Assuming that all response functions and transfer functions $\mathcal{H}_k$ have been solved for, the correlation functions satisfy the closed set of linear equations.

$$\begin{aligned}
\mathcal{C}_0(\omega,\omega') &= \frac{1}{M}\sum_k \lambda_k \mathcal{H}_k(\omega)\mathcal{H}_k(\omega')\left[(w_k^\star)^2 + \frac{1}{\nu}\mathcal{C}_3(\omega,\omega') + \frac{1}{\alpha}\lambda_k \mathcal{R}_3(\omega)\mathcal{R}_3(\omega')\mathcal{C}_1(\omega,\omega')\right] \\
\mathcal{C}_1(\omega,\omega') &= \mathcal{R}_1(\omega)\mathcal{R}_1(\omega')\mathcal{C}_0(\omega,\omega') \\
\mathcal{C}_2(\omega,\omega') &= \frac{1}{M}\sum_k \lambda_k \mathcal{H}_k(\omega)\mathcal{H}_k(\omega')\left[\frac{1}{\alpha}(i\omega)(i\omega')\mathcal{C}_1(\omega,\omega') + \lambda_k \mathcal{R}_1(\omega)\mathcal{R}_1(\omega')\left((w_k^\star)^2 + \frac{1}{\nu}\mathcal{C}_3(\omega,\omega')\right)\right] \\
\mathcal{C}_3(\omega,\omega') &= \mathcal{R}_3(\omega)\mathcal{R}_3(\omega')\mathcal{C}_2(\omega,\omega')
\end{aligned} \tag{35}$$

These equations can be efficiently solved for all pairs of ω, ω' after the response functions have been identified. Then one can take an inverse Fourier transform in both indices.

B. Field Theoretic Derivation of DMFT Equations

In this section, we derive the field theoretic description of our model. We will derive this using both the Martin-Siggia-Rose (MSR) path integral method (Martin et al., 1973) and the dynamical cavity method. For a recent review of these topics in the context of neural networks, see (Helias & Dahmen, 2020).

B.1. Statistical Assumptions for DMFT

The DMFT that we derive in the next few sections requires some assumptions on the structure of $\boldsymbol{A}$ and $\boldsymbol{\Psi}$. To carry out the classic MSR path integral computation, we assume that the entries of both matrices are Gaussian with mean zero and covariance

$$\langle A_{ij} A_{kl}\rangle = \delta_{ik}\delta_{jl} \; , \; \langle \Psi_{\mu k}\Psi_{\nu l}\rangle = \delta_{\mu\nu}\delta_{kl}\lambda_k. \tag{36}$$

These are sufficient conditions for the DMFT description to hold and we will take them as our primary assumptions. However, we note that these restrictions are not strictly necessary and can be relaxed. In general, a more flexible cavity derivation in Appendix B.3 shows that independent entries from any well behaved distribution which admits a central limit theorem for sums of independent draws would also have the same DMFT description of the proportional limit. Prior works on DMFT of M-estimators with random data have demonstrated universality for any data matrix $\boldsymbol{\Psi}$ with a covariance that has bounded spectral norm (Gerbelot et al., 2022).

B.2. Path Integral Derivation

With the MSR formalism, we evaluate the moment generating functional for the field $\boldsymbol{v}^0(t)$:

$$Z[\{\boldsymbol{j}(t)\}] = \left\langle \int \mathcal{D}\boldsymbol{v}^0(t)\,\delta\left(\dot{\boldsymbol{v}}^0(t) + \frac{1}{NP}\boldsymbol{A}^\top\boldsymbol{A}\boldsymbol{\Psi}^\top\boldsymbol{\Psi}\boldsymbol{v}^0(t)\right)\exp\left(\int dt\, \boldsymbol{j}(t)\cdot\boldsymbol{v}^0(t)\right)\right\rangle_{\boldsymbol{A},\boldsymbol{\Psi}}. \tag{37}$$

Note that at zero source, we have the important identity that

$$Z[0] = 1. \tag{38}$$

We insert a Dirac delta functions to enforce the definitions of each of the fields $\{\boldsymbol{v}^1, \boldsymbol{v}^2, \boldsymbol{v}^3, \boldsymbol{v}^4\}$ as in equation 8.

$$\begin{aligned}
Z[\{\boldsymbol{j}(t)\}] = &\int \mathcal{D}[\boldsymbol{v}^0,\ldots,\boldsymbol{v}^4,\hat{\boldsymbol{v}}^1\ldots\hat{\boldsymbol{v}}^4]\,\delta(\dot{\boldsymbol{v}}^0 + \boldsymbol{v}^4)\exp\left(\int dt\,\boldsymbol{j}(t)\cdot\boldsymbol{v}^0(t)\right) \\
&\times\left\langle \exp\left[i\int dt\left[\hat{\boldsymbol{v}}_1(t)\cdot\left(\boldsymbol{v}^1(t) - \frac{1}{\sqrt{M}}\boldsymbol{\Psi}\boldsymbol{v}^0(t)\right) + \hat{\boldsymbol{v}}_2(t)\cdot\left(\boldsymbol{v}^2(t) - \frac{1}{\alpha\sqrt{M}}\boldsymbol{\Psi}^\top\boldsymbol{v}^1(t)\right)\right]\right]\right\rangle_{\boldsymbol{\Psi}} \\
&\times\left\langle \exp\left[i\int dt\left[\hat{\boldsymbol{v}}_3(t)\cdot\left(\boldsymbol{v}^3(t) - \frac{1}{\sqrt{M}}\boldsymbol{A}\boldsymbol{v}^2(t)\right) + \hat{\boldsymbol{v}}_4(t)\cdot\left(\boldsymbol{v}^4(t) - \frac{1}{\nu\sqrt{M}}\boldsymbol{A}^\top\boldsymbol{v}^3(t)\right)\right]\right]\right\rangle_{\boldsymbol{A}}.
\end{aligned} \tag{39}$$

At this stage we can add sources $\tilde{\boldsymbol{j}}$ for each $\hat{\boldsymbol{v}}_i$ variable, yielding a $Z[\boldsymbol{j}(t), \tilde{\boldsymbol{j}}(t)]$. Interpreting each source as modification of the respective evolution equation, we see that this modified moment-generating function remains equal to unity at any value of $\tilde{\boldsymbol{j}}$, $Z[0, \tilde{\boldsymbol{j}}(t)] = 1$. As a consequence, all correlation functions consisting only of $\hat{\boldsymbol{v}}^i$ variables vanish. See (Crisanti & Sompolinsky, 2018) for further details and a worked example.

We now average over the sources of disorder. We assume that the entries of $\boldsymbol{A}$ are i.i.d. with mean zero and variance 1. In the proportional limit, we can replace the entries of $\boldsymbol{A}$ as a draw from a Gaussian $\mathcal{N}(0,1)$ by appealing to Gaussian equivalence. We furhther justify this in the cavity derivation in the next section. This allows us to evaluate the averages over the matrix $\boldsymbol{A}$.

$$\begin{aligned}
&\left\langle \exp\left(-\frac{i}{\sqrt{M}} \mathrm{Tr} \boldsymbol{A}^\top \int dt \left[\hat{\boldsymbol{v}}^3(t) \boldsymbol{v}^2(t)^\top + \nu^{-1} \boldsymbol{v}^3(t) \hat{\boldsymbol{v}}^4(t)^\top\right]\right)\right\rangle_{\boldsymbol{A}} \\
&= \exp\left(-\frac{1}{2}\int dtds [\hat{\boldsymbol{v}}^3(t)\cdot\hat{\boldsymbol{v}}^3(s) \underbrace{\frac{1}{M}\boldsymbol{v}^2(t)\cdot\boldsymbol{v}^2(s)}_{C_2(t,s)} + \nu^{-1}\hat{\boldsymbol{v}}^4(t)\cdot\hat{\boldsymbol{v}}^4(s)\underbrace{\frac{1}{N}\boldsymbol{v}^3(t)\cdot\boldsymbol{v}^3(s)]}_{C_3(t,s)}\right) \\
&\quad \times \exp\left(-\int dtds \underbrace{\frac{1}{N}\hat{\boldsymbol{v}}^3(t)\cdot\boldsymbol{v}^3(s)}_{iR_3(s,t)} \boldsymbol{v}^2(t)\cdot\hat{\boldsymbol{v}}^4(s)\right).
\end{aligned} \tag{40}$$

Similarly, we can calculate the averages over the data, which enters via the design matrices $\boldsymbol{\Psi}$. Again in this proportional limit we can invoke Gaussian equivalence on $\boldsymbol{\Psi}$ to have it take the form $\boldsymbol{\Psi} \sim \boldsymbol{\Phi}\boldsymbol{\Lambda}^{1/2}$ where $\boldsymbol{\Phi}$ has entries drawn from a unit normal. Taking the average then gives us

$$\begin{aligned}
&\left\langle \exp\left(-\frac{i}{\sqrt{M}} \mathrm{Tr}\, \boldsymbol{\Psi}^\top \int dt \left[\hat{\boldsymbol{v}}^1(t) \boldsymbol{v}^0(t)^\top + \alpha^{-1} \boldsymbol{v}^1(t) \hat{\boldsymbol{v}}^2(t)^\top\right]\right)\right\rangle_{\boldsymbol{\Psi}} \\
&= \exp\left(-\frac{1}{2}\int dtds [\hat{\boldsymbol{v}}^1(t)\cdot\hat{\boldsymbol{v}}^1(s) \underbrace{\frac{1}{M}\boldsymbol{v}^0(t)\cdot\boldsymbol{\Lambda}\boldsymbol{v}^0(s)}_{C_0(t,s)} + \alpha^{-1}\hat{\boldsymbol{v}}^2(t)\cdot\boldsymbol{\Lambda}\hat{\boldsymbol{v}}^2(s)\underbrace{\frac{1}{P}\boldsymbol{v}^1(t)\cdot\boldsymbol{v}^1(s)]}_{C_1(t,s)}\right), \\
&\quad \times \exp\left(-\int dtds \underbrace{\frac{1}{P}\hat{\boldsymbol{v}}^1(t)\cdot\boldsymbol{v}^1(s)}_{iR_1(s,t)} \boldsymbol{v}^0(t)\cdot\boldsymbol{\Lambda}\cdot\hat{\boldsymbol{v}}^2(s)\right).
\end{aligned} \tag{41}$$

We now insert delta functions for following bracketed terms: C_0, C_1, C_2, C_3 and R_1, R_3 using the following identity (e.g. for C_0 at times s, t):

$$1 = \int \frac{dC_0(s,t) d\hat{C}_0(s,t)}{4\pi i M^{-1}} \exp\left[\frac{1}{2} M \int dtds\, \hat{C}_0(t,s) \left(C_0(t,s) - \frac{1}{M}\boldsymbol{v}^0(t)\cdot\boldsymbol{\Lambda}\boldsymbol{v}^0(s)\right)\right]. \tag{42}$$

Here the $\hat{C}_i, \hat{R}_i$ integrals are taken over the imaginary axis. This yields a moment generating function (here we'll take $\boldsymbol{j} = 0$):

$$Z = \int \mathcal{D}[C_1, \hat{C}_1, \dots] \exp\left[M S[C_0, C_1, C_2, C_3, R_1, R_3, \hat{C}_0, \hat{C}_1, \hat{C}_2, \hat{C}_3, \hat{R}_1, \hat{R}_3]\right]. \tag{43}$$

The constraint that $Z = 1$ means that $S = 0$ at the saddle point. S here is given by:

$$\begin{aligned}
S[\dots] = &\frac{1}{2}\int dtds \Big[\hat{C}_0(t,s)C_0(t,s) + \alpha\hat{C}_1(t,s)C_1(t,s) + \hat{C}_2(t,s)C_2(t,s) + \nu\hat{C}_3(t,s)C_3(t,s)\Big] \\
&+ \int dtds \Big[-R_1(t,s)\hat{R}_1(s,t) - R_3(t,s)\hat{R}_3(s,t)\Big] \\
&+ \alpha \log \mathcal{Z}_1 + \nu \log \mathcal{Z}_3 + \frac{1}{M}\sum_k \log \mathcal{Z}_{0,2,4;k}.
\end{aligned} \tag{44}$$

We have chosen to take $\hat{R}_i(s,t)$ to have a different sign and s,t ordering convention than the $\hat{C}_i$ to simplify our notation later on. We have also used that Equations (40), (41) factorize over their respective indices, so each $\mathcal{Z}$ is a partition function over a single index. The individual $\mathcal{Z}_i$ are given by:

$$\begin{aligned}\mathcal{Z}_1 = \int \mathcal{D}[v^1,\hat{v}^1] \exp\left[i\int dtds\left(\delta(t-s) - \alpha^{-1}\hat{R}_1(s,t)\right) v^1(t)\hat{v}^1(s)\right] \\ \times \exp\left[-\frac{1}{2}\int dtds (\hat{v}^1(t)\hat{v}^1(s)C_0(t,s) + v^1(t)v^1(s)\hat{C}_1(t,s))\right],\end{aligned} \tag{45}$$

$$\begin{aligned}\mathcal{Z}_3 = \int \mathcal{D}[v^3,\hat{v}^3] \exp\left[i\int dtds\left(\delta(t-s) - \nu^{-1}\hat{R}_3(s,t)\right) v^3(t)\hat{v}^3(s)\right] \\ \times \exp\left[-\frac{1}{2}\int dtds (\hat{v}^3(t)\hat{v}^3(s)C_2(t,s) + v^3(t)v^3(s)\hat{C}_3(t,s))\right],\end{aligned} \tag{46}$$

$$\begin{aligned}\mathcal{Z}_{0,2,4;k} = \int \mathcal{D}[v^{0,2,4},\hat{v}^{0,2,4}] \exp\left[-\frac{1}{2}\int dtds\left(\alpha^{-1}\lambda_k\hat{v}_k^2(t)\hat{v}_k^2(s)C_1(t,s) + \nu^{-1}\hat{v}_k^4(t)\hat{v}_k^4(s)C_3(t,s)\right)\right] \\ \times \exp\left[-\frac{1}{2}\int dtds\left(\lambda_k v_k^0(t)v_k^0(s)\hat{C}_0(t,s) + v_k^2(t)v_k^2(s)\hat{C}_2(t,s) + v_k^4(t)v_k^4(s)\hat{C}_4(t,s)\right)\right] \\ \times \exp\left[-i\int dtds\left(R_3(t,s)v_k^2(s)\hat{v}_k^4(t) + \lambda_k R_1(t,s)v_k^0(s)\hat{v}_k^2(t)\right)\right].\end{aligned} \tag{47}$$

In the large M limit we evaluate this integral via saddle point. The saddle point equations give:

$$\begin{aligned} C_0(t,s) &= \frac{1}{M}\sum_k \lambda_k \left\langle v_k^0(t)v_k^0(s)\right\rangle \\ C_i(t,s) &= \langle v^\ell(t)v^\ell(s)\rangle, \quad \ell=\{1,2,3,4\} \\ \\ R_1(t,s) &= -i\langle v^1(t)\hat{v}^1(s)\rangle \\ R_3(t,s) &= -i\langle v^3(t)\hat{v}^3(s)\rangle \\ \hat{R}^1(t,s) &= -i\frac{1}{M}\sum_k \lambda_k\hat{v}_k^2(t)v_k^0(s) \equiv R_{0,2}(t,s) \\ \hat{R}^3(t,s) &= -i\frac{1}{M}\sum_k \lambda_k\hat{v}_k^4(t)v_k^2(s) \equiv R_{2,4}(t,s). \end{aligned} \tag{48}$$

Here $\langle\cdot\rangle$ denotes an average taken with respect to the statistical ensemble given by the corresponding partition function $\mathcal{Z}_i$. Lastly, the saddle point equations for the $\hat{C}_i(t,s)$ variables are all quadratic functions of the variables $\{\hat{v}^0,\hat{v}^1,\hat{v}^2,\hat{v}^3\}$ which vanish under the average defined by $\mathcal{Z}$ (Helias & Dahmen, 2020). Following the discussion below Equation 39, we take $\hat{C}_i(t,s)=0$, which will enforce $\langle\hat{v}_i(t)\hat{v}_i(s)\rangle = 0$ and lead to the correct dynamical equations.

To evaluate the remaining, we can integrate out the $\hat{v}^i$ variables. First let us look at $\mathcal{Z}_1$. Using the Hubbard-Stratonovich trick we can write the action in terms linear in $\hat{v}^1$. This gives

$$\begin{aligned}\mathcal{Z}_1 = \int \mathcal{D}[v^1,\hat{v}^1,u^1] \exp\left[i\int dtds\, \hat{v}^1(t)\left[\delta(t-s)(v^1(s)-u(s)) - \alpha^{-1}\hat{R}_1(t,s)v^1(s)\right]\right] \\ \times \exp\left[-\frac{1}{2}\int dtds\, u(t)u(s)C_0^{-1}(t,s) + v^1(t)v^1(s)\hat{C}_1(t,s))\right]\end{aligned} \tag{49}$$

We now replace $\hat{C}_1$ by its saddle point value of 0 and $\hat{R}_1$ by $R_{0,2}$. Integrating over $\hat{v}$ gives a delta function:

$$v^1(t) = u^1(t) + \frac{1}{\alpha}\int dsR_{0,2}(t,s)v_1(s), \quad u^1(t) \sim \mathcal{GP}(0,C_0). \tag{50}$$