POLITECNICO DI TORINO
Repository ISTITUZIONALE

Inverse problem for multi-body interaction of nonlinear waves

Original
Inverse problem for multi-body interaction of nonlinear waves / Marruzzo, Alessia; Tyagi, Payal; Antenucci, Fabrizio; Pagnani, Andrea; Leuzzi, Luca. - In: SCIENTIFIC REPORTS. - ISSN 2045-2322. - ELETTRONICO. - 7:1(2017), p. 3463. [10.1038/s41598-017-03163-4]

Availability:
This version is available at: 11583/2675237 since: 2017-06-27T16:35:37Z

Publisher:
Nature Publishing Group

Published
DOI:10.1038/s41598-017-03163-4

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

23 April 2024

SCIENTIFIC REPORTS

OPEN

Inverse problem for multi-body interaction of nonlinear waves

Alessia Marruzzo[1], Payal Tyagi[1], Fabrizio Antenucci[1], Andrea Pagnani[2,3] & Luca Leuzzi[1,4]

Received: 11 November 2016
Accepted: 25 April 2017
Published online: 14 June 2017

The inverse problem is studied in multi-body systems with nonlinear dynamics representing, e.g., phase-locked wave systems, standard multimode and random lasers. Using a general model for four-body interacting complex-valued variables we test two methods based on pseudolikelihood, respectively with regularization and with decimation, to determine the coupling constants from sets of measured configurations. We test statistical inference predictions for increasing number of sampled configurations and for an externally tunable *temperature*-like parameter mimicing real data noise and helping minimization procedures. Analyzed models with phasors and rotors are generalizations of problems of real-valued spherical problems (e.g., density fluctuations), discrete spins (Ising and vectorial Potts) or finite number of states (standard Potts): inference methods presented here can, then, be straightforward applied to a large class of inverse problems. The high versatility of the exposed techniques also concerns the number of expected interactions: results are presented for different graph topologies, ranging from sparse to dense graphs.

Multi-body inference turns out to be essential whenever non-linear response is crucial for a system properties. Light mode interaction in ultra-fast multimode lasers[1–6], random lasers[7–10], multi-variable clause constrained problems[11,12], error correcting codes[13,14], effective interaction among density fluctuations in heterogeneous frustrated glassy systems[15–20] and fish shoals behavior[21,22] are significant diverse examples of direct problems in which nonlinearity plays a non-perturbative role in determining the system behavior. Nevertheless, to our knowledge, not many studies of the inverse problem have been performed so far in the field. In this work we aim at filling in this gap presenting a detailed analysis based on pseudolikelihood maximization (PLM) techniques for the statistical inference in models with multi-body interactions.

Inverse problems consist in determining the interaction couplings among system variables from measurements of variable configurations or correlations. As an instance, in the optical waves framework, this means quantitatively inferring the nonlinear interaction strengths given the wave emissions. Once the theoretical model is designed, assuming an effective equilibrium (true thermodynamic equilibrium or stationary conditions), one has to maximize the likelihood functional with respect to the coupling parameters. The likelihood functional is defined as the probability of a variable configuration given the values of the interaction couplings. For large systems it is numerically intractable but one can resort to the so-called pseudolikelihood functional defined as the probability of one variable conditional to all other variables and to the values of the couplings[23].

Based on pseudolikelihood maximization, we adopt two methods to determine the interactions: the well known ℓ_1-regularization[24,25], that we have improved with a hypothesis testing procedure based on the evaluation of the eigenvalues of the Fisher information matrix[26], and the most recent decimation technique[27]. In order to test the methods, we considered both the phasor and the XY-spin models, generating the data by means of Monte Carlo numerical simulations. Among the simulated networks we analyze both sparse graphs, in which the number of interacting quadruplets N_q scales like the number of variables, $N_q \propto N$, and dense graphs, in which $N_q \propto N^3$ [This is a diluted dense graph: not all quadruplets are present, though their number per variable node scales with N, unlike in sparse graphs. A complete dense graph would contain $O(N^4)$ interacting quadruplets]. We stress that the techniques here reported might be applied to any wave system with non-linear collective behavior, such as phase-locking, breathers and synchronization[28–31], including the prototype Fermi-Pasta-Ulam model[32]. Further on, the methodology can be translated to simpler cases, e.g., discrete variables models like the p-clock model[33,34],

[1]NANOTEC-CNR, Institute of Nanotechnology, Soft and Living Matter Laboratory, Rome, Piazzale A. Moro 2, I-00185, Roma, Italy. [2]Department of Applied Science and Technology, Politecnico di Torino, 10129, Torino, Italy. [3]Human Genetics Foundation, Molecular Biotechnology Center, 10126, Torino, Italy. [4]Dipartimento di Fisica, Università di Roma "Sapienza", Piazzale A. Moro 2, I-00185, Roma, Italy. Correspondence and requests for materials should be addressed to A.M. (email: alessia.marruzzo@uniroma1.it)

in which rotators only take p discrete values. Properly modifying the mode interaction these p-clock models can, eventually, represent multi-body Potts models[35], as well.

Results

Test models. Our first test model consists of N phasors a_k with a global constraint $\sum_{k=1}^{N} |a_k|^2 = \text{const} \times N$, hereafter termed Spherical Model (SM), with Hamiltonian[36]

$$\mathcal{H} = -\frac{1}{8}\sum_{jklm}^{\text{d.i.}} J_{jklm}\, a_j a_k^* a_l a_m^* + \text{c. c.} \tag{1}$$

The a_k's represent, e.g., the complex amplitudes of the normal modes expansion of the electromagnetic field[1]

$$\widetilde{E}(\boldsymbol{r}, t) = \sum_k a_k(t)\boldsymbol{E}_k(\boldsymbol{r})e^{i\omega_k t} + \text{c. c.} \tag{2}$$

characterizing the light modes in the $\boldsymbol{E}_k(\boldsymbol{r})$ basis. The amplitude $a_k(t)$ is the slowly varying coefficient of the normal mode $\boldsymbol{E}_k$ of frequency ω_k and varies on time scales much slower than ω_k^{-1}. We adopt it as test model because it is the lowest order of nonlinearity satisfying time reversal symmetry of light, as occurring, e.g., in centrometric crystals with symmetric atomic potentials[37]. The laser transition can be represented as a phase transition in statistical mechanical theory. This turns out to be possible both in ordered multimode mode-locked lasers[3,5,6,34,38–40] and in random lasers[10,36,41,42]. Considering further orders of the interaction does not change the critical behavior and the onset of the lasing regime, nor the qualitative features of the laser in the high pumping regime. We stress that, simply in order to focus the presentation, also lower order interactions (pairwise and three body) are not considered here: the sum with superscript "d.i." in Eq. (1) is intended solely over quadruplets with distinct indices.

According to multimode laser theory[1–3,37,43] modes do interact nonlinearly if and only if their frequencies satisfy a frequency matching condition[5], i.e., given any four modes j, k, l, m of typical line-width γ, their angular frequencies are such that

$$\left|\omega_j - \omega_k + \omega_l - \omega_m\right| \lesssim \gamma \tag{3}$$

at least in one permutation of their indices.

With equipartite magnitudes ($|a_k| \simeq 1, \forall\ k$) or with quenched ones ($|a_k(t)| = A_k(0)$) Eq. (1) for phasors reduces to the so-called XY model for rotators, $a_j = A_j e^{\iota\phi_j} \to e^{\iota\phi_j}$, with Hamiltonian

$$\begin{aligned}\mathcal{H} \;=\; & -\frac{1}{8}\sum_{jklm}^{\text{d.i.}} [J^R_{jklm}\cos(\phi_j - \phi_k + \phi_l - \phi_m) \\ & + J^I_{jklm}\sin(\phi_j - \phi_k + \phi_l - \phi_m)]\end{aligned} \tag{4}$$

where $J^{R,I}$ are, respectively, real and imaginary parts of the coupling constants. The $4XY$ model is our second test model. Besides being an approximation of the SM model, having locally constrained variables allows for testing the inference techniques also on sparse graphs at low temperature [Indeed, it can be seen that for bond-disordered SM's, if the node connectivity does not increase at least with N^2, all the power condensates into one single quadruplet below threshold[6]]. Furthermore, terming $\delta\omega$ the frequency spacing among the modes, we considered both strict frequency matching conditions, cf. Eq. (3), based on comb-like[44] single mode resonance distributions ($\gamma \ll \delta\omega$), as well as *narrow-band* conditions ($\gamma > \delta\omega$). In the latter case Eq. (3) does not play any role and the node frequencies have no influence on the structure of the graphs. On the other hand, in graphs built considering $\gamma \ll \delta\omega$ frequencies do play an important role. These will be called Mode-Locked (ML) graphs.

We infer data within the Boltzmann-Gibbs equilibrium hypothesis (see Suppl. Mat. for data generation). Then, the probability of a configuration $\boldsymbol{a}$, given a set $\boldsymbol{J}$, i.e., the likelihood functional, reads:

$$P(\boldsymbol{a}|\boldsymbol{J}) = \frac{1}{Z[\boldsymbol{J}]}\exp\{-\beta\mathcal{H}[\boldsymbol{a}|\boldsymbol{J}]\} \tag{5}$$

Computing $Z[\boldsymbol{J}]$ is very hard in general. To circumvent this bottleneck one first defines the single variable pseudo-likelihood[23] of the values of a_i biased by all other $\boldsymbol{a}_{\setminus i}$ values, and by $\boldsymbol{J}$ (see details in Suppl. Mat.)

$$P_i(a_i|\boldsymbol{a}_{\setminus i}, \boldsymbol{J}) = \frac{1}{Z_i[\boldsymbol{a}_{\setminus i}, \boldsymbol{J}]}\exp\{a_i H_i[\boldsymbol{a}_{\setminus i}, \boldsymbol{J}] + \text{c. c.}\} \tag{6}$$

with

$$H_j[\boldsymbol{a}_{\setminus j}, \boldsymbol{J}] = \sum_{klm\neq j}^{\text{d.i.}} J_{jklm}\mathcal{F}_{klm} \tag{7}$$

$$\mathcal{F}_{klm} = \frac{1}{3}[a_k^* a_l a_m^* + a_k a_l^* a_m^* + a_k^* a_l^* a_m] \tag{8}$$

$$Z_i[\boldsymbol{a}_{\setminus i}, \boldsymbol{J}] \equiv \sum_{a_i} \exp\{a_i H_i[\boldsymbol{a}_{\setminus i}, \boldsymbol{J}] + \text{c. c.}\} \tag{9}$$

Further on, if one considers M independent configurations $\{\boldsymbol{a}^{(\mu)}\}$, with $\mu = 1$, $\mu = 1, \ldots, M$, the pseudolikelihood of all the single node variables $\{a_i^{(\mu)}\}$, given all the others $\{\boldsymbol{a}_{\setminus i}^{\mu}\}$ and the couplings $\boldsymbol{J}$, factorizes. In order to deal with sums instead of products, one usually evaluates the log-pseudolikelihood that, thus, reads

$$\begin{aligned} \mathcal{L}_i^{(0)}(\{\boldsymbol{a}_{\setminus i}^{\mu}\}, \boldsymbol{J}) \;=\; & \sum_{\mu=1}^{M} (a_i^{\mu} H_i[\boldsymbol{a}_{\setminus i}^{\mu}, \boldsymbol{J}] + \text{c. c.}) \\ & - \sum_{\mu=1}^{M} \ln Z_i[\boldsymbol{a}_{\setminus i}^{\mu}, \boldsymbol{J}] \end{aligned} \tag{10}$$

The $\boldsymbol{J}$'s maximizing $\mathcal{L}_i^{(0)}$s are considered as the most probable couplings that originate the $\{\boldsymbol{a}^{(\mu)}\}$ configurations.

We analyze data from systems whose coupling values are randomly generated with a bimodal distribution $P(J) = 1/2[\delta(J - \hat{J}) + \delta(J + \hat{J})]$, where $\hat{J} = 1/N^{(z-1)/2}$, when the total number of quadruplets scales as $N_q \sim N^z$. This is the case, e.g., of the frustrated glassy random lasers[10, 45–49], but the methods here exposed also work for the simpler cases of uniform couplings, like in standard mode-locking lasers[2, 3, 5, 34] and random couplings with a relative small fraction of negative values, e.g., random unfrustrated lasers[34, 42, 50].

Data Analysis

Decimation and ℓ_1-regularization. We compare inference predictions obtained by different implementations of PLM. The ℓ_1-regularization consists in adding to Eq. (10) a regularizing term $\mathcal{L}_i^{(0)} - \lambda\|J\|_1$, penalizing large $\boldsymbol{J}$ values and it is known to be particularly useful in retrieving sparse systems[25]. The *decimation* procedure, instead, iteratively removes the smallest couplings (cf., Suppl. Mat.). In this procedure one maximizes the *total* log-pseudolikelihood, summed over all the modes, i.e.,

$$\mathcal{L}(\boldsymbol{J}) \equiv \mathcal{L} = \frac{1}{N} \sum_{i}^{N} \mathcal{L}_i^{(0)} \tag{11}$$

It is important to underline that, by maximizing each $\mathcal{L}_i^{(0)}$ separately, cf. Eq. (10), each coupling J_{ijkl} turns out to be inferred four times with, generally, four different estimates. The mean value is, then, usually taken as best reconstructed value. By maximizing the total $\mathcal{L}$, instead, each coupling J_{ijkl} is inferred only once.

Data size and external tuning. We consider the effects of varying the size M of data sets, as well as, the *temperature*-like parameter T that determines the strength of the interaction. T resembles real data noise[51] or it is used to drive the system to a phase transition, if present. As it will be shown, functioning of PLM's qualitatively change at criticality and in different thermodynamic phases.

Quality indicators. To evaluate the performances of the techniques we will consider the following quality indicators: (i) the True Positive Rate (TPR), that is the fraction of true bonds also appearing in the inferred set of bonds, (ii) the True Negative Rate (TNR), that is the fraction of missing bonds also absent in the inferred set of bonds, and (iii) the reconstruction error

$$\text{err}_J \equiv \sqrt{\frac{\sum_q (J_q - J_q^*)^2}{\sum_q J_q^2}} \tag{12}$$

yielding how far the inferred values J_q^* of the distinct quadruplets $q \equiv \{i, j, k, l\}$ are from the true values J_q. Exclusively for the decimation PLM, in order to reconstruct the number of non-zero couplings, i.e., the number of quadruplets actually present in the system, we analyze also the behavior of the *tilted* pseudolikelihood function defined as:

$$\mathcal{L}_t \equiv \mathcal{L}(x) - x\mathcal{L}_{\max} - (1 - x)\mathcal{L}_{\min} \tag{13}$$

where x is the number of non-decimated, i.e., non-erased, couplings. $\mathcal{L}_{\max}$ is the maximum of the total log-pseudolikelihood, Eq. (11), at the beginning of the decimation procedure, when all possible couplings are contemplated, while $\mathcal{L}_{\min}$ is evaluated on a graph with no links. $\mathcal{L}(x)$ is the maximum with respect to the x fraction of all possible couplings that are still considered to be important parameters of the problem. Erasing irrelevant couplings does not affect $\mathcal{L}(x)$ so that a plateau occurs in $x > x^*$ until important couplings start to be decimated and $\mathcal{L}(x)$ starts to decrease. In order to ease the identification of the optimal number of fitting parameters x^*, $\mathcal{L}(x)$ is tilted: the optimal value x^*, corresponding to the amount of couplings in the true network, is determined looking at the maximum of $\mathcal{L}_t$.

In Fig. 1, using data from a $4XY$ model on Erdos-Renyi (ER)-like sparse graph, we show how the TNR/TPR ratio increases to 1 as the parameter λ used for the ℓ_1-regularization, i.e., $\mathcal{L}_i^{(0)} - \lambda\|J\|_1$, is increased. Further on a δ-threshold criterion[25] is adopted for *model selection*, i.e., the ability to reduce the number of parameters to the relevant ones. Within the δ-threshold criterion, couplings which are inferred, in absolute value, to be less than δ are considered to be irrelevant and are set to zero. The value for δ is, however, chosen *a priori* and the choice might

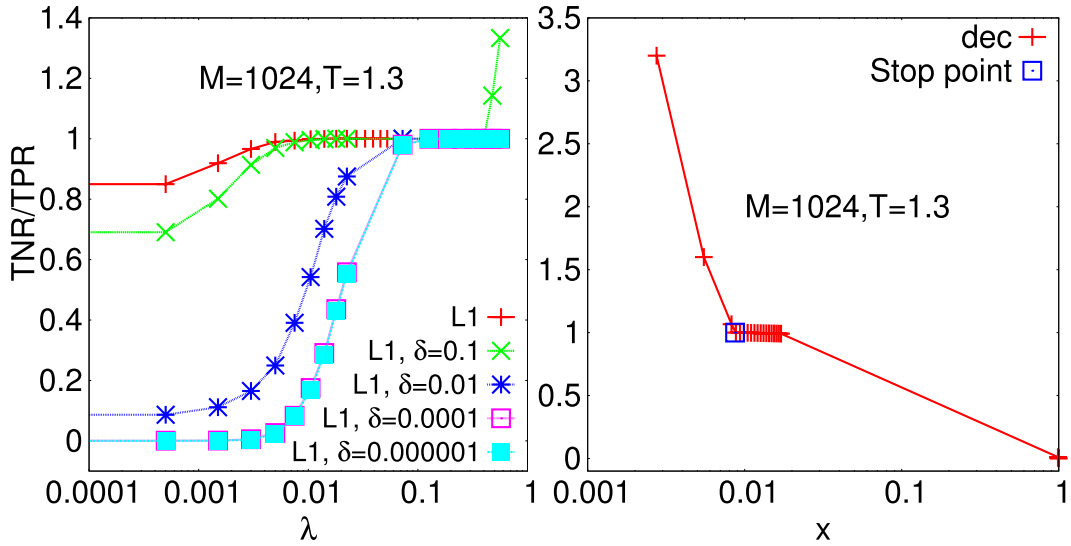

Figure 1. The TNR/TPR ratio vs. the regularizer λ used for the ℓ_1-regularization (left) and vs. the fraction x of undecimated couplings for the PLM with decimation. The stopping point indicates the maximum of $\mathcal{L}_t$, Eq. (13), where the decimation procedure stops. In the first case two different criteria are chosen to eliminate small bonds: the *a-priori* δ thresholding or the *a posteriori* inferred bond distribution thresholding based on the Fisher information matrix (see Suppl. Mat. for details). Data are taken from a 4*XY* model on a sparse Erdos-Renyi random graph with $N_q = N$, $N = 16$, $T = 1.3$, $M = 1024$. In this case the finite size proxy for the critical temperature is $T_c(16) \simeq 1.34$.

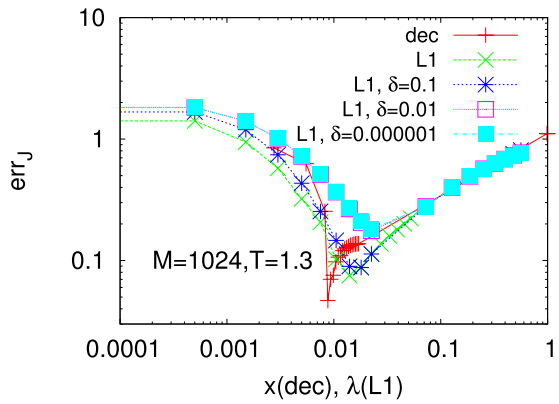

Figure 2. Reconstruction error for the 4*XY* model on sparse Erdos-Renyi graph with $N_q = N = 16$, $M = 1024$ at $T = 1.3$. The error obtained following various ℓ_1-regularized PLMs is displayed vs. λ; the decimation PLM reconstruction error is plotted against the fraction of non-decimated couplings x.

be delicate when there is not a clear gap in the distribution of the inferred couplings[27]. Moreover, as λ is small, we see that the smaller the δ the less precise the network reconstruction. On the other hand, the smaller the λ the less perturbed the original pseudolikelihood (PL). Indeed, increasing λ the chance of globally underestimate the couplings increases.

If the probability distributions of the estimators are known, the issues related to an *a priori* fixing of a δ threshold might be overcome through a more accurate *hypothesis testing* procedure. Indeed, it can be seen that, as $M \to \infty$, the probability distribution of the maximum PL estimator is a Gaussian with variance given by the diagonal element of the inverse of the Fisher information matrix[26]. Therefore, as detailed in Suppl. Mat., we can construct a confidence interval for each estimated value and verify whether it is compatible with the hypothesis "being a zero coupling". If it is the case, it is considered as an irrelevant parameter and erased. As we can see from Figs 1 and 2 this criterion for model selection outperforms, for every value of λ, the δ-threshold method. Moreover, as detailed in the Methods, this criterion provides a method to determine the best value for the regularizer λ. λ is usually chosen through Cross-Validation (CV) or Generalized CV techniques (see, e.g., ref. 52). CV techniques are, however, much more computational demanding and the number of samples used to fit the model and infer the interaction couplings is further reduced in order to have a validation set (see Methods for more details on the application of the CV technique on this model inference). Always in Fig. 1 (right) we display the TNR/TPR ratio obtained with the decimation PLM as the fraction of non-decimated couplings x decreases (from

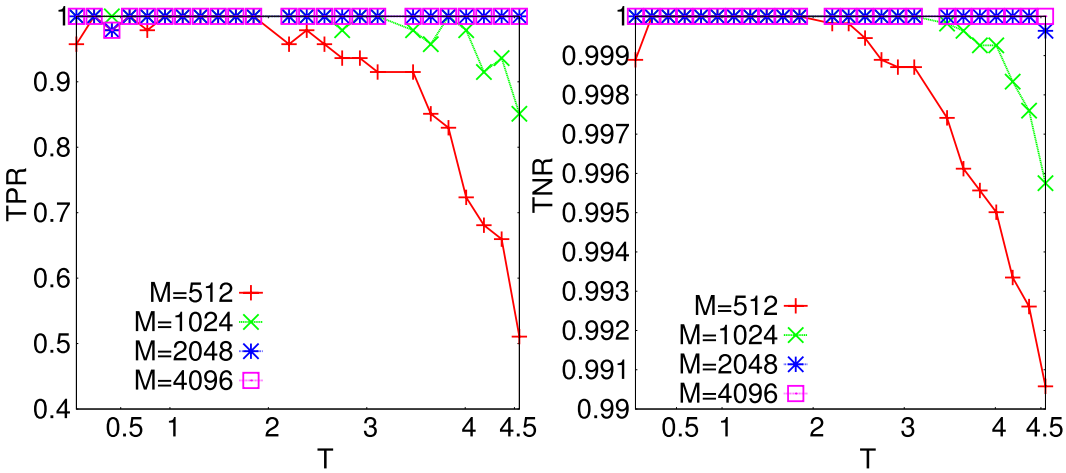

Figure 3. TPR (Left) and TNR (Right) for decimated networks at the maximum x^M of the tilted pseudolikelihood $\mathcal{L}_t$ vs T at different data-set sizes M for the $4XY$ model on sparse Mode-Locked graphs with $N=16$, $N_q=47$, $T_c(N) \simeq 0.50$ ($T_c(\infty)=0$).

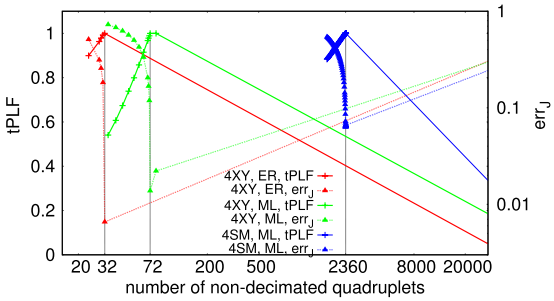

Figure 4. Tilted pseudolikelihood tPLF (normalized to its maximum) and reconstruction error for different models. Concave curves are tPLF, convex curves are err_J. Three apart models on different random graphs are considered. Red (left) curves: $4XY$ model on sparse Erdos-Renyi graph with $N=32$, $N_q=32$, $M=65000$, at $T=1.2$ ($T_c(32) \simeq 1.39$). Green (mid) curves: $4XY$ model on sparse Mode-Locked graph with $N=32$, $N_q=72$, $M=65000$, at $T=1.8$ ($T_c(32) \simeq 0.72$). Blue (right) curves: 4SM model on dense Mode-Locked graph with $N=32$, $N_q=2360$, $M=65000$, at $T=6.2$ ($T_c(32) \simeq 0.91$).

fully connected limit $x=1$ to non-interacting graph $x=0$). At $x=1$ the TPR is always one, for any M and T, whereas the TNR $=0$. As the fraction of non-decimated couplings decreases but remains greater than or equal to the true one ($x^*=2/15$ in the original model analyzed in the right panel of Fig. 1) the TPR does not decrease and the TNR increases towards one. Eventually, more couplings than those of the original network are decimated: the TPR starts decreasing and the ratio TNR/TPR consequently grows above 1 as $x \to 0$. The blue square indicates the stopping point of the decimation procedure determined, instead, as the maximum of the $\mathcal{L}_t$, cf. Eq. (13). It can be observed that in this case it perfectly reconstructs the network of interactions since the TPR $=$ TNR $=1$.

Decimation PLM. When comparing the performances of the most efficient regularization method with the decimation one, we observe that the network reconstruction in terms of true and false couplings is very adequate with both methods. However, the order of magnitude of the reconstruction error, testing also the quality of the inferred *values* of the couplings, is smaller in the decimation PLM, cf. Fig. 2, when the fraction of decimated couplings $(1-x)$ equals the one of the true network. It is important to underline that, within the decimation PLM, no parameters are determined *a priori*: the optimum value of x is determined maximizing the $\mathcal{L}_t$. Moreover, exact fraction of relevant parameters and best estimate of their values are simultaneously inferred, which is not always true in the PLM with ℓ_1-regularization since even the smallest λ for network reconstruction might induce a too high global underestimation of the couplings. We, thus, deepen the analysis of the decimation PLM.

In Fig. 3 we display TPR (left) and TNR (right) vs. T and M for the decimated network at the maximum point, x^M, of the $\mathcal{L}_t$ for the $4XY$ model on a Mode-Locked-like sparse graph with $N_q=47$ number of quadruplets and $N=16$ nodes. For large enough M the reconstruction is optimal for all temperatures, whereas for small M it is guaranteed only in a T interval around the *finite size* proxy to the critical temperature (see Suppl. Mat.). Indeed, we observed that, tuning the external temperature-like parameter for each system studied, one can identify a "critical" T interval where the reconstruction error is minimal, even orders of magnitude smaller than outside

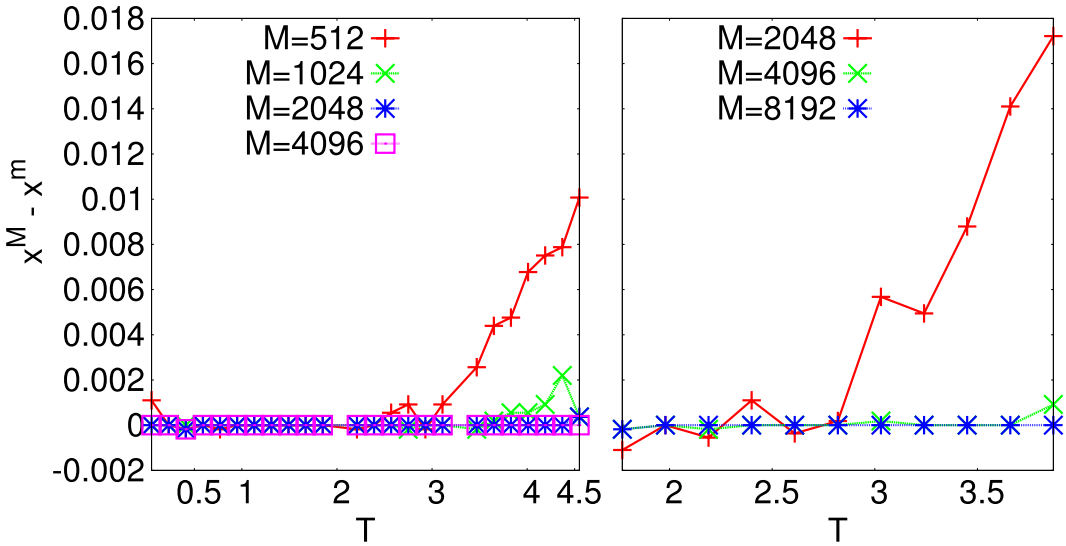

Figure 5. Plot of the difference between the maximum point of $\mathcal{L}_t$, x^{M}, and the minimum point of the reconstruction error, x^{m}, vs T at different M for systems of $N = 16$ variables. Left: 4-XY model on sparse Mode-locked graph with $N_q = 47$ ($T_c(16) \simeq 0.50$). Right: 4SM model on dense Mode-Locked graph with $N_q = 252$ ($T_c \simeq 1.07$).

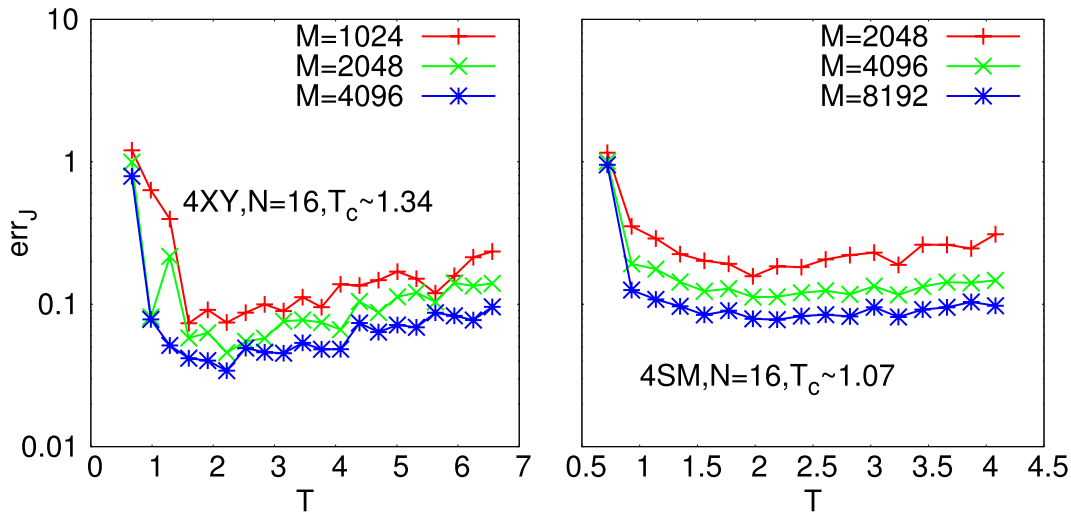

Figure 6. Reconstruction error at its minimum $x^{\min} = x^{\mathrm{true}}$ in the decimation procedure vs. T for various M for $N = 16$ variables systems with bimodal random values of the coupling constants. Left: 4XY model on sparse Erdos-Renyi graph with $N = 16$, $N_q = 16$. Right: 4SM model on dense Mode-Locked graph with $N = 16$, $N_q = 252$.

such interval, and, more in general, the system is better and easier reconstructed. In Fig. 4, the behavior of the $\mathcal{L}_t$ versus x is compared to $\mathrm{err}_J(x)$ for three different systems, the 4XY-model on ML and ER sparse graphs and the 4SM-model on a dense ML graph. All systems have $N = 32$ variable nodes while the number of interaction quadruplets is $N_q = 32$, 72 and 2360, respectively. The number of configurations in all cases is $M = 65000$. It is clearly observed that, given a large enough M and/or a critical-like T the maximum point of the $\mathcal{L}_t$, x^M, coincides with the minimum point of err_J, x^m. The decimation PLM gives then a criterion to determine the number of interaction couplings in the system from measurements data without any *a priori* chosen parameters. We notice that, as M is small and T far from the critical region, the maximum point of $\mathcal{L}_t$ and minimum point of err_J can be mismatched, as shown in Fig. 5. In Fig. 6 the T dependence of err_J is plotted for the 4XY model on a ER sparse graph and for the 4SM model on a ML dense graph at different values of M. As detailed in Suppl. Mat. the critical T interval turns out to be identified by the (finite size) critical temperature estimate of the phase transition point of the direct statistical mechanical problem.

Discussion

We have been applying and improving Pseudo-Likelihood Maximization(PLM) techniques to the inverse problem in multi-body models, representing systems with nonlinerar response in generic theories. Firstly, we have quantitatively measured the quality of the network reconstruction showing that both PLM methods analyzed allow to obtain an optimal reconstruction for these systems in all thermodynamic phases, when large enough number of samples M are available. Decreasing M the optimal reconstruction is better achieved around the finite size proxy to the critical temperature. Performing then a deeper analysis of the reconstruction error, which gives information on how far the inferred couplings are from the true couplings, reveals that with the decimation PLM the inferred values are closer to those ones of the original systems. Our analysis has been motivated by the study of lasers in the framework of statistical mechanics, though, looking at the models employed, Eqs (1 and 4), its range of applicability is more widespread and potentially involves many problems in which both nonlinear and multi-body contributions turn out to be relevant in determining the system behavior, see, e.g., refs 2, 9, 12–14, 17, 21. Focusing on optics, data from experiments would allow to identify active and passive mode-locking in multimode lasers in terms of mode-coupling coefficients. When more modes on the network graph are connected by a non-vanishing coupling they are matched in frequency, cf. Eq. (3), and therefore, beyond some critical point, they will be locked in phase. Configurations of magnitudes and phases can be obtained from the Fourier analysis of the pulses in ultrafast multimode lasers[53–55], pulses known to occur because of mode-locking, and parameters like the self amplitude modulation coefficients of saturable absorbers and the Kerr parameter can be inferred. In presence of relevant light scattering, instead, occurring in random lasers, no direct measurements of phases has been carried out so far, to our knowledge, but only spectral intensities, i.e., modes magnitudes. Acquiring also phases configurations, our inference technique would give the possibility to determine the strength of the interaction among the modes in the systems and to discriminate whether or not self-starting mode-locking occurs in random lasers.

References

1. Sargent III, M., O'Scully, M. & Lamb, W. E. *Laser Physics* (Addison Wesley Publishing Company, 1978).
2. Haus, H. A. Mode-Locking of Lasers. *IEEE J. Quantum Electron.* **6**, 1173–1185 (2000).
3. Gordon, A. & Fischer, B. Phase Transition Theory of Many-Mode Ordering and Pulse Formation in Lasers. *Phys. Rev. Lett.* **89**, 103901 (2002).
4. Katz, M., Gordon, A., Gat, O. & Fischer, B. Statistical theory of passive mode locking with general dispersion and Kerr effect. *Phys. Rev. Lett.* **97**, 113902 (2006).
5. Antenucci, F., Ibáñez Berganza, M. & Leuzzi, L. Statistical physical theory of mode-locking laser generation with a frequency comb. *Phys. Rev. A* **91**, 043811 (2015).
6. Antenucci, F., Ibañez Berganza, M. & Leuzzi, L. Statistical physics of nonlinear wave interaction. *Phys. Rev. B* **92**, 014204 (2015).
7. Lawandy, N. M., Balachandran, R. M., Gomes, A. S. L. & Sauvain, E. Laser action in strongly scattering media. *Nature* **368**, 436–438 (1994).
8. Cao, H. *et al.* Ultraviolet lasing in resonators formed by scattering in semiconductor polycrystalline films. *Appl. Phys. Lett.* **73**, 3656 (1998).
9. Wiersma, D. S. The physics and applications of random lasers. *Nature Physics* **4**, 359 (2008).
10. Antenucci, F., Conti, C., Crisanti, A. & Leuzzi, L. General Phase Diagram of Multimodal Ordered and Disor dered Lasers in Closed and Open Cavities. *Phys. Rev. Lett.* **114**, 043901 (2015).
11. Monasson, R. & Zecchina, R. Statistical mechanics of the random K-satisfiability model. *Phys. Rev. E* **56**, 1357 (1997).
12. Mézard, M., Parisi, G. & Zecchina, R. *Science* **297**, 812 (2002).
13. David, J. C. Mac Kay. *Information Theory*, *Inference*, *and Learning Algorithms* chap. 2 (Cambridge University Press (Cambridge, UK), 2003).
14. Mézard, M. & Montanari, A. *Information*, *Physics*, *and Computation* chap. 6 (Oxford University Press, 2009).
15. Kirkpatrick, T. R., Thirumalai, D. & Wolynes, P. G. Scaling concepts for the dynamics of viscous liquids near an ideal glassy state. *Phys. Rev. A* **40**, 1045 (1989).
16. Crisanti, A. & Sommers, H. The spherical p-spin interaction spin-glass model - the statics. *Z. Phys. B* **87**, 341 (1992).
17. Götze, W. *Complex Dynamics of Glass-Forming Liquids*: *A Mode-Coupling Theory* chap. 4 (Oxford University Press (Oxord, UK), 2009).
18. Franz, S., Parisi, G., Ricci-Tersenghi, F. & Rizzo, T. Field theory of fluctuations in glasses. *The European Physical Journal E* **34**, 1–17 (2011).
19. Caltagirone, F. *et al.* Critical Slowing Down Exponents of Mode Coupling Theory. *Phys. Rev. Lett.* **108**, 085702 (2012).
20. Ferrari, U., Leuzzi, L., Parisi, G. & Rizzo, T. Two-step relaxation next to dynamic arrest in mean-field glasses: Spherical and Ising p-spin model. *Phys. Rev. B* **86**, 014204 (2012).
21. Katz, Y., Tunstrom, K., Ioannou, C. C., Huepe, C. & Couzin, I. D. Inferring the structure and dynamics of interactions in schooling fish. **108**, 18720 (2011).
22. Herbert-Read, J. E. *et al.* Inferring the rules of interaction of shoaling fish. **108**, 18726 (2011).
23. Barber, D. *Bayesian Reasoning and Machine Learning* chap. 9 (Cambridge University Press (Cambridge, UK), 2012).
24. Ravikumar, P., Wainwright, M. J. & Lafferty, J. D. High-dimensional Ising model selection using l_1 regularized logistic regression. *Ann. Statist.* **38**, 1287–1319 (2010).
25. Aurell, E. & Ekeberg, M. Inverse Ising Inference Using All the Data. *Phys. Rev. Lett.* **108**, 090201 (2012).
26. Wasserman, L. *All of Statistics*: *A concise course in statistical inference* (Springer, New York, 2003).
27. Decelle, A. & Ricci-Tersenghi, F. Pseudolikelihood Decimation Algorithm Improving the Inference of the Inter action Network in a General Class of Ising Models. *Phys. Rev. Lett.* **112**, 070603 (2014).
28. Kuramoto, Y. Int. Symp. on Mathematical problems in theoretical physics. *Lect. N. Phys.* **39**, 420–422 (1975).
29. Antoni, M. & Ruffo, S. Clustering and relaxation in Hamiltonian long-range dynamics. *Phys. Rev. E* **52**, 2361 (1995).
30. Acebrón, J. A., Bonilla, L. L., Pérez Vicente, C. J., Ritort, F. & Spigler, R. The Kuramoto model: A simple paradigm for synchronization phenomena. *Rev. Mod. Phys.* **77**, 137 (2005).
31. Gupta, S., Campa, A. & Ruffo, S. Kuramoto model of synchronization: equilibrium and nonequilibrium aspects. *J. Stat. Mech.* R08001 (2014).
32. Fermi, E., Pasta, J. & Ulam, S. M. Studies of nonlinear problems. *LANL Report*. **1940** (1955).
33. Ortiz, G., Cobenera, E. & Nussinov, Z. In *40 Years of Berezinskii-Kosterlitz-Thouless Theory* (ed. José, J. V.) 93–134 (World Scientific Publisher, Singapore, 2013).

34. Marruzzo, A. & Leuzzi, L. Nonlinear XY and p-clock models on sparse random graphs: Mode-locking transition of localized waves. *Phys. Rev. B* **91**, 054201 (2015).
35. Potts, R. Some generalized order - disorder transformations. *Proc. Camb. Phil. Soc.* **48**, 106 (1952).
36. Antenucci, F., Crisanti, A. & Leuzzi, L. Complex spherical 2 + 4 spin glass: A model for nonlinear optics in random media. *Phys. Rev. A* **91**, 053816 (2015).
37. Boyd, R. W. *Nonlinear Optics* 2nd ed. Chap. 1 (Academic Press, New York).
38. Gordon, A. & Fischer, B. Phase transistion theory of pulse formation in passively mode-locked lasers with dispersion and Kerr nonlinearity. *Opt. Comm.* **223**, 151–156 (2003).
39. Weill, R., Rosen, A., Gordon, A., Gat, O. & Fischer, B. Critical Behavior of Light in Mode-Locked Lasers. *Phys. Rev. Lett.* **95**, 013903 (2005).
40. Weill, R., Fischer, B. & Gat, O. Light-Mode Condensation in Actively-Mode-Locked Lasers. *Phys. Rev. Lett.* **104**, 173901 (2010).
41. Angelani, L., Conti, C., Ruocco, G. & Zamponi, F. Glassy behavior of light in random lasers. *Phys. Rev. B* **74**, 104207 (2006).
42. Leuzzi, L., Conti, C., Folli, V., Angelani, L. & Ruocco, G. Phase Diagram and Complexity of Mode-Locked Lasers: From Order to Disorder. *Phys. Rev. Lett.* **102**, 083901 (2009).
43. Svelto, O. *Principles of lasers* (Springer, 1998).
44. Hugi, A., Vullare, G., S., B., Liu, H. C. & Fair, J. Mid-infrared frequency comb based on a quantum cascade laser. *Nature* **492**, 229–233 (2012).
45. Antenucci, F., Crisanti, A. & Leuzzi, L. The glassy random laser: replica symmetry breaking in the intensity fluctuations of emission spectra. *Scientific Reports* **5**, 16792 (2015).
46. Marruzzo, A. & Leuzzi, L. Multi-body quenched disordered *XY* and *p*-clock models on random graphs. *Phys. Rev. B* **93**, 094206 (2016).
47. Ghofraniha, N. *et al.* Experimental evidence of replica symmetry breaking in random lasers. *Nat. Commun.* **6**, 6058 (2014).
48. Gomes, A. S. L. *et al.* Random lasers, Lévy statistics and spin glasses: Synergy between photonics and complex systems. *arXiv:1509.00276* (2015).
49. Pincheira, P. I. R. *et al.* Replica symmetry breaking in specially designed *TiO*2 nanoparticles-based dye-colloidal random laser. *arXiv:1511.03087* (2015).
50. El-Dardiry, R. G. S., Mosk, A. P., Muskens, O. L. & Lagendijk, A. *Phys. Rev. A* **81**, 043830 (2010).
51. Tyagi, P., Marruzzo, A., Pagnani, A., Antenucci, F. & Leuzzi, L. Regularization and decimation pseudolikelihood approaches to statistical inference in XY spin models. *Phys. Rev. B* **94**, 024203 (2016).
52. Fisher, C. K. Variational Pseudolikelihood for Regularized Ising Inference. *arXiv:1409.7074* (2014).
53. Kane, D. J. & Trebino, R. Single-shot measurement of the intensity and phase of an arbitrary ultrashort pulse by using frequency-resolved optical gating. *Opt. Lett.* **18**, 823 (1993).
54. Trebino, R. *et al.* Measuring Ultrashort Laser Pulses in the Time-Frequency Domain Using Frequency-Resolved Optical Gating. *Rev. Sci. Instr.* **68**, 3277–3295 (1997).
55. Trebino, R. *Frequency-Resolved Optical Gating*: *The Measurement of Ultrashort Laser Pulses* (Springer, 2002).

Acknowledgements

We thank Federico Ricci Tersenghi for fruitful discussions. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program, Project *LoTGlasSy*, Grant Agreement No. 694925 and from the Italian Ministry for Education, University and Research (MIUR) under the PRIN 2015, Project *Statistical Mechanics and Complexity*, CINECA code 2015K7KK8L_005.

Author Contributions

All authors substantially contributed to this work. All authors read and approved the final manuscript. A.M., P.T., F.A., A.P. and L.L. work on the programs solving the inverse problems and on the data analysis as well as on the Monte Carlo simulations and on the graphs analysis. A.M., P.T., F.A., A.P. and L.L. reviewed and approved the final version of this manuscript.

Additional Information

Supplementary information accompanies this paper at doi:10.1038/s41598-017-03163-4

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2017