Analogously for v^3 we get

$$v^3(t) = u^3(t) + \frac{1}{\nu}\int dsR_{2,4}(t,s)v^3(s), \quad u^3(t) \sim \mathcal{GP}(0, C_2) \tag{51}$$

For $\mathcal{Z}_{0,2,4;k}$ after replacing $\hat{C}_0, \hat{C}_2, \hat{C}_4$ with their saddle point values we get:

$$\begin{aligned}\mathcal{Z}_{0,2,4;k} = \exp&\left[-\frac{1}{2}\int dtds\left(\alpha^{-1}\lambda_k\hat{v}_k^2(t)\hat{v}_k^2(s)C_1(t,s) + \nu^{-1}\hat{v}_k^4(t)\hat{v}_k^4(s)C_3(t,s)\right)\right]\\ \times\exp&\left[-\int dtds\left(iR_3(t,s)v_k^2(s)\hat{v}_k^4(t) + i\lambda_kR_1(t,s)v_k^0(s)\hat{v}_k^2(t)\right)\right]\end{aligned} \tag{52}$$

Using the same Hubbard-Stratonovich trick on $\hat{v}_k^2$ gives:

$$v_k^2(t) = u_k^2(t) + \lambda_k\int dsR_1(t,s)v_k^0(s), \quad u_k^2(t) \sim \mathcal{GP}\left(0, \frac{1}{\alpha}\lambda_kC_1\right). \tag{53}$$

On $\hat{v}_k^4$ we similarly get:

$$v_k^4(t) = u_k^4(t) + \int dsR_3(t,s)\cdot v_k^2(s), \quad u_k^4(t) \sim \mathcal{GP}\left(0, \frac{1}{\nu}C_3\right) \tag{54}$$

Lastly, the equations of motion for v_k^0 in terms of v_k^4 are known:

$$\partial_t v_k^0(t) = -v_k^4(t). \tag{55}$$

One can easily add momentum by replacing $\partial_t v_k^0(t)$ with $(\beta\partial_t^2 + \partial_t)v_k^0(t)$ without changing anything else about the derivation.

B.2.1. Interpretation of the Response Functions

Following (Crisanti & Sompolinsky, 2018; Helias & Dahmen, 2020), we can understand the $\langle\hat{v}^a(t)v^b(s)\rangle$ correlators by adding in the single-site moment generating function (e.g. Equation (49)) a source $\tilde{j}^b(s)$ that couples to $\hat{v}^b$ at time s. As in the discussion below equation 39, differentiating $\langle v^a(t)\rangle$ with respect to that source corresponds to its response to a kick in the dynamics of v^b at time s. We denote this by:

$$R_{i,j}(t,s) = \left\langle\frac{\delta v^i(t)}{\delta v^j(s)}\right\rangle. \tag{56}$$

B.3. Cavity Derivation

The cavity derivation relies on Taylor expanding the dynamics upon the addition of a new sample or feature. We will work through each cavity step one at a time by considering the influence of a single new base feature, new sample, and new projected feature. In each step, the goal is to compute the marginal statistics of the added variables. This requires tracking the linear response to all other variables in the system.

Adding a Base Feature Upon addition of a base feature with eigenvalue λ_0 so that there are $M+1$ instead of M features $\{v_k^0, v_k^2, v_k^4\}$ for $k \in \{0, 1, ..., M\}$, we have a perturbation to both $v_\mu^1(t)$ and $v_n^3(t)$. Denote the perturbed versions of the dynamics upon addition of the $M+1$st feature as $\tilde{v}_\mu^1(t)$ and $\tilde{v}_n^3(t)$. At large M we can use linear-response theory to relate the dynamics at $M+1$ features to the dynamics of the original M feature system

$$\begin{aligned}\tilde{v}_\mu^1(t) &\sim v_\mu^1(t) + \frac{1}{\sqrt{M}}\sum_{\nu=1}^{P}\int ds\frac{\partial v_\mu^1(t)}{\partial v_\nu^1(s)}\psi_0^\nu v_0^0(s)\\ \tilde{v}_n^3(t) &\sim v_\mu^1(t) + \frac{1}{\sqrt{M}}\sum_{m=1}^{N}\int ds\frac{\partial v_n^3(t)}{\partial v_m^3(s)}A_{m0}\, v_0^2(s)\end{aligned} \tag{57}$$

The next order corrections have a subleading influence on the dynamics. Now, inserting these perturbed dynamics into the dynamics for the new $(M+1)$st set of variables $\{v_0^2(t), v_0^4(t)\}$. For $v_0^2(t)$, we have

$$v_0^2(t) \sim \frac{1}{\alpha\sqrt{M}} \sum_{\mu=1}^{P} \psi_0^\mu v_\mu^1(t) + \frac{1}{\alpha M} \sum_{\mu,\nu=1}^{P} \int ds\ \psi_0^\mu \frac{\partial v_\mu^1(t)}{\partial v_\nu^1(s)} \psi_0^\nu v_0^0(s) \tag{58}$$

There are now two key steps in simplifying the above expression in the proportional limit:

1. By the fact that the $v_\mu^1(t)$ dynamics are statistically independent of the new feature ψ_0^μ, we can invoke a central limit theorem for the first term which is mean zero and variance $\mathcal{O}(1)$.
2. Similarly, we can invoke a law of large numbers for the second term, which has $\mathcal{O}(1)$ mean and variance on the order of $\mathcal{O}(M^{-1})$. Therefore in the asymptotic limit it can be safely approximated by its mean.

We note in passing that neither of these steps require the ψ_0^μ variables to be Gaussian. Thus we obtain the following asymptotic statistical description of the $v_0^2(t)$ random variable

$$\begin{aligned} & v_0^2(t) \sim u_0^2(t) + \int ds R_1(t,s) v_0^0(s) \\ & u_0^2(t) \sim \mathcal{GP}(0, \alpha^{-1}\lambda_0 C_1)\ ,\ R_1(t,s) \equiv \frac{1}{P}\sum_{\mu=1}^{P} \left\langle \frac{\partial v_\mu^1(t)}{\partial v_\mu^1(s)} \right\rangle . \end{aligned} \tag{59}$$

Following an identical argument for $v_0^4(t)$ we have

$$\begin{aligned} v_0^4(t) & \sim \frac{1}{\nu\sqrt{M}} \sum_{n=1}^{M} A_{n0} v_n^3(t) + \frac{1}{\nu M}\sum_{nm} \int ds\ A_{n0} \frac{\partial v_n^3(t)}{\partial v_m^3(s)} A_{m0} v_0^2(s) \\ & \sim u_0^3(t) + \int ds R_3(t,s) v_0^2(s) \\ & u_0^3(t) \sim \mathcal{GP}(0, \nu^{-1} C_3)\ ,\ R_3(t,s) = \frac{1}{N}\sum_{n=1}^{N} \left\langle \frac{\partial v_n^3(t)}{\partial v_n^3(s)} \right\rangle \end{aligned} \tag{60}$$

Adding a Sample Next, we can consider the influence of adding a new data point. We will aim to characterize a $P+1$ data point system in terms of the dynamics when P points are present. Upon the addition of a new data point $\boldsymbol{\psi}^0$ the field $v_k^0(t)$ will be perturbed to $\tilde{v}_k^0(t)$. Again invoking linear response theory, we can expand the perturbed value around the P-sample dynamics

$$\tilde{v}_k^0(t) \sim v_k^0(t) + \frac{1}{\alpha\sqrt{M}} \sum_{\ell=1}^{M} \int ds\ \frac{\partial v_k^0(t)}{\partial v_\ell^2(s)} \psi_\ell^0 v_0^1(s) \tag{61}$$

Now, computing the dynamics of the new random variable $v_0^1(t)$

$$\begin{aligned} v_0^1(t) & \sim \frac{1}{\sqrt{M}} \sum_{k=1}^{M} \psi_k^0 v_k^0(t) + \frac{1}{\alpha M} \int ds \sum_{k\ell} \psi_k^0\ \frac{\partial v_k^0(t)}{\partial v_\ell^2(s)}\ \psi_\ell^0\ v_0^1(s) \\ & \sim u_0^1(t) + \frac{1}{\alpha} \int ds R_{0,2}(t,s) v_0^1(s) \\ u_0^1(t) & \sim \mathcal{GP}\left(0, \frac{1}{M}\sum_k \lambda_k C_k^0\right)\ ,\ R_{0,2}(t,s) = \frac{1}{M}\sum_{k=1}^{M} \lambda_k \left\langle \frac{\partial v_k^0(t)}{\partial v_k^2(s)} \right\rangle \end{aligned} \tag{62}$$

Adding a Projected Feature Now, we finally consider the effect of introducing a single new projected feature so that instead of N we now have $N+1$ projected features. This causes a perturbation to $\{v_k^2(t)\}$ which we

$$\tilde{v}_k^2(t) \sim v_k^2(t) + \frac{1}{\nu\sqrt{M}} \sum_{\ell=1}^{M} \int ds\ \frac{\partial v_k^2(t)}{\partial v_\ell^4(s)} A_{0\ell}\ v_0^3(s) \tag{63}$$

Now, we compute the dynamics for the added variable $v_0^3(t)$

$$\begin{aligned} v_0^3(t) &\sim \frac{1}{\sqrt{M}}\sum_{k=1}^{M} A_{0k} v_k^2(t) + \frac{1}{\nu M}\sum_{k\ell}\int ds A_{0k}\frac{\partial v_k^2(t)}{\partial v_\ell^4(s)} A_{0\ell}\, v_0^3(s) \\ &\sim u_0^3(t) + \frac{1}{\nu}\int ds R_{2,4}(t,s) v_0^3(s) \\ u_0^3(t) &\sim \mathcal{GP}(0, C_2)\ ,\ R_{2,4}(t,s) = \frac{1}{M}\sum_{k=1}^{M}\left\langle \frac{\partial v_k^2(t)}{\partial v_k^4(s)}\right\rangle \end{aligned} \tag{64}$$

Putting it all together Now, using the information gained in the previous sections, we can combine all of the dynamics for each field into a closed set of stochastic processes. This recovers the DMFT equations of Appendix A.2.

C. Final Losses (the $t \to \infty$ Limit of DMFT)

In this section we work out exact expressions for the large time limit of DMFT. By comparing with prior computations of the mean-field statics of this problem computed in (Atanasov et al., 2023; Zavatone-Veth & Pehlevan, 2023; Ruben & Pehlevan, 2023; Maloney et al., 2022; Simon et al., 2021), we show that the large time and large M limits commute, specifically that $\lim_{M,N,P\to\infty}\lim_{t\to\infty}\mathcal{L}(M,N,P,t) = \lim_{t\to\infty}\lim_{M,N,P\to\infty}\mathcal{L}(M,N,P,t)$. We invoke the final value theorem and use the response functions as before.

Final Value Theorem We note that for functions which vanish at $t = -\infty$, that

$$\lim_{\omega\to 0} i\omega\ \mathcal{H}(\omega) = -\lim_{\omega\to 0}\int_{-\infty}^{\infty} d\tau \left[\frac{\partial}{\partial\tau}e^{-i\omega\tau}\right] H(\tau) = \lim_{\omega\to 0}\int_{-\infty}^{\infty} d\tau\left[\frac{\partial}{\partial\tau}H(\tau)\right] e^{-i\omega\tau} = \lim_{\tau\to\infty} H(\tau) \tag{65}$$

where we invoked integration by parts and used the assumption that $\lim_{\tau\to-\infty} H(\tau) = 0$, a condition that is satisfied for the correlation and response functions in our theory. We can therefore use the identity $\lim_{\tau\to\infty} H(\tau) = \lim_{\omega\to 0} i\omega\mathcal{H}(\omega)$ to extract the final values of our order parameters.

$$\lim_{\tau\to\infty} H_k(\tau) = \lim_{\omega\to 0} i\omega\mathcal{H}_k(\omega) = \lim_{\omega\to 0}\frac{1}{1+\lambda_k (i\omega)^{-1}\mathcal{R}_1(\omega)\mathcal{R}_3(\omega)}. \tag{66}$$

We also need to invoke a similar relationship for the final values of the correlation functions

$$\begin{aligned} \lim_{t,s\to\infty} C(t,s) &= \lim_{\omega,\omega'\to 0}(i\omega)(i\omega')\,\mathcal{C}(\omega,\omega') \\ \mathcal{C}(\omega,\omega') &= \int dt\int ds\ e^{-i(\omega t+\omega' s)} C(t,s) \end{aligned} \tag{67}$$

where $\mathcal{C}$ is the two-variable Fourier transform. The final value of the test loss is $\lim_{t\to\infty}\mathcal{L}(t) = \lim_{t,s\to\infty} C_0(t,s)$.

C.1. General Case (Finite ν, α)

Before working out the solution to the response functions, we note that the following condition is always satisfied

$$\nu(1-\mathcal{R}_3(\omega)) = \alpha(1-\mathcal{R}_1(\omega)). \tag{68}$$

For $\nu = \alpha$, this equation implies that $\mathcal{R}_1 = \mathcal{R}_3$. For $\nu \neq \alpha$, we can have either $\mathcal{R}_1 \to 0$ or $\mathcal{R}_3 \to 0$ but not both. We consider each of these cases below.

Over-parameterized Case $\nu > \alpha$: In this case, the response function $\mathcal{R}_1 \sim \mathcal{O}(i\omega)$ as $\omega \to 0$ and $\mathcal{R}_3 \sim 1 - \frac{\alpha}{\nu}$ as $\omega \to 0$. We thus define

$$r \equiv \lim_{\omega\to 0}(i\omega)^{-1}\mathcal{R}_1(\omega)\mathcal{R}_3(\omega) \tag{69}$$