M. Numerical Recipes

M.1. Iteration of DMFT Equations on Time $\times$ Time matrices

The simplest way to solve the DMFT equations is to iterate them from a reasonable initial condition (Mignacco et al., 2020; Bordelon & Pehlevan, 2022b). We solve in discrete time for $T \times T$ matrices $\{\boldsymbol{R}_{0,2}, \boldsymbol{R}_1, \boldsymbol{R}_{2,4}, \boldsymbol{R}_3, \boldsymbol{C}_0, \boldsymbol{C}_1, \boldsymbol{C}_2, \boldsymbol{C}_3\}$ which have entries $[\boldsymbol{R}]_{t,s} = R(t,s), [\boldsymbol{C}]_{t,s} = C(t,s)$, etc. We let $\boldsymbol{\Theta}(t,s) = \eta\Theta(t-s)$ where η is the learning rate.

1. Solve for the response functions by updating the closed equations as matrices by iterating the equations.

$$\begin{aligned}
&\boldsymbol{R}_{0,2,k} \leftarrow -\left[\boldsymbol{\Theta}^{-1} + \lambda_k \boldsymbol{R}_3 \boldsymbol{R}_1\right]^{-1} \boldsymbol{R}_3, \\
&\boldsymbol{R}_{0,2} \leftarrow \frac{1}{M}\sum_k \lambda_k \boldsymbol{R}_{0,2,k}, \\
&\boldsymbol{R}_1 \leftarrow \left[\boldsymbol{I} - \alpha^{-1}\boldsymbol{R}_{0,2}\right]^{-1}, \\
&\boldsymbol{R}_{2,4,k} \leftarrow -\lambda_k \left[\boldsymbol{I} + \lambda_k \boldsymbol{R}_1 \boldsymbol{\Theta} \boldsymbol{R}_3\right]^{-1} \boldsymbol{R}_1 \boldsymbol{\Theta}\ ,\ \boldsymbol{R}_{2,4} = \frac{1}{M}\sum_k \boldsymbol{R}_{2,4,k}, \\
&\boldsymbol{R}_3 \leftarrow \left[\boldsymbol{I} - \nu^{-1}\boldsymbol{R}_{2,4}\right]^{-1}.
\end{aligned} \tag{146}$$

2. Once these response functions have converged, we can iterate the equations for the correlation functions

$$\begin{aligned}
&\boldsymbol{C}_{0,k} \leftarrow \left[\boldsymbol{I} + \lambda_k \boldsymbol{\Theta}\boldsymbol{R}_3\boldsymbol{R}_1\right]^{-1}\left[(w_k^\star)^2 \boldsymbol{1}\boldsymbol{1}^\top + \boldsymbol{\Theta}\left(\nu^{-1}\boldsymbol{C}_3 + \frac{\lambda_k}{\alpha}\boldsymbol{R}_3\boldsymbol{C}_1\boldsymbol{R}_3^\top\right)\boldsymbol{\Theta}^\top\right]\left[\boldsymbol{I} + \lambda_k \boldsymbol{\Theta}\boldsymbol{R}_3\boldsymbol{R}_1\right]^{-1\top}, \\
&\boldsymbol{C}_0 \leftarrow \frac{1}{M}\sum_{k=1}^{M}\lambda_k \boldsymbol{C}_{0,k}, \\
&\boldsymbol{C}_1 \leftarrow \boldsymbol{R}_1\boldsymbol{C}_0\boldsymbol{R}_1^\top, \\
&\boldsymbol{C}_{2,k} \leftarrow \left[\boldsymbol{I} + \lambda_k \boldsymbol{R}_1\boldsymbol{\Theta}\boldsymbol{R}_3\right]^{-1}\left(\frac{\lambda_k}{\alpha}\boldsymbol{C}_1 + \boldsymbol{R}_1\left[(w_k^\star)^2\lambda_k^2 \boldsymbol{1}\boldsymbol{1}^\top + \frac{\lambda_k^2}{\nu}\boldsymbol{\Theta}\boldsymbol{C}_3\boldsymbol{\Theta}^\top\right]\boldsymbol{R}_1^\top\right)\left[\boldsymbol{I} + \lambda_k \boldsymbol{R}_1\boldsymbol{\Theta}\boldsymbol{R}_3\right]^{-1\top}, \\
&\boldsymbol{C}_2 \leftarrow \frac{1}{M}\sum_k \boldsymbol{C}_{2,k}.
\end{aligned} \tag{147}$$

After iterating these equations, one has the discrete time solution to the DMFT order parameters and any other observable can then be calculated.

M.2. Fourier Transform Method

To accurately compute the Fourier transforms in the model/data bottleneck regime ($\alpha < 1$ or $\nu < 1$) we have that $\mathcal{R}_1(\omega)\mathcal{R}_3(\omega) \sim i\omega r$ as $\omega \to 0$ so we must resort to analyzing the principal part and the delta-function contribution to the integral. Construct a shifted and non-divergent version of the function $\mathcal{H}(\omega)$.

$$\begin{aligned}
&\mathcal{H}(\omega) = \tilde{\mathcal{H}}(\omega) + \frac{1}{\epsilon + i\omega(1+r)} \\
&\tilde{\mathcal{H}}(\omega) = \frac{1}{\epsilon + i\omega + \mathcal{R}_1(\omega)\mathcal{R}_3(\omega)} - \frac{1}{\epsilon + i\omega(1+r)} = \frac{i\omega r - \mathcal{R}_1(\omega)\mathcal{R}_3(\omega)}{(\epsilon + i\omega + \mathcal{R}_1(\omega)\mathcal{R}_3(\omega))(\epsilon + i\omega(1+r))},
\end{aligned} \tag{148}$$

where $r = \lim_{\omega\to 0}\frac{1}{i\omega}\mathcal{R}_1(\omega)\mathcal{R}_3(\omega)$. We see that rather than diverging like $\mathcal{H}(\omega)$, this function $\tilde{\mathcal{H}}(\omega)$ vanishes as $\omega \to 0$. We therefore numerically perform Fourier integral against $\tilde{H}(\omega)$ and then add the singular component which can be computed separately.

$$H(\tau) = \int \frac{d\omega}{2\pi} e^{i\omega\tau}\tilde{\mathcal{H}}(\omega) + \int \frac{d\omega}{2\pi}\frac{e^{i\omega\tau}}{\epsilon + i\omega(1+r)} = \int \frac{d\omega}{2\pi} e^{i\omega\tau}\tilde{\mathcal{H}}(\omega) + \frac{1}{1+r} \tag{149}$$

where we used the fact that

$$\frac{1}{\epsilon + i\omega(1+r)} = \frac{\pi}{1+r}\delta(\omega) - \frac{i}{1+r}\mathcal{P}(\omega^{-1})\ ,\ \epsilon \to 0 \tag{150}$$

The Dirac mass is trivial to integrate over giving $\frac{1}{2(1+r)}$. Lastly, we must perform an integral of the type

$$-\frac{i}{1+r}\mathcal{P}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\frac{e^{i\omega\tau}}{\omega} = \frac{1}{\pi(1+r)}\int_0^{\infty} d\omega\frac{\sin(\omega\tau)}{\omega} = \frac{1}{2(1+r)} \tag{151}$$

Adding these two terms together, our transfer function has the form

$$H(\tau) = \frac{1}{1+r} + \int_{-\infty}^{\infty}\frac{d\omega}{2\pi}e^{i\omega\tau}\tilde{\mathcal{H}}(\omega). \tag{152}$$

The last integral can be performed numerically, giving a more stable result.

N. Compute Optimal Scaling from Sum of Power-Laws

We suppose that the loss scales as (neglecting irrelevant prefactors)

$$\mathcal{L} = t^{-r_t} + N^{-r_N} + P^{-r_P} + \mathcal{L}_\infty \tag{153}$$

Our goal is to minimize the above expression subject to the constraint that compute $C = Nt$ is fixed. Since C is fixed we can reduce this to a one-dimensional optimization problem

$$\min_N \left[C^{-r_t}N^{r_t} + N^{-r_N}\right] \tag{154}$$

The optimality condition $\partial_N L = 0$ is

$$\begin{aligned} &r_t C^{-r_t}N^{r_t-1} - r_N N^{-r_N-1} = 0 \\ &\implies N \propto C^{\frac{r_t}{r_t+r_N}} \implies t \propto C^{\frac{r_t}{r_t+r_N}} \end{aligned} \tag{155}$$

From this last expression one can evaluate the loss at the optimum

$$\mathcal{L}_\star(C) \propto C^{-\frac{r_t r_N}{r_t+r_N}}. \tag{156}$$