

Chapter 5

Numerical Tests

Our theoretical results and reconstruction method are tested using numerical data from simulating dynamical systems with different networks and non-linear dynamics. Besides, we also study some additional systems for which our theoretical derivation for Eqs. (3.10) and (3.20) does not directly apply in order to explore the general applicability of our method. This chapter introduces the numerical tests for these purposes.

5.1 Networks studied

Our reconstruction method is applicable to networks with general structure. We have tested the method using five different networks, including bidirectional and directed networks with coupling strength in different distributions. (i) The first network studied is a 100-node directed weighted random network (DWR1) [52] in which a link has a fixed connection probability of 0.2 to be placed from one node to another, and coupling strength of links is Gaussian distributed with mean 10 and standard deviation 2. (ii) The second network DWR1s has the same adjacency matrix as DWR1, and has Gaussian distributed coupling strength with mean and

standard derivation equal to 10. DWR1s network involves both excitatory ($g_{ij} > 0$) and inhibitory ($g_{ij} < 0$) links. (iii) Another random network DWR2 has a connection probability of 0.2 and contains only unidirectional links with $A_{ij} = 0$ for all $i > j$. We also study scale-free networks with node degree in power-law distribution. (iv) WSF network is a 1000-node bidirectional scale-free network with power-law distributed node degree and coupling strength, while (v) DWSF network is a directed weighted scale-free network constructed by converting some bidirectional links of WSF into directed links while keeping out-degree in power-law distribution. Table 5.1 shows the details of link and coupling strength of all networks studied.

Network	N	N_B	N_U	N_L	ρ	g_{ij}
DWR1	100	186	1678	2050	0.207	$N(10, 2)$
DWR1s	100	186	1678	2050	0.207	$N(10, 10)$
DWR2	100	0	1035	1035	0.105	$N(10, 2)$
WSF	1000	4985	0	9970	0.00998	$P(g_{ij}) \sim g_{ij}^{-6.6}$
DWSF	1000	3120	3730	9970	0.00998	$P(g_{ij}) \sim g_{ij}^{-6.6}$

Table 5.1: List of network studied. N_B and N_U are the number of bidirectionally-linked and unidirectionally-linked pairs of nodes respectively. N_L is total number of links given by $2N_B + N_U$ and ρ is link density given by $N_L/[N(N-1)]$.

5.1.1 Algorithm of generating WSF

We follow the generative model in Ref. [53] with parameters $(m_0, m, \omega_0, \delta) = (5, 5, 10, 2)$ to construct the WSF network. This algorithm exhibits the preferential attachment dynamics in growing network. Procedure is as follows:

1. It is started with m_0 nodes which are connected to each other with coupling strength ω_0 .
2. A new node is added at each step and it makes m new links with coupling

strength ω_0 connecting to existing nodes. The connection probability to one existing node i is proportional to its node strength $s_i = \sum_k g_{ik} A_{ik}$.

3. Whenever an existing node i is connected by a new link, then the coupling strength from each of its existing links are changed from g_{ik} to $g_{ik}(1 + \delta/s_i)$.

5.1.2 Algorithm of generating DWSF

DWSF network is generated from bidirectional WSF network by converting $A_{ij} = 1$ for $i > j$ to be zero with a chance of 0.5. Whenever a conversion is done, a $A_{kj} = 0$ for $k < j$ is randomly selected and it is set to be 1 with coupling strength assigned to be same as original g_{ij} . In this way, the power-law distributions of out-degree and coupling strength are same as ones in WSF network.

5.2 Dynamics studied

We consider intrinsic dynamics f_i given by non-linear logistic function with parameter r_i

$$f_i(x) = r_i x(1 - x) \quad (5.1)$$

and two types of coupling functions as shown below.

$$h^{\text{diff}}(x, y) = y - x \quad (5.2)$$

$$h^{\text{syn}}(x, y) = (1/\beta_1)\{1 + \tanh[\beta_2(y - y_0)]\} \quad (5.3)$$

The linear diffusive coupling function $h^{\text{diff}}(x, y)$ is a common model for gap junction coupling [54], while $h^{\text{syn}}(x, y)$ is a generalisation [39] of a model for synaptic coupling [55] with parameters β_1 , β_2 and y_0 chosen which make X_i close to y_0 .

These systems studied have the noise-free steady state and the dynamical equation can be linearised about the noise-free steady state as described by Eq. (3.4). Also the dynamics of each node i ($1 \leq i \leq N$) is described by the state variable x_i . To explore the applicability of our method to more general dynamics, we study three types of additional cases which cannot be linearised, or do not have noise-free steady state or its state variable is higher-dimensional. The first type is $f_i = 0$ with $h^{\text{cubic}}(x, y) = (y - x)^3$ which does not have linear approximation. The second type of dynamics studied is **FitzHugh-Nagumo (FHN) dynamics** [56]. Each node has a **two-dimensional state variable** $(x_i(t), y_i(t))$ and dynamics is governed by

$$\dot{x}_i = \frac{1}{\epsilon}(x_i - x_i^3/3 - y_i) + \sum_{j \neq i} g_{ij} A_{ij} h(x_i, x_j) + \eta_i \quad (5.4)$$

$$\dot{y}_i = x_i + \alpha \quad (5.5)$$

with parameters $\epsilon = 0.01$ and α . This models the neuron spike generation of membrane potential $x_i(t)$. In our studied cases, FHN dynamics approaches a stable steady state when $\alpha > 1$. When $\alpha < 1$, its noise-free solution is time-dependent and Fig. 5.1(a) shows one studied case of DWR1 network with FHN dynamics $\alpha = 0.95$ and h^{diff} having periodic spiking. The third type of dynamics studied is non-linear Rössler dynamics [57] where the nodes have a three-dimensional state variable $(x_i(t), y_i(t), z_i(t))$ and dynamics is governed by

$$\dot{x}_i = -y_i - z_i + \sum_{j \neq i} g_{ij} A_{ij} h(x_i, x_j) + \eta_i \quad (5.6)$$

$$\dot{y}_i = x_i + a y_i + \sum_{j \neq i} g_{ij} A_{ij} h(y_i, y_j) \quad (5.7)$$

$$\dot{z}_i = b + z_i(x_i - c) + \sum_{j \neq i} g_{ij} A_{ij} h(z_i, z_j) \quad (5.8)$$

where $a = b = 0.2$ and $c = 9$. The noise-free solution in this case is time-dependent (see Fig. 5.1(b)).

For these additional cases where the derivation for Eqs. (3.10) and (3.20) does not apply, we study numerically how off-diagonal elements M_{ij} relate to A_{ij} . In the cases with FHN and Rössler dynamics, only time series of $x_i(t)$, whose dynamics is perturbed by noise directly, are used to evaluate M_{ij} .

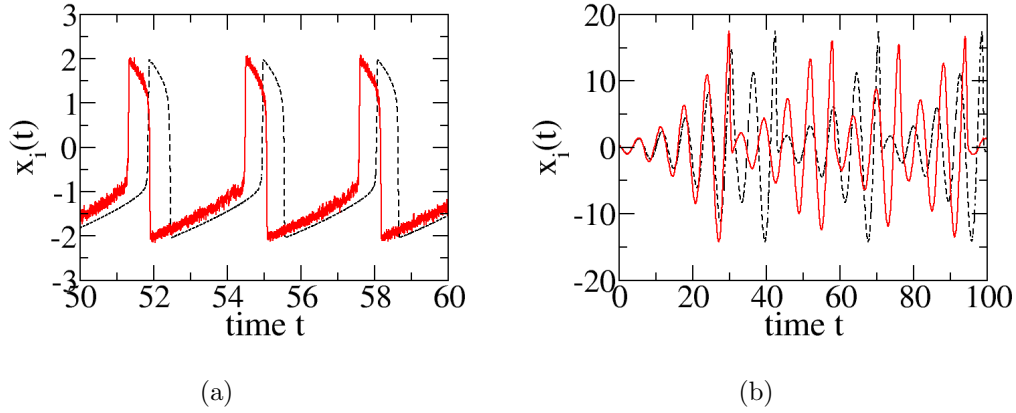


Figure 5.1: Time series plot (red line) of one node for DWR1 network with (a) FHN dynamics $\alpha = 0.95$ and h^{diff} driven by noise with $\tau_n = 10^{-4}$ and $D_{ij} = \delta_{ij}$; (b) Rössler dynamics and h^{diff} driven by noise with $\tau_n = 10^{-1}$ and $D_{ij} = \delta_{ij}$. Noise-free time series is indicated by black dashed line.

5.3 Simulation scheme

We integrate the equations of motion using the Euler-Maruyama method [58], which is applied in the following form:

$$dx(t) = A(x, t)dt + B(x, t)dW(t), \quad x(t_0) = c, \quad (5.9)$$

where A and B are functions of x and t , t_0 is initial time with initial state c , $W(t)$ represents the Wiener process whose time derivative is a white noise [44]. Time series

of $x_i(t)$ is approximated by $x_i(n)$, $n = 1, 2, \dots$, with corresponding time step t_n and the integration time step $\Delta t = t_{n+1} - t_n$. When the dynamical system subject to a correlated noise is considered (Eqs. (3.1) and (3.3)), $x_i(n)$ is evaluated by

$$x_i(n+1) = x_i(n) + \left[f_i(x_i(n)) + \sum_{j \neq i} g_{ij} A_{ij} h(x_i(n), x_j(n)) + \eta_i(n) \right] \Delta t \quad (5.10)$$

$$\eta_i(n+1) = -\frac{\Delta t}{\tau_n} \eta_i(n) + \frac{\sqrt{2\Delta t}}{\tau_n} \sum_{j=1}^N L_{ij} z_j(n), \quad (5.11)$$

where $z_i(n)$ is a random number chosen from a standard normal distribution $N(0, 1)$, and L_{ij} is an element of a lower triangular matrix \mathbf{L} satisfying $\mathbf{D} = \mathbf{L}\mathbf{L}^T$. This Cholesky decomposition is guaranteed as the noise strength \mathbf{D} is a symmetric and positive semidefinite matrix [59]. When the dynamical system is subject to a white noise with noise strength $2\mathbf{D}$, then $x_i(n)$ is evaluated by

$$\begin{aligned} x_i(n+1) = x_i(n) + & \left[f_i(x_i(n)) + \sum_{j \neq i} g_{ij} A_{ij} h(x_i(n), x_j(n)) \right] \Delta t \\ & + \sqrt{2\Delta t} \sum_{j=1}^N L_{ij} z_j(n). \end{aligned} \quad (5.12)$$

This integration algorithm has a strong order of convergence 1/2 and weak order of convergence 1 [58].

Time series measurements of dynamical data is obtained to evaluate \mathbf{K}_0 , \mathbf{K}_τ , $\mathbf{K}_{2\tau}$ and $\mathbf{K}_{3\tau}$ with a time average approximating the ensemble average. For stationary dynamics, it is a good approximation. Unless otherwise specified, we collect time series data at a sampling interval $\delta t = 5 \times 10^{-4}$ and total number of data points $N_{data} = 2 \times 10^6$, and take $\tau = \delta t$. Table 5.2 shows a list of all cases studied in this thesis.

Case	Network	Dynamics	$h(x, y)$	σ_i^{2*}	τ_n
1	DWR1	logistic $r_i = 10$	h^{diff}	1	0
2	DWR1	logistic $r_i = 10$	h^{diff}	$\sigma_i \in N(1, 0.2)$	0
3	DWR1	logistic $r_i \in U(1, 50)$	h^{diff}	1	0
4	DWR1	logistic $r_i = 10$	$h^{\text{syn}}, (\beta_1, \beta_2, y_0) = (2, 0.5, 4)$	1	0
5	DWR1s	logistic $r_i = 10$	h^{diff}	1	0
6	DWR2	logistic $r_i = 50$	h^{diff}	1	0
7	DWR2	logistic $r_i = 50$	$h^{\text{syn}}, (\beta_1, \beta_2, y_0) = (0.1, 0.5, 4)$	1	0
8	DWSF	logistic $r_i = 100$	h^{diff}	1	0
9	DWSF	logistic $r_i = 100$	$h^{\text{syn}}, (\beta_1, \beta_2, y_0) = (0.5, 0.5, 4)$	1	0
10	DWR1	logistic $r_i = 10$	h^{diff}	2	10^{-4}
11	DWSF	logistic $r_i = 100$	h^{diff}	2	5×10^{-5}
12	DWSF	logistic $r_i = 100$	h^{diff}	2	10^{-4}
13	WSF	logistic $r_i = 1$	h^{diff}	2	5×10^{-5}
14	WSF	logistic $r_i = 1$	h^{diff}	2	10^{-4}
15	DWR1	logistic $r_i = 10$	h^{diff}	2	10^{-3}
16	DWR1	logistic $r_i = 10$	h^{diff}	2	10^{-2}
17	DWR1	logistic $r_i = 10$	h^{diff}	2	10^{-1}
18	DWR1	logistic $r_i = 10$	h^{diff}	2	1
19	DWR1	logistic $r_i = 10$	h^{diff}	2	5
20	DWR1	logistic $r_i = 10$	h^{diff}	2	10
21	DWR1	logistic $r_i = 10$	h^{diff}	*non-diagonal \mathbf{D}	10^{-1}
22	DWR1	logistic $r_i = 10$	$h^{\text{syn}}, (\beta_1, \beta_2, y_0) = (2, 0.5, 4)$		10^{-1}
23	DWSF	logistic $r_i = 100$	h^{diff}		10^{-3}
24	DWSF	logistic $r_i = 100$	h^{diff}	2	10^{-2}
25	DWSF	logistic $r_i = 100$	h^{diff}	2	10^{-1}
26	DWSF	logistic $r_i = 100$	h^{diff}	2	1
27	DWSF	logistic $r_i = 100$	$h^{\text{syn}}, (\beta_1, \beta_2, y_0) = (0.5, 0.5, 4)$	2	10^{-1}
28	WSF	logistic $r_i = 1$	h^{diff}	2	10^{-2}
29	WSF	logistic $r_i = 1$	h^{diff}	2	1
30	DWR1	$f_i = 0$	h^{cubic}	1	0
31	DWR1	FHN $\alpha = 1.05$	h^{diff}	1	0
32	DWR1	FHN $\alpha = 2$	$h^{\text{syn}}, (\beta_1, \beta_2, y_0) = (0.1, 2, -1)$	10^{-2}	0
33	DWR1	Rössler	h^{diff}	1	0
34	DWR2	FHN $\alpha = 1.05$	h^{diff}	10^{-2}	0
35	DWR2	Rössler	h^{diff}	1	0
36	DWSF	FHN $\alpha = 1.05$	h^{diff}	10^{-2}	0
37	DWSF	Rössler	h^{diff}	1	0
38	DWR1	FHN $\alpha = 0.95$	h^{diff}	2	10^{-4}
39	DWR1	Rössler	h^{diff}	2	10^{-1}
40	DWR1	FHN $\alpha = 1.05$	h^{diff}	2	10^{-1}

Table 5.2: List of all cases studied. Cases 1-29 are cases for which our theoretical derivation of Eqs. (3.10) and (3.20) holds. Cases 30-40 are additional cases for which our theoretical derivation does not hold.

* $D_{ij} = (\sigma_i^2/2)\delta_{ij}$ except for case 21, where D_{ii} taken from a uniform distribution $U(5, 10)$ from 5 to 10 and D_{ij} for $i \neq j$ taken from a Gaussian distribution $N(0.1, 0.1)$.

Chapter 6

Reconstruction Results for Networks in the White Noise Limit

We start with the reconstruction of networks in the white noise limit which includes cases 1-14 and 30-38 as shown in Table 5.2 having $\Delta < \delta$.

6.1 Validity of Eq. (4.21)

Figures 6.1 and 6.2 show the comparison of off-diagonal and diagonal elements of \mathbf{M} and \mathbf{Q} respectively for cases that have white noise ($\tau_n = 0$) and our theoretical derivation for Eqs. (3.10) and (3.20) holds. These cases include inhomogeneous f_i and non-uniform noise strength. Data points in each plot scatter along the diagonal line $y = x$ which demonstrates a good agreement of M_{ij} with Q_{ij} .

The scatters are due to finite N_{data} taken. We study how $M_{ij} - Q_{ij}$ depends on N_{data} using case 1 as an example. Figure 6.3 shows a comparison of off-diagonal elements M_{ij} with Q_{ij} for different N_{data} . It can be seen that the extent of scatters decreases with N_{data} . The standard deviation of off-diagonal elements $M_{ij} - Q_{ij}$,

which is used to quantify the data scatters, is found to be approximately proportional to $1/\sqrt{N_{data}}$ as shown in Fig. 6.4.

We approximate ensemble average by time average. Therefore, we would like to check how good this approximation is. We study one case (i.e. case 1) and calculate by taking a time average as discussed above and by an ensemble average respectively. For the ensemble averaging, dynamical data at specific times t^* and $t^* + \tau$ are taken for 2×10^6 different realisations of the noise. For taking a time average, covariance matrices are evaluated using $N_{data} = 2 \times 10^6$ dynamical data from time series for one realisation of the noise. Figure 6.5 shows that there is not much difference on M_{ij} obtained from both methods and both agree with Q_{ij} well. This supports our approximation of the ensemble average by a time average.

For systems with non-linear dynamics, our theory requires that the fluctuations in the presence of noise are small enough for the linearisation to hold. Hence we also check the validity of Eq. (4.21) for different noise strength. Another case with dynamics same as case 1 but different noise strength with $\sigma_i = \sigma = 5$ is also studied. Figure 6.6 shows that M_{ij} obtained in this case also agrees with Q_{ij} well. Hence our method works well for different noise strength as long as the system exhibits stationary fluctuations about noise-free steady state.

In addition, M_{ij} agrees with Q_{ij} well for different values of τ (with $\tau = \delta t$ accordingly) as shown in Fig. 6.7. From our result Eq. (4.1): $\mathbf{K}_\tau = e^{\tau \mathbf{Q}} \mathbf{K}_0$, eigenvalues of $\mathbf{K}_\tau \mathbf{K}_0^{-1}$ tend to be closer to zero for larger τ value because of \mathbf{Q} having all eigenvalues with negative real parts. Hence to evaluate \mathbf{M} for a larger value of τ (e.g. $\tau = 2 \times 10^{-2}$), we need to use a larger $N_{data} = 2.5 \times 10^7$ (i.e. Fig. 6.8(a)) in order to have enough precision, otherwise one will get an inaccurate \mathbf{M} which deviates much from \mathbf{Q} when using $N_{data} = 2 \times 10^6$ as shown in Fig. 6.8(b), or even an erroneous complex \mathbf{M} numerically when using $N_{data} = 10^6$.

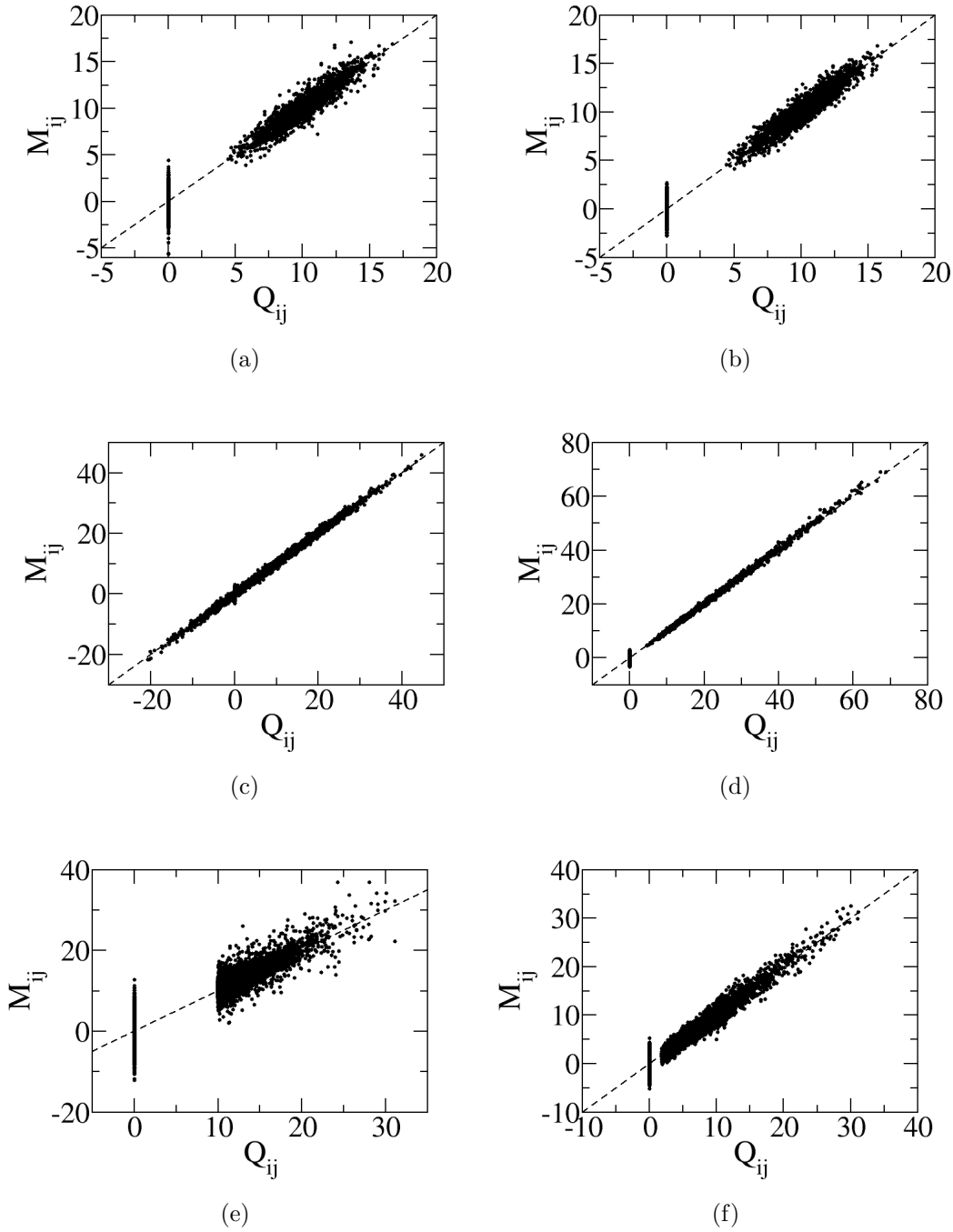


Figure 6.1: Plot of off-diagonal elements M_{ij} versus Q_{ij} for different cases with white noise ($\tau_n = 0$). (a) case 2; (b) case 3; (c) case 5; (d) case 7; (e) case 8; (f) case 9. Dashed line is $y = x$.

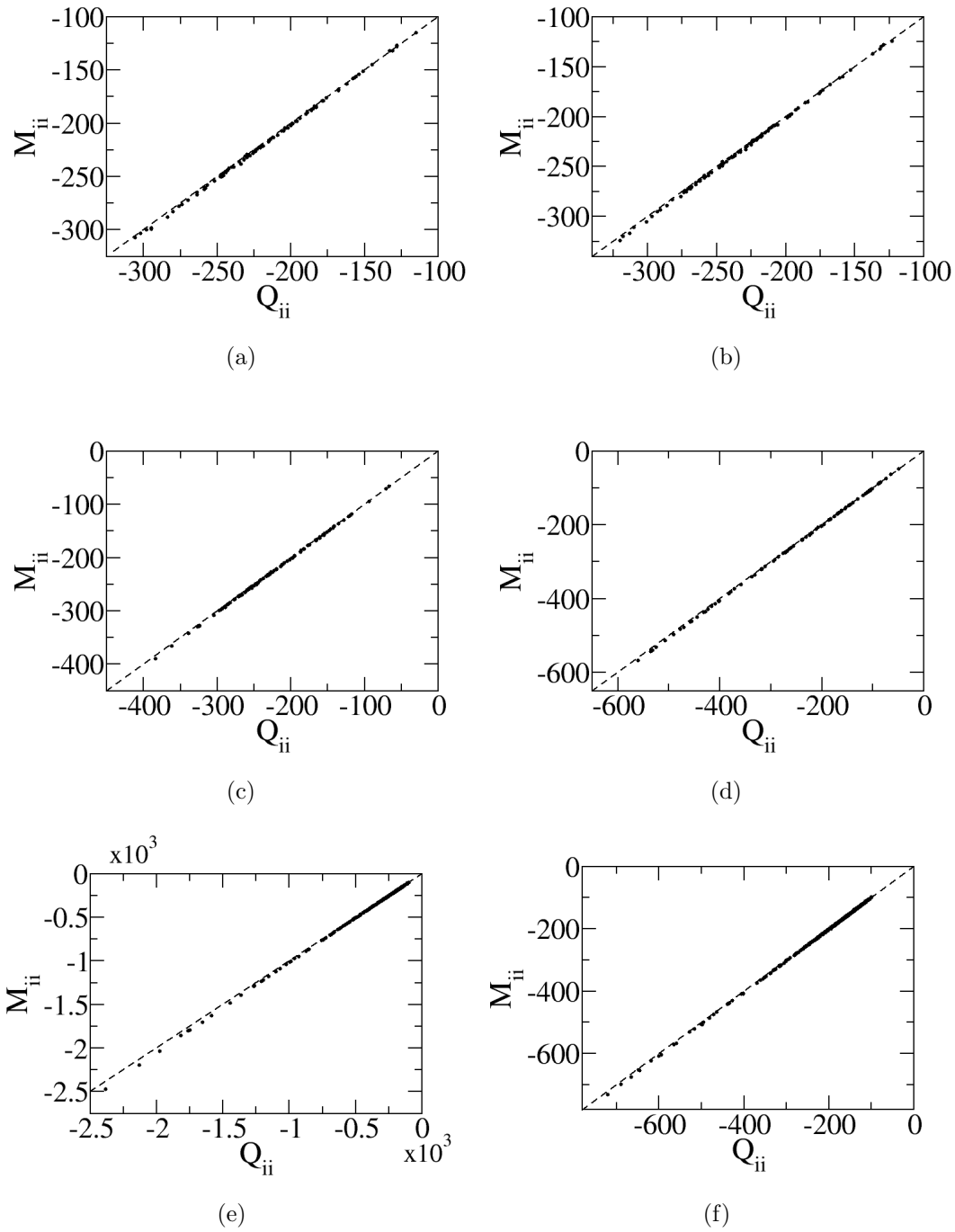


Figure 6.2: Plot of diagonal elements M_{ii} versus Q_{ii} for corresponding cases in Fig. 6.1.

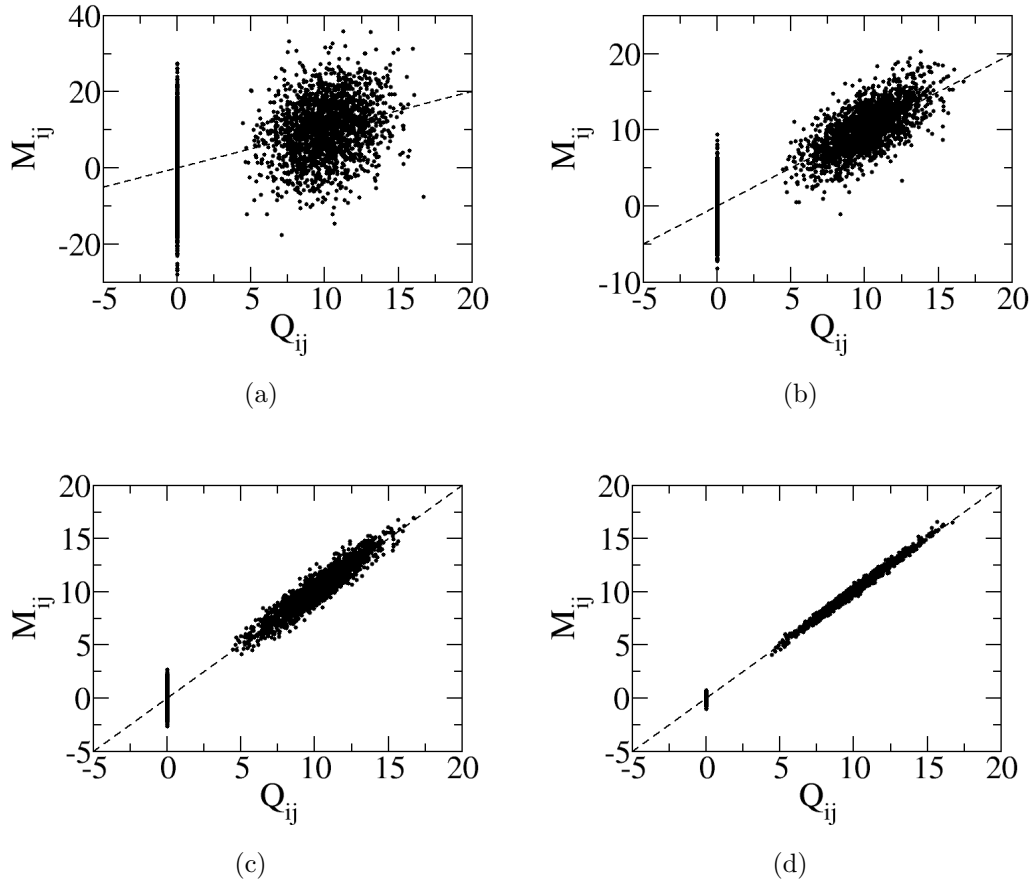


Figure 6.3: Plot of off-diagonal elements M_{ij} versus Q_{ij} for case 1 with $N_{data} =$ (a) 2×10^4 ; (b) 2×10^5 ; (c) 2×10^6 ; (d) 2×10^7 . Dashed line is $y = x$.

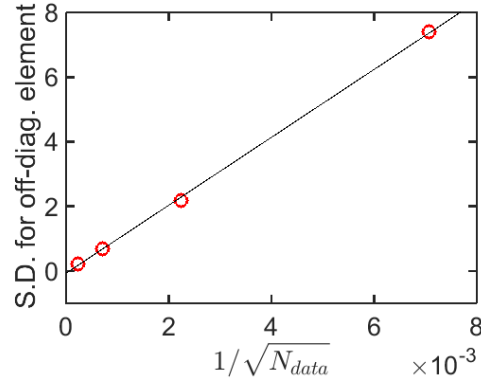


Figure 6.4: Plot of standard deviation of off-diagonal elements $M_{ij} - Q_{ij}$ obtained from each case in Fig. 6.3 versus corresponding $1/\sqrt{N_{data}}$. Black line represents a fitted straight line $y = 1.1 \times 10^3 x - 6.7 \times 10^{-2}$.

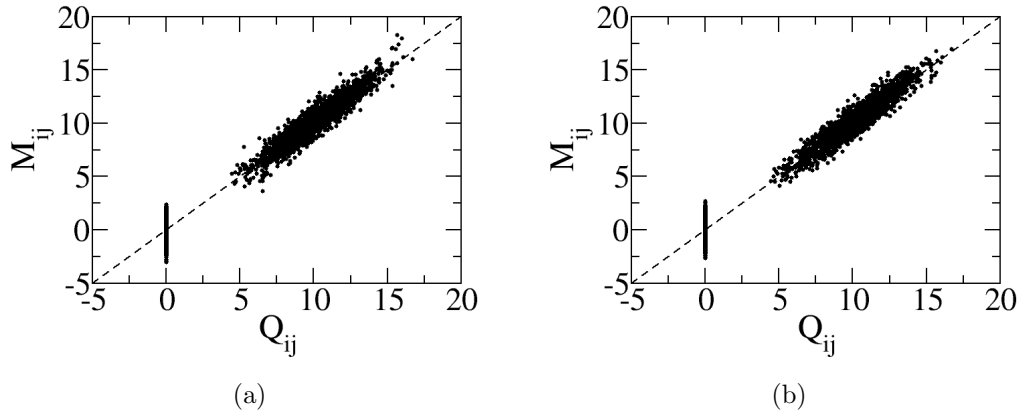


Figure 6.5: Plot of off-diagonal elements M_{ij} versus Q_{ij} obtained by (a) taking an ensemble average and (b) taking a time average respectively for case 1. Dashed line is $y = x$.

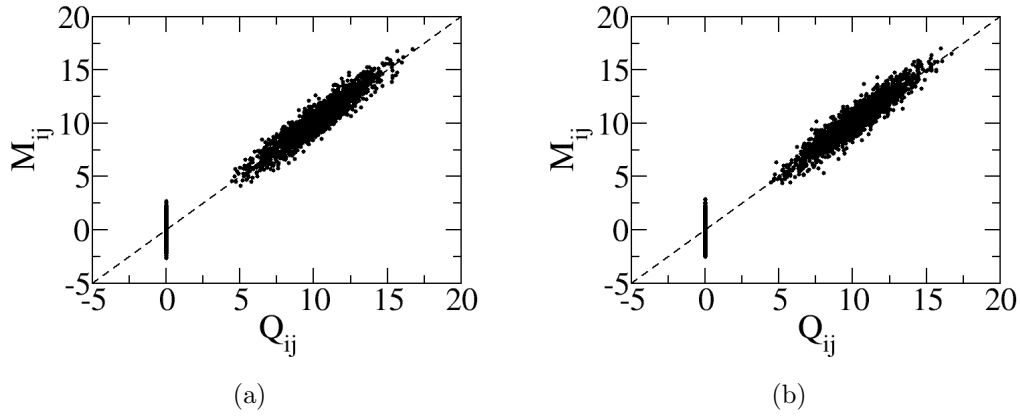


Figure 6.6: Plot of off-diagonal elements M_{ij} versus Q_{ij} for (a) case 1 with $\sigma = 1$ and (b) another case with dynamics same as case 1 but with $\sigma = 5$. Dashed line is $y = x$.

M_{ij} corresponding to $A_{ij} = 0$ is found to be approximately Gaussian distributed with mean around zero as demonstrated in Fig. 6.9 with one example from case 1 and another example from case 9. In the latter example which DWSF network has 1000 nodes, it is seen that M_{ij} corresponding to $A_{ij} = 0$ for different fixed j is in Gaussian distribution. As discussed in Section 4.4, this supports our clustering analysis using Gaussian mixture model. Also since we expect that there is a spread of M_{ij} corresponding for $A_{ij} = 0$ around zero, we only consider relative discrepancies between diagonal elements F_{ii} and H_{ii} when evaluating Δ in Eq. (4.19) as in the white noise limit, both \mathbf{F} and \mathbf{H} give $\tau\mathbf{Q}$.

After studying M_{ij} for cases with $\tau_n = 0$, now we compare \mathbf{M} with \mathbf{Q} for cases with small $\tau_n > 0$ such that $\Delta < \delta$ as shown in Fig. 6.10. Since τ_n is non-zero and these cases are taken as a white noise approximation, we expect that \mathbf{M} would deviate from \mathbf{Q} . Larger deviations between M_{ii} and Q_{ii} indeed are found but off-diagonal elements M_{ij} are approximately proportional to Q_{ij} with proportional constant close

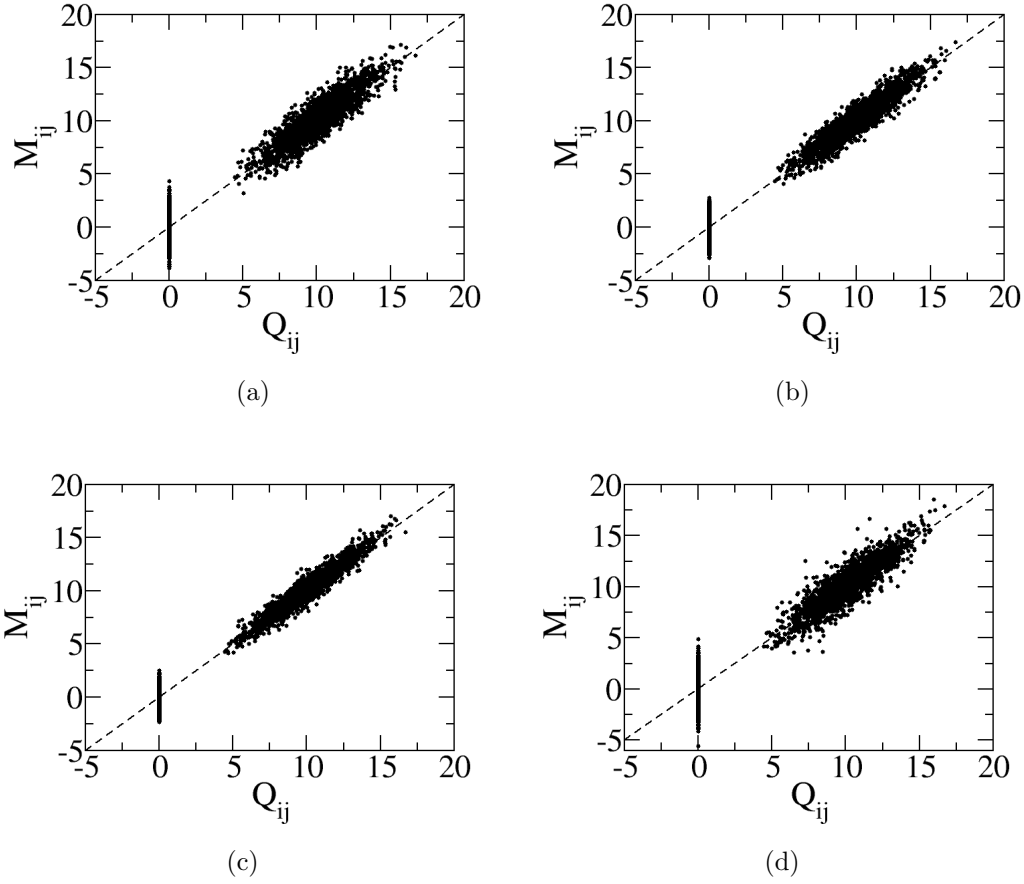


Figure 6.7: Plot of off-diagonal elements M_{ij} versus Q_{ij} for case 1 taking $N_{data} = 10^6$ and $\tau = \delta t$ equal to (a) 5×10^{-4} ; (b) 10^{-3} ; (c) 5×10^{-3} and (d) 10^{-2} . Dashed line is $y = x$.

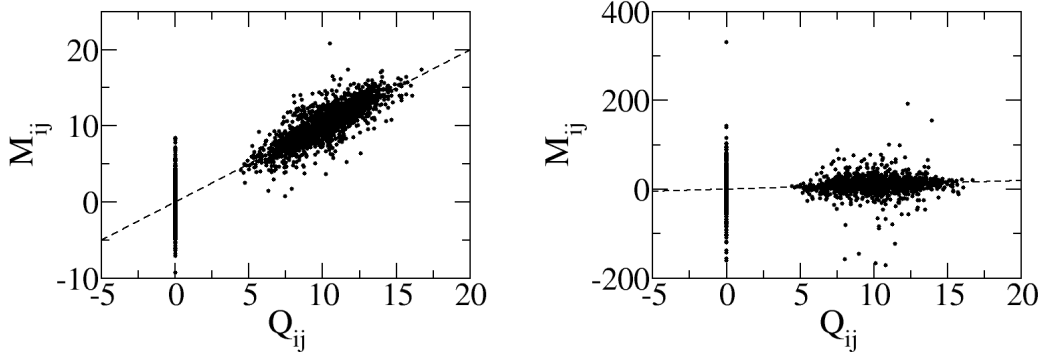


Figure 6.8: Plot of off-diagonal M_{ij} versus Q_{ij} for case 1 with $\tau = \delta t = 2 \times 10^{-2}$ and N_{data} equal to (a) 2.5×10^7 ; (b) 2×10^6 .

to one. Hence our clustering algorithm which applies to the off-diagonal elements M_{ij} only can also reconstruct links well for these cases which will be shown in the next section.

For the additional cases including cases 30-38, it is found that off-diagonal elements M_{ij} are approximately proportional to $g_{ij}A_{ij}$ (see Fig. 6.11):

$$M_{ij} \approx \alpha g_{ij} A_{ij}, \quad i \neq j, \quad (6.1)$$

where α is some constant. For instance, α is found to be 0.11 for case 30 which has $f_i = 0$ and h^{cubic} . Therefore for these additional cases for which our theoretical results do not directly apply, M_{ij} data still separate into two clusters based on A_{ij} values and network links can be reconstructed using our method with clustering analysis on off-diagonal elements M_{ij} . Relative coupling strength G_{ij} can also be reconstructed using Eq. (4.24).

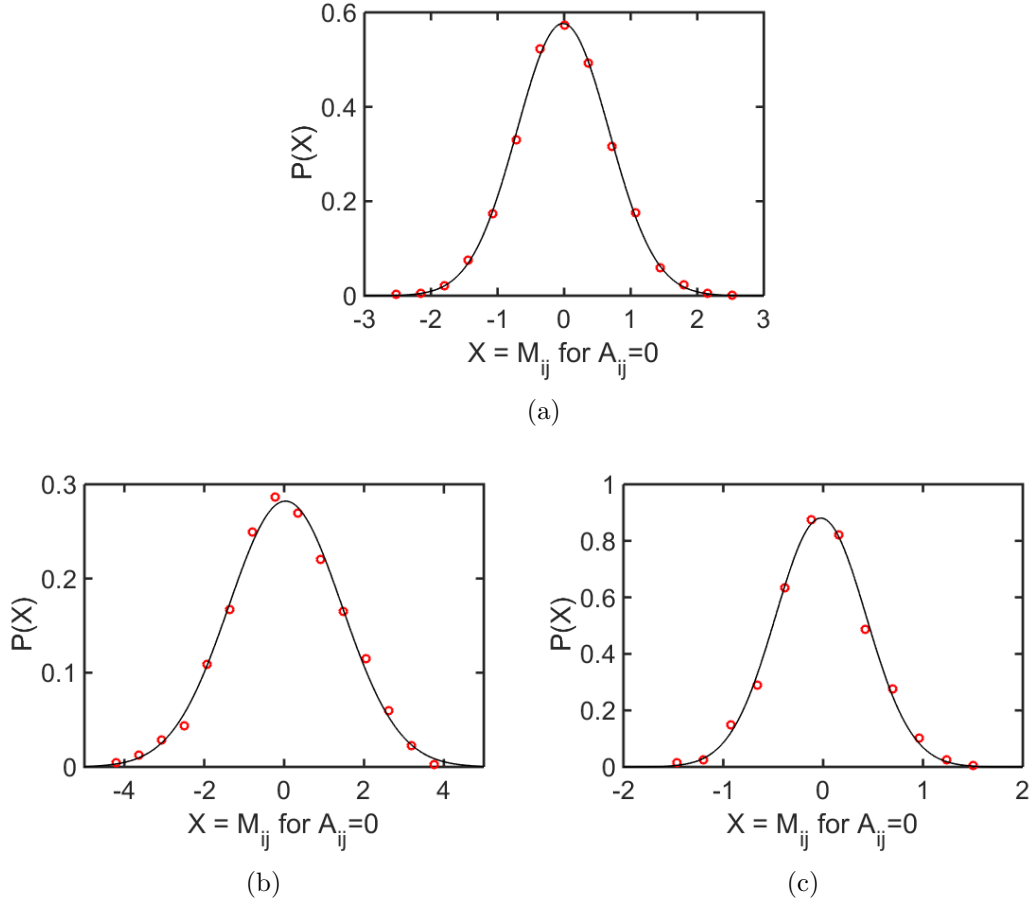


Figure 6.9: (a) case 1 with DWR1 network: normalised distribution of M_{ij} corresponding to $A_{ij} = 0$; (b)-(c) case 9 with DWSF network: Normalised distributions of M_{ij} corresponding to $A_{ij} = 0$ for two fixed values of j . Black line in each plot represents the fitted probability density function of Gaussian distribution.

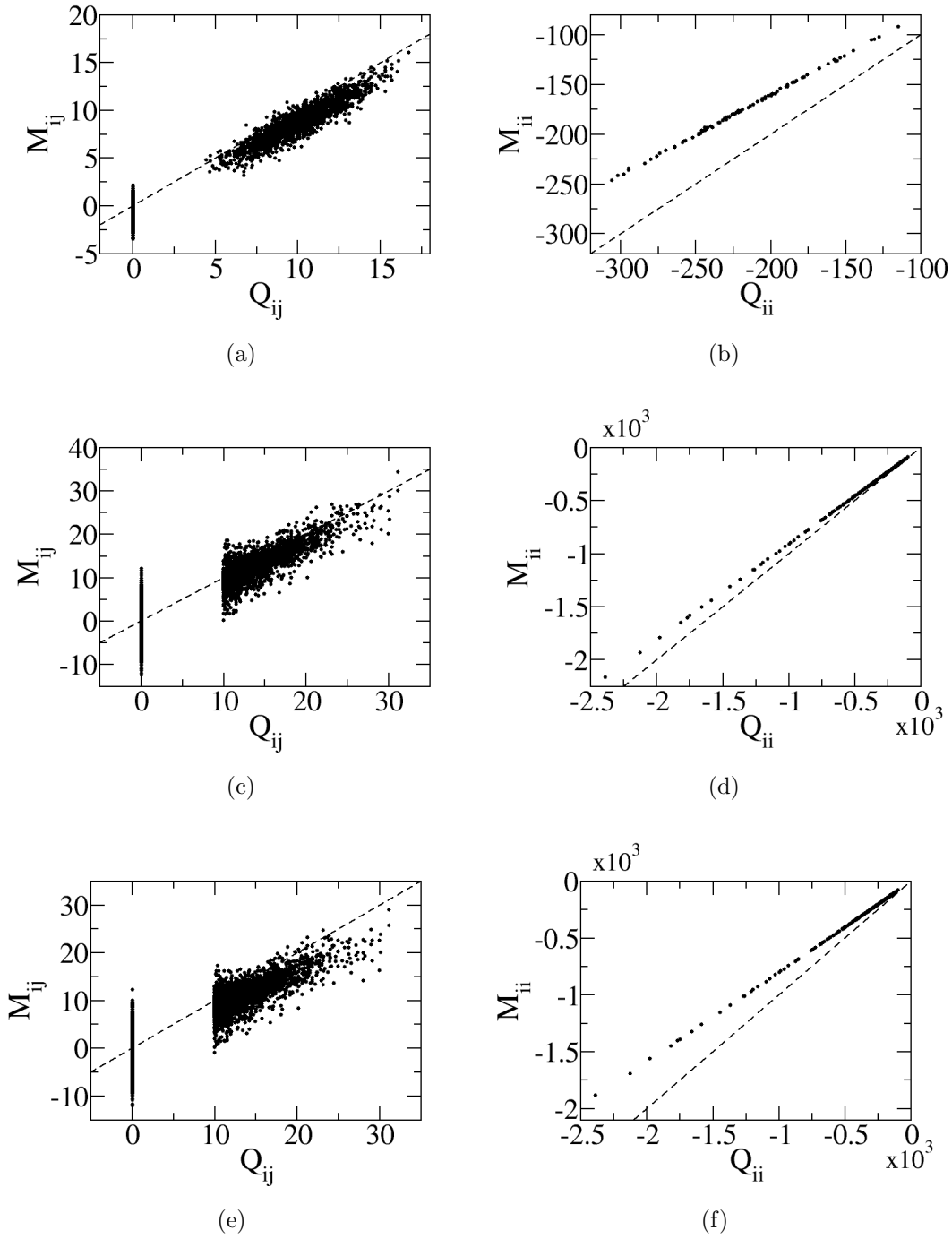


Figure 6.10: Plot of \mathbf{M} versus \mathbf{Q} for cases which have $\tau_n > 0$ and $\Delta < \delta$. Case 10 for (a) off-diagonal and (b) diagonal elements; and case 11 for (c) off-diagonal and (d) diagonal elements; case 12 for (e) off-diagonal and (f) diagonal elements.

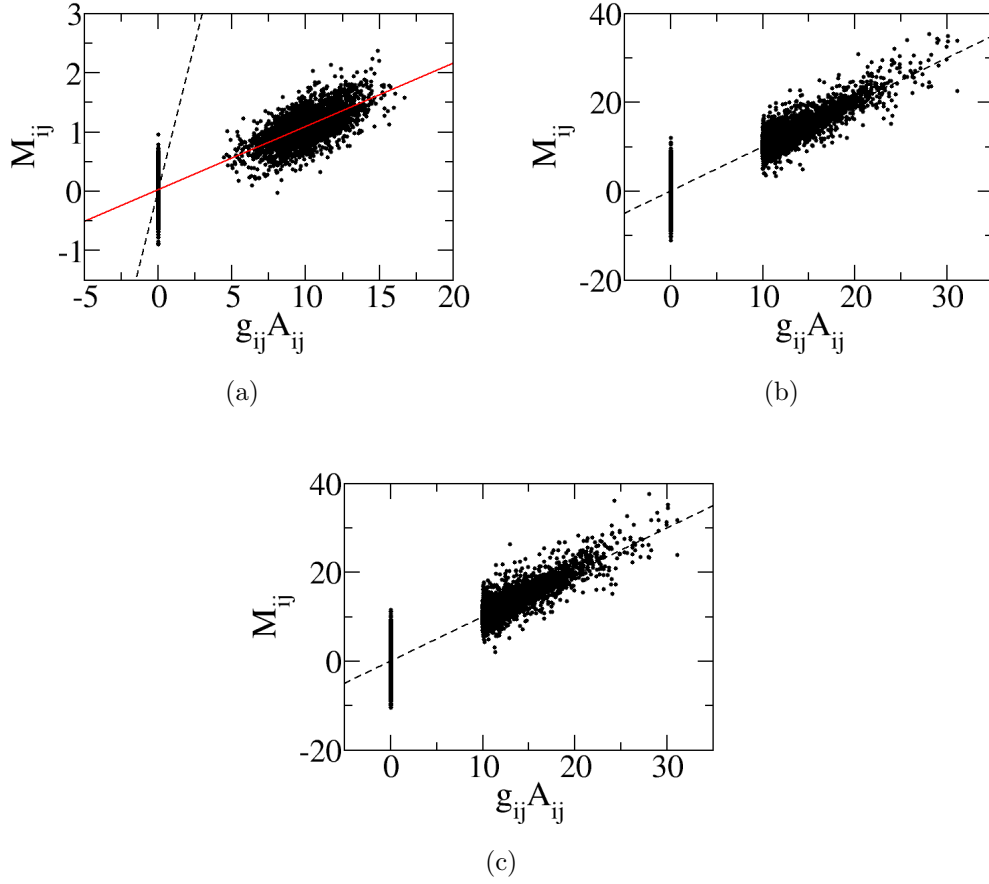


Figure 6.11: Plot of off-diagonal elements M_{ij} versus $g_{ij}A_{ij}$ for (a) case 30. Red line represents a straight line $y = 0.11x + 0.02$ obtained by fitting data of M_{ij} for $A_{ij} = 1$; (b) case 36; (c) case 37; Dashed line is $y = x$.

6.2 Performance of reconstructing network connectivity

Network links are reconstructed by performing clustering analysis on off-diagonal elements M_{ij} . To quantify the accuracy of link reconstruction, we measure the error rates FN/N_L (%) and FP/N_L (%), where false negative (FN) and false positive (FP) are the number of links missed and the number of incorrectly inferred links respectively, and N_L is the total number of links in a network. Sensitivity and specificity are another measures commonly found in literatures (e.g. [32, 37]) for quantifying the accuracy of link reconstruction. By definition, sensitivity P_{sen} is the fraction of existing links which are correctly inferred (i.e. true positive TP) compared to total number of links, while specificity P_{spec} is the fraction of non-existing links which are correctly inferred (i.e. true negative TN) compared to total number of non-existing links. Relations between P_{sen} , P_{spec} and the error rates are

$$P_{sen} \equiv \frac{\text{TP}}{N_L} = 1 - \frac{\text{FN}}{N_L} \quad (6.2)$$

$$P_{spec} \equiv \frac{\text{TN}}{N(N-1) - N_L} = 1 - \frac{\text{FP}}{N_L} \left(\frac{\rho}{1 - \rho} \right) \quad (6.3)$$

where $\rho \equiv N_L/[N(N-1)]$ is the link density. Hence when the network is sparse, i.e. small ρ , then P_{spec} becomes less sensitive to FP so that the reconstructed network structure can be significantly distorted from the actual one even when P_{spec} is high. Therefore we use the error rates FN/N_L and FP/N_L to measure the performance of our method in reconstructing links instead of P_{sen} and P_{spec} .

Table 6.1 lists the error rates of reconstruction of network connectivity for cases which are described by our theory, as well as additional cases for which our theoretical derivation of Eqs. (3.10) and (3.20) does not hold. As listed in the table, these cases have Δ smaller than δ . In the cases with white noise, i.e. $\tau_n = 0$, the value of Δ

is very close to zero because $\mathbf{F} = \mathbf{H}$ in Eq. (4.18) is held. For example, we get $\Delta = 4.8 \times 10^{-7}$ in case 1. It is expected that other cases with $\tau_n = 0$ also get a very small value of Δ which we do not show in the table.

Results shown in the table verify that our method can reconstruct links well with error rates smaller than 10% for all cases studied (case 9 requires $N_{data} = 4 \times 10^6$ for obtaining error rates smaller than 10%). Reconstruction errors are produced due to the overlap between connected and unconnected groups of M_{ij} data. For instance, among all cases studied, cases 5 and 9 get high error rates when $N_{data} = 2 \times 10^6$ because there are small M_{ij} data corresponding to $A_{ij} = 1$ in these two cases as shown in Fig. 6.1(c) and (f). In case 5, DWR1s network has some weak links due to g_{ij} taken from $N(10, 10)$, while case 9 with h^{syn} has small values of Q_{ij} for some links.

Our reconstruction is also able to reconstruct the node degrees well. Figure 6.12(a) and (b) shows that reconstructions of both in-degree and out-degree are good for case 5 and case 9 which have the highest error rates among all cases studied for random and scale-free networks respectively. Also our method can capture the power-law distributed out-degree for case 9 as shown in Fig. 6.12(c).

Ref. [60] has presented a method inferring network subject to white noise. The method is based on results derived from linearised dynamical equation Eq. (3.4) and involving matrices $\langle \delta \dot{\mathbf{x}} \delta \mathbf{x}^T \rangle$ and $\langle \delta \mathbf{x} \delta \mathbf{x}^T \rangle$. In the paper, time-derivative data is computed by finite difference approximation using a small sampling interval and the result is equivalent to

$$\mathbf{Q} \approx (\mathbf{K}_\tau \mathbf{K}_0^{-1} - \mathbf{I})/\tau, \quad (6.4)$$

which is actually our result in Eq. (4.2) with small τ limit. We have compared the

Case	Δ	FN/ N_L (%)	FP/ N_L (%)	e_G (%)
1	4.8×10^{-7}	0	0	5.7
2	—	0	0.63	6.0
3	—	0	0	6.0
4	—	2.98	2.00	14.2
5	—	9.32	1.90	9.8
6	—	0	0	4.3
7	—	0	3.38	4.5
8	—	0.50	1.11	6.9
9	—	0.65	2.84	10.3
	—	(9.20)	(21.60)	(13.1)
10	0.015	0	0	6.9
11	2.8×10^{-3}	0.73	1.43	7.2
12	0.015	1.10	1.79	8.3
13	2.8×10^{-3}	0.32	0.90	5.5
14	0.014	0.48	1.13	5.9
30	—	4.24	2.78	16.6
31	—	0	0	5.7
32	—	0	0	4.9
33	—	0	0	6.9
34	—	0	0	3.3
35	—	0	0.10	5.4
36	—	0.32	1.20	5.8
37	—	0.35	1.58	6.1
38	0.013	1.46	0.73	12.4

Table 6.1: Accuracy of our reconstruction as measured by the error rates FN/ N_L and FP/ N_L (in %) for the cases studied with values of Δ presented. Case 9: $N_{data} = 4 \times 10^6$ with results for $N_{data} = 2 \times 10^6$ given in parentheses. Cases 1-14 are cases that our theoretical derivation of Eqs. (3.10) and (3.20) applies. Cases 30-38 are additional cases for which our theoretical derivation does not hold. Cases 10-14 and 38 have noise with finite τ_n which is taken as a white noise approximation. e_G (in %) is the average percentage error of the reconstructed relative coupling strength \hat{G}_{ij} or $\hat{G}_j^{\text{out}}(i)$ defined by Eq. (6.5).

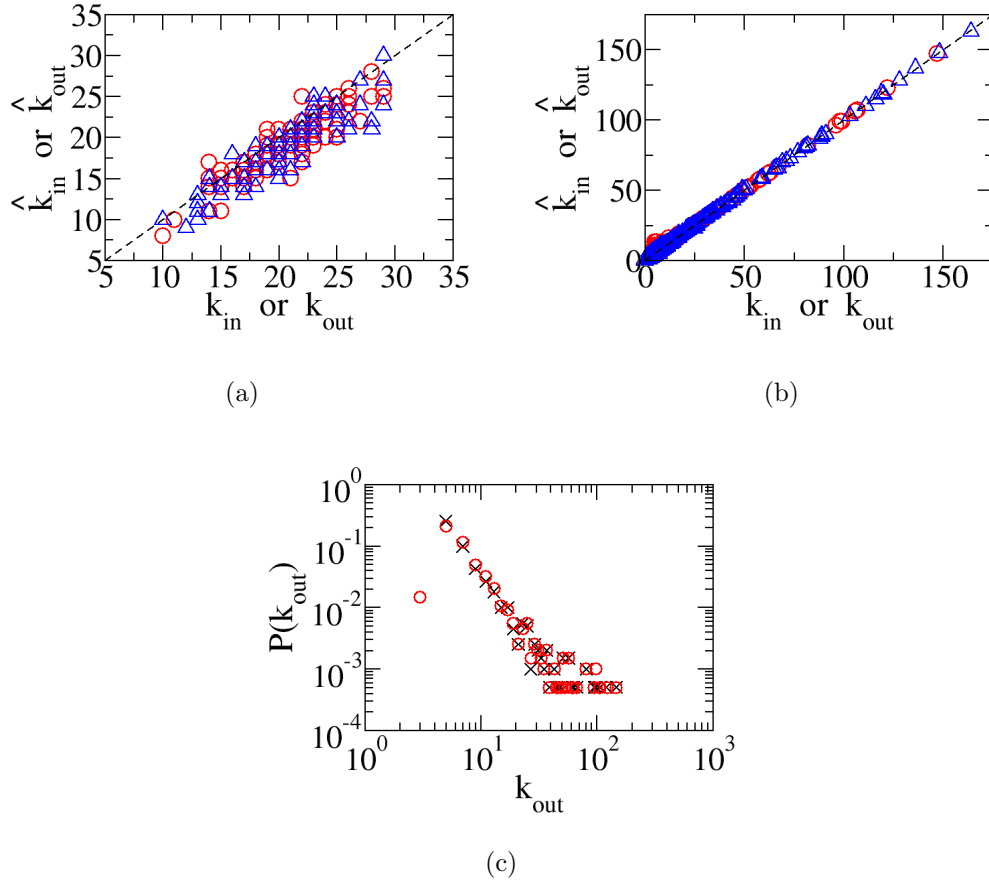


Figure 6.12: Plot of the reconstructed \hat{k}_{in} (blue triangle) and \hat{k}_{out} (red circle) versus the actual values for (a) case 5; (b) case 9. (c) Distribution of reconstructed \hat{k}_{out} (red circle) and actual values (black cross) for case 9 in plot (b).

reconstruction performance of our method and the method from Ref. [60]. Dynamical data from case 1 is collected with $N_{data} = 10^6$ and different sampling intervals δt (with $\tau = \delta t$ accordingly). Figure 6.13 shows that Eq. (6.4) does not hold well for a larger value of $\tau = \delta t = 10^{-2}$ and connected and unconnected groups of data overlap much, while M_{ij} agrees with Q_{ij} well. Table 6.2 clearly shows that the performance using Eq. (6.4) to reconstruct links is getting worse with larger sampling interval while our method maintains a good performance for different sampling intervals.

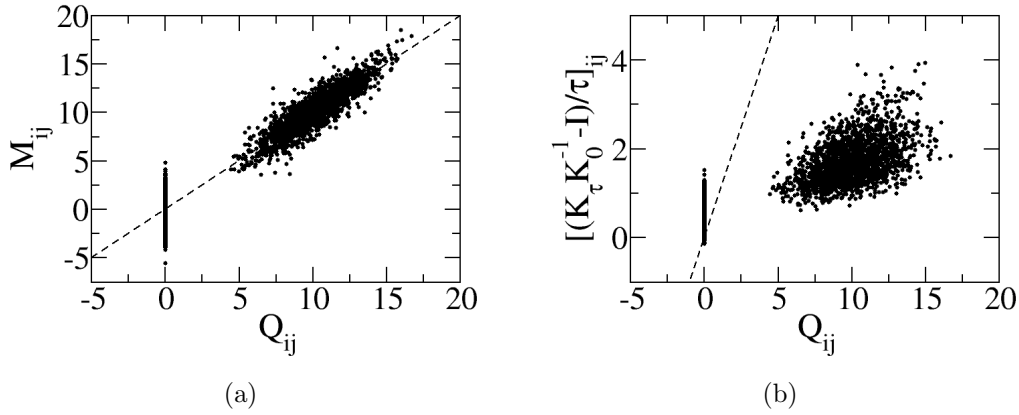


Figure 6.13: Plot of off-diagonal elements (a) M_{ij} and (b) $[(\mathbf{K}_\tau \mathbf{K}_0^{-1} - \mathbf{I})/\tau]_{ij}$ versus Q_{ij} for case 1 with $\tau = \delta t = 10^{-2}$. Dashed line is $y = x$.

$\tau = \delta t$	FN/ N_L (%)	FP/ N_L (%)	e_G (%)
10^{-3}	0 (0)	0 (0)	6.2 (6.4)
5×10^{-3}	0 (0)	0 (0.15)	4.7 (11.5)
10^{-2}	0 (2.44)	0.68 (9.66)	8.2 (20.0)

Table 6.2: Comparison of the accuracy of reconstruction of links and relative coupling strength for case 1 with $N_{data} = 10^6$ and $\tau = \delta t$ using our method and Eq. (6.4) based on Ref. [60]. Results for the latter are given in parentheses.

6.3 Result of reconstruction of relative coupling strength

Our method can reconstruct the relative coupling strength of links whenever the coupling function h satisfies some properties as stated in Section 4.5, or additional cases having M_{ij} approximately proportional to $g_{ij}A_{ij}$. Accuracy of reconstructing relative coupling strength is indicated by e_G (%) as shown in Table 6.1, which evaluates the average percentage error excluding missed and incorrectly predicted links as defined by

$$e_G = \frac{1}{\text{TP}} \sum_{\substack{i,j \\ \hat{A}_{ij}=A_{ij}=1}} \left| \frac{\hat{G}_{ij} - G_{ij}}{G_{ij}} \right| \times 100\% \quad \text{or} \quad \frac{1}{\text{TP}} \sum_{\substack{i,j \\ \hat{A}_{ji}=A_{ji}=1}} \left| \frac{\hat{G}_j^{\text{out}}(i) - G_j^{\text{out}}(i)}{G_j^{\text{out}}(i)} \right| \times 100\% \quad (6.5)$$

Our method gets e_G less than 17% for all cases studied. It is seen that our method can reconstruct \hat{G}_{ij} or \hat{G}_j^{out} which agrees with actual values well for different cases which follows our theoretical derivation (see Fig. 6.14), as well as additional cases (see Fig. 6.15).

For cases with random networks, the reconstructed \hat{G}_{ij} follows the Gaussian distribution well as shown in one example in Fig. 6.16(a) from case 2 with $e_G = 6.0\%$. In some cases with connected and unconnected parts of M_{ij} data overlapped much like case 30 with $f_i = 0$ and h^{cubic} , some missed weak links result in skewed reconstructed distribution of \hat{G}_{ij} and the Gaussian distribution of G_{ij} can be recovered when using a larger N_{data} as shown in Fig. 6.16(b). In DWSF cases, the reconstructed \hat{G}_{ij} for most links follows the power-law distribution as shown in one example of case 8 ($e_G = 6.9\%$) as shown in Fig. 6.16(c), while some links have too small values of relative coupling strength which are due to the scatters of M_{ij} around Q_{ij} .

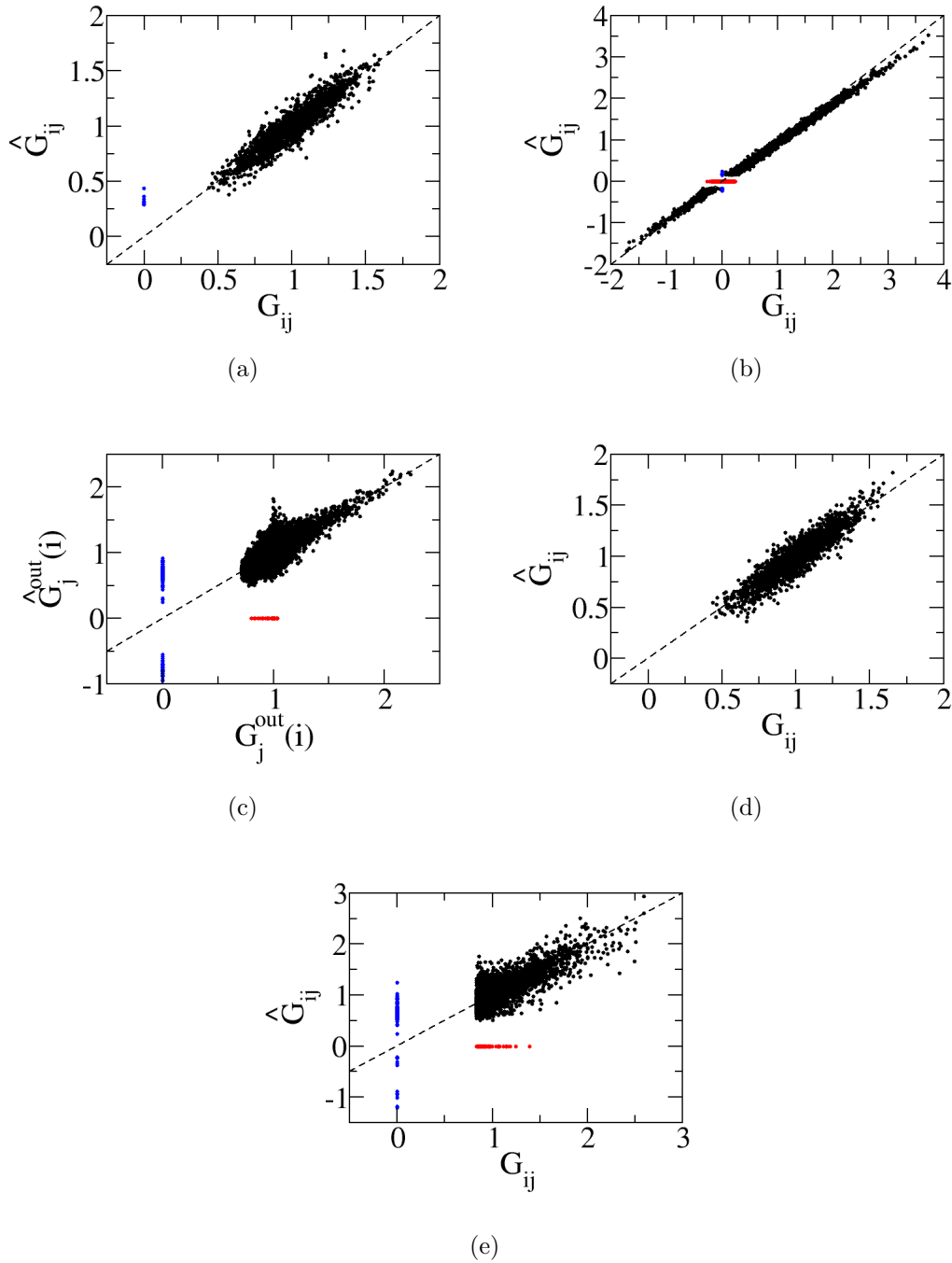


Figure 6.14: Plot of reconstructed \hat{G}_{ij} or $\hat{G}_j^{\text{out}}(i)$ versus actual values for (a) case 2; (b) case 5; (c) case 9; (d) Case 10; (e) Case 12. Blue and red dots represent incorrectly predicted links and missed links respectively. Dashed line is $y = x$.

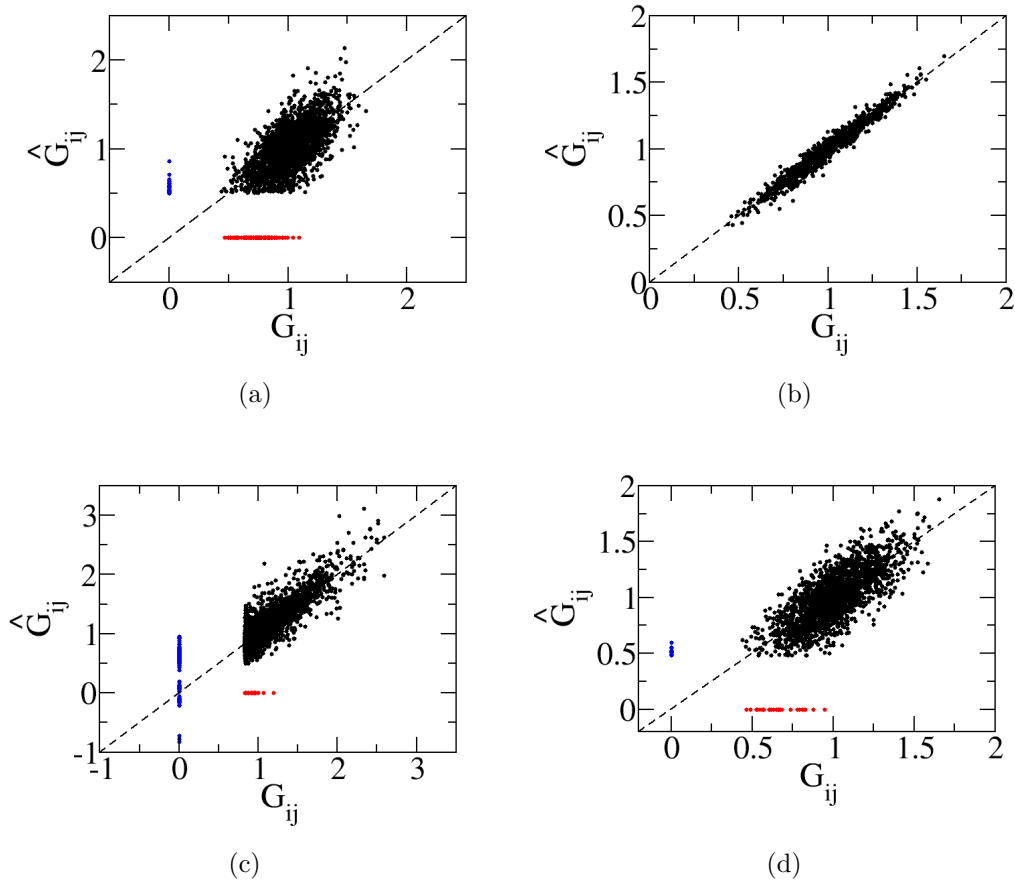


Figure 6.15: Plot of reconstructed \hat{G}_{ij} versus actual values for additional cases including (a) case 30; (b) case 34; (c) case 37; (d) Case 38. Blue and red dots represent incorrectly predicted links and missed links respectively. Dashed line is $y = x$.

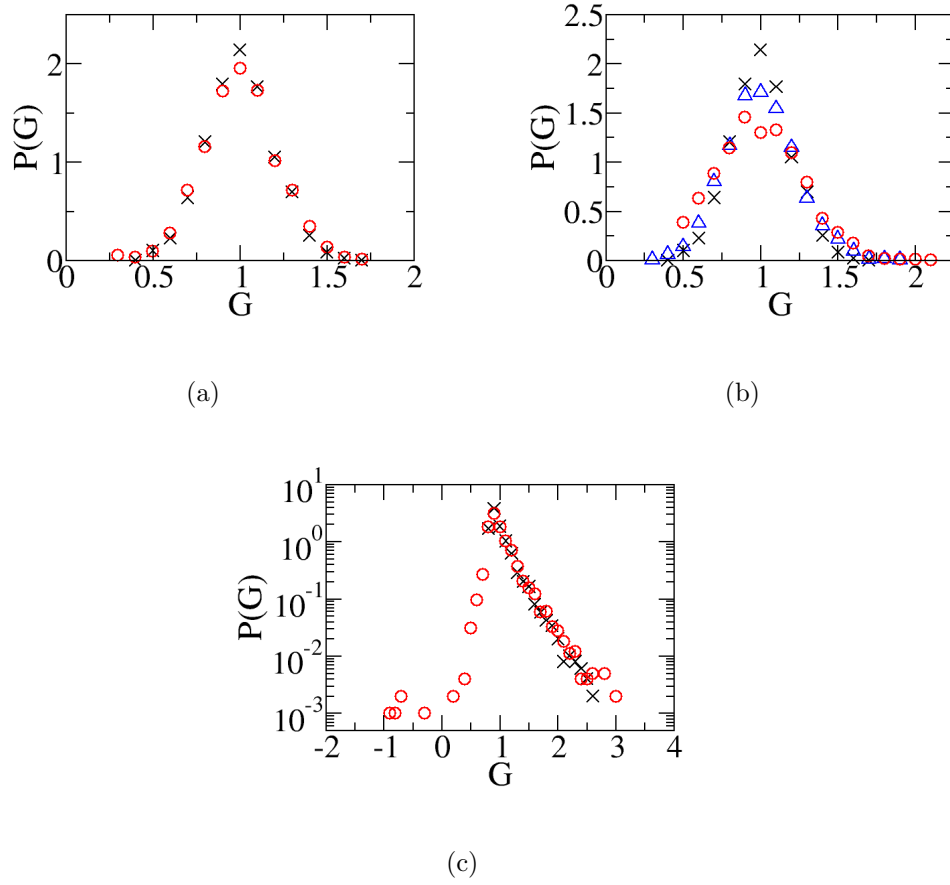


Figure 6.16: Distributions of reconstructed \hat{G}_{ij} (red circle) and actual values (black cross) for (a) case 2; (b) case 30 (Blue triangles denote distribution obtained from $N_{data} = 10^7$); (c) case 8.