Statistics and Probability Letters 171 (2021) 109016

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Exponential decay of pairwise correlation in Gaussian graphical models with an equicorrelational one-dimensional connection pattern

Guillaume Marrelec [a,b,*], Alain Giron [a,b], Laura Messio [c,d]

[a] Sorbonne Université, CNRS, INSERM, Laboratoire d'imagerie biomédicale, LIB, F-75006, Paris, France
[b] Centre de recherches et d'études en sciences des interactions (CRÉSI) – Center for Interaction Science (CIS), F-75006, Paris, France
[c] Sorbonne Université, CNRS, Laboratoire de Physique Théorique de la Matière Condensée, LPTMC, F-75005 Paris, France
[d] Institut Universitaire de France, IUF, F-75005 Paris, France

ARTICLE INFO

Article history:
Received 16 June 2020
Received in revised form 17 November 2020
Accepted 25 November 2020
Available online 13 December 2020

Keywords:
Gaussian graphical model
Multivariate normal distributions
Conditional independence graph
Equicorrelational one-dimensional connection pattern
Tridiagonal matrix
Gaussian free fields

ABSTRACT

We consider a Gaussian graphical model associated with an equicorrelational and one-dimensional conditional independence graph. We show that pairwise correlation decays exponentially as a function of distance. We also provide a limit when the number of variables tend to infinity and quantify the difference between the finite and infinite cases.

1. Introduction

Let $\boldsymbol{X} = (X_1, \ldots, X_n)$ be an n-dimensional variable. A conditional independence graph on $\boldsymbol{X}$ is a graphical representation of $\boldsymbol{X}$ which emphasizes the relationships of conditional independence between the X_i's (Whittaker, 1990). More precisely, there is *no* link between nodes i and j if X_i and X_j are conditionally independent given $\boldsymbol{X}_{[n]\setminus\{i,j\}}$, denoted as $X_i \perp\!\!\!\perp X_j | \boldsymbol{X}_{[n]\setminus\{i,j\}}$. In the particular case where $\boldsymbol{X}$ is a multivariate normal distribution, we refer to Gaussian graphical models (Uhler, 2017). Let then $\boldsymbol{X}$ be a Gaussian graphical model characterized by its covariance matrix $\boldsymbol{\Sigma} = (\Sigma_{ij})$, or, equivalently, its precision (or concentration) matrix $\boldsymbol{\Upsilon} = (\Upsilon_{ij}) = \boldsymbol{\Sigma}^{-1}$. Two other key quantities are the pairwise correlation matrix $\boldsymbol{\Omega} = (\Omega_{ij})$, defined as $\Omega_{ij} = \Sigma_{ij}/\sqrt{\Sigma_{ii}\Sigma_{jj}}$ for $i \neq j$ and $\Omega_{ii} = 1$, as well as the partial correlation matrix $\boldsymbol{\Pi} = (\Pi_{ij})$, defined as $\Pi_{ij} = -\Upsilon_{ij}/\sqrt{\Upsilon_{ii}\Upsilon_{jj}}$ for $i \neq j$ and $\Pi_{ii} = 1$. Then, for $i \neq j$, the relationship of conditional independence $X_i \perp\!\!\!\perp X_j | \boldsymbol{X}_{[n]\setminus\{i,j\}}$ is equivalent to $\Upsilon_{ij} = 0$ and $\Pi_{ij} = 0$ (Whittaker, 1990, Chap. 6).

Our interest in Gaussian graphical models originates from statistical mechanics, where the Ising model and its various extensions (Potts model, XY model, Heisenberg model, n-vector model, ϕ^4 model) are used to investigate the behavior of variables related through various connection patterns. One extension of the Ising model to continuous real variables

* Corresponding author at: Sorbonne Université, CNRS, INSERM, Laboratoire d'imagerie biomédicale, LIB, F-75006, Paris, France.
E-mail address: guillaume.marrelec@inserm.fr (G. Marrelec).

https://doi.org/10.1016/j.spl.2020.109016

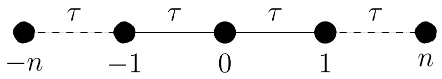

Fig. 1. A conditional independence graph whose limit when $n \to \infty$ yields the one-dimensional Gaussian free field.

with noncompact support is the so-called Gaussian free field model (Friedli and Velenik, 2017, Chap. 8). In this case, each vertex $i \in \mathbb{Z}^d$ is associated with a real-valued variable x_i and the corresponding Hamiltonian is of the form

$$\frac{\beta}{4d} \sum_{i,j\in\mathbb{Z}^d:\|i-j\|_2=1} (x_i - x_j)^2 + \frac{m^2}{2} \sum_{i\in\mathbb{Z}^d} x_i^2,$$

where $\beta \geq 0$ is the inverse temperature and $m \geq 0$ is the mass. In massive models ($m > 0$), pairwise correlation is known to decrease exponentially with distance (Friedli and Velenik, 2017, Prop. 8.30).

While this result is shown in the "thermodynamic limit", that is, for an infinite-dimensional variable (i.e., on $\mathbb{Z}^d$), we are here interested in the finite case. The reason for this interest is twofold. First, a main way to approach statistical mechanics is through simulations, which only deal with finite case scenarios. It is therefore important to understand what the expected behavior of the system should be in such cases. Does pairwise correlation also decay exponentially? Also, we would like to gain a sense of how convergence from the finite to the infinite case occurs through some results regarding the speed of convergence.

In the present study, we focus on the unidimensional case ($d = 1$) and consider the particular case of a (finite) Gaussian graphical model on $\boldsymbol{X}^{(n)}$ with an equicorrelational one-dimensional connection pattern between the $X_i^{(n)}$'s, as represented in Fig. 1. Such a conditional independence graph entails that the Gaussian graphical model has a tridiagonal partial correlation matrix with an off-diagonal element τ that can be related to the parameters of the one-dimensional Gaussian free field by

$$\tau = \frac{\frac{\beta}{4d}}{\frac{2\beta}{4d} + \frac{m^2}{2}}. \tag{1}$$

We here restrict ourselves to the case $\tau > 0$ and only consider diagonally dominant matrices, leading to $0 \leq \tau < 1/2$ (which corresponds to the massive case, $m > 0$). Under these assumptions, we show that $\Omega_{ij}^{(n)}$, the pairwise correlation between any two variables $X_i^{(n)}$ and $X_j^{(n)}$, decreases exponentially with the distance $|j - i|$ between variables, with a rate given by

$$\lambda = \arg\cosh\left(\frac{1}{2\tau}\right). \tag{2}$$

More specifically, we show the following theorem.

Theorem 1. *Let $\boldsymbol{X}^{(n)}$ be a Gaussian graphical model with conditional independence graph given by Fig. 1. Then the following results yield:*

- $0 < \Omega_{ij}^{(n)} < e^{-|j-i|\lambda}$ *for all n;*
- $\Omega_{ij}^{(n)} \to e^{-|j-i|\lambda}$ *when $n \to \infty$;*
- *The absolute error $\Omega_{ij}^{(n)} - e^{-|j-i|\lambda}$ is $O\left[e^{-2(n+1)\lambda}\right]$ when $n \to \infty$;*
- *The relative error $\Omega_{ij}^{(n)} e^{|j-i|\lambda} - 1$ is equal to*

$$-\left\{\sinh[2\max(i,j)\lambda] - \sinh[2\min(i,j)\lambda]\right\} e^{-2(n+1)\lambda} + o\left[e^{-2(n+1)\lambda}\right]$$

when $n \to \infty$.

Here, $O(\cdot)$ and $o(\cdot)$ are the usual big-O and little-o Bachmann–Landau notations, respectively, with

$$u_n = O(v_n) \qquad \Leftrightarrow \qquad \exists n_0, c \quad |u_n| < c\,|v_n| \ \forall n > n_0$$

and

$$u_n = o(v_n) \qquad \Leftrightarrow \qquad \frac{u_n}{v_n} \overset{n\to\infty}{\to} 0.$$

2. Proof of theorem

We start by expressing pairwise correlation in the case of the simpler model of an n-dimensional Gaussian graphical model $\boldsymbol{Y}^{(n)}$ with conditional independence graph given by Fig. 2. We then relate the pairwise correlations for both models and derive the results for $\boldsymbol{X}^{(n)}$.

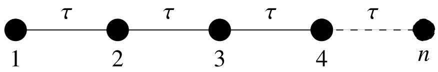

Fig. 2. Conditional independence graph of the Gaussian graphical model $\boldsymbol{Y}^{(n)}$.

2.1. Partial correlation matrix

Assume that $\boldsymbol{Y}^{(n)}$ is a Gaussian graphical model with conditional independence graph given by Fig. 2. The corresponding partial correlation matrix is then given by the following n-by-n symmetric tridiagonal matrix

$$\Pi_{ij}^{(n)} = \begin{cases} 1 & \text{if } i = j \\ \tau & \text{if } |j-i| = 1 \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

2.2. From partial to pairwise correlation

Letting $\boldsymbol{I}_n$ be the n-by-n identity matrix and setting $\boldsymbol{\Upsilon}^{(n)} = 2\boldsymbol{I}_n - \boldsymbol{\Pi}^{(n)}$, the (pairwise) correlation matrix $\boldsymbol{\Psi}^{(n)} = (\Psi_{ij}^{(n)})$ corresponding to the distribution can be obtained in two steps:

1. Invert $\boldsymbol{\Upsilon}^{(n)}$ to obtain $\boldsymbol{\Sigma}^{(n)} = \boldsymbol{\Upsilon}^{(n)^{-1}}$;
2. Decompose $\boldsymbol{\Sigma}^{(n)} = (\Sigma_{ij}^{(n)})$ using the correlation transform:

$$\boldsymbol{\Sigma}^{(n)} = \boldsymbol{\Delta}^{(n)}\boldsymbol{\Psi}^{(n)}\boldsymbol{\Delta}^{(n)},$$

where $\boldsymbol{\Delta}^{(n)} = (\Delta_{ij}^{(n)})$ is a diagonal matrix with $\Delta_{ii}^{(n)} = \sqrt{\Sigma_{ii}^{(n)}}$.

2.3. Expression of $\Psi_{ij}^{(n)}$

If $\boldsymbol{\Pi}^{(n)}$ has the form of Eq. (3), then $\boldsymbol{\Upsilon}^{(n)}$ is also a tridiagonal matrix with off-diagonal element equal to $-\tau$. Defining λ as in Eq. (2) and applying results from Hu and O'Connell (1996), we obtain that

$$\Sigma_{ij}^{(n)} = \frac{1}{\tau}\,\frac{\cosh[(n+1-|j-i|)\lambda] - \cosh[(n+1-i-j)\lambda]}{2\sinh(\lambda)\,\sinh[(n+1)\lambda]}.$$

Using a basic identity of hyperbolic functions (Gradshteryn and Ryzhik, 2007, §1.314),

$$\cosh(x) - \cosh(y) = 2\sinh\left(\frac{x+y}{2}\right)\sinh\left(\frac{x-y}{2}\right),$$

we obtain

$$\Sigma_{ij}^{(n)} = \frac{1}{\tau}\,\frac{\sinh\left[\frac{2(n+1)-i-j-|j-i|}{2}\lambda\right]\sinh\left[\frac{i+j-|j-i|}{2}\lambda\right]}{\sinh(\lambda)\,\sinh[(n+1)\lambda]}.$$

In particular, the diagonal elements read

$$\Sigma_{ii}^{(n)} = \frac{1}{\tau}\,\frac{\sinh[(n+1-i)\lambda]\,\sinh(i\lambda)}{\sinh(\lambda)\,\sinh[(n+1)\lambda]}.$$

This leads to the following expression for the correlation coefficient

$$\Psi_{ij}^{(n)} = \frac{\sinh\left[\frac{2(n+1)-i-j-|j-i|}{2}\lambda\right]\sinh\left[\frac{i+j-|j-i|}{2}\lambda\right]}{\sqrt{\sinh[(n+1-i)\lambda]\,\sinh(i\lambda)}\sqrt{\sinh[(n+1-j)\lambda]\,\sinh(j\lambda)}}.$$

In the following, we will restrict our attention to $i < j$ without loss of generality. For $j < i$, we can then use the symmetry identity $\Psi_{ij}^{(n)} = \Psi_{ji}^{(n)}$. So, if $j > i$, the previous result can be simplified to yield

$$\begin{aligned}\Psi_{ij}^{(n)} &= \frac{\sinh[(n+1-j)\lambda]\,\sinh(i\lambda)}{\sqrt{\sinh[(n+1-i)\lambda]\,\sinh(i\lambda)}\sqrt{\sinh[(n+1-j)\lambda]\,\sinh(j\lambda)}} \\ &= \sqrt{\frac{\sinh[(n+1-j)\lambda]\,\sinh(i\lambda)}{\sinh[(n+1-i)\lambda]\,\sinh(j\lambda)}} \\ &= e^{-\lambda(j-i)}\sqrt{\frac{1-e^{-2[(n+1-j)]\lambda}}{1-e^{-2[(n+1-i)]\lambda}}\,\frac{1-e^{-2i\lambda}}{1-e^{-2j\lambda}}}.\end{aligned} \tag{4}$$

2.4. Connection between $\boldsymbol{Y}^{(n)}$ and $\boldsymbol{X}^{(n)}$

Gaussian free fields can be obtained as the limit when $n \to \infty$ of a $(2n+1)$-dimensional variables $\boldsymbol{X}^{(n)} = (X_{-n}, \dots, X_{-1}, X_0, X_1, \dots, X_n)$ with a conditional independence graph given by Fig. 1. Results regarding this model can be derived from the previous model and calculations by replacing n with $2n+1$ and considering pairwise correlations of the form $\Omega_{ij}^{(n)} \equiv \Psi_{n+1+i,n+1+j}^{(2n+1)}$. In this perspective, Eq. (4) leads to, for $i < j$,

$$\Omega_{ij}^{(n)} = e^{-(j-i)\lambda} \sqrt{\frac{1-e^{-2[(n+1-j)]\lambda}}{1-e^{-2[(n+1-i)]\lambda}} \frac{1-e^{-2(n+1+i)\lambda}}{1-e^{-2(n+1+j)\lambda}}}. \tag{5}$$

2.5. Bounds

From Eq. (5), it is straightforward to see that $\Omega_{ij}^{(n)}$ is always strictly positive. Also, since $u \mapsto 1-e^{-2(n+1-u)\lambda}$ is a strictly increasing function of u, and $u \mapsto 1-e^{-2(n+1-u)\lambda}$ a strictly decreasing function of u, we obtain for $i < j$

$$\sqrt{\frac{1-e^{-2[(n+1-j)]\lambda}}{1-e^{-2[(n+1-i)]\lambda}}} < 1 \quad \text{and} \quad \sqrt{\frac{1-e^{-2(n+1+i)\lambda}}{1-e^{-2(n+1+j)\lambda}}} < 1,$$

so that

$$0 < \Omega_{ij}^{(n)} < e^{-(j-i)\lambda}$$

for all n.

2.6. Asymptotics

We can now provide the limit of $\Omega_{ij}^{(n)}$ when $n \to \infty$. Using the fact that $1-e^{-2[(n+1-u)]\lambda}$ tends to 1 when $n \to \infty$ for a given u, Eq. (5) leads to

$$\Omega_{ij}^{(n)} \overset{n\to\infty}{\to} e^{-(j-i)\lambda}. \tag{6}$$

Besides, using the following Taylor expansion for $u \to 0$,

$$(1+u)^k = 1 + ku + o(u), \tag{7}$$

we can express $\Omega_{ij}^{(n)}/e^{-(j-i)\lambda}$ as

$$\begin{aligned}
\Omega_{ij}^{(n)} e^{(j-i)\lambda} &= \left[1-e^{-2(n+1-j)\lambda}\right]^{\frac{1}{2}} \left[1-e^{-2(n+1-i)\lambda}\right]^{-\frac{1}{2}} \left[1-e^{-2(n+1+i)\lambda}\right]^{\frac{1}{2}} \left[1-e^{-2(n+1+j)\lambda}\right]^{-\frac{1}{2}} \\
&= \left\{1 - \frac{1}{2} e^{-2(n+1-j)\lambda} + o\left[e^{-2(n+1)\lambda}\right]\right\} \left\{1 + \frac{1}{2} e^{-2(n+1-i)\lambda} + o\left[e^{-2(n+1)\lambda}\right]\right\} \\
&\quad \times \left\{1 - \frac{1}{2} e^{-2(n+1+i)\lambda} + o\left[e^{-2(n+1)\lambda}\right]\right\} \left\{1 + \frac{1}{2} e^{-2(n+1+j)\lambda} + o\left[e^{-2(n+1)\lambda}\right]\right\} \\
&= 1 - \frac{1}{2}\left[e^{2j\lambda} - e^{2i\lambda} + e^{-2i\lambda} - e^{-2j\lambda}\right] e^{-2(n+1)\lambda} + o\left[e^{-2(n+1)\lambda}\right] \\
&= 1 - \left[\sinh(2j\lambda) - \sinh(2i\lambda)\right] e^{-2(n+1)\lambda} + o\left[e^{-2(n+1)\lambda}\right].
\end{aligned}$$

We therefore have that

$$\Omega_{ij}^{(n)} e^{(j-i)\lambda} = 1 + O\left[e^{-2(n+1)\lambda}\right],$$

so that

$$\Omega_{ij}^{(n)} - e^{-(j-i)\lambda} = e^{-(j-i)\lambda}\left[\Omega_{ij}^{(n)} e^{(j-i)\lambda} - 1\right] = O\left[e^{-2(n+1)\lambda}\right].$$

2.7. General results

All results were proved for $i < j$. As mentioned earlier, the case $j < i$ can be solved by using the symmetry identity $\Omega_{ij}^{(n)} = \Omega_{ji}^{(n)}$. The most general results can therefore be expressed by replacing i with $\min(i,j)$, j with $\max(i,j)$, and $j-i$ with $|j-i|$, leading to

- Bounds: $0 < \Omega_{ij}^{(n)} < e^{-|j-i|\lambda}$ for all n;
- Limit: $\Omega_{ij}^{(n)} \to e^{-|j-i|\lambda}$ when $n \to \infty$;

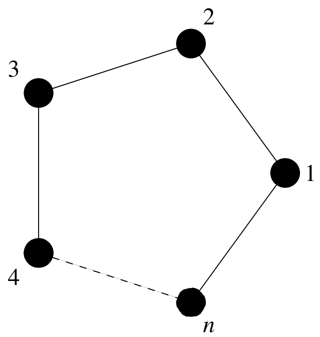

Fig. 3. Another instance of conditional independence graph whose limit when $n \to \infty$ yields the one-dimensional Gaussian free field. Such graph corresponds to a symmetric circulant partial correlation matrix.

- Asymptotic expansion: the absolute error $\Omega_{ij}^{(n)} - e^{-|j-i|\lambda}$ is $O\left[e^{-2(n+1)\lambda}\right]$, and the relative error is given by

$$\Omega_{ij}^{(n)} e^{(j-i)\lambda} - 1 = -\left\{\sinh[2\max(i,j)\lambda] - \sinh[2\min(i,j)\lambda]\right\} e^{-2(n+1)\lambda} + o\left[e^{-2(n+1)\lambda}\right].$$

3. Discussion

In the present manuscript, we considered a (finite-dimensional) Gaussian graphical model with the conditional independence graph depicted in Fig. 1. We proved that the pairwise correlation decays exponentially at a rate given by λ of Eq. (2). We also provided bounds for pairwise correlation as well as asymptotic expansions of the absolute and relative errors.

These results are in line with what is known about the one-dimensional Gaussian free field. Indeed, setting $\beta = 1$, pairwise correlation is known to be of the form $\exp(-\xi_m|j-i|)$ with (Friedli and Velenik, 2017, Th. 8.33)

$$\xi_m = \ln(1 + m^2 + \sqrt{2m^2 + m^4}).$$

Using the relationship between τ and (β, m) of Eq. (1) as well as the expression of arg cosh in terms of logarithm (Gradshteryn and Ryzhik, 2007, §1.622), it can be shown that ξ_m corresponds to our λ.

Another quantity of interest is $\alpha = e^{-\lambda}$, which can be expressed using again the expression of arg cosh in terms of logarithm (Gradshteryn and Ryzhik, 2007, §1.622), leading to

$$\frac{1}{\alpha} = \frac{1 + \sqrt{1 - 4\tau^2}}{2\tau},$$

or equivalently

$$\alpha = \frac{1 - \sqrt{1 - 4\tau^2}}{2\tau}. \tag{8}$$

From the definition, it is obvious that $\alpha \in [0, 1)$, and that pairwise correlation decreases as $\alpha^{|j-i|}$. α appears naturally in the case where the Gaussian graphical model has a partial correlation matrix that is circulant instead of tridiagonal (see below).

One could wonder what the results are for the Gaussian graphical model $\boldsymbol{Y}^{(n)}$ with conditional independence graph of Fig. 2 that was used to derive our main results. The corresponding results are given in §1 of the online supplement. They are more complex due to the proximity of the boundary point 0 to i and j.

Another finite pattern of conditional independence that would lead to one-dimensional Gaussian free fields is the one given in Fig. 3. In this case, the partial correlation matrix is symmetric circulant and it can be shown that the pairwise correlation still decays exponentially with the same rate λ (see online supplement, §2). However, we were able to provide neither bounds nor an asymptotic expansion in that particular case.

Our results show that pairwise correlation in (finite-dimensional) Gaussian graphical models behaves in a manner very similar to one-dimensional (infinite-dimensional) Gaussian free fields, the difference between both cases decreasing exponentially with n. As a consequence, computer simulations can be trusted to provide precise approximations for the behavior of one-dimensional Gaussian free fields.

Beyond pairwise correlation, a measure that we think would be relevant to quantify the global level of dependence within the system is a multivariate generalization of mutual information known as total correlation (Watanabe, 1960), multivariate constraint (Garner, 1962), δ (Joe, 1989), or multiinformation (Studeny, 1998). In the case of multivariate normal distributions, this measure has a simple expression in terms of the covariance matrix. While we were able to provide neither the closed form expression in the case of tridiagonal nor circulant partial correlation matrices, we believe that such expressions might be helpful to understand the global behavior of the system.

Now that we have solved the case $d = 1$, we would like to investigate more general cases with more complex connectivity patterns, still in the case of a finite n. Note that a major advantage of multivariate normal distributions is that their structures of conditional independence can be read off their precision matrices. For instance, moving from a one-dimensional to a two-dimensional model simply implies to change from a tridiagonal partial correlation matrix to a partial correlation matrix with more non-zero bands. More complex connectivity patterns with specific features (e.g., random or small world) simply translate into different patterns in the precision matrix which can then be investigated either analytically or through computer simulations. And, again, multiinformation could provide interesting insight into the global behavior of the system.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.spl.2020.109016.

References

Friedli, S., Velenik, Y., 2017. Statistical Mechanics of Lattice Systems: A Concrete Mathematical Introduction. Cambridge University Press, Cambridge.
Garner, W.R., 1962. Uncertainty and Structure As Psychological Concepts. John Wiley & Sons, New York.
Gradshteryn, I.S., Ryzhik, I.M., 2007. Table of Integrals, Series, and Product, seventh ed. Academic Press.
Hu, G.Y., O'Connell, R.F., 1996. Analytical inversion of symmetric tridiagonal matrices. J. Phys. A: Math. Gen. 29, 1511.
Joe, H., 1989. Estimation of entropy and other functionals of a multivariate density. Ann. Inst. Statist. Math. 41, 683–697.
Studeny, M., 1998. Complexity of structural models. In: Proccedings of the Joint Session of the 6th Prague Conference on Asymptotic Statistics and the 13th Prague Conference on Information Theory, Statistical Decision Functions and Random Processes, pp. 23–28.
Uhler, C., 2017. Gaussian graphical models: an algebraic and geometric perspective. arXiv:1707.04345 [math.ST].
Watanabe, S., 1960. Information theoretical analysis of multivariate correlation. IBM J. Res. Dev. 4, 66–82.
Whittaker, J., 1990. Graphical Models in Applied Multivariate Statistics. J. Wiley and Sons, Chichester.