ORIGINAL PAPER

Optimizing optimization: accurate detection of hidden interactions in active body systems from noisy data

Chun-Wang Su · Zi-Gang Huang · Wen-Xu Wang · Jue Wang · Xiao-Fan Wang · Ying-Cheng Lai

Abstract Given deficient and noisy movement data from a pedestrian crowd—a class of active body systems, is it possible to uncover the hidden group interaction patterns or connections? Yes, it is possible. Here, we develop a general framework based on an optimal combination of the conventional compressive sensing (L_1 minimization) and L_2 optimization procedure to achieve optimal detection of the contact network embedded in pedestrian crowd under the data shortage conditions. Different from previous publications, in our framework, the optimal weights of the L_1 and L_2 components in the combination can be determined specifically from the noisy data, which can obtain more accurate detection for the corresponding system. To detect hidden interaction patterns from spatiotemporal data has broader applications, and our optimized compressive sensing-based framework provides a practically viable solution. In addition, we provide a relative entropy perspective to facilitate more general theoretical and technological extensions of the framework.

Keywords Active body system · Compressive sensing · L_1-regularized least squares · Optimal detection

C.-W. Su
Institute of Computational Physics and Complex Systems, Lanzhou University, Lanzhou 730000, Gansu, China

C.-W. Su · Z.-G. Huang (✉) · J. Wang
The Key Laboratory of Biomedical Information Engineering of Ministry of Education, The Key Laboratory of Neuro-informatics & Rehabilitation Engineering of Ministry of Civil Affairs, School of Life Science and Technology, Xi'an Jiaotong University, 710049 Xi'an, China
e-mail: huangzg@xjtu.edu.cn

W.-X. Wang
The State Key Laboratory of Cognitive Neuroscience and Learning, School of Systems Science, Beijing Normal University, Beijing 100875, China

X.-F. Wang
Department of Automation, Shanghai University, Shanghai 200444, China

Y.-C. Lai
School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ 85287, USA

Y.-C. Lai
Department of Physics, Arizona State University, Tempe, AZ 85287, USA

1 Introduction

Movements of people or pedestrian crowd constitute a complex spatiotemporal dynamical system, where the output data are time series of individuals' instantaneous positions and velocities. Such data sets have become widespread, but the available analysis tools have been limited. For example, a task of interest is to uncover any hidden connection patterns or the interaction networks among the individuals. The pedestrian interaction systems represent a prototypical class of active body

systems, in which the interactions among the individuals can be described by certain forms of social forces [1–4]. The resulting mutual interactions among the individuals lead to continuous updates of the underlying dynamical processes that can be measured and characterized by the velocities and positions of the individuals. For example, for a group of friends or acquaintances, the dynamical patterns of their movements can be quite distinct from those with complete strangers. Often, individuals in a pedestrian crowd tend to self-organize into mesoscopic groups with certain interaction structure. To decipher the mesoscopic interactions and then to map out the interaction network of the whole system are essential to detecting and analyzing the emergence of collective behaviors in the underlying complex system.

The discovery of the special groups and the interaction rules has generated considerable interest in active body systems [5–13], which are relevant to real-world systems such as fish shoals and insect swarms [14–17]. Most previous works were on developing and analyzing various statistical techniques that often require a large amount of data. In complex dynamical systems, recent years have witnessed a growing interest in solving a variety of inverse problems based on data, where the central goal is to uncover the equations governing the system dynamics and the connection topology among the components of the system [18–22]. A general class of methods that have proven to be effective in detecting system equations and network structures [22–35] is those based on sparse optimization paradigms such as compressive sensing [36–40] that typically require only limited amount of data. In this paper, to solve the problem of group and pattern detection in active body systems, we develop a general framework based on an optimal combination of compressive sensing (L_1) and the L_2 optimization procedure. The optimal weights of the L_1 and L_2 components can be determined solely from data through the basic optimization principle. Using pedestrian crowds as a prototypical system and based solely on data, we demonstrate that the framework can uncover the underlying interaction network, group structures, and collective behaviors efficiently and accurately, even when the available data are noisy and with small amount. Active body systems are ubiquitous not only in social systems, but also in physical, chemical, and biological systems [3–6,41]. The completely data-driven framework developed represents an effective approach to understanding, predicting, and even controlling active body systems.

2 Sparse signal detection based on compressive sensing

To solve the dynamical network reconstruction problem associated with active body system using compressive sensing, it is necessary to formulate the problem into the standard linear form: $Y = M \cdot a$, where $Y \in \mathbb{R}^{T \times 1}$ is a vector that can be obtained through data with T being the number of discrete data points sampled from the original time series data, $M \in \mathbb{R}^{T \times N}$ is a matrix that can also be calculated from the observed data, $a \in \mathbb{R}^{N \times 1}$ is the coefficient vector to be solved, whose N entries constitute a faithful representation (e.g., power series or Fourier expansion) of the original dynamical system that is assumed to be nonlinear and complex. The vector a is sparse when most of its elements are zero (Fig. 1a). In this case, application of compressive sensing [36–38] through minimizing the corresponding L_1 norm can lead to a unique solution for a. In particular, the solutions of $Y = M \cdot a$ form an $(N-T)$-dimensional hyperplane in the N-dimensional space. The norm contour surface of a is a polyhedron with vertices located on the axes. Let the polyhedron grow up from the origin, we get $\hat{a}$ with minimum L_1 norm at the first contact of the polyhedron with the solution plane: $\hat{a} \approx \arg \min_{Y=M \cdot a} ||a||_1$, as shown in Fig. 1b, even if the matrix M has a low rank in the sense that it has far more columns than rows.

The L_1 norm optimization method performs well for noise-free systems, while for noisy data, a more effective variant of compressive sensing named L_1- regularized least squares (L_1-RLS) [34,42] can be adopted. The L_1-RLS takes into account a term of L_2 norm to enhance the generalization ability of compressive sensing framework especially for systems with noise, which reads

$$\hat{a} \approx \arg \min \{||Y - M \cdot a||_2 + \lambda ||a||_1\}. \qquad (1)$$

Note that, under noise, we have $Y = M \cdot a + \xi$, for which the solution a is located somewhere inside a quasi-3D block, just like the shaded block in Fig. 1c, rather than the 2D plane $Y = M \cdot a$. The thickness of the 3D block is determined by the weight parameter λ in Eq. (1), where a larger value of λ makes the 3D block thicker. The weight parameter λ plays an important role
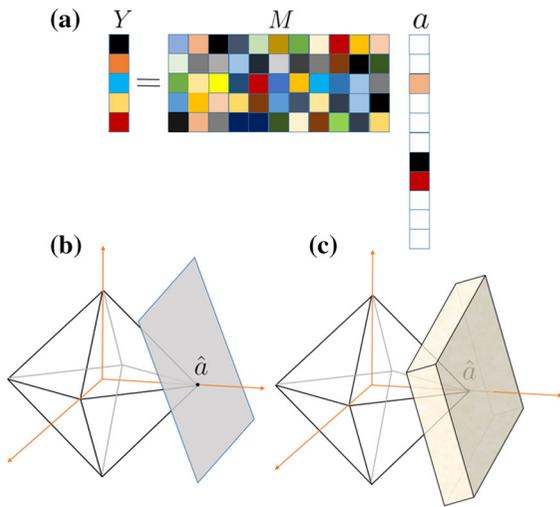
Fig. 1 (Color online) Sparse solution estimation by L_1 minimization and L_1-RLS. **a** Sketch map for the underdetermined linear inverse problem with the expected solution a to be sparse. **b** For a noise-free system with $Y = Ma$, L_1 minimization leads to a sparse solution on the solution plane. **c** Extracting the sparse solution by L_1-RLS from the noisy system data that $Y = Ma + \xi$, with ξ denoting the effect of noise. L_1-RLS defaults to assume the noise to be Gaussian

in adjusting the relative weights of the L_1 and L_2 terms in the objective function. If the value of λ is too large, e.g., as λ exceeds a threshold value λ_{th}, since the origin has already been contained in the 3D block, one would always obtain $\hat{a} = 0$ from Eq. (1). However, if λ is too small, the 3D block would be too thin, which leads to larger errors. The L_1-RLS optimization approaches the least squares procedure for further decreases of λ, say $\lambda \to 0$. Thus, a suitable weight parameter λ in the L_1-RLS is quite necessary.

3 L_1-RLS reconstruction for relationship network in active body system

Casting the reconstruction problem for an active body system into the standard form for compressive sensing requires some minimal rules about the underlying dynamics. As the system is intrinsically extremely complex, we resort to a physically intuitive approach. First, the interactions among the acquaintances are different from those among strangers. The system can then be regarded as a friendship network, where each link connects a pair of friends, and the number of links that agent i has is the degree k_i. Next, consider a pair of

friends (nodes): i and j. If they are apart, they tend to be attracted toward each other. If they are too close, they tend to repel each other to avoid invasion of personal space. The consideration suggests that the interaction between a pair of friends be modeled through the following form of spring type of "social force":

$$F_f(i, j) = \kappa (1 - \frac{r_0}{||r_{ji}||}) \frac{r_{ji}}{||r_{ji}||}, \qquad (2)$$

which is a function of the relative position between nodes i and j, i.e., $r_{ji} = r_j - r_i$ in the plane $r_i = (x_i, y_i)$. The parameters κ and r_0 are the elastic coefficient and the "comfortable" distance between two friends, respectively. Finally, for a pair of strangers, the interaction between them can be described by the empirical pedestrian interaction model [4] extracted from real data:

$$F_s(i, j) = - \left[\frac{ke^{-\tau/\tau_0}}{|v_{ij}|^2 \tau^2} \left(\frac{2}{\tau} + \frac{1}{\tau_0} \right) \right] \left[v_{ij} \right.$$
$$\left. - \frac{|v_{ij}|^2 x_{ij} - (v_{ij} \cdot x_{ij}) v_{ij}}{\sqrt{(v_{ij} \cdot x_{ij})^2 - |v_{ij}|^2 [|x_{ij}|^2 - (R_i + R_j)^2]}} \right], \qquad (3)$$

where $v_{ij} = v_i - v_j$ is the relative velocity of the two nodes, R is the equivalent radius of an agent, τ is the mean "free time" for the current movement state, which corresponds to the time duration between two collisions with other agents in the network. The parameters k and τ_0 are set as constants.

Let $r(t) = (x(t), y(t))$ and $v(t) = (v_x(t), v_y(t))$ be the measured position and velocity of agent i at a set of discretely sampled time: $t = t_1, t_2, \ldots, t_n$. The increment in the x-component of the velocity can be written as

$$\Delta v_{ix} = v_{ix}(t + \Delta t) - v_{ix}(t)$$
$$= \sum_{j \neq i} \left[a_{ij} \frac{F_{fx}(i, j)(t)}{m} \right.$$
$$\left. + (1 - a_{ij}) \frac{F_{sx}(i, j)(t)}{m} \right] \Delta t$$
$$= \sum_{j \neq i} a_{ij} M_{ij}(t) + R_i(t), \qquad (4)$$

where v_{ix}, F_{fx}, and F_{sx} are the x components of v_i, F_f, and F_s, respectively, m is a parameter characterizing *inertia*, Δt is the time step of the dynamics determined by the evolution rate of the system, $M_{ij}(t) = [F_{fx}(i, j) - F_{sx}(i, j)] \Delta t / m$, and $R_i(t) =$

Fig. 2 (Color online) Hidden connections in a dynamical pedestrian crowd. Agents are specified by circles or squares, and the interactions between friends, as characterized by Eq. (2), are represented by links. For example, the focal agent i (square) in the sketch map has five friends. Red and yellow links correspond to exclusion and attraction, respectively. The line width indicates the associated strength of the connection or interaction "force"
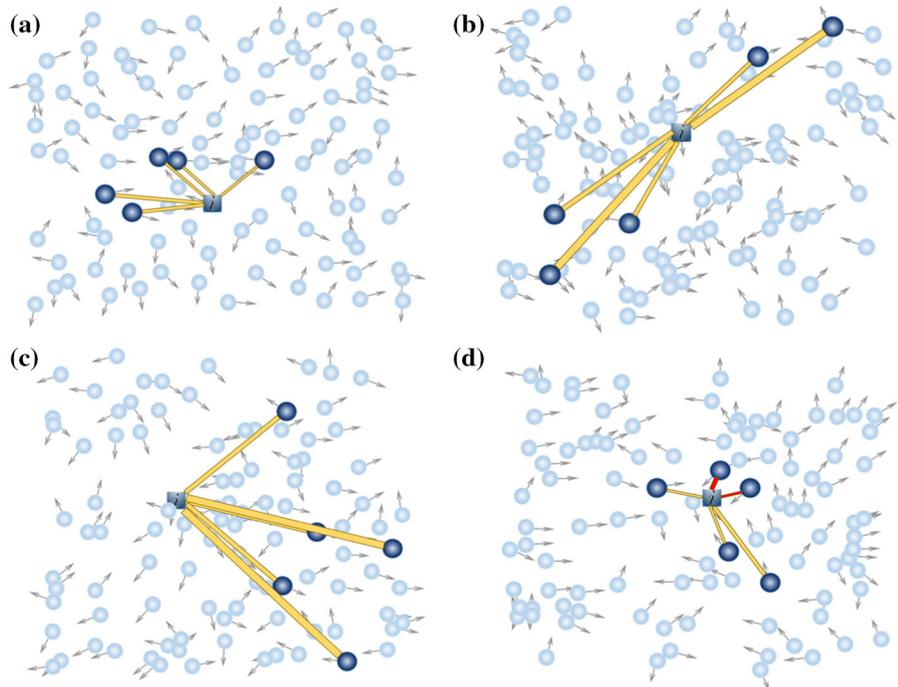
$\sum_{j \neq i} F_{sx}(i, j) \Delta t / m$. The matrix element a_{ij} indicates the possible existence of a link between agents i and j: If i and j are friends, i.e., they interact with each other according to Eq. (2), we have $a_{ij} = 1$. Otherwise, $a_{ij} = 0$. What is needed for Eq. (4) is pairs of successive data points with Δt apart, and these pairs do not need to be consecutive. Equation (4) can be written in matrix form as $\Delta v = M \cdot a + R$, where $Y \equiv \Delta v - R = M \cdot a$.

For any agent, the connection coefficient vector a characterizing the agent's relationship with other agents in the crowd is sparse because the expected number of friends k_i is generally much smaller than the system size N, making it applicable for the compressive sensing framework. Once Y and M have been calculated from data, the vector a for each agent can be estimated through L_1-RLS optimization. Matching the local connection structures of all agents gives the topology of the underlying network.

As a concrete example, we consider a crowd of $N = 100$ agents that move according to the interaction rules in Eqs. (2) and (3) in a 2D domain with periodic boundary conditions. Figure 2a–d shows examples of the friendship network of a focal agent revealed by our method. To characterize the performance, we define the following prediction errors for existent and nonexistent links: $\varepsilon_1 = (1/k_i) \sum_{j \in} |\widehat{a_{ij}} - a_{ij}|$ and $\varepsilon_0 = [1/(N - k_i - 1)] \sum_{j \notin} |\widehat{a_{ij}} - a_{ij}|$, respectively, where Ω is the set of k_i neighbors of node i. Figure 3a shows the errors of the L_1-RLS method versus the movement noise, whereas for comparison, the results from the conventional compressive sensing method are displayed in Fig. 3b. We see that the prediction error increases monotonically with $\tilde{\sigma}$. Due to noise, the conventional compressive sensing method leads to large errors, but those associated with the L_1-RLS method are much smaller. The movement noise refers to the noise or stochastic fluctuations in agents' velocity, with strength given by $\tilde{\sigma} \equiv \sigma / \langle|\Delta v|\rangle$, where σ is the standard deviation of Gaussian noise, and $\langle|\Delta v|\rangle$ is the absolute velocity increment averaged over all agents in the system. Figure 3c, d shows ε_1 and ε_0 versus the data amount T/N, respectively. In general, in the presence of noise, the errors with L_1-RLS are smaller than those with L_1 method.

4 Optimal detection based on L_1-RLS

We now address the issue of optimal detection. The key parameter in the L_1-RLS method (Eq. 1) is λ, on which the performance depends, especially when strong noise is present. Figure 4a shows the prediction error ε_1 ver-
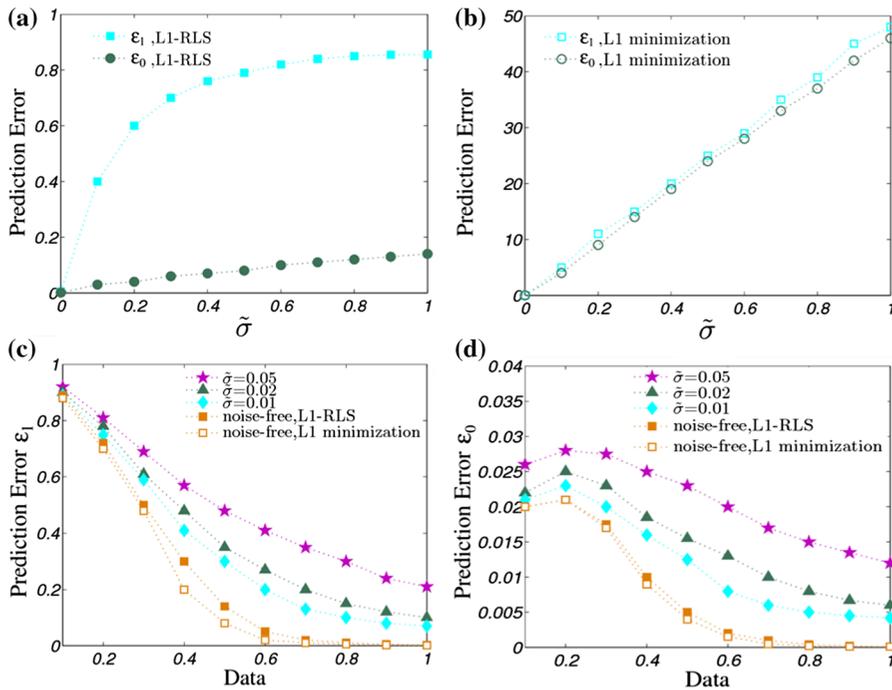
Fig. 3 (Color online) Compressive sensing-based reconstruction of hidden relationship in a pedestrian crowd. **a, b** Prediction error as a function of the relative noise strength $\tilde{\sigma}$, where ε_1 and ε_0 are the errors for the existent ($a_{ij} = 1$, friend) and nonexistent ($a_{ij} = 0$, stranger) links, respectively. The normalized amount of data for prediction is $T/N = 0.9$. **c, d** Errors ε_1 and ε_0 ver-

sus the normalized data amount T/N, respectively, for a number of values of $\tilde{\sigma}$ (0, 0.01, 0.02, and 0.05). The parameter setting is $N = 100$, $\kappa = 2$, $k = 0.5$, $\tau_0 = 0.5$, $R = 0.01$, $m = 60$, $\Delta t = 0.1$, and $\lambda = 10^{-3}$ for L_1-RLS. All the results are averaged over 1000 independent network realizations

sus $\eta \equiv -\log_{10} \lambda$, where the value of λ varies over several orders of magnitude. For large values of λ, the method yields $\hat{a} = 0$. As a result, the values of ε_1 are about one in the region of small η values. As η is increased from one, a non-monotonic behavior in ε_1 arises. The overall behavior of the error indicates the existence of an value of $\lambda_{\text{opt}} = 10^{-\eta_0}$, at which ε_1 is minimized. For values of η much larger than η_0, the error is saturated, where the contribution of $||a||_1$ item in Eq. (1) becomes negligible and the method degenerates to the conventional least squares.

From the result in the inset of Fig. 4d, we see that the optimal value λ_{opt} increases as the noise strength $\tilde{\sigma}$ becomes larger. This average result suggests that if you want to apply the L_1-RLS method to a system under stronger noise, you should take a larger value of λ. However, a complication is that the value of λ_{opt} may fluctuate in a wide range even for the same noise level. To determine the actual value of λ_{opt} is thus an issue of practical significance. One possible approach

is to find the zero point in the gradient of ε_1. As shown in Fig. 4b, the zero point can be obtained numerically based on the increment $\Delta \varepsilon_1$ as a function of η. However, in order to achieve this goal, the prior knowledge of the connection structure of the system as characterized by a is needed for calculating $\Delta \varepsilon_1$ and the value of λ_{opt}. But in realistic applications, the true vector a is not available and thus this way cannot get through. Is it possible to extract available information from the data itself? Now let us answer this question. The deviation of $\hat{a}$ from the actual vector due to the variance of λ (or equivalently, the variance of η) can be characterized by the following quantity:

$$\delta \hat{A}(\eta) \equiv \frac{1}{N-1} \sum_{j=1}^{N-1} |\widehat{a_{ij}}(\eta) - \widehat{a_{ij}}(\eta + \delta \eta)|, \qquad (5)$$

where $\widehat{a_{ij}}(\eta)$ is the L_1-RLS prediction for a specific η value, and $\delta \eta$ is a small positive variation. Figure 4c shows the value of $\delta \hat{A}(\eta)$ as a function of η, where it can be seen that the optimal parameter λ_{opt} for which
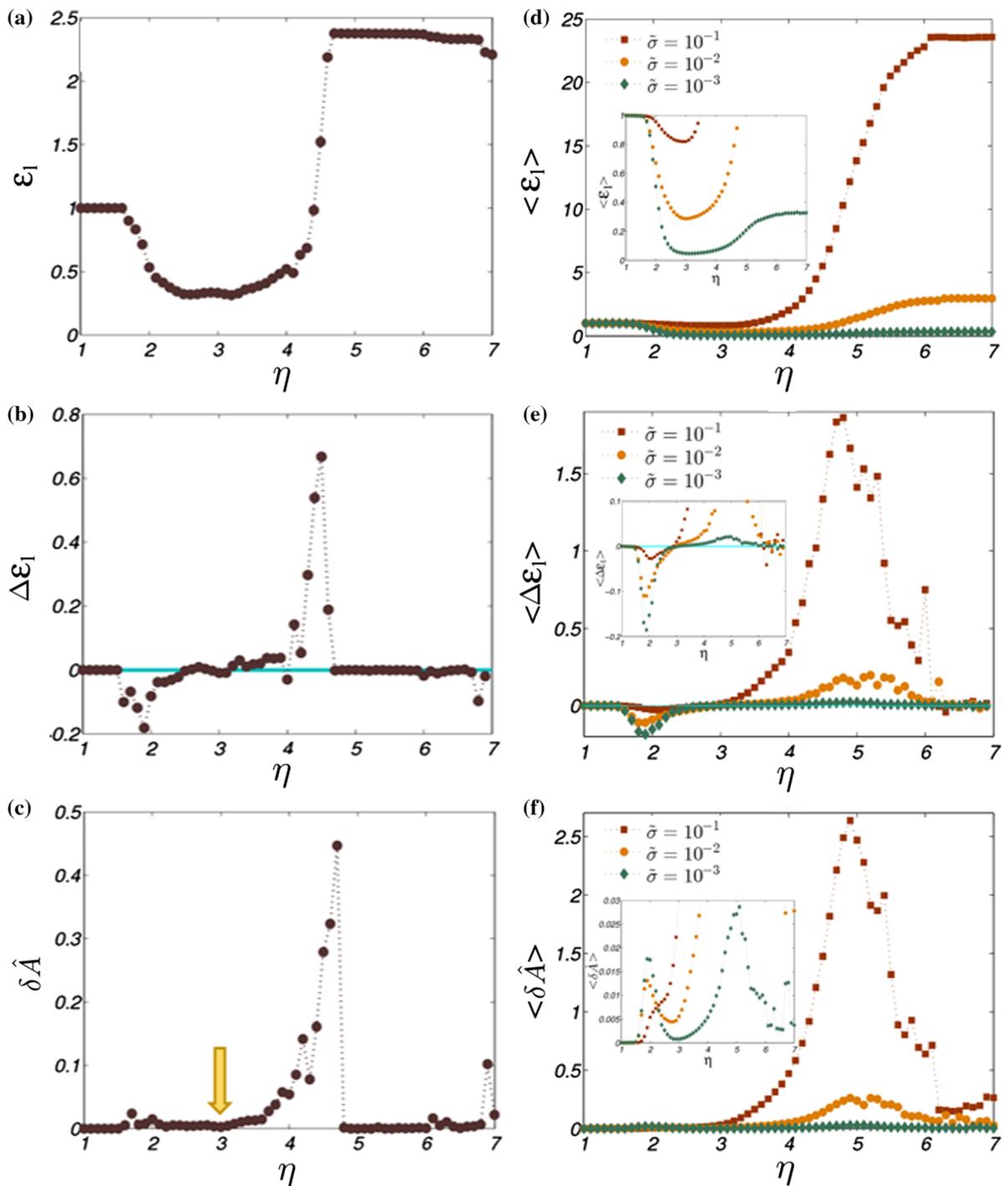
Fig. 4 (Color online) Optimizability of L_1- RLS: emergence of optimal weighting parameter λ_{opt}. **a** Prediction error ε_1 and **b** the corresponding increment $\Delta\varepsilon_1$ versus $\eta \equiv -\log_{10}\lambda$. **c** Value of $\delta\hat{A}$ versus η, in which the beginning point (marked by the arrow) of the large peak corresponds to the value η_0 for which ε_1 is minimized. Other parameters are $T = 0.9N$ and $\tilde{\sigma} = 10^{-2}$. **d–f** The average quantities corresponding to those in **a–c**, respectively, over 500 realizations

ε_1 is minimized (Fig. 4a) is related to the starting point of the large peak, as marked by the yellow arrow in Fig. 4c. The value of λ_{opt} can then be estimated through the following steps: (I) To calculate Y, M, and then $\hat{a}$ based on Eq. (1), (II) to plot $\delta \hat{A}(\eta)$ as a function of η (cf., Fig. 4c), and (III) to identify the starting point of the large peak. In addition, for weak noise, two large peaks of similar size can emerge. In this case, the second peak can be chosen to determine the value of λ_{opt}. The existence of an optimal value of λ thus makes L_1-RLS a powerful method for solving the data-based inverse problem of active body systems in a noisy environment.

5 Relative entropy interpretation for L_1-RLS

Theoretical insights into the emergence and determination of the optimal parameter λ_{opt} can be gained through the concept of relative entropy [43]. In particular, Eq. (1) can be rewritten as

$$\hat{a} \approx \arg \min \{||Y - M \cdot a||_2^2 + \lambda ||a||_1\}. \tag{6}$$

This form of objective function allows us to obtain an explicit relation between optimizing L_1-RLS and the relative entropy minimization. Assume that the elements a_j of the vector a follow the Laplace distribution [44]: $a_j \sim \mathrm{Laplace}(0, b)$ for $j = 1, 2, \ldots, N-1$. Further assume Gaussian distribution for the elements ξ_m of the bias vector $\xi = Y - M \cdot a$: $\xi_m \sim \mathbb{N}(0, \sigma^2)$ for $m = 1, 2, \ldots, T$. The weighted relative entropy for the empirical distributions $P(a)$ and $P(\xi)$ with respect to the model Laplace and Gaussian distributions [$P_L(a)$ and $P_G(\xi)$, respectively] can be written as

$$
\begin{aligned}
H(a) &= N \cdot \mathbb{E}[\log P(a) - \log P_L(a)] \\
&\quad + T \cdot \mathbb{E}[\log P(\xi) - \log P_G(\xi)] \\
&= -\{N \cdot \mathbb{E}[\log P_L(a)] + T \cdot \mathbb{E}[\log P_G(\xi)]\} + C \\
&= -\left[\sum_j \log P_L(a_j) + \sum_m \log P_G(\xi_m)\right] + C \\
&= [1/(2\sigma^2)](||Y - Ma||_2^2 + (2\sigma^2/b)||a||_1) + C
\end{aligned}
$$

where C is a constant and $\mathbb{E}[\cdot]$ denotes the expectation value. The weights N and T are the numbers of samples for a and ξ, respectively. The parameter b characterizes the sparsity of a: A sparser a corresponds to a smaller value of b. Remarkably, the form of the relative entropy $H(a)$ is identical to the objective function in Eq. (6), suggesting that minimization of $H(a)$ has a

similar effect as L_1-RLS optimization in Eq. (1). This similarity leads to $\lambda_{\mathrm{opt}} \sim 2\sigma^2/b$, a qualitative relation of the optimal value λ_{opt} with the sparsity parameter b and the noise strength σ. While the scaling relation for λ_{opt} is obtained under the assumptions of $P_{\mathrm{Laplace}}(a)$ and $P_{\mathrm{Gauss}}(\xi)$, these distributions are the basic approximations employed in the development of the compressive sensing framework. The scaling relation indicates that for sparser connection with larger noise variance σ^2, the optimal value λ_{opt} generally takes on larger values, which coincides qualitatively with the numerical results.

6 Summary and discussion

To summarize, we have developed an optimal compressive sensing-based framework for uncovering hidden connections (relationships) in pedestrian crowds based on individual movement data. The framework naturally inherits the appealing virtue of compressive sensing, which means the low data requirement. The remarkable feature is, by incorporating L_2 optimization into L_1-based compressive sensing, an optimal relative weighting of the two types of optimization arises, at which the framework can accurately predict the hidden networking relationships even in the presence of strong noise. We provide a practical method to estimate the optimal weights. This overcomes a fundamental difficulty in compressive sensing-based reconstruction method [22] that noise causes severe performance degradation. Active body systems arise in various fields of science and engineering and are a class of complex dynamical systems that have not been extensively studied. Our optimized compressive sensing framework represents the first attempt to address the inverse or reverse engineering problem for such systems.

Our relative entropy perspective excavates the intrinsic connotation for the L_1-RLS optimization, and the relationship $\lambda_{\mathrm{opt}} \sim 2\sigma^2/b$ we obtained has practical significance in promoting the feasibility of compressive sensing for real-world systems with noise. In the L_1-RLS optimization as Eq. (1) shows, the L_1 term is rooted in the fact that the solution to the problem is known a priori to be *sparse* and can be roughly described by Laplacian distribution, while the L_2 term is presented based on the central limit theorem which states that the system noise can be approximated as Gaussian. Actually, the framework we proposed for the given original inverse problem can also be extended to

more general cases, rather than limited to the form of Eq. (1). It depends on what the specific prior knowledge is. What is more, as the objective function is designed through combining different characteristics of the system according to the particular prior knowledge and the minimum relative entropy principle, our optimization technique can be adopted based on an appropriate weight parameter to explore the better optimal solution, namely to optimize optimization.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Helbing, D., Molnár, P.: Social force model for pedestrian dynamics. Phys. Rev. E **51**, 4282 (1995)
2. Lakoba, T.-I., Finkelstein, N.-M., Finkelstein, N.-M.: Modifications of the Helbing–Molnr–Farkas–Vicsek Social Force Model for Pedestrian Evolution. Society for Computer Simulation International, San Diego (2005)
3. Parisi, D.-R., Gilman, M., Moldovan, H.: A modification of the social force model can reproduce experimental data of pedestrian flows in normal conditions. Physica A **388**(17), 3600–3608 (2009)
4. Karamouzas, I., Skinner, B., Guy, S.-J.: Universal power law governing pedestrian interactions. Phys. Rev. Lett. **113**, 238701–238701 (2014)
5. Ryoo, M.S., Aggarwal, J.K.: Stochastic representation and recognition of high-level group activities. Int. J. Comput. Vis. **93**, 183–200 (2011)
6. Zhang, S.-P., Zhang, J.-Q., Huang, Z.-G., Guo, B.-H., Wu, Z.-X., Wang, J.: Collective behavior of AI population: transition from optimization to game. Nonlinear Dyn (2019). https://doi.org/10.1007/s11071-018-4649-4
7. Arnold, W., Vamsi, M., Wageeh, B., Prasad, Y.: A suspicious behaviour detection using a context space model for smart surveillance systems. Comput. Vis. Image Underst. **116**, 194–209 (2012)
8. Guo, P., Miao, Z.-J., Zhang, X.-P., Shen, Y., Wang, S.: Coupled observation decomposed hidden markov model for multiperson activity recognition. IEEE Trans. Circuits Syst. Video Technol. **22**, 1306–1320 (2012)
9. Li, R.-N., Chellappa, R., Zhou, S.-K.: Recognizing interactive group activities using temporal interaction matrices and their riemannian statistics. Int. J. Comput. Vis. **101**, 305–328 (2013)
10. Choi, W., Savarese, S.: Understanding collective activities of people from videos. IEEE Trans. Pattern Anal. Mach. Intell. **36**, 1242–1257 (2014)
11. Liu, L., Peng, Y.-X., Wang, S., Liu, M., Huang, Z.-G.: Complex activity recognition using time series pattern dictionary learned from ubiquitous sensors. Inf. Sci. **340–341**, 41–57 (2016)
12. Liu, L., Wang, S., Peng, Y.-X., Huang, Z.-G., Liu, M., Hu, B.: Mining intricate temporal rules for recognizing complex activities of daily living under uncertainty. Pattern Recognit. **60**, 1015–1028 (2016)
13. Liu, L., Wang, S., Su, G.-X., Huang, Z.-G., Liu, M.: Towards complex activity recognition using a Bayesian network-based probabilistic generative framework. Pattern Recognit. **68**, 295–309 (2017)
14. Szabó, G., Czárán, T.: Defensive alliances in spatial models of cyclical population interactions. Phys. Rev. E **64**, 042902 (2001)
15. Dobrinevski, A., Alava, M., Reichenbach, T., Frey, E.: Mobility-dependent selection of competing strategy associations. Phys. Rev. E **89**, 012721 (2014)
16. Szabó, P., Czárán, T., Szabó, G.: Competing associations in bacterial warfare with two toxins. J. Theor. Biol. **248**, 736–744 (2007)
17. Cheng, H.-Y., Yao, N., Huang, Z.-G., Park, J., Do, Y., Lai, Y.-C.: Mesoscopic interactions and species coexistence in evolutionary game dynamics of cyclic competitions. Sci. Rep. **4**, 7486 (2014)
18. Bertero, M., Bocacci, P.: Introduction to inverse problem in imaging. Opt. Photonics News **12**(10), 46–47 (1998)
19. Beck, A., Teboulle, M.: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J. Imaging Sci. **2**(1), 183–202 (2009)
20. Tropp, J.-A., Wright, S.-J.: Computational methods for sparse solution of linear inverse problems. Proc. IEEE **98**(6), 948–958 (2010)
21. Giménez, J., Botella-Estrada, R., Hernández, D., Carbonell, M., Martínez, M.-A., Guillén, C., Vázquez, C.: The convex geometry of linear inverse problems. Found. Comp. Math. **12**(6), 805–849 (2010)
22. Wang, W.-X., Lai, Y.-C., Grebogi, C.: Data based identification and prediction of nonlinear and complex dynamical systems. Phys. Rep. **644**, 1–76 (2016)
23. Babacan, S.-D., Molina, R., Katsaggelos, A.-K.: Bayesian compressive sensing using Laplace priors. IEEE Trans. Image Process. **19**(1), 53–63 (2009)
24. Ji, S., Dunson, D., Carin, L.: Multitask compressive sensing. IEEE Trans. Signal Process. **57**(1), 92–106 (2009)
25. Wright, J., Yang, A.-Y., Ganesh, A., Sastry, S.-S., Ma, Y.: Robust face recognition via sparse representation. IEEE Trans. Pattern Anal. Mach. Intell. **31**, 210–227 (2009)
26. Potter, L.-C., Ertin, E., Parker, J.-T., Cetin, M.: Sparsity and compressed sensing in radar imaging. Proc. IEEE **98**(6), 1006–1020 (2010)

27. Zibulevsky, M., Elad, M.: L1–L2 optimization in signal and image processing. IEEE Sig. Process. Mag. **27**(3), 76–88 (2010)

28. Wang, W.-X., Yang, R., Lai, Y.-C., Kovanis, V., Grebogi, C.: Predicting catastrophes in nonlinear dynamical systems by compressive sensing. Phys. Rev. Lett. **106**, 154101 (2011)

29. Wang, W.-X., Yang, R., Lai, Y.-C., Kovanis, V., Harrison, M.A.F.: Time-series-based prediction of complex oscillator networks via compressive sensing. Europhys. Lett. **94**, 48006 (2011)

30. Wang, W.-X., Lai, Y.-C., Grebogi, C., Ye, J.-P.: Network reconstruction based on evolutionary-game data via compressive sensing. Phys. Rev. X **1**, 1–7 (2011)

31. Walk, P., Jung, P.: Compressed sensing on the image of bilinear maps. In: IEEE International Symposium on Information Theory Proceedings, pp. 1291–1295 (2012)

32. Duarte, M.-F., Baraniuk, R.-G.: Spectral compressive sensing. Appl. Comput. Harmon. Anal. **35**(1), 111–129 (2013)

33. Tang, S.-Q., Shen, Z.-S., Wang, W.-X., Di, Z.-R.: Uncovering transportation networks from traffic flux by compressed sensing. Euro. Phys. J. B **88**, 1–7 (2015)

34. Han, X., Shen, Z.-S., Wang, W.-X., Di, Z.-R.: Robust reconstruction of complex networks from sparse data. Phys. Rev. Lett. **114**, 028701 (2015)

35. Su, R.-Q., Wang, W.-X., Wang, X., Lai, Y.-C.: Data-based reconstruction of complex geospatial networks, nodal positioning and detection of hidden nodes. R. Soc. Open Sci. **3**, 150577 (2016)

36. Donoho, D.-L.: Compressed sensing. IEEE Trans. Inf. Theory **52**, 1289–1306 (2006)

37. Candes, E.-J., Wakin, M.-B.: An introduction to compressive sampling. IEEE Signal Process. Mag. **25**, 21–30 (2008)

38. Duarte, M.-F., Eldar, Y.-C.: Structured compressed sensing: from theory to applications. IEEE Trans. Signal Process. **59**, 4053–4085 (2011)

39. Gan, H.-P., Li, Z., Li, J., Wang, X., Cheng, Z.-F.: Compressive sensing using chaotic sequence based on Chebyshev map. Nonlinear Dyn. **78**, 2429–2438 (2014)

40. Abdechiri, M., Faez, K., Amindavar, H., Bilotta, E.: The chaotic dynamics of high-dimensional systems. Nonlinear Dyn. **87**, 1–14 (2016)

41. Zhang, S.-P., Dong, J.-Q., Liu, L., Huang, Z.-G., Huang, L., Lai, Y.-C.: Reinforcement learning meets minority game: toward optimal resource allocation. Phys. Rev. E, to be published

42. Kim, S.-J., Koh, K., Lustig, M., Boyd, S.: An interior-point method for large-scale l_1-regularized least squares. J. Mach. Learn. Res. **8**, 1519–1555 (2007)

43. Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. MIT Press, Cambridge (2016)

44. Zhao, Q., Meng, D.-Y., Xu, Z.-B., Zuo, W.-M., Yan, Y.: l_1-norm low-rank matrix factorization by variational Bayesian method. IEEE Trans. Neural Netw. Learn. Syst. **26**, 825–839 (2015)