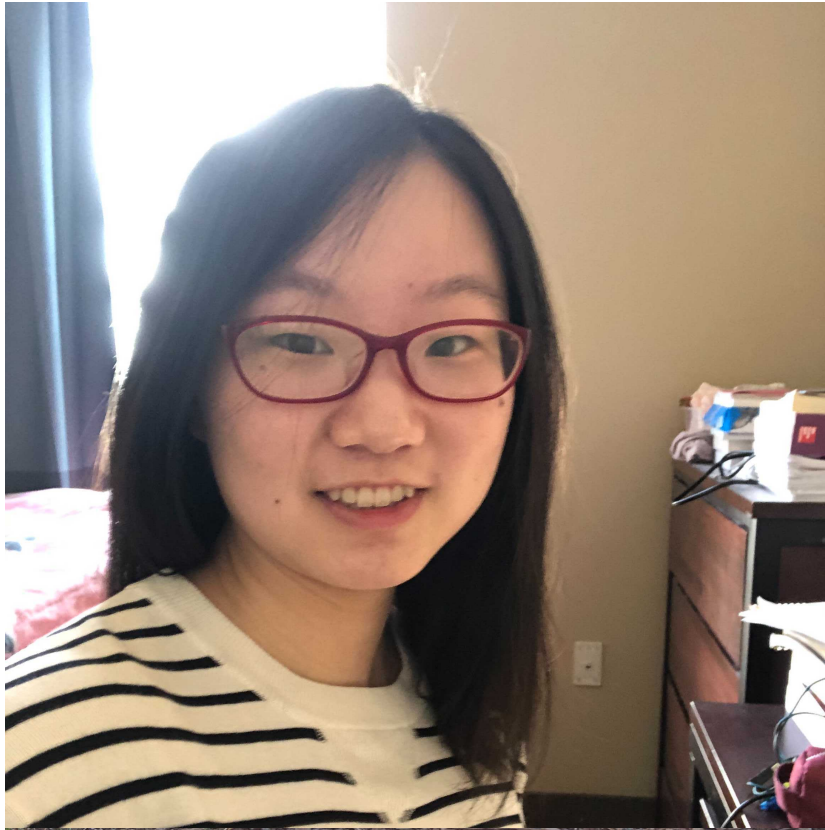
Jiani Fei, Chia-Nan Yeh, Dominika Zgid

SIMONS FOUNDATION

Phys. Rev. Lett. 126, 056402 (2021)

Phys. Rev. B 104, 165111 (2021)

Collaborators



Jiani Fei
UMich undergraduate
-> Stanford PhD



Chia-Nan Yeh
Michigan PhD->
Flatiron

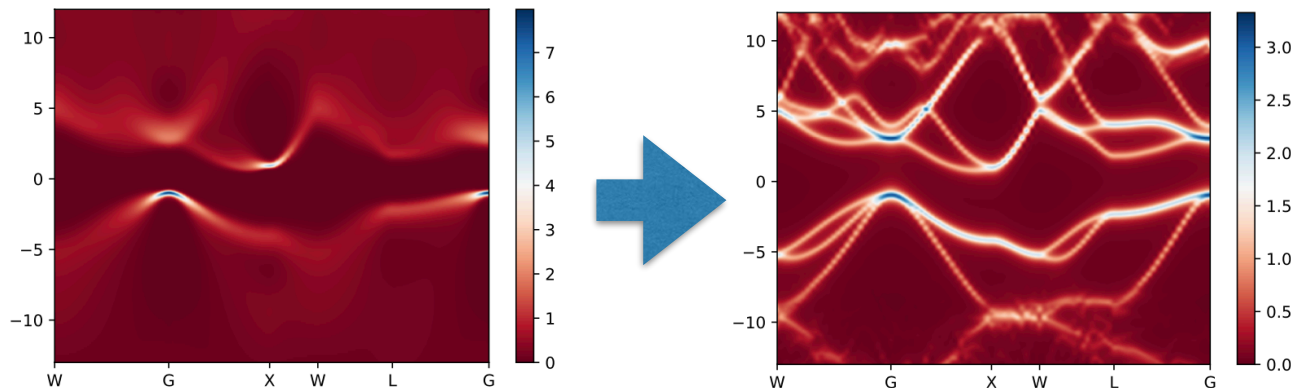
This talk

- Analytic Continuation: methods, common challenges.
- Nevanlinna Theory and connection to Analytic Continuation
- Pick matrices
- Hamburger Moments
- Carathéodory Generalizations

- Results for model systems
- Results for GW calculations
- Outlook and future prospects

Not in this talk

- Physics
- Real-world applications



Analytic Continuation

Finite-temperature simulations

$$Z = \text{Tr} e^{-\beta H}$$

$$G(\tau) = -\frac{1}{Z} \text{Tr} [e^{-(\beta-\tau)H} c e^{-\tau H} c^\dagger]$$

$$G(i\omega_n) = \int_0^\beta d\tau e^{-i\omega_n \tau} G(\tau)$$

$$A(\omega) = \frac{-1}{\pi} \text{Im} G(\omega)$$

Spectral function

Analytic continuation

$$G(i\omega_n) = -\frac{1}{\pi} \int \frac{\text{Im} G(\omega) d\omega}{i\omega_n - \omega}$$

$$G(\tau) = -\frac{1}{\pi} \int \frac{\text{Im} G(\omega) e^{-\tau\omega} d\omega}{1 + e^{-\beta\omega}}$$

Operational problem: inversion of the kernel K ,

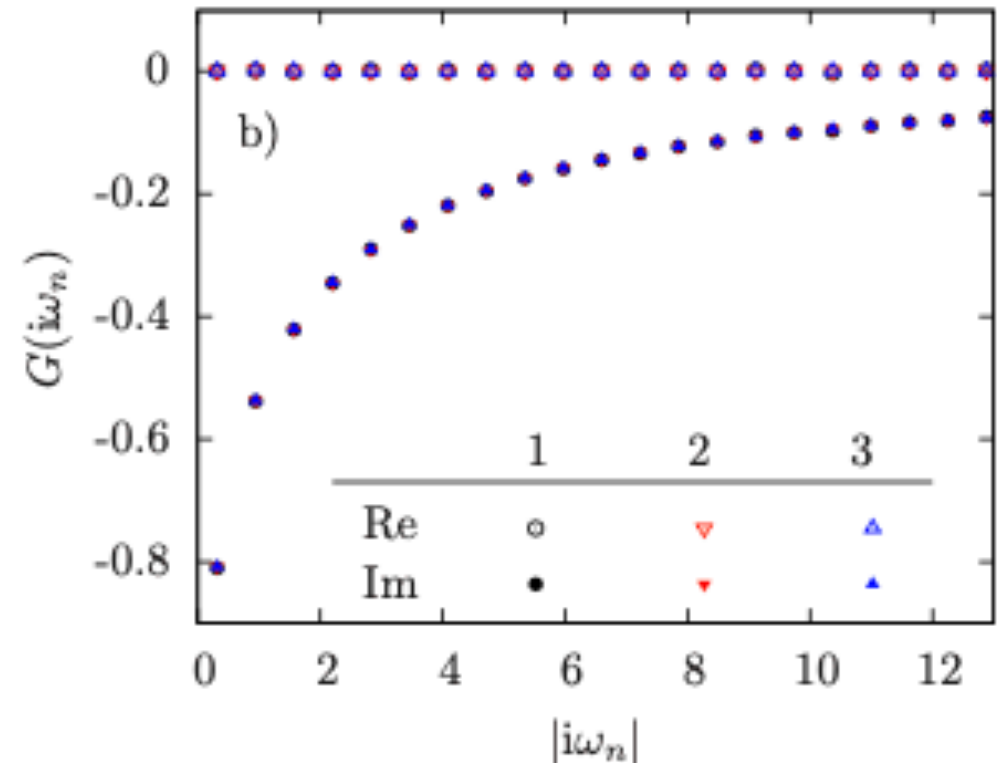
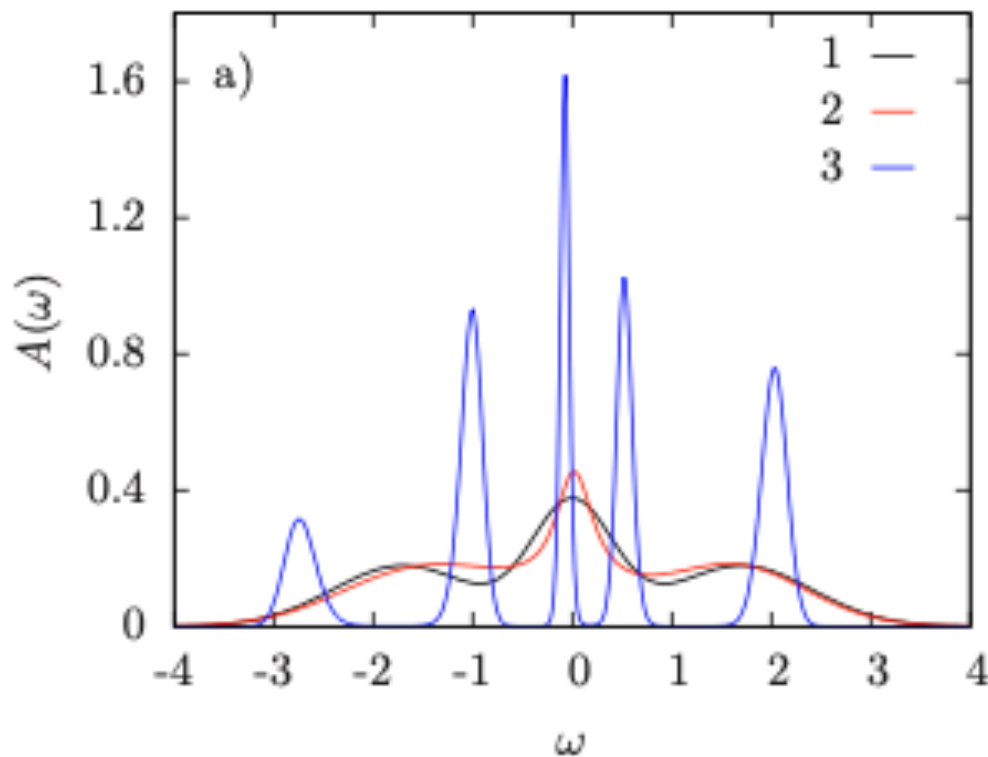
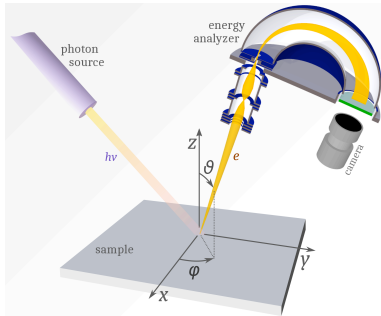
$$G(i\omega_n) = K(i\omega_n, \omega) G(\omega)$$

↓

$$G(\omega) = [K(i\omega_n, \omega)]^{-1} G(i\omega_n)$$

...is ill conditioned. Small changes on the imaginary axis cause large changes on the real axis.

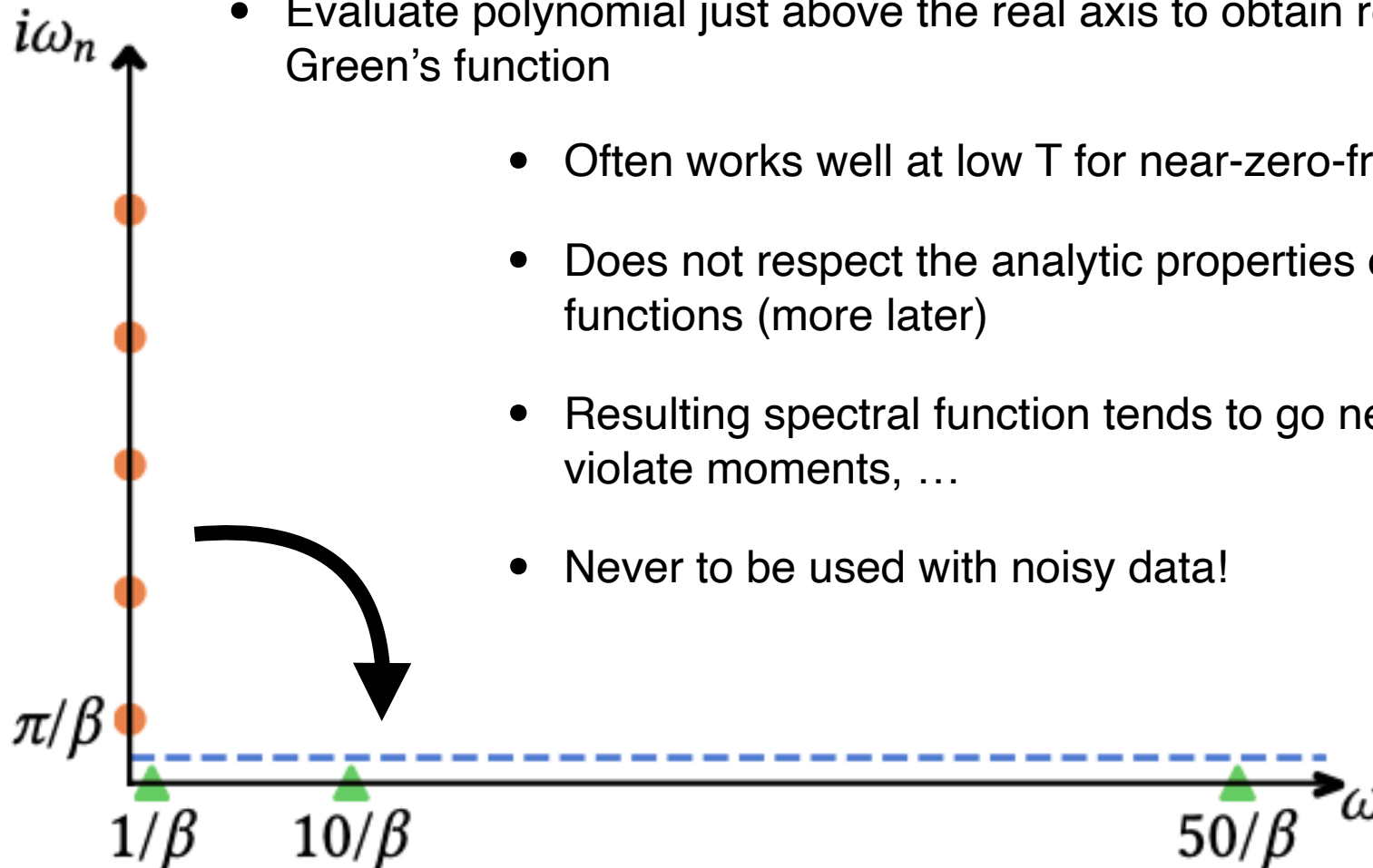
Analytic Continuation



S. Fuchs, PhD thesis,
Göttingen, 2010

Algorithm I – Rational Function Interpolation

- Given $2n$ points on the Matsubara axis, find a rational polynomial interpolant (“Padé”) with degree $n+1$ in denominator, degree n in numerator, and $\sim 1/i\omega_n$ high frequency decay
- Evaluate polynomial just above the real axis to obtain retarded Green’s function



- Often works well at low T for near-zero-frequency properties
- Does not respect the analytic properties of the Green’s functions (more later)
- Resulting spectral function tends to go negative, oscillate, violate moments, ...
- Never to be used with noisy data!

Algorithm II – Maximum Entropy Method

- Define a functional Q that, for a given spectral function A , balances the effect of deviations from a default model with the desire to fit imaginary time data as accurately as possible.

$$Q = \alpha S - \frac{1}{2} \chi^2$$

Deviation from a default model

$$S[A] = - \int d\omega \left[A(\omega) \log \frac{A(\omega)}{m(\omega)} \right]$$

Deviation from the input data

$$\chi^2 = \sum_{ij}^L (G(i\omega_n) - \bar{G}(i\omega_n))_i (C^{-1})_{ij} (G(i\omega_n) - \bar{G}(i\omega_n))_j$$

Minimize Q . First term will keep spectral functions smooth, second term will fit imaginary time data as well as possible.

A practitioner's guide to continuation...

- All of this is a sign of the ill-conditioned continuation matrix and an intrinsic problem of the finite-temperature Green's function formalism
- The higher the frequency, the less trustworthy your data
- Trust integrals over wide areas
- Trust the first gap / peak but nothing behind it. Trust evolutions of features (emergence of gap/peak etc with control parameters)
- You can always trade a bit of a shoulder for a bit of peak height
- Careful with bosonic functions: continuations are even less reliable due to multiplication with bosonic frequency.
- If you find that your results depend on the type of continuation method used, the type of default model used, the choice of alpha, the precise noise statistics, etc, they are probably not reliable
- Best practice: only interpret when you can see a clear signature on the imaginary axis



Our motivation for revisiting AC

- Lots of experience with AC in Quantum Monte Carlo

- New data:

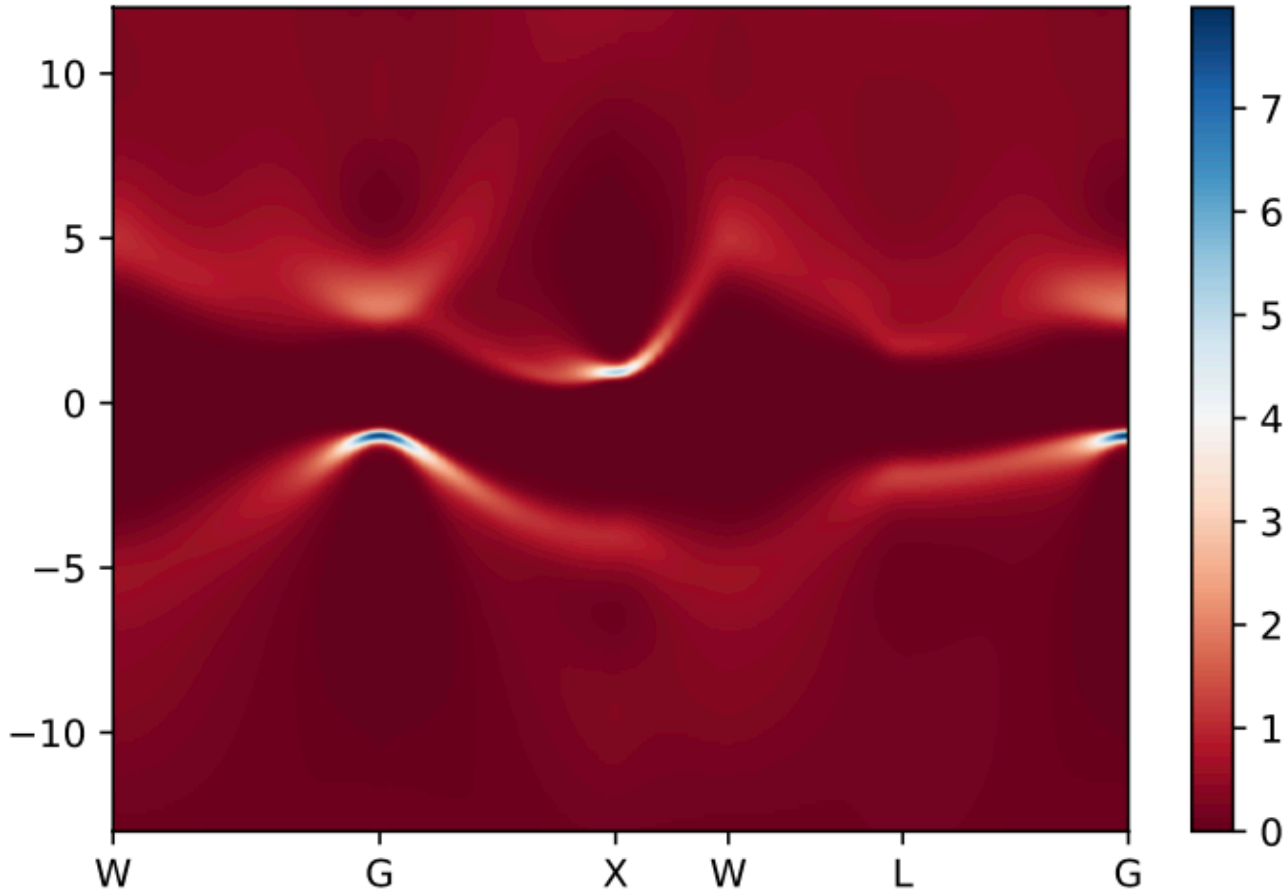
- Is very
- Spans
- MaxEn
otherw

- What do o

- Take b.
energy

- Downf

- Use 'q



osed,
=S washed out

art of self-

Maxent, orbital- and k-resolved

fects uncertain

- Missing capabilities: Off-diagonal entries of self-energies & GF, real-frequency Dyson equation, optics

Green's functions & Lehmann Representation

Lehmann representation

$$G_\gamma(z) = \frac{1}{Z} \sum_{m,n} \frac{|\langle m|c_\gamma^\dagger|n\rangle|^2}{z + E_n - E_m} (e^{-\beta E_n} + e^{-\beta E_m})$$

G coincides with Matsubara Green's function on imaginary axis, with retarded Green's function just above real axis. Define

$$A = \frac{1}{Z} |\langle m|c_\gamma^\dagger|n\rangle|^2 (e^{-\beta E_n} + e^{-\beta E_m}) > 0$$

For

$$z = x + iy$$

$$S = \frac{A}{(x + E_n - E_m) + iy} = \frac{A(x + E_n - E_m - iy)}{(x + E_n - E_m)^2 + y^2}$$

$$\text{Im}S = \frac{-Ay}{(x + E_n - E_m)^2 + y^2}$$

And therefore for any Green's function, independent of the system:

$$\text{Im} G_\gamma(z) \leq 0 \quad \text{for } z \in \mathbb{C}^+$$

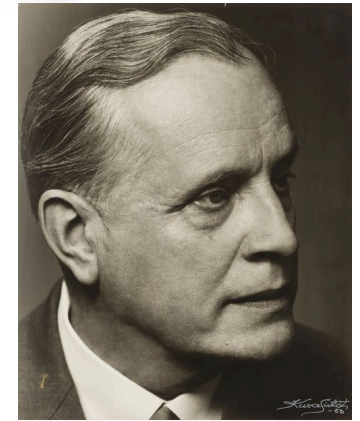
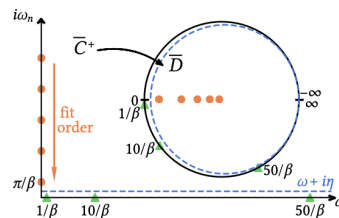
Nevanlinna and Schur functions

Nevanlinna functions are functions with a positive imaginary part on the upper half of the complex plane.

$$\mathcal{N}G = -G \quad \text{Is a Nevanlinna function}$$

The invertible Möbius transform h maps the upper half plane to the unit disk

$$h(z) : z \rightarrow \frac{z - i}{z + i}$$



Rolf Nevanlinna

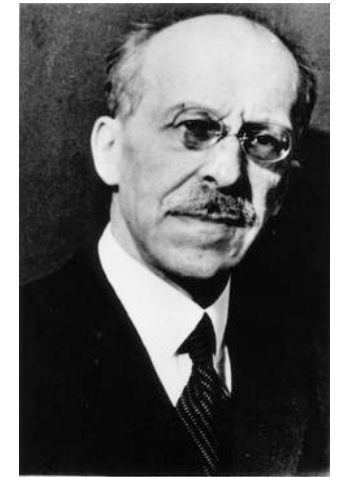
$$\mathcal{D} = \{z : |z| < 1\};$$

$$\overline{\mathcal{D}} = \{z : |z| \leq 1\}$$

Nevanlinna functions can be mapped onto Schur functions: Schur functions map the open unit disk \mathcal{D} to the closed unit disk $\overline{\mathcal{D}}$ ('contractive' functions). Every Schur function has a continued fraction expansion. Given a set of interpolation points, a Schur function can be constructed using a recursive algorithm.

Combine mapping to contractive functions with Schur's continued fraction expansion to obtain an intrinsically causal expansion for Green's functions.

The Schur algorithm



I. Schur

Issai Schur

Input data $f(Y_i) = C_i \quad i = 1, 2, \dots, M \quad Y_i = i\omega_n \in \mathcal{C}^+$ and $C_i \in \mathcal{C}^+$

Contractive interpolant $\theta(Y_i) = \lambda_i = h(C_i) = \frac{C_i - i}{C_i + i} \quad i = 1, 2, \dots, M$

Start the interpolation by constructing an interpolant through Y_1 . Express this contractive interpolant as a function that is zero at Y_1 , and a constant λ_1 :

We want $\theta(Y_1) = \lambda_1 \quad |\lambda_1| < 1$

Functional form $\theta(z) = \frac{\phi(z) + \lambda_1}{\lambda_1^* \phi(z) + 1}$ Where $\phi(z) = \frac{z - Y_1}{z - Y_1^*} \theta_1(z)$

J. Schur, Über potenzreihen, die im innern des einheitskreises beschränkt sind, *Journal für die reine und angewandte Mathematik* **1918**, 122 (1918).

Such that $\phi \in \mathcal{B}$ and $\phi(Y_1) = 0$

Note that $\theta_1(z)$ is now an arbitrary contractive function. Express it as a sum of a function that is λ_2 at Y_2 and an arbitrary contractive function. Express that one as the sum of a function that is λ_3 at Y_3 and an arbitrary contractive function, iterate and repeat for all interpolation points.

This will result in an expression for **all possible interpolants** in terms of a remaining arbitrary Schur/Nevanlinna function. We will use this freedom later.

The Pick criterion: existence of interpolants

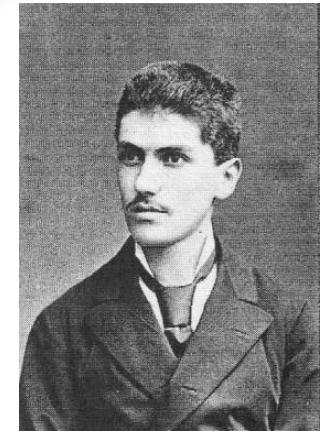
If $g(x_i) = y_i \quad (x_i \in \mathcal{D}, y_i \in \overline{\mathcal{D}})$

Then a Schur interpolant to g can be found iff the Pick matrix is positive semi-definite. It has a unique solution if furthermore the Pick matrix is singular.

$$P_{ij} = \begin{bmatrix} 1 - y_i y_j^* \\ 1 - x_i x_j^* \end{bmatrix}$$

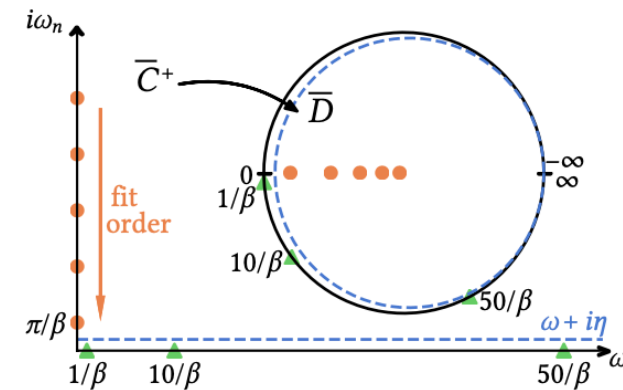
Provides a straightforward check on any input data. Transform the data to the unit circle, evaluate Pick matrix, check if it has negative eigenvalues. If it does, there WILL NOT be a positive spectral function.

Interesting observation: Monte Carlo data never fulfills this criterion. GW data only if very well converged and not too many interpolation points. Synthetic benchmark data shows very high precision at high frequency needed to make it work. Sign of the very constrained nature of Nevanlinna/Schur function space.



Georg A. Pick

G. Pick, Über die beschränkungen analytischer funktionen, welche durch vorgegebene funktionswerte bewirkt werden, *Math. Ann.* **78**, 270 (1917).



The Hamburger moment problem

The (truncated) Hamburger moment problem aims to construct a measure $\sigma(\omega)$ on the real axis such that

$$h_k = \int \omega^k d\sigma(\omega) \quad b = (h_0, h_1, h_2, \dots)$$



Hans Ludwig
Hamburger

In our context think of $d\sigma(\omega) = A(\omega)d\omega$

H. Hamburger, *Math. Ann.* **81**, 235 (1920).

The Hamburger-Nevanlinna theorem establishes a one-to-one correspondence between solutions to the moment problem and a subset of Nevanlinna functions.

$$\begin{aligned} \mathcal{N}G(i\omega_n) &= -G(i\omega_n) \\ &= - \int_0^\beta d\tau G(\tau) e^{i\omega_n \tau} \\ &= - \sum_{k=0}^{\infty} \frac{(-1)^{k+1} (G^{(k)}(\beta) + G^{(k)}(0))}{(i\omega_n)^{k+1}} \\ &= - \frac{h_0}{i\omega_n} - \frac{h_1}{(i\omega_n)^2} - \frac{h_2}{(i\omega_n)^3} - \dots \end{aligned}$$

It is possible to combine the moment with the interpolation problem to both enforce moments and interpolation values. This further constrains the solution.

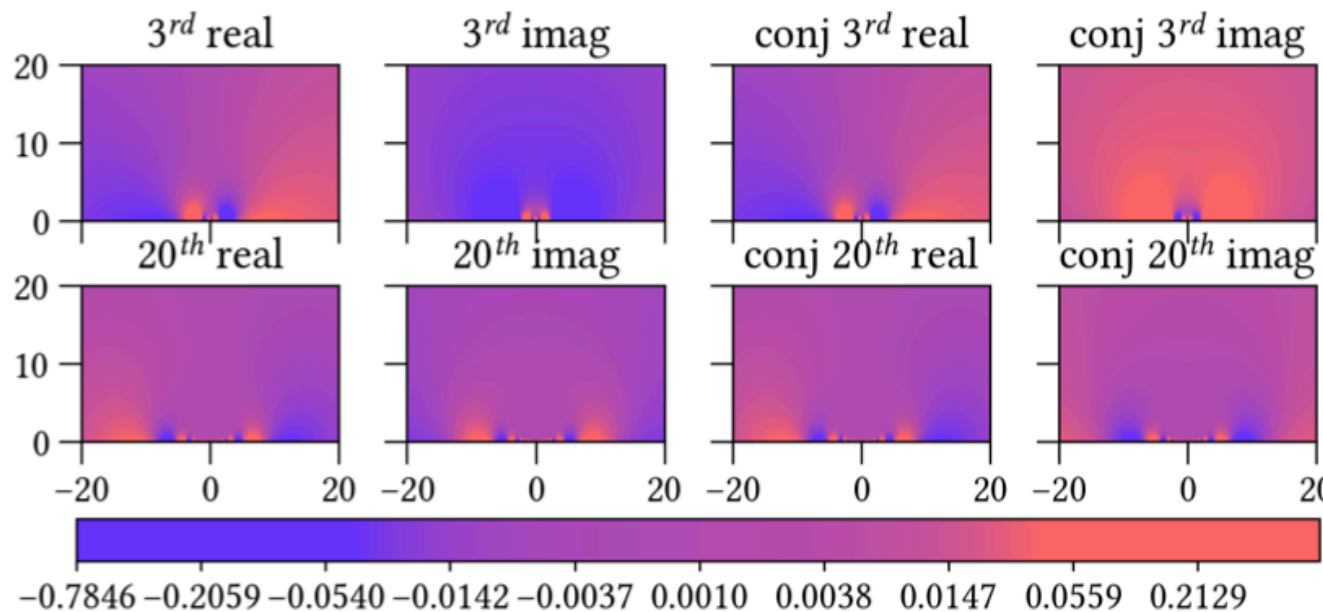
The Hamburger moments are routinely computed in a high-frequency tail analysis of the Green's functions, where they typically supplement Green's functions outside of the interval where data is available.

Hardy function optimization

A positive definite Pick matrix guarantees an infinite number of solutions. In the Schur algorithm the freedom appears as an arbitrary Nevanlinna function that can be added to the solution. For any such function, the resulting interpolant will still be a Nevanlinna function and interpolate all possible solutions.

The freedom of choosing an additional function can be used to impose additional properties of the interpolant, *e.g.* impose smoothness on the real axis.

We chose Hardy functions – other choices are possible $B^k(z) = \left\{ \frac{1}{\sqrt{\pi}(z+i)} \left(\frac{z-i}{z+i} \right)^k \right\}_{k \in \mathbb{N}}$



Real and imaginary parts of the 3rd and 20th Hardy function mapped to the upper complex plane

Minimize F , optimize smoothness while respecting norm

$$F[A_{\theta_M}(\omega)] = \left| 1 - \int A_{\theta_M}(\omega) d\omega \right|^2 + \lambda \left\| \frac{d^2 A_{\theta_M}(\omega)}{d\omega^2} \right\|^2$$

Results for synthetic systems

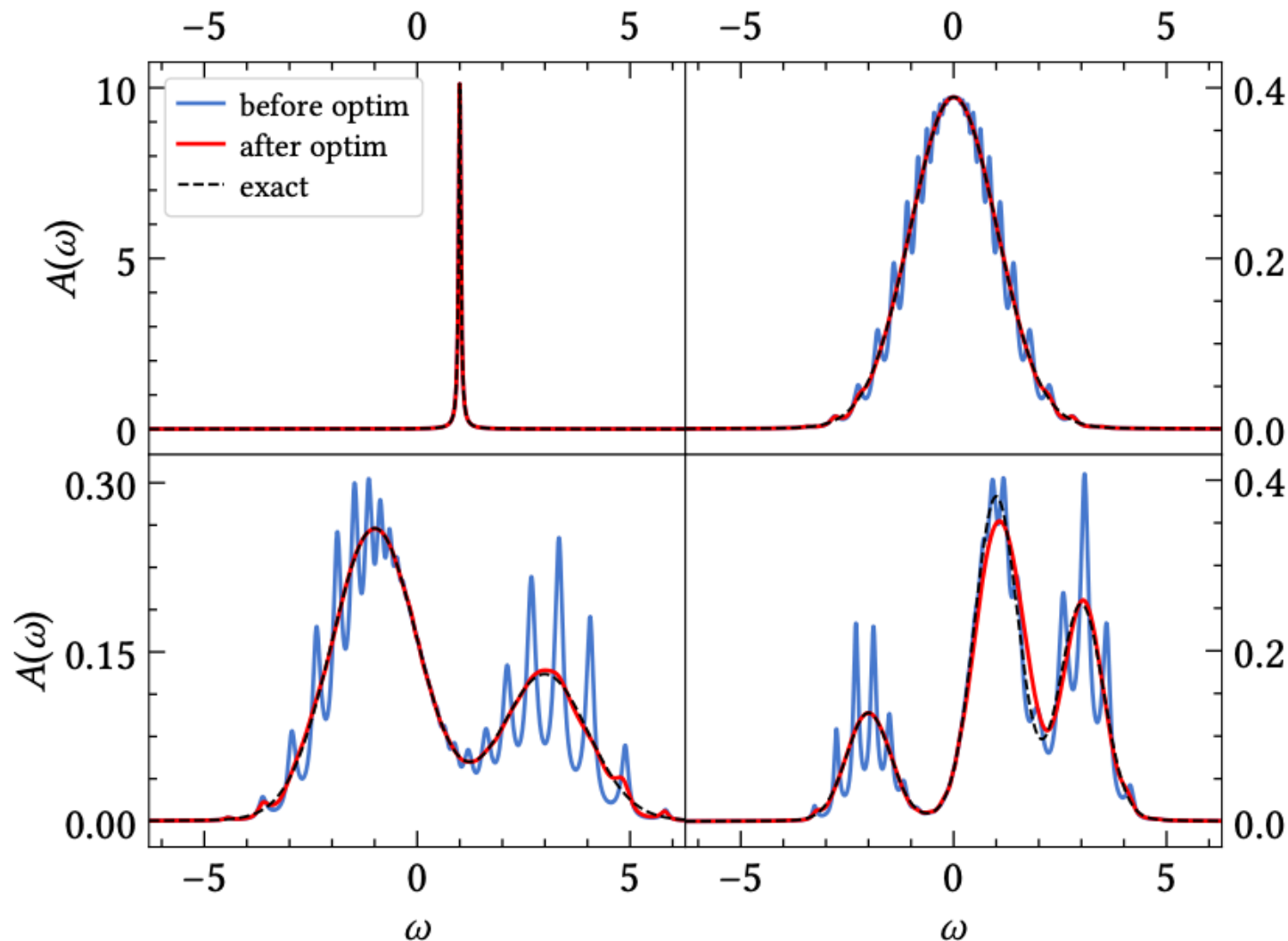


Figure 1.4: Continuation with and without Hardy function optimization. Off-centered δ peak (top left), Gaussian (top right), two-peak scenario (bottom left), and a three-peak scenario (bottom right). $\beta = 100$, IR grid [13, 14] with 36 Matsubara positive frequency points.

Hardy function contribution

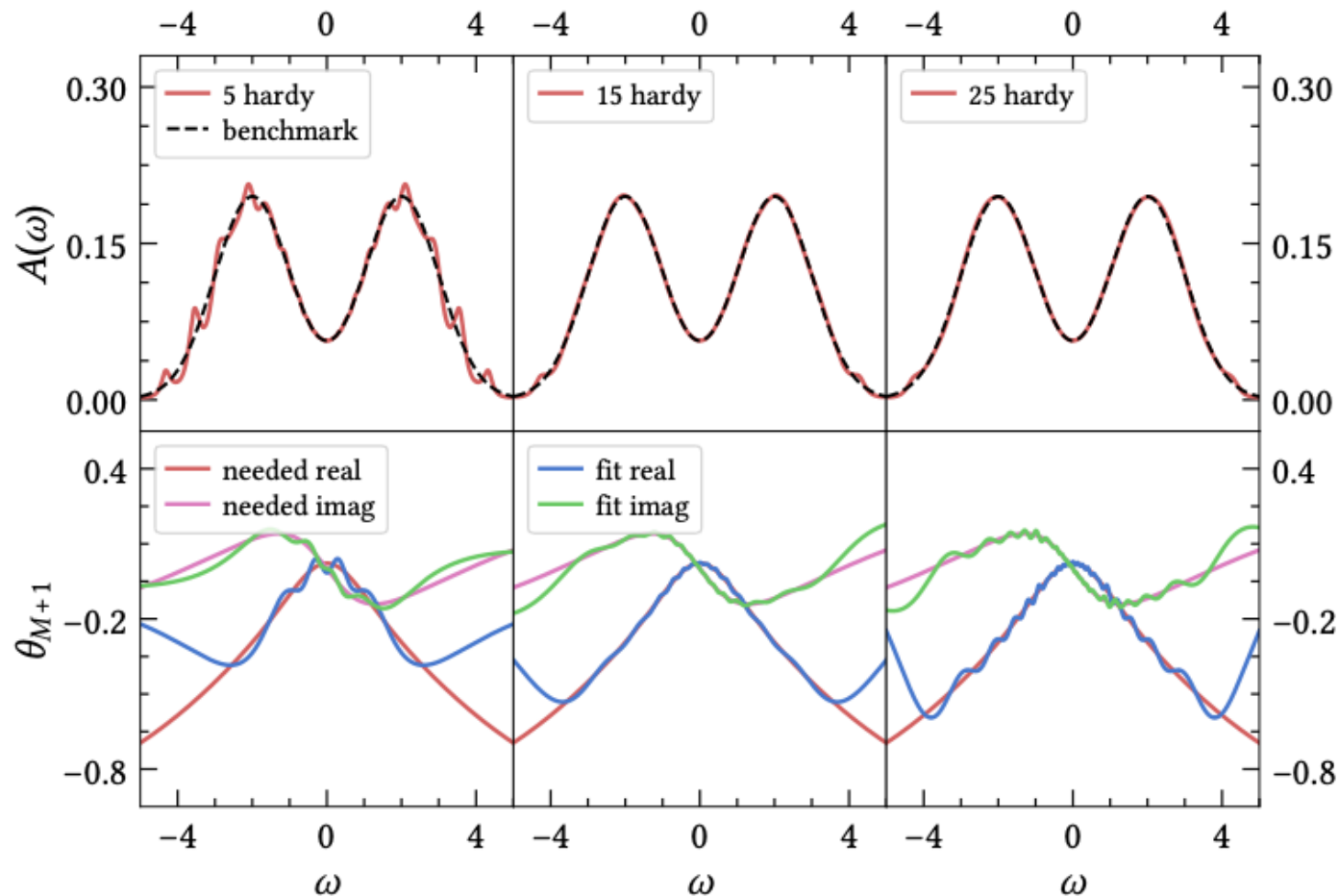


FIG. 1. Optimization with 5, 15 and 25 Hardy basis functions. Top panels: resulting spectral functions $A(\omega)$. Bottom panels: real and imaginary part of the exact and fitted parametric functions θ_{M+1} ($\theta_{M+1} : \mathcal{C}^+ \rightarrow \overline{\mathcal{D}}$). The needed θ_{M+1} is what would restore our synthetic input $A(\omega)$ when plugged into the final interpolant (Eq. 7 in the manuscript). The fitted θ_{M+1} is expanded by the hardy functions with the optimized coefficients a_{ks} and b_{ks} . Note that 15 hardy functions would give a better result than 25, as the function optimization of the smoothing norm works better.

Effects of temperature, grid points, noise

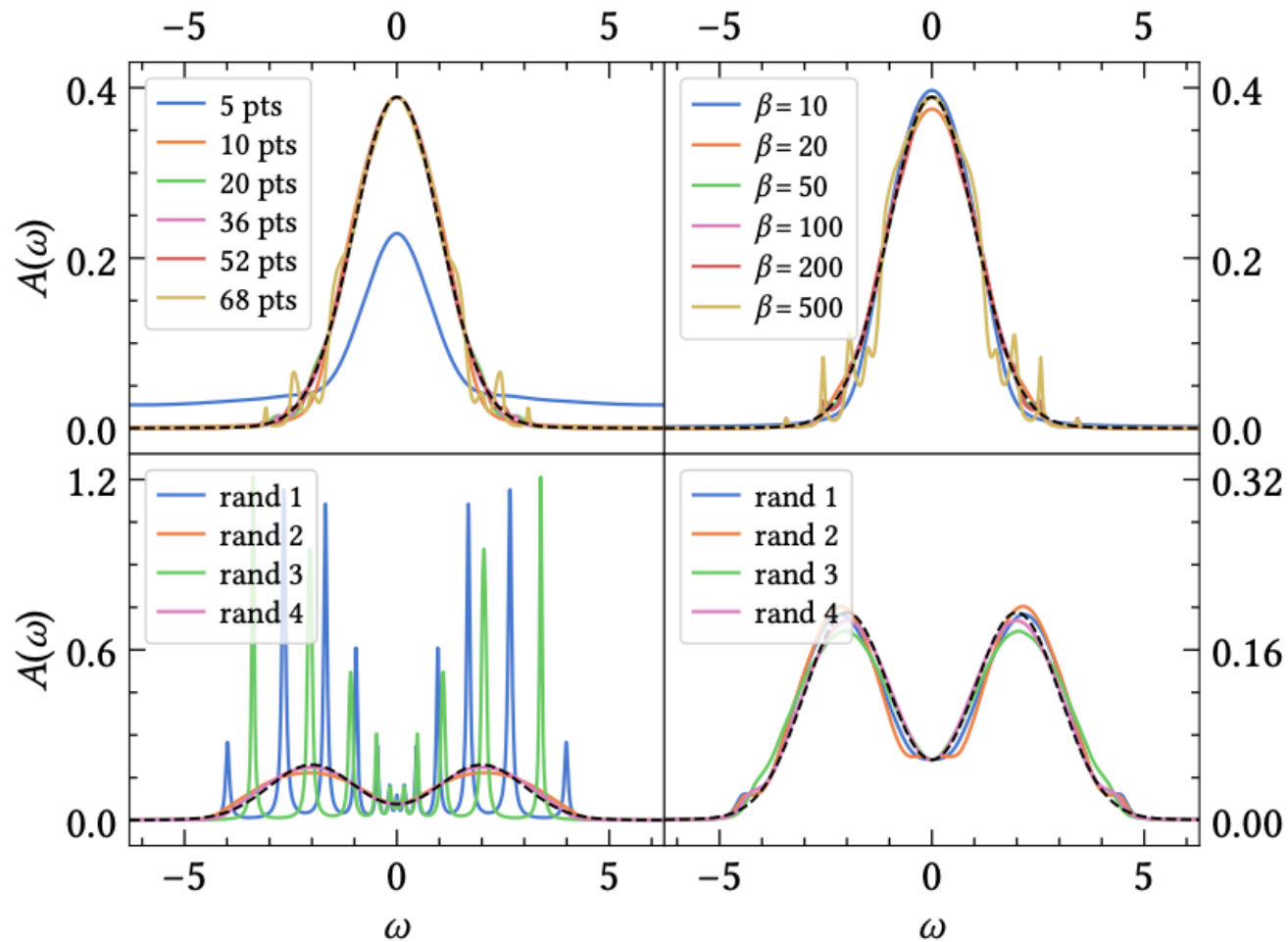
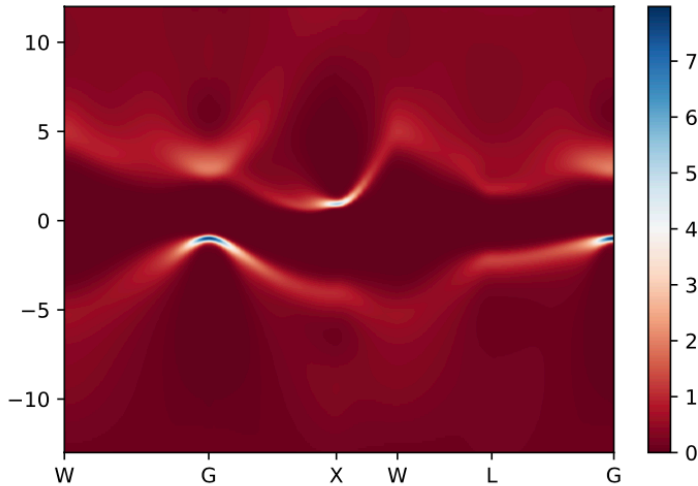


Figure 1.5: Top left: effect of the number of interpolation nodes (IR grid [13, 14]) used. Top right: effect of Matsubara spacing (inverse temperature), uniform grid $\omega_n = (2n + 1)\pi/\beta$ used. Bottom two panels: effect of independent Gaussian noise with relative standard deviation of 10^{-4} (left) and 10^{-6} (right), four sample curves each.

Comparison to Maximum Entropy

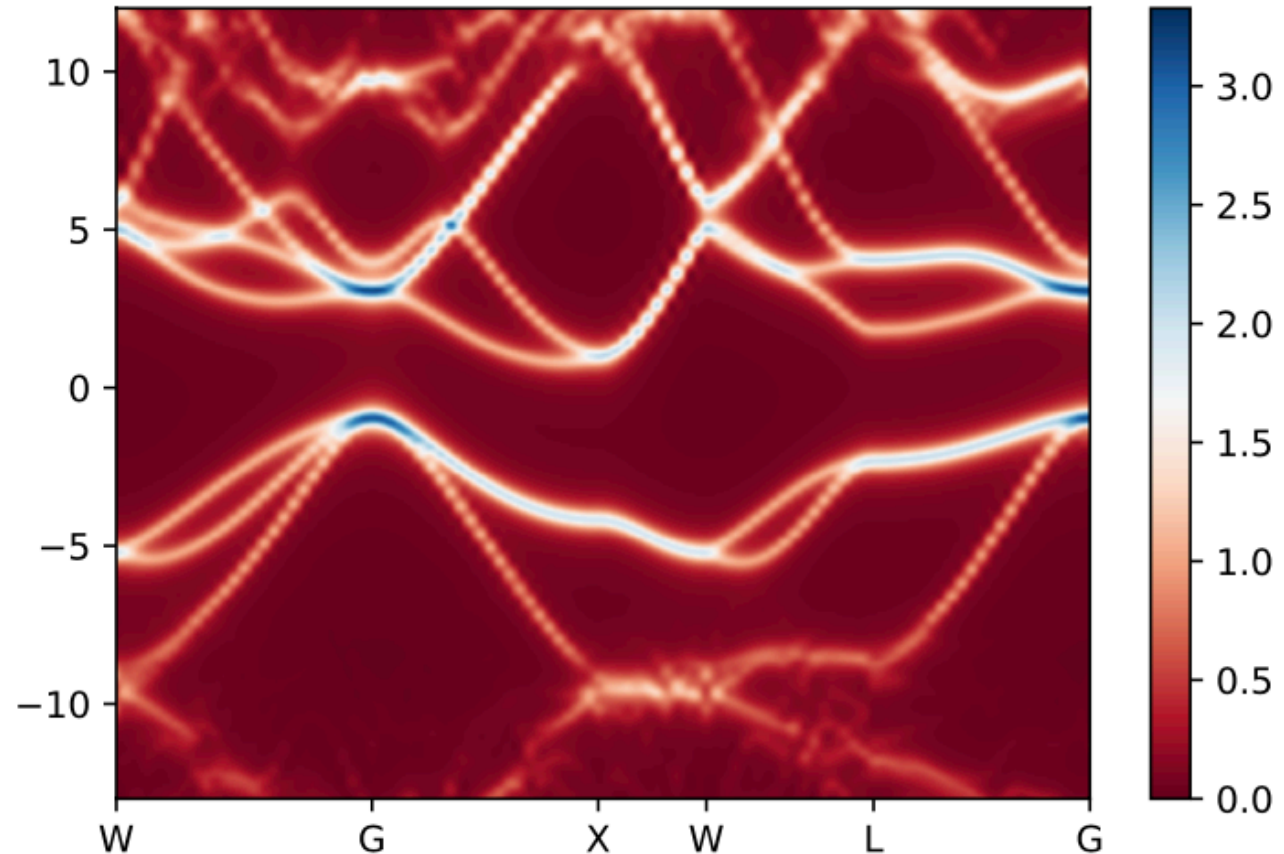


Maxent, orbital- and k-resolved

Fully self-consistent GW of Si, no quasiparticle or similar approximations, analytic continuation of fully interacting Green's function.

Band structure is visible, individual bands can be separated

Both continuations operating on same input data!



Nevanlinna, orbital- and k-resolved

Comparison to Maximum Entropy

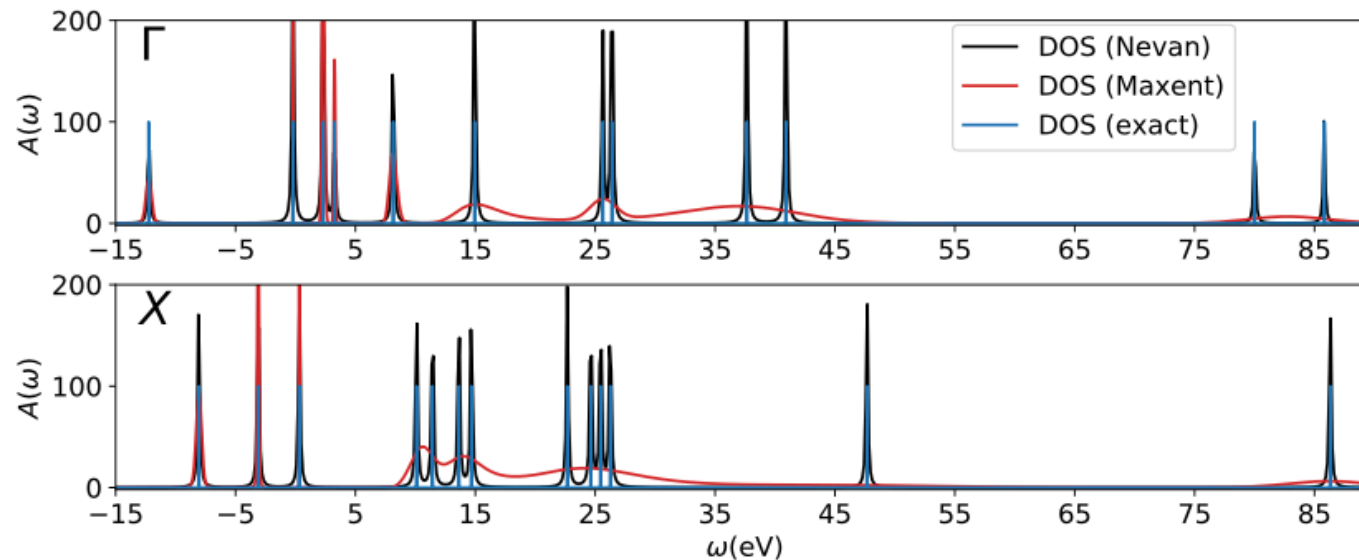


Figure 1.6: LDA band structure (Kohn Sham eigenvalues, DOS) of solid Si (green) at the Γ and the X point, as well as Nevanlinna (blue) and MaxEnt (orange) continuations of the corresponding Green's functions. $T=316$ K, 52 non-uniform [14] IR Basis [13] Matsubara positive frequency points.

Comparison to Real Time propagation

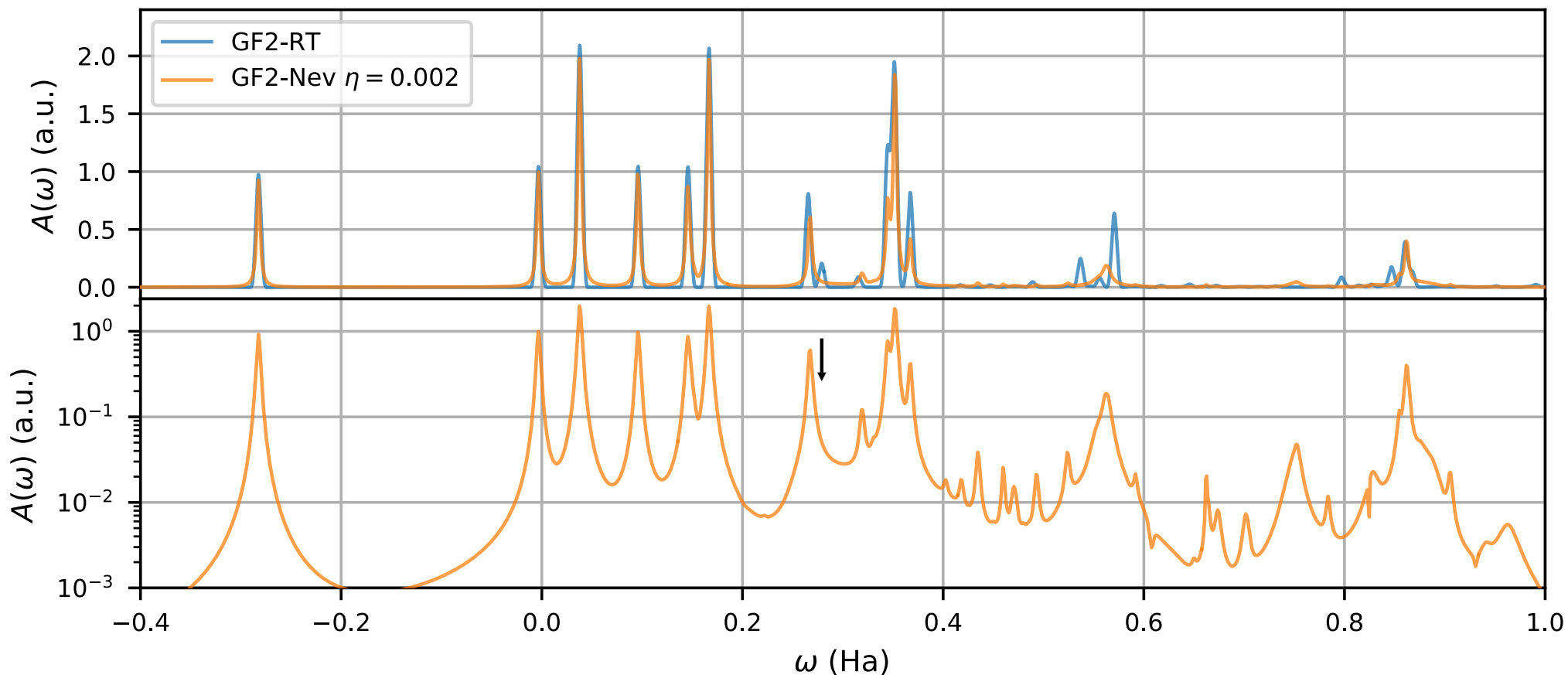


FIG. 10. GF2 spectral function (blue) from real-time propagation compared to the spectral function obtained by analytical continuation of the GF2 solution in imaginary time using the Nevanlinna method (orange) [26], for LiH in the cc-pVDZ basis at $\beta = 200$ and equilibrium inter-atomic distance $r_0 = 1.62 \text{ \AA}$, on a linear scale (upper panel) and logarithmic y-axis (lower panel). The spectral functions are scaled so that a non-degenerate single-particle state has a peak height of unity.

Matrix-valued Carathéodory generalization

The Carathéodory class of matrix-valued analytic functions in the unit disk (or: upper half plane) is defined as

$$C = \{M(z) : M(z) + M^\dagger(z) \geq 0 \quad \forall |z| < 1\}$$

Note that $M(z) + M^\dagger(z) \geq 0 \Leftrightarrow \text{Re}\{x^\dagger M(z)x\} \geq 0$

i.e. the real part of M is positive semidefinite.

$-iG^<(\omega)$ is Carathéodory:

C. Carathéodory, Über den Variabilitätsbereich der Koeffizienten von Potenzreihen, die gegebene Werte nicht annehmen, *Math. Ann.* **64**, 95 (1907).



Κωνσταντίνος
Καραθεοδωρή

$$G_{ij}^<(\omega) = 2\pi i \sum_{mn} \frac{e^{-\beta E_n}}{Z} \langle n | c_j^\dagger | m \rangle \langle m | c_i | n \rangle \delta(\omega - E_n + E_m)$$

Insert x, do the Math:

$$\langle x | -iG^<(\omega) | x \rangle = 2\pi \sum_{mn} \frac{e^{-\beta E_n}}{Z} |\langle m | \sum_i c_i x_i^* | n \rangle|^2 \delta(\omega - E_n + E_m)$$

Carathéodory in many-body

$-iG^<(\omega)$ is Carathéodory / has positive semidefinite real part on upper half plane

$iG^>(\omega)$ is Carathéodory / has positive semidefinite real part on upper half plane

$iG(i\omega_n)$ is Carathéodory / has positive semidefinite real part on upper half plane

Without proof here: if we split the Self-energy into a Hartree term and a dynamical term, and the non-interacting Hamiltonian is quadratic, then

$i\Sigma(i\omega_n)$ is Carathéodory / has positive semidefinite real part on upper half plane

Define the ‘cumulant’ as

$$M^{-1}(z) = G^{-1}(z) + F$$

(The ‘Green’s function without the Fock matrix’)

$iM(z)$ is Carathéodory / has positive semidefinite real part on upper half plane

Carathéodory mapping

The interpolation of Carathéodory functions is a well-known problem. Proceed by Möbius transform to unit disc, Schur algorithm, evaluation on the boundary, and Möbius transform back. Guaranteed to give intrinsically PSD interpolants with full off-diagonal structure!

Conformal mapping of Matsubara points:

$$h : \mathcal{C}^+ \rightarrow \mathcal{D}, z \rightarrow \frac{z - i}{z + i}$$

Conformal mapping of function values
(all quantities are matrices):

$$\Psi(z) = [I - F(z)][I + F(z)]^{-1}$$

...where F is the PSD function to be interpolated, i.e.

$$F(z) = i\Sigma(z)$$

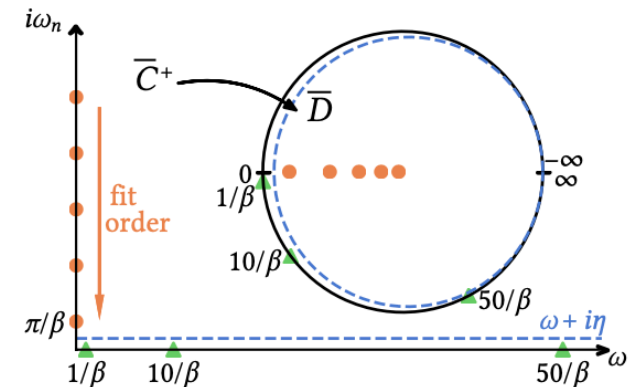
$$F(z) = iG(z)$$

Set $F(x_i) = Y_i$

$$\Psi(z_i) = [I - Y_i][I + Y_i]^{-1}$$

With Psi a Schur class function on the unit disc

Phys. Rev. B 104, 165111 (2021)



Generalized Pick criterion

Carathéodory interpolants for the original problem exist if and only if the Pick matrix is positive semi definite, a unique solution only exists if the Pick matrix is singular.

$$P_C = \left[\frac{Y_k + Y_l^*}{1 - z_k^* z_l} \right]_{(mn) \times (mn)}$$

Each matrix Y has size m , we have n interpolation points, *i.e.* this is a rather large matrix.

Observation with numerical data: Pick matrix is extremely sensitive to noise in the input data.

Note that the Pick matrix elements are linear in the Matsubara values. ‘Pickifying’ is a semidefinite program

The Hubbard Dimer

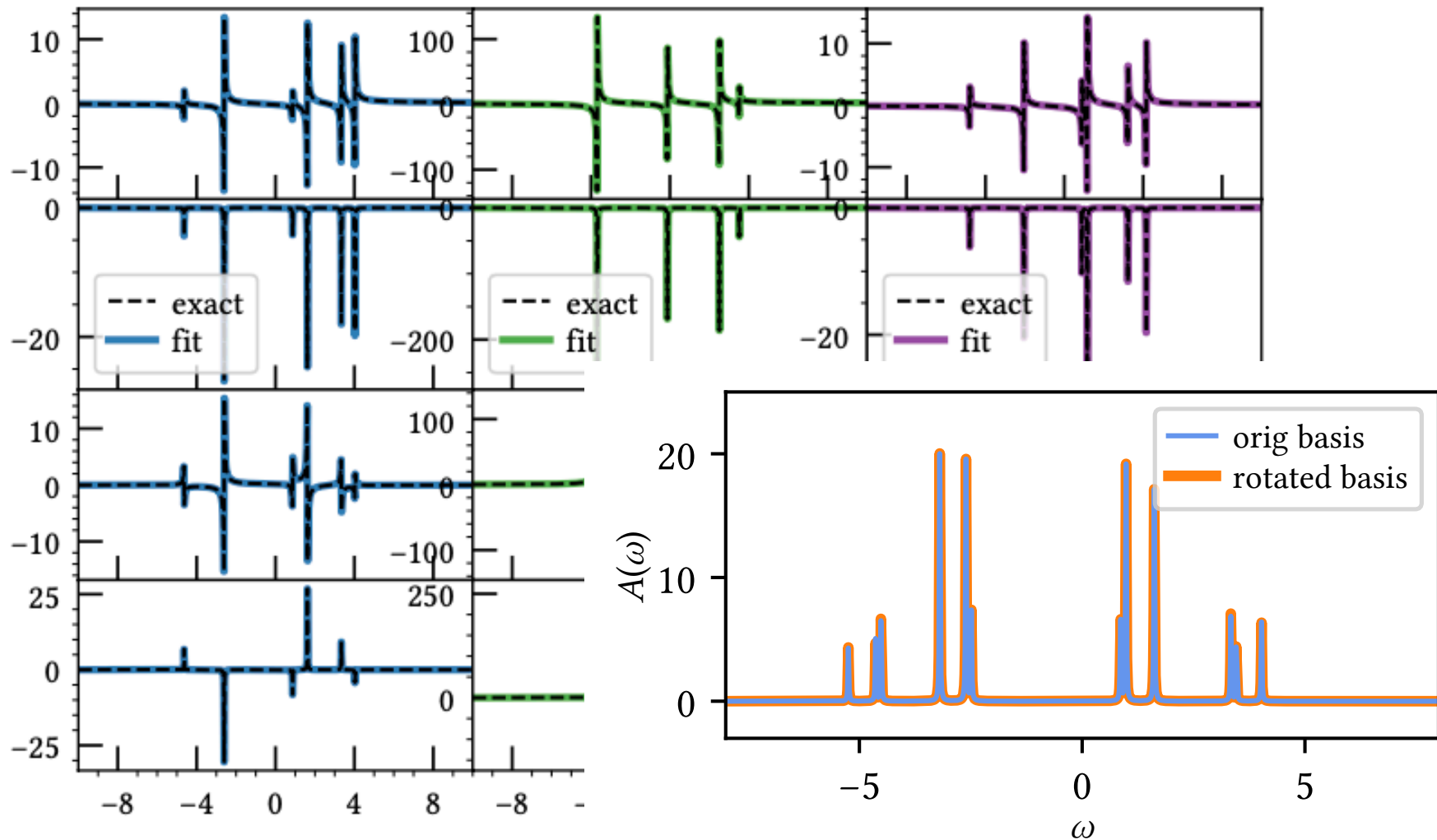
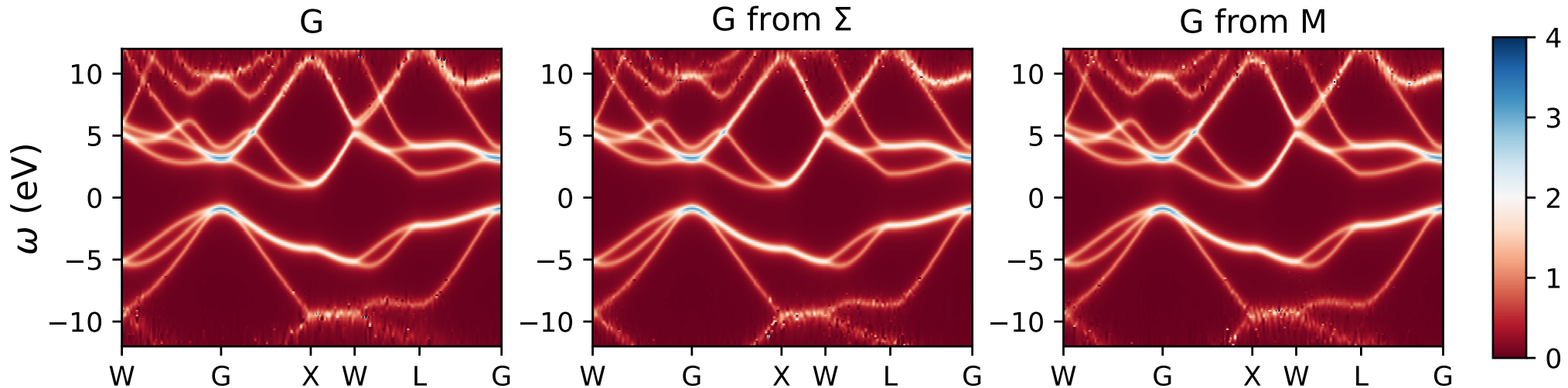
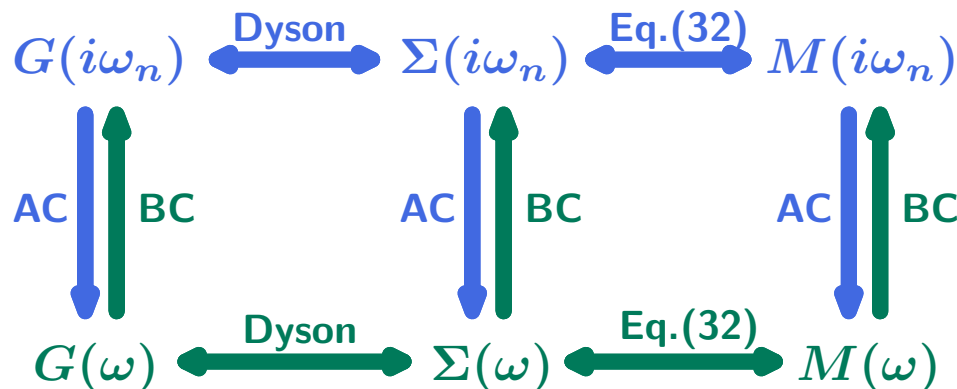


Figure 2.2: The Hubbard dimer model interpolation, 2 random non-zero entries [rows are real(1), imaginary(1), real(2), imaginary(2) part] of the 4×4 matrices, on the real axis. Using input data with 10^{-7} standard deviation Gaussian noise. Left are Green's functions, middle are self-energies, right are cumulants. The exact data comes from analytic model formula. Fitted data comes from continuing matrices from the imaginary axis, which we get from analytic formula and disturbed with noise.

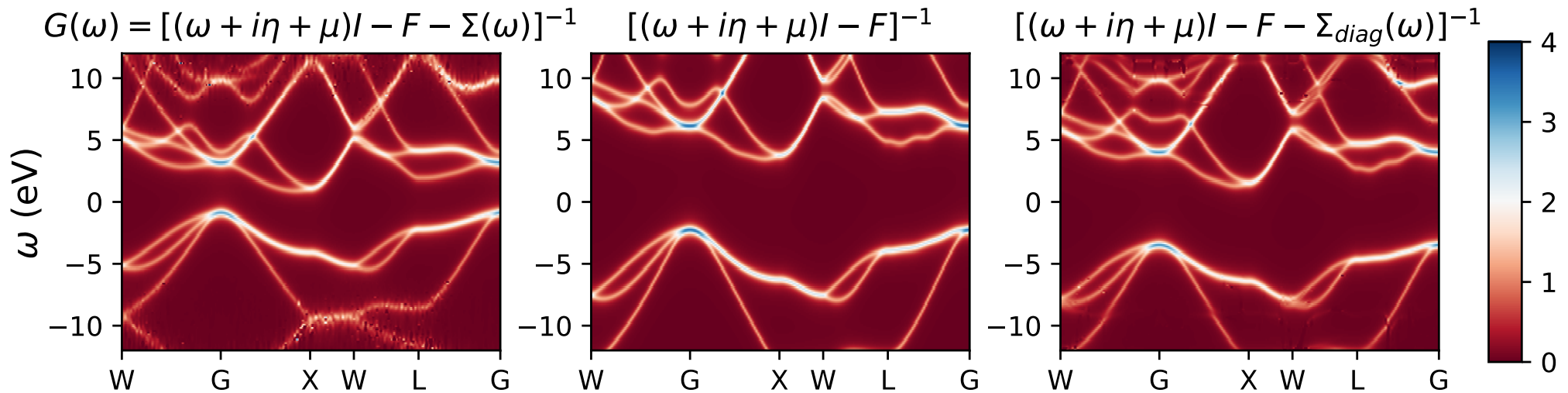
Dyson commutes with continuation



Band structure of crystalline Silicon (26 orbitals per unit cell), simulation in Gaussian orbitals with fully self-consistent GW; matrix valued continuation followed by Dyson equation.



Careful with approximations to Sigma!



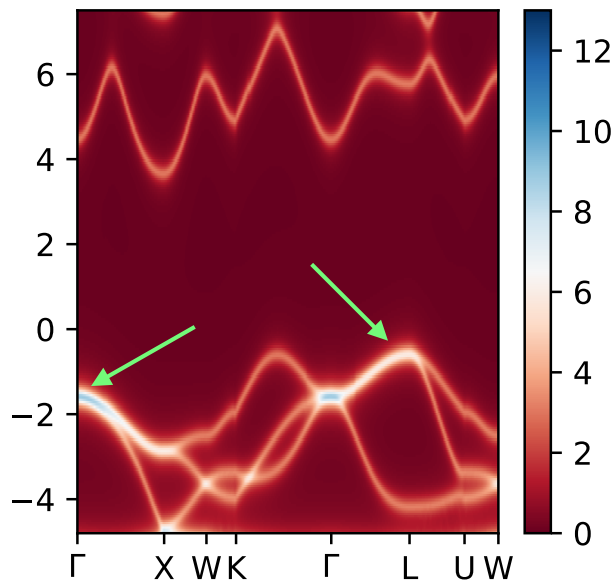
Truncation of the dynamical part of the self-energy (just the Fock matrix of the fully interacting system); or truncation of the self-energy to just diagonal parts.

‘Diagonal’ approximations to the self-energy have a large effect on the band structure. Truncation is often done in LDA+DMFT-type ‘real materials’ calculations.

Careful with (self-consistent/non-self-consistent) quasiparticle approximation (diagonal truncation & lowest frequency only)

Revealing ab-initio spin-orbit & correlations

- Spin-orbit coupling emerges ab-initio from the solution of the Dirac, rather than the Schrödinger equation. Electrons and Positrons!
- Recent development in quantum chemistry: exact two-component relativistic approach. In the x2c1e approximation, the two-body integrals remain non-relativistic
 - Neglects relativistic corrections to the two-body integrals
 - Diagrammatic structure remains unchanged, interaction vertices remain unchanged, bare propagators pick up relativistic contributions. (Identical impurity solvers!)

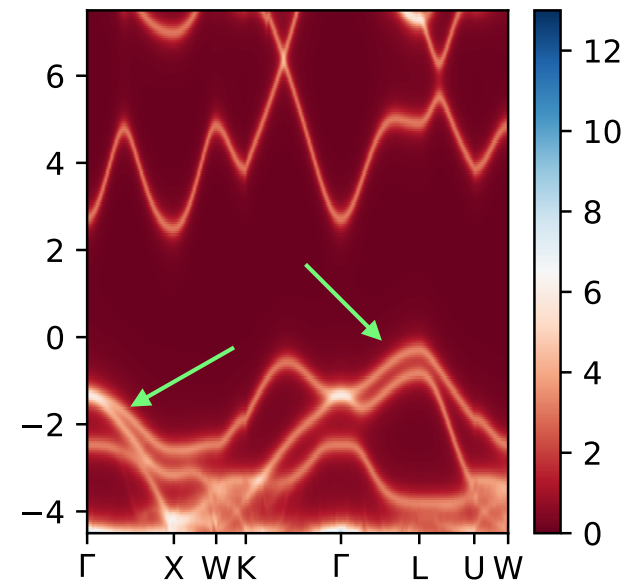


Non-relativistic

Example: self-consistent GW simulation of AgI in the x2c1e approximation

Parameter free ab-initio **spin-orbit coupling** splits bands

No way to see this w/o Nevanlinna techniques



Relativistic x2c1e

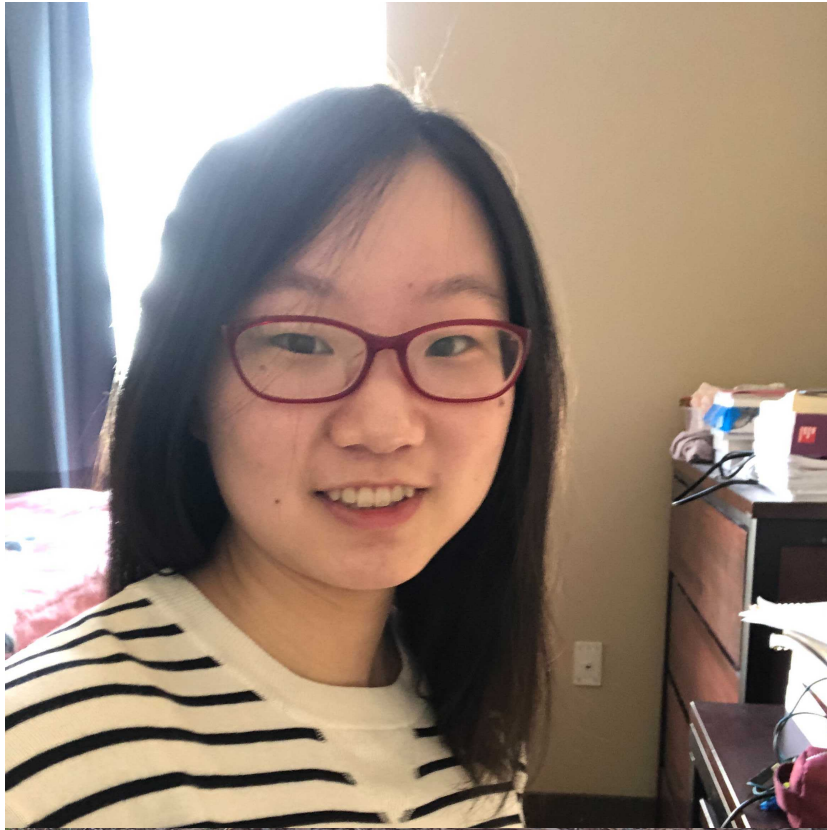
A few remarks

- This math is done by non-experts! (i.e. tell us we're wrong, we won't be offended... ..tell us how to do better! We know there's much more out there...)
- Most of the 'deep' insights come from 1910-1930; technology has not been used in this context at all, most of the theory seems to be forgotten.
- I showed some successes, but there are challenges/questions/problems:
 - Can we find a basis for Nevanlinna functions? (...so we can expand in it/project to it...?)
 - Plenty of stability issues (everything done in multi precision arithmetics). Is there a way to avoid this?
 - Data that is non-causal: would like to find the 'closest' Nevanlinna function to a given set of input data.

In conclusion

- Analytic Continuation remains difficult. No way to bring information back that isn't in the data. Careful with 'p-hacking' analytic continuation
- ...but by using the appropriate Math we can obtain much more accurate continuations
 - Build in causality
 - Build in moments
- ...New capabilities: continuation of off-diagonal terms that respect the analytic structure of the Green's function
- ...New capabilities: continuation of moments and self-energies
- Complex Analysis is very powerful!

Many Thanks to Jiani & Chia-Nan!



Jiani Fei



Chia-Nan Yeh

