

Statistical Field Theory for Neural Networks

Mathematical Statistics and Field-Theoretic Frameworks (Ch. 6-10)

Ziliang Shen

2026 年 8 月 26 日

School of Statistics and Data Science
Shanghai University of Finance and Economics

Functional Preliminaries

Functional Formulation of SDEs

Ornstein-Uhlenbeck Process: Free Gaussian Theory

Perturbation Theory for SDEs

Dynamic Mean-Field Theory for Random Networks

Functional Preliminaries

Overview: This section summarizes the basic rules of functional calculus required in subsequent chapters.

Definition

Assume $f : C \mapsto \mathbb{R}$ is a functional mapping from a smooth function space C to real numbers. The functional derivative, which measures the sensitivity of the functional to the parameter at $x(t)$, is defined as:

$$\frac{\delta f[x]}{\delta x(t)} := \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} (f[x + \epsilon \delta(\cdot - t)] - f[x]) \quad (1)$$

Overview: This section summarizes the basic rules of functional calculus required in subsequent chapters.

Definition

Assume $f : C \mapsto \mathbb{R}$ is a functional mapping from a smooth function space C to real numbers. The functional derivative, which measures the sensitivity of the functional to the parameter at $x(t)$, is defined as:

$$\frac{\delta f[x]}{\delta x(t)} := \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} (f[x + \epsilon \delta(\cdot - t)] - f[x]) \quad (1)$$

Equivalently, evaluated via standard calculus:

$$= \left. \frac{d}{d\epsilon} F(\epsilon) \right|_{\epsilon=0}, \quad \text{where } F(\epsilon) := f[x + \epsilon \delta(\cdot - t)]$$

The linearity of the derivative is obvious.

Product Rule

Since functional derivatives reduce to ordinary derivatives, conventional rules apply:

$$\frac{\delta}{\delta x(t)}(f[x]g[x]) = \frac{\delta f[x]}{\delta x(t)}g[x] + f[x]\frac{\delta g[x]}{\delta x(t)} \quad (2)$$

Basic Rules of Functional Derivatives

Product Rule

Since functional derivatives reduce to ordinary derivatives, conventional rules apply:

$$\frac{\delta}{\delta x(t)}(f[x]g[x]) = \frac{\delta f[x]}{\delta x(t)}g[x] + f[x]\frac{\delta g[x]}{\delta x(t)} \quad (2)$$

Chain Rule

For $g : C \mapsto C$, by discretizing the t -axis and applying the multi-dimensional chain rule in the infinitesimal limit, we obtain:

$$\frac{\delta}{\delta x(t)}f[g[x]] = \int ds \left. \frac{\delta f[y]}{\delta y(s)} \right|_{y=g} \frac{\delta g[x](s)}{\delta x(t)} \quad (3)$$

Special Case: Fourier Transform

Fourier Transform (傅里叶变换)

The relationship between the derivative of a functional and its Fourier transform exhibits **opposite transformation properties**. This is manifested by the opposite sign of ω and the factor $1/2\pi$.

Special Case: Fourier Transform

Fourier Transform (傅里叶变换)

The relationship between the derivative of a functional and its Fourier transform exhibits **opposite transformation properties**. This is manifested by the opposite sign of ω and the factor $1/2\pi$.

By applying the chain rule, we obtain an expression that remains invariant under Fourier transformation:

$$\begin{aligned}\int \frac{\delta f[x]}{\delta x(s)} y(s) ds &= \int \frac{1}{2\pi} \int e^{i\omega s} \underbrace{\frac{\delta f[x]}{\delta x(s)}}_{=\frac{\delta \hat{f}}{\delta X(\omega)}} ds Y(\omega) d\omega \\ &= \int \frac{\delta \hat{f}}{\delta X(\omega)} Y(\omega) d\omega\end{aligned}\tag{4}$$

Note: We will exploit this property when evaluating Feynman diagrams in the Fourier domain.

Functional Taylor Series

Perturbative methods often require the action to be expressed as an algebraic functional of the fields. This is achieved through a functional expansion:

$$f[x] = \sum_{n=0}^{\infty} \int dt_1 \cdots \int dt_n \frac{1}{n!} \frac{\delta^n f}{\delta x(t_1) \cdots \delta x(t_n)} \Big|_{x=0} \prod_{i=1}^n x(t_i) \quad (5)$$

Functional Taylor Series

Perturbative methods often require the action to be expressed as an algebraic functional of the fields. This is achieved through a functional expansion:

$$f[x] = \sum_{n=0}^{\infty} \int dt_1 \cdots \int dt_n \frac{1}{n!} \frac{\delta^n f}{\delta x(t_1) \cdots \delta x(t_n)} \Big|_{x=0} \prod_{i=1}^n x(t_i) \quad (5)$$

Generalizations:

- **Shifting the Expansion Point:** To expand around a point other than $x \equiv 0$, simply replace x with $x - x^0$.
- **Multiple Fields:** For functionals dependent on multiple fields, apply the functional Taylor expansion to each dependence individually.

Bridge: From 0D Variables to 1D Fields (Time-Series)

Recap of Ch. 1-5: We treated random variables $x \in \mathbb{R}^N$ as "0D fields". The physics Action $S(x)$ is simply the negative **Log-Likelihood** of a static distribution.

The Paradigm Shift to Ch. 6-9: We now introduce **Time (t)**. The random vector becomes a continuous **Stochastic Process** (or time-series) $x(t)$.

Concept	0D (Static Statistics)	1D (Dynamic Statistics)
Statistical Object	Random Vector $x \in \mathbb{R}^N$	Time-Series / Stochastic Process $x(t)$
Probability	Joint PDF $p(x)$	Path Measure $\mathcal{P}[x(t)]$
The Action S	Log-Likelihood $\ln p(x)$	Dynamic Log-Likelihood of a trajectory
Calculus	Partial Derivative $\frac{\partial}{\partial x_i}$	Functional Derivative $\frac{\delta}{\delta x(t)}$
Expectation	Multiple Integral $\int \dots dx$	Path Integral $\int \mathcal{D}x$ (Sum over histories)

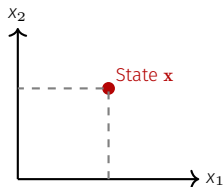
Goal: We need functional calculus to generalize statistical moments and cumulants (via Feynman diagrams) from static variables to continuous-time Stochastic Differential Equations (SDEs).

Conceptualizing the Path Measure: From Points to Trajectories

Static: The Point Measure

We evaluate the likelihood of discrete variables landing in a specific state (a "point" in space).

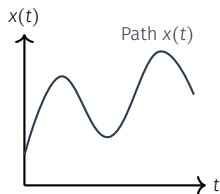
$$\rho(\mathbf{x}) \propto \exp(-S_{\text{static}}(\mathbf{x}))$$



Dynamic: The Path Measure

We evaluate the likelihood of the system realizing an **entire continuous curve** over time.

$$\mathcal{P}[x(t)] \propto \exp(-S_{\text{dynamic}}[x(t)])$$



The SDE as a Translation Engine

An SDE ($\dot{x} = f(x) + \xi$) acts as a **stochastic change of variables**.

It maps the known, simple Gaussian measure of the white noise $\mathcal{P}[\xi(t)]$ directly onto the trajectory.

The resulting Action in the exponent is exactly the **Trajectory Log-Likelihood**.

Functional Formulation of SDEs

Stochastic Differential Equations (SDEs)

We consider the SDE with an initial value $x(0+) = a$:

$$dx(t) = \underbrace{f(x)dt}_{\text{Drift}} + \underbrace{g(x)dW(t)}_{\text{Diffusion}}, \quad (6)$$

where W is a Wiener process.

SDEs are defined as the continuous-time limit of dynamics on a discrete time lattice with spacing h . For additive noise where $g(x) = \text{const}$, different discretization conventions for the drift term still lead to different path-integral representations.

Stochastic Differential Equations (SDEs)

We consider the SDE with an initial value $x(0+) = a$:

$$dx(t) = \underbrace{f(x)dt}_{\text{Drift}} + \underbrace{g(x)dW(t)}_{\text{Diffusion}}, \quad (6)$$

where W is a Wiener process.

SDEs are defined as the continuous-time limit of dynamics on a discrete time lattice with spacing h . For additive noise where $g(x) = \text{const}$, different discretization conventions for the drift term still lead to different path-integral representations.

Both conventions can be treated simultaneously by introducing $\alpha \in [0, 1]$:

$$x_l - x_{l-1} = f(\alpha x_l + (1 - \alpha)x_{l-1})h + a\delta_{l-1} + \xi_l \quad (7)$$

where ξ_l : Gaussian noise ($\xi_l \sim \mathcal{N}(0, Dh)$), $a\delta_{l-1}$: Initial state $x_0 = a$.

Discretization: Ito vs. Stratonovich

Both conventions are unified by the interpolation parameter $\alpha \in [0, 1]$:

$$x_l - x_{l-1} = f(\alpha x_l + (1 - \alpha)x_{l-1})h + a\delta_{l-1} + \xi_l \quad (8)$$

Property	Ito ($\alpha = 0$)	Stratonovich ($\alpha = \frac{1}{2}$)
Evaluation	Past	Midpoint
Chain Rule	Ito's Lemma	Standard Calculus
Jacobian in Action	Trivial ($= 1$)	Adds $-\frac{1}{2}f'(x)$ term
Eq-Time $\chi(t, t)$	0 (Loops vanish)	1/2 (Loops exist)

The Markov Property

The joint density $p(x_1, \dots, x_M | a)$ depends only on the preceding time step. This step-by-step factorization ensures the total path probability measure is rigorously normalized to 1.

For Gaussian white noise with variance $\langle \xi_i \xi_j \rangle = \delta_{ij} Dh$, taking the continuous limit $M \rightarrow \infty$ and $h \rightarrow 0$ yields the path integral:

$$p[x|x(0+) = a] \mathcal{D}x := \exp \left(-\frac{1}{2D} \int_0^T (\partial_t x - f(x) - a\delta(t))^2 - \alpha f'(x) dt \right) \mathcal{D}x \quad (9)$$

Drawback

The action contains a squared term. This complicates analytical treatments significantly, particularly when attempting to calculate quenched averages (淬火平均) in disordered systems.

7.3 MSRDJ Path Integral

To remove the troublesome squared term $(\dot{x} - f(x))^2$ from the Onsager-Machlup action, the MSRDJ formalism introduces an auxiliary **response field (响应场)** $\tilde{x}$.

Instead of integrating out the noise early, we start with the exact delta constraint of the SDE: $\delta(\partial_t x - f(x) - \xi) = 0$. We represent this Dirac δ -functional via its Fourier representation:

$$\delta(\dots) \propto \int \mathcal{D}\tilde{x} \exp \left(\int dt \tilde{x}(t) [\partial_t x(t) - f(x(t)) - \xi(t)] \right) \quad (10)$$

7.3 MSRDJ Path Integral

To remove the troublesome squared term $(\dot{x} - f(x))^2$ from the Onsager-Machlup action, the MSRDJ formalism introduces an auxiliary **response field (响应场)** $\tilde{x}$.

Instead of integrating out the noise early, we start with the exact delta constraint of the SDE: $\delta(\partial_t x - f(x) - \xi) = 0$. We represent this Dirac δ -functional via its Fourier representation:

$$\delta(\dots) \propto \int \mathcal{D}\tilde{x} \exp \left(\int dt \tilde{x}(t) [\partial_t x(t) - f(x(t)) - \xi(t)] \right) \quad (10)$$

Averaging over the noise ξ now only affects the $e^{-\tilde{x}\xi}$ term. By definition of the cumulant-generating function W_ξ :

$$\langle e^{-\int \tilde{x}\xi dt} \rangle_\xi = \exp(W_\xi[-\tilde{x}]) \xrightarrow{\text{Gaussian Noise}} \exp \left(\int dt \frac{D}{2} \tilde{x}^2(t) \right) \quad (11)$$

Result: The action is now **strictly linear** in $\partial_t x$ and $f(x)$, moving the non-linearity to a mathematically manageable position!

7.4 Moment-Generating Functional

By adding a source term $j(t)$, we define the generating functional $Z[j]$ in a compact vectorial notation (using inner product $x^T y := \int x(t)y(t)dt$):

$$Z[j] = \int \mathcal{D}x \int \mathcal{D}_{2\pi i} \tilde{x} \exp \left(\tilde{x}^T (\partial_t x - f(x) - a\delta(t)) + j^T x - \alpha \mathbf{1}^T f'(x) + W_\xi[-\tilde{x}] \right) \quad (12)$$

Moments of the path are simply extracted via functional derivatives:

$$\langle x(t) \dots x(s) \rangle = \frac{\delta^n}{\delta j(t) \dots \delta j(s)} Z[j] \Big|_{j=0}$$

7.5 Response Function (响应函数)

To compute the system's response to an external perturbation $\tilde{j}(t)$, we add $\tilde{j}(t-)\tilde{x}(t)$ to the action. $\tilde{j}$ couples to $\tilde{x}$, validating $\tilde{x}$ as the "response field".

The linear response function $\chi(t, s)$ is given by:

$$\chi(t, s) := \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \left(\langle x(t) \rangle_{\tilde{j} = -\epsilon \delta(\cdot - s)} - \langle x(t) \rangle_{\tilde{j} = 0} \right) = -\langle x(t) \tilde{x}(s) \rangle \quad (13)$$

7.5 Response Function (响应函数)

To compute the system's response to an external perturbation $\tilde{j}(t)$, we add $\tilde{j}(t-)\tilde{x}(t)$ to the action. $\tilde{j}$ couples to $\tilde{x}$, validating $\tilde{x}$ as the "response field".

The linear response function $\chi(t, s)$ is given by:

$$\chi(t, s) := \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \left(\langle x(t) \rangle_{\tilde{j} = -\epsilon \delta(\cdot - s)} - \langle x(t) \rangle_{\tilde{j} = 0} \right) = -\langle x(t) \tilde{x}(s) \rangle \quad (13)$$

Equal-Time Response Advantages:

- In the **Ito** convention ($\alpha = 0$), $\chi(s, s) = 0$. This is highly advantageous because it rigorously leads to vanishing loop contributions in Feynman diagrams for equal time points.

Ornstein-Uhlenbeck Process: Free Gaussian Theory

8.1 Definition

This section identifies a Gaussian solvable system. The Ornstein-Uhlenbeck (OU) process serves as the **free Gaussian theory (自由高斯理论)** because its action is bilinear in the fields x and $\tilde{x}$.

The SDE for the OU process is:

$$dx = mxdt + dW, \quad \langle dW_i(t)dW_j(s) \rangle = D_{ij}\delta_{t,s}dt \quad (14)$$

where $m \in \mathbb{R}^{N \times N}$. Generalizing the MSRDJ formalism gives the action:

$$S[x, \tilde{x}] = \int \tilde{x}^T(t) (\partial_t - m) x(t) + \tilde{x}(t)^T \frac{D}{2} \tilde{x}(t) dt \quad (15)$$

We can write the action in a more symmetric form by introducing the compound field $y(t) = \begin{pmatrix} x(t) \\ \tilde{x}(t) \end{pmatrix}$:

$$S[y] = -\frac{1}{2} \iint y^T(t) A(t, s) y(s) dt ds$$

where the operator $A(t, s)$ is explicitly defined as:

$$A(t, s) = \begin{pmatrix} 0 & \partial_t + m^T \\ -\partial_t + m & -D \end{pmatrix} \delta(t - s)$$

Note: The minus sign in the upper right entry follows from integration by parts acting on $\tilde{x}^T(t) \partial_t x(t)$, assuming vanishing boundary terms.

8.2 Propagators in Time Domain

The moment-generating functional takes the Gaussian form $Z[\bar{j}] = \exp(\frac{1}{2}\bar{j}^T \Delta \bar{j})$. We determine the propagators (Green's functions) Δ as the inverse of A :

$$\Delta = A^{-1}, \quad \text{i.e.,} \quad \int A(s, t) \Delta(t, u) dt = \text{diag}(\delta(s - u)) \quad (16)$$

8.2 Propagators in Time Domain

The moment-generating functional takes the Gaussian form $Z[\bar{j}] = \exp(\frac{1}{2}\bar{j}^T \Delta \bar{j})$. We determine the propagators (Green's functions) Δ as the inverse of A :

$$\Delta = A^{-1}, \quad \text{i.e.,} \quad \int A(s, t) \Delta(t, u) dt = \text{diag}(\delta(s - u)) \quad (16)$$

The matrix elements represent the correlation and response functions:

$$\Delta(t, u) = \begin{pmatrix} \Delta_{xx}(t, u) & \Delta_{x\tilde{x}}(t, u) \\ \Delta_{\tilde{x}x}(t, u) & \Delta_{\tilde{x}\tilde{x}}(t, u) \end{pmatrix} = \begin{pmatrix} \langle x(t)x(u) \rangle & \langle x(t)\tilde{x}(u) \rangle \\ \langle \tilde{x}(t)x(u) \rangle & \langle \tilde{x}(t)\tilde{x}(u) \rangle \end{pmatrix}$$

8.3 Propagators in Fourier Domain

Because the OU process is time-translation invariant, the operator $A(\partial_t)$ becomes purely algebraic in the Fourier domain ($\partial_t \rightarrow -i\omega$). The action matrix is:

$$A(\omega) = \begin{pmatrix} 0 & -i\omega + m \\ i\omega + m & -D \end{pmatrix} \quad (17)$$

8.3 Propagators in Fourier Domain

Because the OU process is time-translation invariant, the operator $A(\partial_t)$ becomes purely algebraic in the Fourier domain ($\partial_t \rightarrow -i\omega$). The action matrix is:

$$A(\omega) = \begin{pmatrix} 0 & -i\omega + m \\ i\omega + m & -D \end{pmatrix} \quad (17)$$

The propagator matrix $\Delta(\omega)$ is simply the inverse $A^{-1}(\omega)$:

$$\Delta(\omega) = \frac{1}{\omega^2 + m^2} \begin{pmatrix} D & i\omega - m \\ -i\omega - m & 0 \end{pmatrix} \quad (18)$$

Immediate Insights:

- $\Delta_{\tilde{x}\tilde{x}}(\omega) = 0$ (Response fields never correlate with themselves).
- $\Delta_{xx}(\omega) = \frac{D}{\omega^2 + m^2}$ (The classic Lorentzian power spectrum).
- $\Delta_{x\tilde{x}}(\omega) = \frac{1}{-i\omega - m}$ (Inverse mapping back to time domain via residue theorem yields the causal step function $H(t)$).

Similarly, the response function relies on the Heaviside step function $H(t - s)$ to preserve causality (因果性):

$$\Delta_{x\tilde{x}}(t, s) = \langle x(t)\tilde{x}(s) \rangle = -H(t - s)e^{m(t-s)} \quad (19)$$

Diagrammatic Formulation (图解表述)



- **Response function:** Represented by a straight line with a single arrow indicating time flow. It is strictly causal ($\langle x(t)\tilde{x}(s) \rangle = 0$ if $t \leq s$).
- **Correlation function:** Represented by a line with two outgoing arrows.
- **Vanishing moments:** $\langle \tilde{x}(t)\tilde{x}(s) \rangle = 0$.

Perturbation Theory for SDEs

Abstract

This chapter introduces perturbation theory to SDEs. Non-linearities correspond to the non-Gaussian part of the action requiring perturbative corrections.

9.1 Vanishing Moments of Response Fields

Consider an SDE with a non-linearity $f(x)$ and a source $\tilde{j}$:

$$dx(t) = f(x(t))dt - \tilde{j}(t)dt + dW(t) \quad (20)$$

Exact Property

Because Z belongs to a properly normalized density for any right-hand side, we have exactly $Z[0, \tilde{j}] \equiv 1 \quad \forall \tilde{j}$. Consequently, all functional derivatives with respect to $\tilde{j}$ must vanish:

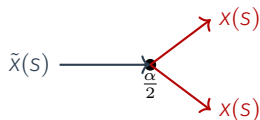
$$\langle \tilde{x}(t_1) \cdots \tilde{x}(t_n) \rangle \equiv 0 \quad \forall n > 0$$

This holds independent of any perturbation expansion.

9.2 Feynman Rules for SDEs in Time Domain

Consider the non-linear SDE: $dx + xdt - \frac{\alpha}{2}x^2dt = dW$. The Taylor expansion of the action introduces an **interaction vertex** (相互作用顶点):

$$V[x, \tilde{x}] = \frac{\alpha}{2} \int ds \tilde{x}(s) x^2(s)$$



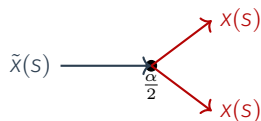
To compute the first-order correction to the mean $\langle x(t) \rangle$, we pull down one vertex V :

$$\langle x(t) \rangle_{1st} = \frac{\alpha}{2} \int ds \underbrace{\langle x(t) \tilde{x}(s) x(s) x(s) \rangle_0}_{\text{Wick Contraction}}$$

9.2 Feynman Rules for SDEs in Time Domain

Consider the non-linear SDE: $dx + xdt - \frac{\alpha}{2}x^2dt = dW$. The Taylor expansion of the action introduces an **interaction vertex (相互作用顶点)**:

$$V[x, \tilde{x}] = \frac{\alpha}{2} \int ds \tilde{x}(s) x^2(s)$$



To compute the first-order correction to the mean $\langle x(t) \rangle$, we pull down one vertex V :

$$\langle x(t) \rangle_{1st} = \frac{\alpha}{2} \int ds \underbrace{\langle x(t) \tilde{x}(s) x(s) x(s) \rangle_0}_{\text{Wick Contraction}}$$

By Wick's theorem, we must pair the fields. Since $\langle \tilde{x}\tilde{x} \rangle = 0$, the only valid non-zero pairing is:

$$\langle x(t) \rangle_{1st} = \frac{\alpha}{2} \int ds \underbrace{\langle x(t) \tilde{x}(s) \rangle_0}_{\text{Straight Line}} \underbrace{\langle x(s) x(s) \rangle_0}_{\text{Loop}} = \frac{\alpha D}{4|m|} \quad (21)$$

Diagrammatically, the observation point t connects to the vertex at s (response $\Delta_{x\tilde{x}}$), and the remaining two $x(s)$ legs contract with each other to form a loop (Δ_{xx}).

9.2 Feynman Rules in Frequency Domain

For time-translation invariant problems, the Fourier domain is simpler. The interaction vertex enforces **frequency conservation** (frequencies sum to zero):

$$V_{\text{freq}} \propto \iiint d\omega_1 d\omega_2 d\omega_3 \delta(\omega_1 + \omega_2 + \omega_3) \tilde{X}(\omega_1) X(\omega_2) X(\omega_3) \quad (22)$$

9.2 Feynman Rules in Frequency Domain

For time-translation invariant problems, the Fourier domain is simpler. The interaction vertex enforces **frequency conservation** (frequencies sum to zero):

$$V_{\text{freq}} \propto \iiint d\omega_1 d\omega_2 d\omega_3 \delta(\omega_1 + \omega_2 + \omega_3) \tilde{X}(\omega_1) X(\omega_2) X(\omega_3) \quad (22)$$

The first-order correction evaluates an integral over the internal frequency:

$$\begin{aligned} \langle X(0) \rangle_{1\text{st}} &= \frac{\alpha}{2!} \int \frac{d\omega_1}{2\pi} (-i\omega_1 - 1)^{-1} D(i\omega_1 - 1)^{-1} \\ &= \frac{\alpha D}{4} \end{aligned} \quad (23)$$

This correction only affects the stationary expectation value at $\omega = 0$.

9.3 Diagrams with Multiple External Legs

- To correct the variance (2nd cumulant), take the second functional derivative of Z , which isolates symmetric **1PI (One-Particle Irreducible)** contributions.
- Multiple poles in the frequency domain are solved via Cauchy's differential formula.

9.3 Diagrams with Multiple External Legs

- To correct the variance (2nd cumulant), take the second functional derivative of Z , which isolates symmetric **1PI (One-Particle Irreducible)** contributions.
- Multiple poles in the frequency domain are solved via Cauchy's differential formula.

9.5 Vanishing Response Loops

Interaction vertices couple fields at equal times.

- **Ito Convention:** $\langle x(t)\tilde{x}(s) \rangle = 0$ for $t \leq s$. Any diagram forming a closed response loop must travel forward in time, hence equal-time loops rigorously vanish.
- **Stratonovich Convention:** $\langle x(t)\tilde{x}(t) \rangle = 1/2$. Closed loops exactly cancel with the functional determinant $-\frac{1}{2}\mathcal{F}'(x)$ arising from the discretization limit.

Dynamic Mean-Field Theory for Random Networks

Sompolinsky-Crisanti-Sommers (1988) Model

We consider a large network of continuously interacting neurons. The state $x_i(t)$ of neuron i evolves according to the stochastic rate equation:

$$dx_i(t) + x_i(t)dt = \sum_{j=1}^N J_{ij}\phi(x_j(t))dt + dW_i(t) \quad (24)$$

Quenched Disorder (淬火无序): The synaptic weights J_{ij} are frozen (constant in time) and randomly drawn from a Gaussian distribution:

$$J_{ij} \sim \mathcal{N}\left(0, \frac{g^2}{N}\right) \quad \text{for } i \neq j \quad (25)$$

where g represents the global coupling strength, and $\phi(x) = \tanh(x)$ is a non-linear gain function.

Goal: Solve the macroscopic dynamics of this system as $N \rightarrow \infty$ to rigorously study the transition to chaos.

The Disorder Average via MSRDJ

To find the macroscopic behavior generic to the ensemble, we construct the MSRDJ generating functional and average it over the quenched disorder J_{ij} . Because J_{ij} appears linearly in the action, integrating over its Gaussian distribution yields a four-field interaction term:

$$\left\langle \exp \left(-J_{ij} \int \tilde{x}_i \phi(x_j) dt \right) \right\rangle_J = \exp \left(\frac{g^2}{2N} \left(\int \tilde{x}_i \phi(x_j) dt \right)^2 \right) \quad (26)$$

We decouple this using auxiliary macroscopic fields Q_1 and Q_2 . In the large N limit, we apply a **saddle-point approximation**:

- $Q_1^*(s, t) = g^2 \langle \phi(x(s)) \phi(x(t)) \rangle = g^2 C_{\phi\phi}(s, t).$
- $Q_2^*(s, t) = 0$ (due to vanishing moments of response fields).

Dynamic Mean-Field Theory (DMFT)

The saddle-point approximation reduces the highly coupled N -neuron network into N identical, non-interacting units. Each follows a single effective equation:

$$(\partial_t + 1)x(t) = \eta(t) \quad (27)$$

Self-Consistency Condition (自洽条件)

The noise $\eta(t) = \zeta(t) + \dot{W}(t)$ is generated by the rest of the network! It is a Gaussian process with an autocorrelation determined by the system itself[cite: 1]:

$$\langle \eta(t)\eta(s) \rangle = g^2 C_{\phi\phi}(t, s) + D\delta(t - s) \quad (28)$$

Insight: The network acts as a self-generated, self-consistent noise bath for each individual neuron[cite: 1].

Autocovariance as a Particle in a Potential[cite: 1]

The stationary statistics are completely determined by the autocovariance $c(t-s) = \langle x(t)x(s) \rangle$ [cite: 1]. It obeys a differential equation mathematically equivalent to a particle in a potential $V(c; c_0)$ [cite: 1]:

$$\partial_\tau^2 c(\tau) = -V'(c(\tau); c_0) - D\delta(\tau) \quad (29)$$

- $V(c; c_0) = -\frac{1}{2}c^2 + g^2 f_\Phi(c(\tau), c_0) - \text{const}$ [cite: 1].
- The initial value $c_0 = c(0)$ is fixed by "energy conservation":
 $V(c_0; c_0) + E_{\text{kin}}^{(0)} = 0$ [cite: 1].

For $g < 1$, the parabolic part dominates, bending the potential downwards and leading to a silent state[cite: 1]. For $g > 1$, the curvature at the origin changes, allowing non-trivial chaotic solutions[cite: 1].

The Edge of Chaos[cite: 1]

To assess stability, we consider two identical replicas of the system and measure their mean-squared distance $d(t)$ [cite: 1]. The exponential growth rate maps to the ground state E_0 of a **Schrödinger equation**[cite: 1]:

$$(-\partial_\tau^2 - V''(c(\tau); c_0))\psi(\tau) = E_0\psi(\tau) \quad (30)$$

The maximum Lyapunov exponent is $\lambda_{\max} = -1 + \sqrt{1 - E_0}$ [cite: 1]. A necessary condition for the transition to the chaotic state is[cite: 1]:

$$1 \leq g^2 \langle \phi'(x)^2 \rangle \quad (31)$$

Conclusion: SFT proves that macroscopic chaos emerges purely from the variance of random synaptic weights g^2 , separated by a sharp phase transition[cite: 1].