

**SUPPLEMENTARY MATERIAL OF:  
HETEROGENEOUS MULTISCALE MULTIVARIATE AUTOREGRESSIVE  
MODEL: EXISTENCE, SPARSE ESTIMATION AND APPLICATION TO  
FUNCTIONAL CONNECTIVITY IN NEUROSCIENCE**

BY STEFANO SPAZIANI<sup>1,a</sup>, GABRIELLE GIRARDEAU<sup>2,b</sup>,  
INGRID BETHUS<sup>3,c</sup> AND PATRICIA REYNAUD-BOURET<sup>4,d</sup>

<sup>1</sup>Université Côte d'Azur, CNRS, LJAD, France, <sup>a</sup>Stefano.Spaziani@univ-cotedazur.fr

<sup>2</sup>Institut du Fer-à-Moulin, Inserm U1270, Sorbonne Université, France, <sup>b</sup>gabrielle.girardeau@inserm.fr

<sup>3</sup>Université Côte d'Azur, CNRS, IPMC, France, <sup>c</sup>Ingrid.Bethus@univ-cotedazur.fr

<sup>4</sup>Université Côte d'Azur, CNRS, LJAD, France, <sup>d</sup>Patricia.Reynaud-Bouret@univ-cotedazur.fr

## 1. Proofs.

PROOF OF LEMMA 3.3. Let us prove in fact by induction that

$$\mathbb{E}(\|\mathbb{G}_s(\mathbf{0}, \xi_0^s)\|_1^m) < \infty,$$

for  $s = 0, \dots, 2^J - 1$ .

For  $s = 0$ ,

$$\|g_0(0, \xi_0)\|_1 \leq M + \|\mu\|_1 + \|\varepsilon_0\|_1.$$

Since  $\varepsilon_0$  is a finite Gaussian vector, this has finite moment of every order.

For  $s > 0$ , we have that

$$\|\mathbb{G}_s(\mathbf{0}, \xi_0^s)\|_1 = \|\mathbb{G}_{s-1}(\mathbf{0}, \xi_0^{s-1})\|_1 + \|g_s([0, \mathbb{G}_{s-1}(\mathbf{0}, \xi_0^{s-1})], \xi_s)\|_1.$$

So it remains to prove that  $\|g_s([0, \mathbb{G}_{s-1}(\mathbf{0}, \xi_0^{s-1})], \xi_s)\|_1$  has a finite moment of order  $m$ .

But

$$\|g_s([0, \mathbb{G}_{s-1}(\mathbf{0}, \xi_0^{s-1})], \xi_s)\|_1 \leq M + \|\mu\|_1 + \|H^W * \kappa^W[\mathbb{G}_{s-1}(\mathbf{0}, \xi_0^{s-1})]\|_1 + \|\varepsilon_s\|_1.$$

Since  $\varepsilon_s$  is Gaussian, it has finite moment of order  $m$ . So we only need to prove that  $\|H^W * \kappa^W[\mathbb{G}_{s-1}(\mathbf{0}, \xi_0^{s-1})]\|_1$  has finite moment of order  $m$ .

By (??), one can see that

$$\begin{aligned} \|H^W * \kappa^W[0, \mathbb{G}_{s-1}(\mathbf{0}, \xi_0^{s-1})]\|_1 &\leq \max_{k=1, \dots, s} \|H^W(k)\|_{1, \infty} \|\kappa^W[\mathbb{G}_{s-1}(\mathbf{0}, \xi_0^{s-1})]\|_1 \\ &\leq (K \vee 1) \max_{k=1, \dots, s} \|H^W(k)\|_{1, \infty} \|\mathbb{G}_{s-1}(\mathbf{0}, \xi_0^{s-1})\|_1, \end{aligned}$$

by 1-lipschitzianity of the masks. So by induction hypothesis, this quantity has therefore finite moment of order  $m$ , which concludes the proof.  $\square$

PROOF OF LEMMA 3.4. We are again using induction to prove that

$$\mathbb{E}(\|\mathbb{G}_s(X_{-\infty}^{-1}, \xi_0^s) - \mathbb{G}_{2^J-1}(Z_{-\infty}^{-1}, \xi_0^s)\|_1) \leq \sum_{b=1}^{\infty} a_b^s \mathbb{E}(\|X_{-b} - Z_{-b}\|_1),$$

with  $s = 0, \dots, 2^J - 1$ .

For  $s = 0$ ,

$$\|g_0(X_{-\infty}^{-1}, \xi_0) - g_0(Z_{-\infty}^{-1}, \xi_0)\|_1 \leq A^S + A^W,$$

with

$$A^S = \|\mathbb{1}_{U_0 \leq \nu + H^S * \kappa^S(X_{-\infty}^{-1})} - \mathbb{1}_{U_0 \leq \nu + H^S * \kappa^S(Z_{-\infty}^{-1})}\|_1$$

and

$$A^W = \|(\mu + H^W * \kappa^W(X_{-\infty}^{-1}) + \varepsilon_0) \cdot \delta(0) - (\mu + H^W * \kappa^W(Z_{-\infty}^{-1}) + \varepsilon_0) \cdot \delta(0)\|_1.$$

For  $A^S$ , note that for all  $c, a \in [0, 1]$  and  $U$  uniform variable  $\mathbb{E}(|\mathbb{1}_{U \leq a} - \mathbb{1}_{U \leq c}|) = \mathbb{E}(\mathbb{1}_{\min(a, c) \leq U \leq \max(a, c)}) = |a - c|$ . So

$$\mathbb{E}(A^S) = \mathbb{E}(\|H^S * [\kappa^S(X_{-\infty}^{-1}) - \kappa^S(Z_{-\infty}^{-1})]\|_1) \leq (K \vee 1) \sum_{b=1}^{\infty} \max_{k \in \mathbf{I}_b} \|H^S(k)\|_{1, \infty} \mathbb{E}\|X_{-b} - Z_{-b}\|_1,$$

with  $\mathbf{I}_b = (i_{b-1} + 1) : (j_{b-1} + 1)$ .

For  $A^W$ , it is even more straight forward since

$$A^W \leq \|H^W * [\kappa^W(X_{-\infty}^{-1}) - \kappa^W(Z_{-\infty}^{-1})]\|_1 \leq (K \vee 1) \sum_{b=1}^{\infty} \max_{k \in \mathbf{I}_b} \|H^W(k)\|_{1, \infty} \|X_{-b} - Z_{-b}\|_1.$$

And therefore

$$a_b^0 = (K \vee 1) [\max_{k \in \mathbf{I}_b} \|H^S(k)\|_{1, \infty} + \max_{k \in \mathbf{I}_b} \|H^W(k)\|_{1, \infty}].$$

At step  $s > 0$ ,

$$\begin{aligned} \mathbb{E}(\|\mathbb{G}_s(X_{-\infty}^{-1}, \xi_0^s) - \mathbb{G}_s(Z_{-\infty}^{-1}, \xi_0^s)\|_1) &\leq \mathbb{E}(\|\mathbb{G}_{s-1}(X_{-\infty}^{-1}, \xi_0^{s-1}) - \mathbb{G}_{s-1}(Z_{-\infty}^{-1}, \xi_0^{s-1})\|_1) \\ &\quad + \mathbb{E}(\|g_s([X_{-\infty}^{-1}, \mathbb{G}_{s-1}(X_{-\infty}^{-1}, \xi_0^{s-1})], \xi_s) - g_s([Z_{-\infty}^{-1}, \mathbb{G}_{s-1}(Z_{-\infty}^{-1}, \xi_0^{s-1})], \xi_s)\|_1) \end{aligned}$$

Let us concentrate on the last term. We obtain by splitting again as before between the spike part and the wavelet part that

$$\begin{aligned} \mathbb{E}(\|g_s([X_{-\infty}^{-1}, \mathbb{G}_{s-1}(X_{-\infty}^{-1}, \xi_0^{s-1})], \xi_s) - g_s([Z_{-\infty}^{-1}, \mathbb{G}_{s-1}(Z_{-\infty}^{-1}, \xi_0^{s-1})], \xi_s)\|_1) &\leq \\ \sum_{b=1}^{\infty} (K \vee 1) [\max_{k \in \mathbf{I}_b} \|H^S(k+s)\|_{1, \infty} + \max_{k \in \mathbf{I}_b} \|H^W(k+s)\|_{1, \infty}] \mathbb{E}\|X_{-b} - Z_{-b}\|_1 &+ \\ (K \vee 1) [\max_{k=1, \dots, s} \|H^S(k)\|_{1, \infty} + \max_{k=1, \dots, s} \|H^W(k)\|_{1, \infty}] \mathbb{E}\|\mathbb{G}_{s-1}(X_{-\infty}^{-1}, \xi_0^{s-1}) - \mathbb{G}_{s-1}(Z_{-\infty}^{-1}, \xi_0^{s-1})\|_1 & \end{aligned}$$

We have that  $a_b^0$  as well as all the

$$(K \vee 1) [\max_{k \in \mathbf{I}_b} \|H^S(k+s)\|_{1, \infty} + \max_{k \in \mathbf{I}_b} \|H^W(k+s)\|_{1, \infty}]$$

are smaller than  $e_b$ , whereas

$$d \geq (K \vee 1) [\max_{k=1, \dots, s} \|H^S(k)\|_{1, \infty} + \max_{k=1, \dots, s} \|H^W(k)\|_{1, \infty}]$$

So finally we have proved by recursion that

$$\mathbb{E}(\|\mathbb{G}_s(X_{-\infty}^{-1}, \xi_0^s) - \mathbb{G}_s(Z_{-\infty}^{-1}, \xi_0^s)\|_1) \leq \sum_{b=1}^{\infty} a_b^s \mathbb{E}(\|X_{-b} - Z_{-b}\|_1),$$

with  $a_b^s = a_b^{s-1}(1+d) + e_b$ .

By solving the recursion, and using that  $a_b^0 \leq e_b$  we get that if  $d > 0$ ,

$$a_b^s \leq ((1+d)^{s+1} - 1) \frac{e_b}{d},$$

and if  $d = 0$ ,  $a_b^s \leq (s+1)e_b$  which concludes the proof.  $\square$

**PROOF OF THEOREM 3.2.** The condition of Theorem 3.2 on the  $e_b$ 's combined with the result of Lemma 3.4 is exactly the Lipschitz condition that is needed in Theorem 3.1 of [2], whereas Lemma 3.3 with  $m = 1$  is the other condition needed to apply it.

The existence of a  $\tau$ -weakly stationary integrable solution for  $X_b$  is therefore granted and we obtain a block stationary integrable solution (see Lemma 3.1).

It remains to prove the existence of moment of every order for the wavelet part. If the masks are bounded,  $\mathbb{E}(W_{j,k}^p | \mathcal{F}_{\ell_{j,k}-1})$  is bounded by a fixed positive constant, say  $C_{p,j}$  (all the sums converge because  $\sum_b e_b < \infty$ ). So for every positive integer  $m$ ,

$$\mathbb{E}(|W_{j,k}^p|^m) \leq 2^{m-1} C_{p,j}^m + 2^{m-1} \mathbb{E}(|\varepsilon_{p,j}|^m)$$

and  $\mathbb{E}(|W_{j,k}^p|^m)$  is therefore finite.  $\square$

**PROOF OF LEMMA 4.5.** We follow the steps in [3]. For the first inequality, we can assume that  $\|\phi\|_\infty = 1$  because the formula is homogeneous in this quantity. Let for all  $t$ ,

$$M_t = \sum_{s=0}^t \phi(Y_{-\infty}^{s-1}) [S_s^m - \mathbb{E}(S_s^m | \mathcal{F}_{s-1})] \quad \text{and} \quad V_t = \sum_{s=0}^t \phi(Y_{-\infty}^{s-1})^2 \mathbb{E}(S_s^m | \mathcal{F}_{s-1}).$$

Note first that  $M_t$  is a martingale with respect to the filtration and that  $M_T = M_T^\phi$ . We denote

$$A_t^k = \sum_{s=0}^t \mathbb{E}((M_s - M_{s-1})^k | \mathcal{F}_{s-1}).$$

Lemma 3.3. of [4] gives that for all  $\lambda > 0$

$$\mathcal{E}_t = \exp(\lambda M_t - \sum_{k \geq 2} \frac{\lambda^k}{k} A_t^k)$$

is a supermartingale and that in particular  $\mathbb{E}(\mathcal{E}_T) \leq 1$ . But  $A_t^k \leq V_t$  because  $\|\phi\|_\infty \leq 1$  and because  $|S_s^m - \mathbb{E}(S_s^m | \mathcal{F}_{s-1})| \leq 1$ . So if we take  $\psi(\lambda) = e^\lambda - \lambda - 1$ , we have that

$$\mathbb{E}(e^{\lambda M_T - \psi(\lambda) V_T}) \leq 1.$$

In particular, we have by Markov inequality, for every  $x, \lambda > 0$ ,

$$(1) \quad \mathbb{P}(M_T \geq V_T \frac{\psi(\lambda)}{\lambda} + \frac{x}{\lambda}) \leq e^{-x}.$$

But we can apply (1) to

$$\sum_{t=0}^T (-\phi^2(Y_{-\infty}^{s-1})) [S_s^m - \mathbb{E}(S_s^m | \mathcal{F}_{s-1})].$$

So if  $\hat{V}_t = \sum_{s=0}^t \phi(Y_{-\infty}^{s-1})^2 S_s^m$ , this leads us to: for all  $\mu > 0$ ,

$$\mathbb{P}(V_T \geq \hat{V}_T + \frac{\psi(\mu)}{\mu} V_T + \frac{x}{\mu}) \leq e^{-x}.$$

This means that with probability larger than  $1 - e^{-x}$ ,

$$V_T \leq \frac{\mu}{\mu - \psi(\mu)} \hat{V}_T + \frac{x}{\mu - \psi(\mu)},$$

as long as  $\mu - \psi(\mu) > 0$ . We take  $\mu = 1/2$ , which leads to

$$V_T \leq \frac{3}{2} \hat{V}_T + 3x,$$

so that

$$\mathbb{P}(M_T \geq \left(\frac{3}{2} \hat{V}_T + 3x\right) \frac{\psi(\lambda)}{\lambda} + \frac{x}{\lambda}) \leq 2e^{-x}$$

Note also that  $\psi(\lambda) \leq \lambda^2 / (2[1 - \lambda/3])$  for all  $\lambda < 3$ . This is the starting point of a peeling argument similar to [3]. We get therefore that for all  $\epsilon, x > 0$

$$\mathbb{P}(M_T \geq \sqrt{2(1 + \epsilon)} \left(\frac{3}{2} \hat{V}_T + 3x\right) x + x/3 \text{ and } w \leq \left(\frac{3}{2} \hat{V}_T + 3x\right) \leq v) \leq 2 \left(\frac{\log(v/w)}{\log(1 + \epsilon)} + 1\right) e^{-x}.$$

Note that we can always take  $w = 3x$  and  $v = \frac{3}{2}(T+1) + 3x$ . This leads to the first inequality.

Then because we want to look at the absolute value, we apply the same inequality to  $-M_T^\phi$  and then to all  $\phi \in \Phi$ . This leads to

$$\mathbb{P}\left(\exists \phi \in \Phi, |M_T^\phi| \geq \sqrt{3(1 + \epsilon)} \hat{V}_T^\phi x + \left(\frac{1}{3} + \sqrt{6(1 + \epsilon)}\right) \|\phi\|_\infty x\right) \leq 4\#\Phi \left(\frac{\log\left(\frac{T+1}{2x} + 1\right)}{\log(1 + \epsilon)} + 1\right) e^{-x}.$$

To make the upper bound less than  $\alpha$ , we need to take  $x > 1$  and then  $\frac{T+1}{2x} + 1 \leq T$ . So it is sufficient to take  $x$  such that

$$4\#\Phi \left(\frac{\log(T)}{\log(1 + \epsilon)} + 1\right) e^{-x} = \alpha.$$

In particular with  $\epsilon = 0.5$ , one has that  $\log(T)/\log(1 + \epsilon) + 1 \leq 4 \log(T)$ , which gives the final part.  $\square$

**PROOF OF LEMMA 4.6.** Let us assume as before that  $\|\phi\|_\infty = 1$  because the formula is homogeneous.

Let  $R \leq L$  be integers such that

$$M_T^\phi = \sum_{k \in R:L} \phi(Y_{-\infty}^{t-1}) \left[ W_{j,k}^p - \mathbb{E}(W_{j,k}^p | \mathcal{F}_{\ell_{j,k-1}}) \right].$$

Note that for all  $\lambda > 0$

$$\mathbb{E}(e^{\lambda M_T^\phi} | \mathcal{F}_{\ell_{j,L-1}}) = \exp\left(\sum_{k \in R:L-1} \phi(Y_{-\infty}^{t-1}) \left[ W_{j,k}^p - \mathbb{E}(W_{j,k}^p | \mathcal{F}_{\ell_{j,k-1}}) \right] + \frac{\lambda^2 \sigma_{p,j}^2}{2} \left[ \phi(Y_{-\infty}^{\ell_{j,k-1}}) \right]^2\right).$$

So that by recursion, one can easily prove that

$$\mathbb{E}\left(e^{\lambda M_T^\phi - \frac{\lambda^2}{2} V_T}\right) \leq 1.$$

This leads as before by Markov inequality : for all  $x, \lambda > 0$

$$\mathbb{P}(M_T^\phi \geq \frac{\lambda}{2}V_T + \frac{x}{\lambda}) \leq e^{-x}.$$

We are going, as before to use a peeling argument. To do so we need to restrict ourselves to the event  $w \leq V_T \leq v$ . We fix  $\epsilon > 0$ ,  $v_0 = w$ ,  $v_{d+1} = (1 + \epsilon)v_d$  and  $D$  the smallest integer such that  $v_D \geq v$ . For all  $\lambda$  and fix  $d \in 1 : D$ ,

$$\mathbb{P}(M_T^\phi \geq \frac{\lambda}{2}V_T + \frac{x}{\lambda} \text{ and } v_{d-1} \leq V_T \leq v_d) \leq e^{-x}.$$

So

$$\mathbb{P}(M_T^\phi \geq \frac{\lambda}{2}v_d + \frac{x}{\lambda} \text{ and } v_{d-1} \leq V_T \leq v_d) \leq e^{-x}.$$

We choose the optimal  $\lambda = \sqrt{2x/v_d}$  to get

$$\mathbb{P}(M_T^\phi \geq \sqrt{2v_dx} \text{ and } v_{d-1} \leq V_T \leq v_d) \leq e^{-x}.$$

But  $v_d = (1 + \epsilon)v_{d-1} \leq (1 + \epsilon)V_T$  on this event. Hence we get

$$\mathbb{P}(M_T^\phi \geq \sqrt{2(1 + \epsilon)V_Tx} \text{ and } v_{d-1} \leq V_T \leq v_d) \leq e^{-x}.$$

It remains to take a union bound to have

$$\mathbb{P}(M_T^\phi \geq \sqrt{2(1 + \epsilon)V_Tx} \text{ and } w \leq V_T \leq v) \leq De^{-x}.$$

But  $D \leq (\log(v/w)/\log(1 + \epsilon) + 1)$ .

Moreover,  $V_T$  does not have an obvious lower bound, so we can redo all this argument with  $[V_T + \eta\sigma_{p,j}^2]$  instead of  $V_T$  and therefore take  $w = \eta$ . On the other hand, we can always take  $v = (T + 1)\sigma_{p,j}^2 + \eta\sigma_{p,j}^2$ . This leads to the first inequality. The second one is just an union bound and we choose  $x$  such that  $\#\Phi(\log(v/w)/\log(1 + \epsilon) + 1)e^{-x} \leq \alpha$ .

The final part is obtained by taking  $\eta = \epsilon = 0.5$ . In this case, for all  $T > 1$

$$\left( \frac{\log((T + 1)\eta^{-1} + 1)}{\log(1 + \epsilon)} + 1 \right) \leq 9 \log(T).$$

□

**PROOF OF THEOREM 4.8.** The proof is very classical, with a slight complexity due to the weights (see [6, 7, 5, 1] and the references therein. It is there for sake of completeness.

Since  $i$  is fixed, let us drop the index  $i$  to simplify notation and let us also write  $\hat{f} = f_{\hat{a}_i}^i$ .

Let  $\bar{b}_\phi = \langle \phi, f \rangle$ , so that the martingales  $M_T^\phi$  controlled in Lemma 4.5 and 4.6 are in fact the  $b_\phi - \bar{b}_\phi$ . In particular on the event of interest, we have that

$$|b_\phi - \bar{b}_\phi| \leq d_\phi \leq w_\phi.$$

First of all, let us prove that for all  $a \in \mathbb{R}^\Phi$ , we have

$$(2) \quad 2 \langle \hat{f} - f_a, \hat{f} - f \rangle + (\gamma - 2) \sum_{\phi \notin S(a)} w_\phi |\hat{a}_\phi| \leq (\gamma + 2) \sum_{\phi \in S(a)} w_\phi |\hat{a}_\phi - a_\phi|.$$

Let us take an  $s$  such that  $s_\phi = \text{sign}(a_\phi)$  if  $\phi \in S(a)$  and  $s_\phi = \text{sign}(\hat{a}_\phi)$  if  $\phi \notin S(a)$ . By solving the subgradient conditions, the solution  $\hat{a}$  of the Lasso minimisation satisfy that there exists some  $\hat{s}$ , such that  $\hat{s}_\phi = \text{sign}(\hat{a}_\phi)$  if  $\phi \in S(\hat{a})$  and  $\hat{s}_\phi \in [-1, 1]$  if  $\phi \notin S(\hat{a})$  and such that

$$2G\hat{a} - 2b + \gamma w \cdot \hat{s} = 0,$$

with . the term-to-term product. This implies that

$$(2G\hat{a} - 2\bar{b})^\top (\hat{a} - a) + \gamma[w \cdot (\hat{s} - s)]^\top (\hat{a} - a) = -\gamma[w \cdot s]^\top (\hat{a} - a) + 2(b - \bar{b})^\top (\hat{a} - a).$$

But by looking at all cases, one can easily prove that for all  $\phi$ ,  $(\hat{s}_\phi - s_\phi)(\hat{a}_\phi - a_\phi) \geq 0$ .

Next note that

$$[w \cdot s]^\top (\hat{a} - a) = \sum_{\phi \in S(a)} w_\phi \text{sign}(a_\phi)(\hat{a}_\phi - a_\phi) + \sum_{\phi \notin S(a)} w_\phi \text{sign}(\hat{a}_\phi)(\hat{a}_\phi).$$

So

$$-\gamma[w \cdot s]^\top (\hat{a} - a) \leq \gamma \sum_{\phi \in S(a)} w_\phi |\hat{a}_\phi - a_\phi| - \gamma \sum_{\phi \notin S(a)} w_\phi |\hat{a}_\phi|,$$

which gives, since  $(G\hat{a} - \bar{b})^\top (\hat{a} - a) = \langle \hat{f} - f_a, \hat{f} - f \rangle$ , that

$$2 \langle \hat{f} - f_a, \hat{f} - f \rangle + \gamma \sum_{\phi \notin S(a)} w_\phi |\hat{a}_\phi| \leq (\gamma + 2) \sum_{\phi \in S(a)} w_\phi |\hat{a}_\phi - a_\phi| + 2(b - \bar{b})^\top (\hat{a} - a).$$

But

$$2(b - \bar{b})^\top (\hat{a} - a) \leq 2 \sum_{\phi \in \Phi} w_\phi |\hat{a}_\phi - a_\phi| = 2 \sum_{\phi \in S(a)} w_\phi |\hat{a}_\phi - a_\phi| + 2 \sum_{\phi \notin S(a)} w_\phi |\hat{a}_\phi|,$$

which gives (2).

Next note that

$$(3) \quad 2 \langle \hat{f} - f_a, \hat{f} - f \rangle = \|\hat{f} - f\|^2 + \|\hat{f} - f_a\|^2 - \|f_a - f\|^2,$$

so that if  $\langle \hat{f} - f_a, \hat{f} - f \rangle \geq 0$ , the first equation of Theorem 4.8 holds.

If  $\langle \hat{f} - f_a, \hat{f} - f \rangle < 0$ , by (2)

$$(4) \quad (\gamma - 2) \sum_{\phi \notin S(a)} w_\phi |\hat{a}_\phi| \leq (\gamma + 2) \sum_{\phi \in S(a)} w_\phi |\hat{a}_\phi - a_\phi|.$$

So if  $x = \hat{a} - a$  and  $J = S(a)$ , we see that this is exactly the condition in  $\mathbf{RE}(\kappa, c, s)$ , and that therefore

$$(5) \quad (\hat{a} - a)^\top G_i (\hat{a} - a) \geq N\kappa \sum_{\phi \in S(a)} (\hat{a}_\phi - a_\phi)^2.$$

Moreover (2) also gives us that

$$(6) \quad 2 \langle \hat{f} - f_a, \hat{f} - f \rangle \leq (\gamma + 2) \sum_{\phi \in S(a)} w_\phi |\hat{a}_\phi - a_\phi| \leq (\gamma + 2) \sqrt{\sum_{\phi \in S(a)} w_\phi^2} \sqrt{\sum_{\phi \in S(a)} (\hat{a}_\phi - a_\phi)^2}.$$

So (3) gives us that

$$\|\hat{f} - f\|^2 \leq \|f_a - f\|^2 + (\gamma + 2) \sqrt{\sum_{\phi \in S(a)} w_\phi^2} \sqrt{\sum_{\phi \in S(a)} (\hat{a}_\phi - a_\phi)^2} - \|\hat{f} - f_a\|^2.$$

By using (5), noting that  $\|\hat{f} - f_a\|^2 = (\hat{a} - a)^\top G (\hat{a} - a)$ , we get

$$\|\hat{f} - f\|^2 \leq \|f_a - f\|^2 + (\gamma + 2) \sqrt{\sum_{\phi \in S(a)} \frac{w_\phi^2}{\kappa N}} \|\hat{f} - f_a\| - \|\hat{f} - f_a\|^2.$$

It remains to use that  $\beta\eta - \beta^2 \leq \eta^2/4$  for all real numbers  $\beta$  and  $\eta$  to get the result.

If  $a^*$  exists with  $\#S(a^*) \leq s$ , then (6) with  $a = a^*$  gives us directly that

$$(\hat{a} - a^*)^\top G(\hat{a} - a^*) \leq \frac{(\gamma + 2)}{2} \sqrt{\sum_{\phi \in S(a^*)} w_\phi^2} \sqrt{\sum_{\phi \in S(a^*)} (\hat{a}_\phi - a_\phi^*)^2}.$$

So by  $\mathbf{RE}(\kappa, c, s)$ , we get that

$$\sqrt{\sum_{\phi \in S(a)} (\hat{a}_\phi - a_\phi^*)^2} \leq \frac{(\gamma + 2)}{2\kappa N} \sqrt{\sum_{\phi \in S(a^*)} w_\phi^2}.$$

Since  $\langle \hat{f} - f_a, \hat{f} - f \rangle = (\hat{a} - a^*)^\top G(\hat{a} - a^*) \geq 0$ , combining the previous result and (4), we obtain

$$\begin{aligned} \sum_{\phi \in \Phi} w_\phi |\hat{a}_\phi - a_\phi^*| &= \sum_{\phi \in S(a^*)} w_\phi |\hat{a}_\phi - a_\phi^*| + \sum_{\phi \notin S(a^*)} w_\phi |\hat{a}_\phi| \\ &\leq \frac{2\gamma}{\gamma - 2} \sum_{\phi \in S(a^*)} w_\phi |\hat{a}_\phi - a_\phi^*| \\ &\leq \frac{2\gamma}{\gamma - 2} \sqrt{\sum_{\phi \in S(a^*)} w_\phi^2} \sqrt{\sum_{\phi \in S(a)} (\hat{a}_\phi - a_\phi^*)^2} \\ &\leq \frac{2\gamma(\gamma + 2)}{2\kappa(\gamma - 2)} \sum_{\phi \in S(a^*)} \frac{w_\phi^2}{N}, \end{aligned}$$

which leads to the result. □

## 2. Table of performance measures.

Error name	Definition	Formula
$\nu_e$	total error for the spontaneous firing rates	$\nu_e = \sum_{m=1}^M  \hat{\nu}^m - \nu^m $
$\mu_e$	total error for the LFP spontaneous values	$\mu_e = \sum_{p=1}^P \sum_{j=0}^J  \hat{\mu}_j^p - \mu_j^p $
$e_{sS}$	total error for spike-spike interactions	$e_{sS} = \sum_{m,m'=1}^M \sum_{r=1}^R  \hat{h}_{m'}^m(r\delta) - h_{m'}^m(r\delta) $
$e_{sW}$	total error for spike-LFP interactions	$e_{sW} = \sum_{p=1}^P \sum_{j=0}^J \sum_{m'=1}^M \sum_{r=1}^R  \hat{h}_{m'}^{p,j}(r\delta) - h_{m'}^{p,j}(r\delta) $
$e_{wS}$	total error for LFP-spikes interactions	$e_{wS} = \sum_{m=1}^M \sum_{p'=1}^P \sum_{j'=0}^J \sum_{t=t_1}^{t_{\#supph}}  \hat{h}_{p',j'}^m(t) - h_{p',j'}^m(t) $
$e_{wW}$	total error for LFP-LFP interactions	$e_{wW} = \sum_{p,p'=1}^P \sum_{j,j'=0}^J \sum_{t=t_1}^{t_{\#supph}}  \hat{h}_{p',j'}^{p,j}(t) - h_{p',j'}^{p,j}(t) $
DG	=1 if connectivity graph is perfectly reconstructed	$DG = \mathbb{1}_{\nexists h=0   \hat{h} \neq 0 \cap \nexists h \neq 0   \hat{h} = 0}$
S	# of spontaneous values detected	$S = \sum_{m=1}^M \mathbb{1}_{(\nu^m > 0 \cap \hat{\nu}^m > 0) \cup (\nu^m = 0 \cap \hat{\nu}^m = 0)} + \sum_{p=1}^P \sum_{j=0}^J \mathbb{1}_{(\mu_j^p \neq 0 \cap \hat{\mu}_j^p \neq 0) \cup (\mu_j^p = 0 \cap \hat{\mu}_j^p = 0)}$
$c^+_{sS}$	# false positive spike-spike parameters	$c^+_{sS} = \sum_{m,m'=1}^M \sum_{r=1}^R \mathbb{1}_{h_{m'}^m(r\delta) = 0 \cap \hat{h}_{m'}^m(r\delta) \neq 0}$
$c^-_{sS}$	# false negative spike-spike parameters	$c^-_{sS} = \sum_{m,m'=1}^M \sum_{r=1}^R \mathbb{1}_{h_{m'}^m(r\delta) \neq 0 \cap \hat{h}_{m'}^m(r\delta) = 0}$
$c^+_{sW}$	# false positive spike-LFP parameters	$c^+_{sW} = \sum_{p=1}^P \sum_{j=0}^J \sum_{m'=1}^M \sum_{r=1}^R \mathbb{1}_{h_{m'}^{p,j}(r\delta) = 0 \cap \hat{h}_{m'}^{p,j}(r\delta) \neq 0}$
$c^-_{sW}$	# false negative spike-LFP parameters	$c^-_{sW} = \sum_{p=1}^P \sum_{j=0}^J \sum_{m'=1}^M \sum_{r=1}^R \mathbb{1}_{h_{m'}^{p,j}(r\delta) \neq 0 \cap \hat{h}_{m'}^{p,j}(r\delta) = 0}$
$c^+_{wS}$	# false positive LFP-spike parameters	$c^+_{wS} = \sum_{m=1}^M \sum_{p'=1}^P \sum_{j'=0}^J \sum_{t=t_1}^{t_{\#supph}} \mathbb{1}_{h_{p',j'}^m(t) = 0 \cap \hat{h}_{p',j'}^m(t) \neq 0}$
$c^-_{wS}$	# false negative LFP-spike parameters	$c^-_{wS} = \sum_{m=1}^M \sum_{p'=1}^P \sum_{j'=0}^J \sum_{t=t_1}^{t_{\#supph}} \mathbb{1}_{h_{p',j'}^m(t) \neq 0 \cap \hat{h}_{p',j'}^m(t) = 0}$
$c^+_{wW}$	# false positive LFP-LFP parameters	$c^+_{wW} = \sum_{p,p'=1}^P \sum_{j,j'=0}^J \sum_{t=t_1}^{t_{\#supph}} \mathbb{1}_{h_{p',j'}^{p,j}(t) = 0 \cap \hat{h}_{p',j'}^{p,j}(t) \neq 0}$
$c^-_{wW}$	# false negative LFP-LFP parameters	$c^-_{wW} = \sum_{p,p'=1}^P \sum_{j,j'=0}^J \sum_{t=t_1}^{t_{\#supph}} \mathbb{1}_{h_{p',j'}^{p,j}(t) \neq 0 \cap \hat{h}_{p',j'}^{p,j}(t) = 0}$

TABLE 1

Complete formula for the performance measures.



## REFERENCES

- [1] BÜHLMANN, P. and VAN DE GEER, S. (2011). *Statistics for high-dimensional data*. Springer.
- [2] DOUKHAN, P. and WINTENBERGER, O. (2007). Weakly dependent chains with infinite memory. <https://doi.org/10.48550/ARXIV.0712.3231>
- [3] HANSEN, N. R., REYNAUD-BOURET, P. and RIVOIRARD, V. (2015). Lasso and probabilistic inequalities for multivariate point processes. *Bernoulli* **21**. <https://doi.org/10.3150/13-bej562>
- [4] HOUDRÉ, C. and REYNAUD-BOURET, P. (2003). Exponential inequalities, with constants, for U-statistics of order two. *Stochastic inequalities and applications* **56** 55–69.
- [5] HUNT, X. J., REYNAUD-BOURET, P., RIVOIRARD, V., SANSONNET, L. and WILLETT, R. (2019). A data-dependent weighted LASSO under Poisson noise. *IEEE Transactions on Information Theory* **65** 1589–1613.
- [6] OST, G. and REYNAUD-BOURET, P. (2018). Sparse space-time models: Concentration Inequalities and Lasso. <https://doi.org/10.48550/ARXIV.1807.07615>
- [7] WILLETT, R. M. and NOWAK, R. D. (2007). Multiscale Poisson intensity and density estimation. *IEEE Transactions on Information Theory* **53** 3171–3187.