ISSN 1063-4541, Vestnik St. Petersburg University, Mathematics, 2022, Vol. 55, No. 2, pp. 166–173. © Pleiades Publishing, Ltd., 2022.
Russian Text © The Author(s), 2022, published in Vestnik Sankt-Peterburgskogo Universiteta: Matematika, Mekhanika, Astronomiya, 2022, Vol. 9, No. 2, pp. 246–255.

MATHEMATICS

On the Asymptotical Separation of Linear Signals from Harmonics by Singular Spectrum Analysis

N. V. Zenkova[a],* and V. V. Nekrutkin[b],**

[a] *JetBrains Research, St. Petersburg, 197342 Russia*

[b] *St. Petersburg State University, St. Petersburg, 199034 Russia*

**e-mail: natalia.zenkova@jetbrains.com*

***e-mail: vnekr@statmod.ru*

Received September 8, 2021; revised October 24, 2021; accepted December 2, 2021

Abstract—The general theoretical approach to the asymptotic extraction of the signal series from the additively perturbed signal with the help of singular spectrum analysis (SSA) was already outlined in Nekrutkin (2010, Stat. Its Interface **3**, 297–319). In this paper, the example of such an analysis applied to the linear signal and the additive sinusoidal noise is considered. It is proven that, in this case, the so-called reconstruction errors $r_i(N)$ of SSA uniformly tend to zero as the series length N tends to infinity. More precisely, we demonstrate that $\max_i|r_i(N)| = O(N^{-1})$ as $N \to \infty$ if the "window length" L equals $(N + 1)/2$. It is important to mention that a completely different result is valid for the increasing exponential signal and the same noise. As is proven in Ivanova and Nekrutkin (2019, Stat. Its Interface **12**(1), 49–59), no finite number of last terms of the error series tends to any finite or infinite values in this case.

Keywords: signal processing, singular spectral analysis, separability, linear signal, asymptotical analysis

DOI: 10.1134/S1063454122020157

1. INTRODUCTION

Let us first consider the variant of the singular spectrum analysis (SSA), which is discussed in this paper. A detailed description of this method can be found in [1] or [2].

The real "signal" $F = (f_0, ..., f_n, ...)$ is considered. It is assumed that the series F is governed by a linear recurrent formula (LRF) of order d

$$f_n = \sum_{k=1}^{d} a_k f_{n-k}, \quad n \ge d \tag{1}$$

with $a_d > 0$, which is minimal in the sense that there is no LRF of lower order governing the series F.

In addition, "noise" $E = (e_0, ..., e_n, ...)$ is introduced and it is assumed that the series $X_N = F_N + \delta E_N$ is observed, where F_N and E_N are matched segments of length N of the signal and noise, and δ is the formal perturbation parameter. In other words,

$$F_N = (f_0, \ldots, f_{N-1}), \quad E_N = (e_0, \ldots, e_{N-1}) \quad \text{and} \quad X_N = (f_0 + \delta e_0, \ldots, f_{N-1} + \delta e_{N-1}).$$

The general problem is to extract (approximately) the signal F_N from the sum X_N. It is assumed that only the order value d of the LRF (1) is known. The terms "signal" and "noise" emphasize our interest particularly in the series F_N.

1.1. Brief Description of the Method

The SSA method is described in this case as follows.

1. First of all, the *window length* $L < N$ is chosen and the Hankel trajectory matrix $\mathbf{H}(\delta)$ of dimension $L \times K$, $K = N - L + 1$, with elements $\mathbf{H}(\delta)[ij] = x_{i+j-2}$, $0 < i \le L$, $0 < j \le K$ *is constructed from the series* X_N. It is assumed here that $\min(L, K) \ge d$. In [1], this operation is called *embedding*.

If we denote $\mathbf{H}$ and $\mathbf{E}$ as the Hankel matrices obtained from F_N and E_N series by embedding with the same window length L, then, of course, $\mathbf{H}(\delta) = \mathbf{H} + \delta\mathbf{E}$.

2. The matrix $\mathbf{H}(\delta)$ is then subjected to a singular value decomposition and d main (i.e., corresponding to the largest singular values) elementary matrices of this decomposition are summed. The result $\tilde{\mathbf{H}}(\delta)$ of this operation is the best approximation of the matrix $\mathbf{H}(\delta)$ using matrices of rank d in the Frobenius norm.

3. After that, the Hankel matrix $\hat{\mathbf{H}}(\delta)$ is searched for, which is the closest to $\tilde{\mathbf{H}}(\delta)$ in the same Frobenius norm. Explicitly, this means that, on each secondary diagonal $i + j =$ const, all elements of the matrix $\tilde{\mathbf{H}}(\delta)$ are replaced by their average values. Therefore, this operation is called *diagonal averaging* in [1]. Denoting it $\mathcal{S}$ we get that $\hat{\mathbf{H}}(\delta) = \mathcal{S}\tilde{\mathbf{H}}(\delta)$.

4. Finally, applying to $\hat{\mathbf{H}}(\delta)$ the operation inverse to embedding, we obtain the *reconstructed* series

$$F_N(\delta) = (f_0(\delta), \ldots, f_{N-1}(\delta)),$$

which is declared as an approximation to the signal F_N.

A more formalized notation of this variant of the SSA can be found in ([2], p. 128) with $M = 1$. It is natural to name the series

$$R_N(\delta) = (r_0(\delta), \ldots, r_{N-1}(\delta))$$

with $r_i(\delta) = f_i(\delta) - f_i$ *a series of reconstruction errors*, and the matrix $\Delta_\delta(\mathbf{H}) = \hat{\mathbf{H}}(\delta) - \mathbf{H}$ *a matrix of reconstruction errors*.

In this work, we consider linear signal

$$f_n = \theta_1 n + \theta_0, \tag{2}$$

where $\theta_1 \neq 0$, and the noise as a linear combination of harmonics

$$e_n = \sum_{\ell=1}^{r} \tau_\ell \cos(2\pi n\omega_\ell + \varphi_\ell), \tag{3}$$

where $\tau_\ell \neq 0$, $\omega_\ell \neq \omega_p$ at $\ell \neq p$ and $0 < \omega_\ell < 1/2$. Since the signal (2) is governed by LRF $f_n = 2f_{n-1} - f_{n-2}$, then $d = 2$ in this case.

In ([3], paragraph 5.3), a general scheme of asymptotic analysis of reconstruction errors at $N \to \infty$ is proposed. As it is used below, let us give a brief description if it.

1.2. Approach to Analysis of Reconstruction Errors

We are interested in uniform convergence of the residuals $r_i(\delta)$ to zero first of all, i.e., the behavior of the norm $\|F_N(\delta) - F_N\|_{\max} = \max_{0 \le i < N} |r_i(\delta)|$ at $N \to \infty$.

In addition, it is assumed that $\min(L, K) \ge d$. In this work, we use relation $L = (N + 1)/2$, i.e., the length of the series N is assumed odd.

Further, if $\mathrm{U}_0^\perp$ is a linear space generated by columns of the matrix $\mathbf{H}$, then it follows from (1) that the dimension of $\mathrm{U}_0^\perp$ equals d independently on N and L in these conditions.

Let us denote by $\mathbf{P}_0^\perp(\delta)$ orthogonal projector on linear space $\mathrm{U}_0^\perp$ and by $\mathbf{P}_0^\perp(\delta)$ orthogonal projector on linear space generated by columns of the matrix $\tilde{\mathbf{H}}(\delta)$. Then, as shown in ([3], paragraph 5.3),

$$\tilde{\mathbf{H}}(\delta) - \mathbf{H} = (\mathbf{P}_0^\perp(\delta) - \mathbf{P}_0^\perp)\mathbf{H}(\delta) + \delta\mathbf{P}_0^\perp\mathbf{E}. \tag{4}$$

In this work, according to ([3], paragraph 5.3), we use two matrix norms. For matrix $\mathbf{A}$ with size $L \times K$, spectral norm $\|\mathbf{A}\|$ is determined as maximal singular number of this matrix and uniform norm $\|\mathbf{A}\|_{\max}$ as maximum of the modules of the elements of this matrix. Relation between these norms is well-known ([4], paragraph 2.3.2):

$$\|\mathbf{A}\|_{\max} \le \|\mathbf{A}\| \le \sqrt{LK}\,\|\mathbf{A}\|_{\max}. \tag{5}$$

Since $\|\mathcal{S}\mathbf{A}\|_{\max} \le \|\mathbf{A}\|_{\max}$, then the left of the inequalities (5) allows using the spectral norm to study the behavior of reconstruction errors. At the same time, the form of the first term on the right side of (4) shows that it is necessary to pay attention to the difference of the projectors $\mathbf{P}_0^{\perp}(\delta) - \mathbf{P}_0^{\perp}$.

Using the classical results of Kato ([5], chapter 2, paragraph 3), an upper estimate of spectral norm $\|\mathbf{P}_0^{\perp}(\delta) - \mathbf{P}_0^{\perp}\|$ is obtained in ([3], Theorem 2.1), which is used in some so-called subspace methods of signal processing. However, since the difference of the projectors in (4) is multiplied by $\mathbf{H}(\delta)$, this estimate turns out to be insufficient and it is necessary to distinguish the "main part" of the difference $\mathbf{P}_0^{\perp}(\delta) - \mathbf{P}_0^{\perp}$.

This is done as follows. Let us denote maximal and minimal positive eigenvalues of the matrix $\mathbf{H}\mathbf{H}^{\mathrm{T}}$ as $\mu_{\max} = \|\mathbf{H}\|^2$ and $\mu_{\min}$, respectively. In addition, assume that $\mathbf{S}_0$ is the pseudoinverse Moore–Penrose matrix to the matrix $\mathbf{H}\mathbf{H}^{\mathrm{T}}$ with $\|\mathbf{S}_0\| = 1/\mu_{\min}$. Then we denote

$$\mathbf{B}(\delta) = \mathbf{H}(\delta)(\mathbf{H}(\delta))^{\mathrm{T}} - \mathbf{H}\mathbf{H}^{\mathrm{T}} = \delta(\mathbf{H}\mathbf{E}^{\mathrm{T}} + \mathbf{E}\mathbf{H}^{\mathrm{T}}) + \delta^2\mathbf{E}\mathbf{E}^{\mathrm{T}}$$

and

$$\mathbf{W}_1(\delta) = \mathbf{P}_0\mathbf{B}(\delta)\mathbf{S}_0 + \mathbf{S}_0\mathbf{B}(\delta)\mathbf{P}_0, \tag{6}$$

where $\mathbf{P}_0 = \mathbf{I} - \mathbf{P}_0^{\perp}$, and $\mathbf{I}$ is identity $(L \times L)$ matrix. The following assertion then holds (see ([3], Theorem 2.4)).

Theorem 1. *Assume that* $\delta_0 > 0$ *and* $\|\mathbf{B}(\delta)\|/\mu_{\min} < 1/4$ *for all* $\delta \in (-\delta_0; \delta_0)$. *There is then such an absolute constant C that*

$$\left\|\mathbf{P}_0^{\perp}(\delta) - \mathbf{P}_0^{\perp} - \mathbf{W}_1(\delta)\right\| \le C\left(\frac{\|\mathbf{B}(\delta)\|}{\mu_{\min}}\right)^2 \frac{1}{1 - 4\|\mathbf{B}(\delta)\|/\mu_{\min}}. \tag{7}$$

The inequality (7) is used as follows. The equality (4) is rewritten as

$$\tilde{\mathbf{H}}(\delta) - \mathbf{H} = (\mathbf{P}_0^{\perp}(\delta) - \mathbf{P}_0^{\perp} - \mathbf{W}_1(\delta))\mathbf{H}(\delta) + \delta\mathbf{P}_0^{\perp}\mathbf{E} + \mathbf{W}_1(\delta)\mathbf{H}(\delta). \tag{8}$$

If it turns out that, in this case, the following inequality holds at $N \to \infty$

$$\left\|(\mathbf{P}_0^{\perp}(\delta) - \mathbf{P}_0^{\perp} - \mathbf{W}_1(\delta))\mathbf{H}(\delta)\right\| \le \left\|(\mathbf{P}_0^{\perp}(\delta) - \mathbf{P}_0^{\perp} - \mathbf{W}_1(\delta))\right\|\left\|\mathbf{H}(\delta)\right\| \to 0, \tag{9}$$

then it is left to check the asymptotic behavior of the elements of a specific (even possibly complex) residual matrix $\delta\mathbf{P}_0^{\perp}\mathbf{E} + \mathbf{W}_1(\delta)\mathbf{H}(\delta)$.

It is proposed to solve the problem of asymptotic separability of signal (2) from noise (3) in this work exactly in this way. Namely, we first prove that inequality (7) holds for any δ for sufficiently large N, then we arrive at convergence (9) by estimating the right-hand side of (7) from above (see Section 2).

Section 3 presents the proofs of relations $\|\mathbf{W}_1(\delta)\mathbf{H}(\delta)\|_{\max} = O(N^{-1})$ and $\left\|\mathbf{P}_0^{\perp}\mathbf{E}\right\|_{\max} = O(N^{-1})$. This immediately implies convergence of $\|F_N(\delta) - F_N\|_{\max}$ to zero; moreover, this expression is of the order $O(N^{-1})$ at $N \to \infty$.

All these facts are first discussed for a single sinusoidal noise, transition to noise of the general form, and the final results of the work—Theorem 2—are also placed in Section 3.

2. AUXILIARY ASSERTIONS AND PROOF OF CONVERGENCE OF (9)

As was mentioned before, we consider here the signal (2) (it is enough to take $\theta_1 = 1$) and the noise

$$e_n = \cos(2\pi n\omega + \varphi), \quad \omega \in (0, 1/2). \tag{10}$$

In addition, let $L = K := N - L + 1 = (N + 1)/2$ (i.e., the matrices $\mathbf{H}$ and $\mathbf{E}$ are quadratic and symmetrical), while $N \to \infty$.

As noted in ([3], Lemma 3.1), there are such positive constants $C_{\cos}$, $C_{\max}$, and $C_{\min} < C_{\max}$ in these conditions that the following relations hold at $N \to \infty$:

$$\left\|\mathbf{E}\mathbf{E}^{\mathrm{T}}\right\| = \|\mathbf{E}\|^2 \sim C_{\cos}N^2, \quad \mu_{\max} \sim C_{\max}N^4 \quad \text{and} \quad \mu_{\min} \sim C_{\min}N^4. \tag{11}$$

Lemma 1. *At $N \to \infty$ relation $\|\mathbf{HE}^{\mathrm{T}}\|_{\max} = O(N)$ holds.*

Proof. At $1 \le p \le L$ and $1 \le s \le K$, we have

$$\begin{aligned}\mathbf{HE}^{\mathrm{T}}[p,s] &= \sum_{j=0}^{K-1}(p+j)\cos(2\pi(s+j)\omega+\varphi)\\ &= p\sum_{j=0}^{K-1}\cos(2\pi j\omega+\varphi_s)+\sum_{j=0}^{K-1} j\cos(2\pi j\omega+\varphi_s),\end{aligned}$$

where $\varphi_s = 2\pi s\omega + \varphi$.

Since inequalities

$$p\left|\sum_{j=0}^{K-1}\cos(2\pi j\omega+\psi)\right| = p\left|\frac{\sin(\pi K\omega)}{\sin(\pi\omega)}\cos(\pi(K-1)\omega+\psi)\right| \le \frac{p}{\sin(\pi\omega)} = O(N)$$

are true for any ψ and in the notations

$$B_K = \frac{1}{2\sin(\pi\omega)}\sin(\pi(2K-1)\omega+\psi), \quad E_K = \frac{\sin(\pi K\omega)}{2\sin^2(\pi\omega)}\sin(\pi K\omega+\psi)$$

there is a relation

$$\left|\sum_{j=0}^{K-1} j\cos(2\pi j\omega+\psi)\right| = |KB_K - E_K| \le \frac{K}{2\sin(\pi\omega)} + \frac{1}{2\sin^2(\pi\omega)} = O(N),$$

then $\|\mathbf{HE}^{\mathrm{T}}\|_{\max} = O(N)$.

Note 1. Since $LK \sim N^2/4$, and $\mu_{\min} \sim C_{\min}N^4$ at $N \to \infty$, then, applying the right of inequalities (5), we obtain that $\|\mathbf{HE}^{\mathrm{T}}\|/\mu_{\min} = O(N^{-2})$.

Lemma 2. *There is a relation* $\left\|\mathbf{P}_0^{\perp}\mathbf{E}\right\|_{\max} = O(N^{-1})$.

Proof. Let us denote

$$P_L(0) = (1,\ldots,1)^{\mathrm{T}}, \quad P_L(1) = (0,1,\ldots,(L-1))^{\mathrm{T}}.$$

Of course, the pair $P_L(0)$, $P_L(1)$ is the basis of the linear space $\mathrm{U}_0^{\perp}$. Thus, the matrix $\mathbf{P}_0^{\perp}$ can be represented as

$$\begin{aligned}\mathbf{P}_0^{\perp} &= \gamma_{00}^2 P_L(0)P_L^T(0) + (\gamma_{11}P_L(1)-\gamma_{10}P_L(0))(\gamma_{11}P_L^{\mathrm{T}}(1)-\gamma_{10}P_L^{\mathrm{T}}(0))\\ &= (\gamma_{00}^2+\gamma_{10}^2)P_L(0)P_L^{\mathrm{T}}(0) + \gamma_{11}^2 P_L(1)P_L^{\mathrm{T}}(1) - \gamma_{11}\gamma_{10}(P_L(1)P_L^{\mathrm{T}}(0)+P_L(0)P_L^{\mathrm{T}}(1)),\end{aligned} \tag{12}$$

where $(L \times L)$ matrices have the form

$$P_L(0)P_L^{\mathrm{T}}(0) = \begin{pmatrix}1 & \ldots & 1\\ \vdots & \ddots & \vdots\\ 1 & \ldots & 1\end{pmatrix}, \quad P_L(0)P_L^{\mathrm{T}}(1) = \begin{pmatrix}0 & 1 & \ldots & L-1\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 1 & \ldots & L-1\end{pmatrix},$$

$$P_L(1)P_L^{\mathrm{T}}(0) = \begin{pmatrix}0 & \ldots & 0\\ 1 & \ldots & 1\\ \vdots & \ddots & \vdots\\ L-1 & \ldots & L-1\end{pmatrix}, \quad P_L(1)P_L^{\mathrm{T}}(1) = \begin{pmatrix}0 & 0 & \ldots & 0\\ 0 & 1 & \ldots & L-1\\ \vdots & \vdots & \ddots & \vdots\\ 0 & L-1 & \ldots & (L-1)^2\end{pmatrix}$$

and

$$\gamma_{11} = \frac{\sqrt{12}}{\sqrt{L(L^2-1)}}, \quad \gamma_{10} = \frac{\sqrt{3(L-1)}}{\sqrt{L(L+1)}}, \quad \gamma_{00} = 1/\sqrt{L}.$$

Multiplying each term in the right part of (12) by the matrix

$$\mathbf{E} = \begin{pmatrix} \cos(\varphi) & \dots & \cos(2\pi(j-1)\omega+\varphi) & \dots & \cos(2\pi(K-1)\omega+\varphi) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \cos(2\pi(i-1)\omega+\varphi) & \dots & \cos(2\pi(i+j-2)\omega+\varphi) & \dots & \cos(2\pi(K+i-1)\omega+\varphi) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \cos(2\pi(L-1)\omega+\varphi) & \dots & \cos(2\pi(L+j-1)\omega+\varphi) & \dots & \cos(2\pi(N-1)\omega+\varphi) \end{pmatrix}$$

and acting in the same way as in Lemma 1, we obtain the result, omitting absolutely elementary, but cumbersome and numerous calculations.

Lemma 3. *There is a relation* $\|\mathbf{S}_0\mathbf{E}\| = O(N^{-4})$.

Proof. Let us consider singular decomposition of matrices $\mathbf{HH}^{\mathrm{T}}$, $\mathbf{S}_0$, and $\mathbf{H}$:

$$\mathbf{HH}^{\mathrm{T}} = \mu_{\max}U_1U_1^{\mathrm{T}} + \mu_{\min}U_2U_2^{\mathrm{T}}, \quad \mathbf{S}_0 = \mu_{\max}^{-1}U_1U_1^{\mathrm{T}} + \mu_{\min}^{-1}U_2U_2^{\mathrm{T}},$$
$$\mathbf{H} = \mu_{\max}^{1/2}U_1U_1^{\mathrm{T}} + \mu_{\min}^{1/2}U_2U_2^{\mathrm{T}}.$$

In addition, $\mathbf{P}_0^{\perp} = U_1U_1^{\mathrm{T}} + U_2U_2^{\mathrm{T}}$, where U_1 and U_2 are orthonormalized eigenvectors of the matrix $\mathbf{HH}^{\mathrm{T}}$. Then we have

$$\mathbf{HE}^{\mathrm{T}} = \mu_{\min}^{1/2}\left(\frac{\mu_{\max}^{1/2}}{\mu_{\min}^{1/2}}U_1U_1^{\mathrm{T}}\mathbf{E}^{\mathrm{T}} + U_2U_2^{\mathrm{T}}\mathbf{E}^{\mathrm{T}}\right),$$

and, since $\mu_{\min}^{1/2} \sim C_{\max}(N^2)$, $\mu_{\max}^{1/2}/\mu_{\min}^{1/2} \to c > 1$, and $\|\mathbf{HE}^{\mathrm{T}}\|_{\max} = O(N)$, then this implies that

$$\left\|cU_1U_1^{\mathrm{T}}\mathbf{E}^{\mathrm{T}} + U_2U_2^{\mathrm{T}}\mathbf{E}^{\mathrm{T}}\right\|_{\max} = O(N^{-1}).$$

Since $\left\|\mathbf{P}_0^{\perp}\mathbf{E}\right\|_{\max} = \left\|U_1U_1^{\mathrm{T}}\mathbf{E} + U_2U_2^{\mathrm{T}}\mathbf{E}\right\|_{\max} = O(N^{-1})$, then $\left\|U_iU_i^{\mathrm{T}}\mathbf{E}^{\mathrm{T}}\right\|_{\max} = O(N^{-1})$ at $i = 1, 2$, and, therefore, accounting that $\mathbf{E} = \mathbf{E}^{\mathrm{T}}$, we obtain (see (5))

$$\|\mathbf{S}_0\mathbf{E}\|_{\max} = O(N^{-5}) \quad \text{and} \quad \|\mathbf{S}_0\mathbf{E}\| = O(N^{-4}),$$

which was to be proven.

Proposition 1. *Assume that N is odd,* $N \to \infty$ *and* $L = (N+1)/2$. *Then, for any* δ,

$$\left\|(\mathbf{P}_0^{\perp}(\delta) - \mathbf{P}_0^{\perp} - \mathbf{W}_1(\delta))\mathbf{H}(\delta)\right\| = O(N^{-2}).$$

Proof. First, according to Lemma 1 and asymptotics (11), there is such a constant C_1 that

$$\|\mathbf{B}(\delta)\|/\mu_{\min} \le \delta^2\left\|\mathbf{EE}^{\mathrm{T}}\right\|/\mu_{\min} + 2|\delta|\left\|\mathbf{HE}^{\mathrm{T}}\right\|/\mu_{\min} \le C_1(\delta^2N^{-2} + |\delta|N^{-2}) = O(N^{-2}).$$

Therefore, for any δ the inequality (7) is held at large enough N and as a result at $N \to \infty$ there is a relation

$$\left\|(\mathbf{P}_0^{\perp}(\delta) - \mathbf{P}_0^{\perp} - \mathbf{W}_1(\delta))\mathbf{H}(\delta)\right\| \le \left\|(\mathbf{P}_0^{\perp}(\delta) - \mathbf{P}_0^{\perp} - \mathbf{W}_1(\delta))\right\|\|\mathbf{H}(\delta)\|$$
$$\le C\left(\frac{\|\mathbf{B}(\delta)\|}{\mu_{\min}}\right)^2\frac{\|\mathbf{H}(\delta)\|}{1 - 4\|\mathbf{B}(\delta)\|/\mu_{\min}} \sim C\left(\frac{\|\mathbf{B}(\delta)\|}{\mu_{\min}}\right)^2\|\mathbf{H}\|,$$

whence immediately follows the required.

3. STUDYING THE ELEMENTS OF THE RESIDUAL MATRIX AND THE FINAL RESULT

We have to research asymptotic behavior of the elements of matrices $\mathbf{W}_1(\delta)\mathbf{H}(\delta)$ and $\mathbf{P}_0^{\perp}\mathbf{E}$ in conditions of Proposition 1. In this case, as in the previous section, it is assumed that the noise has the form (10).

According to Lemma 2, $\left\|\mathbf{P}_0^{\perp}\mathbf{E}\right\|_{\max} = O(N^{-1})$; therefore, it is necessary to deal with $\mathbf{W}_1(\delta)\mathbf{H}(\delta)$.

Proposition 2. *In conditions of Proposition* 1, $\|\mathbf{W}_1(\delta)\mathbf{H}(\delta)\|_{\max} = O(N^{-1})$.

Proof. Let us begin with some simplifications. According to formula (6), we have

$$\mathbf{W}_1(\delta) = \delta\mathbf{V}_0^{(1)} + \delta^2(\mathbf{P}_0\mathbf{E}\mathbf{E}^{\mathrm{T}}\mathbf{S}_0 + \mathbf{S}_0\mathbf{E}\mathbf{E}^{\mathrm{T}}\mathbf{P}_0),$$

where $\mathbf{V}_0^{(1)} = \mathbf{P}_0\mathbf{E}\mathbf{H}^{\mathrm{T}}\mathbf{S}_0 + \mathbf{S}_0\mathbf{H}\mathbf{E}^{\mathrm{T}}\mathbf{P}_0$. Since (see Lemma 3)

$$\left\|\mathbf{P}_0\mathbf{E}\mathbf{E}^{\mathrm{T}}\mathbf{S}_0 + \mathbf{S}_0\mathbf{E}\mathbf{E}^{\mathrm{T}}\mathbf{P}_0\right\| \le 2\|\mathbf{S}_0\mathbf{E}\|\|\mathbf{E}\| = O(N^{-3})$$

and $\|\mathbf{E}\| = o(\|\mathbf{H}\|)$, then $\left\|(\mathbf{W}_1(\delta) - \delta\mathbf{V}_0^{(1)})\mathbf{H}(\delta)\right\| = O(N^{-1}) \to 0$, and, thus, it is enough to consider the elements of the matrix

$$\mathbf{V}_0^{(1)}\mathbf{H}(\delta) = \mathbf{P}_0\mathbf{E}\mathbf{H}^{\mathrm{T}}\mathbf{S}_0\mathbf{H} + \delta(\mathbf{P}_0\mathbf{E}\mathbf{H}^{\mathrm{T}}\mathbf{S}_0\mathbf{E} + \mathbf{S}_0\mathbf{H}\mathbf{E}^{\mathrm{T}}\mathbf{P}_0\mathbf{E})$$

instead of $\mathbf{W}_1(\delta)\mathbf{H}(\delta)$. Further, since

$$\left\|\mathbf{P}_0\mathbf{E}\mathbf{H}^{\mathrm{T}}\mathbf{S}_0\mathbf{E} + \mathbf{S}_0\mathbf{H}\mathbf{E}^{\mathrm{T}}\mathbf{P}_0\mathbf{E}\right\| \le 2\left\|\mathbf{H}\mathbf{E}^{\mathrm{T}}\right\|\|\mathbf{E}\|\|\mathbf{S}_0\| = O(N^{-1}),$$

then we have to deal with the matrix $\mathbf{P}_0\mathbf{E}\mathbf{H}^{\mathrm{T}}\mathbf{S}_0\mathbf{H}$.

Since $\mathbf{H}^{\mathrm{T}}\mathbf{S}_0\mathbf{H} = \mathbf{Q}_0^{\perp}$, where $\mathbf{Q}_0^{\perp}$ is the matrix of orthogonal projection on the space of rows of the matrix $\mathbf{H}$, then we finally will deal with the elements of the matrix

$$\mathbf{P}_0\mathbf{E}\mathbf{Q}_0^{\perp} = \mathbf{E}\mathbf{Q}_0^{\perp} - \mathbf{P}_0^{\perp}\mathbf{E}\mathbf{Q}_0^{\perp}.$$

Since $\mathbf{Q}_0^{\perp} = \mathbf{P}_0^{\perp}$, then Lemma 2 implies that $\left\|\mathbf{E}\mathbf{Q}_0^{\perp}\right\|_{\max} = O(N^{-1})$. In order to obtain a similar inequality for $\mathbf{P}_0^{\perp}\mathbf{E}\mathbf{Q}_0^{\perp}$, it is enough to calculate this 2×2 matrix explicitly using formula (12) and make sure that each of its element has the order $O(N^{-1})$. Like in Lemma 2, we will omit here this cumbersome and elementary procedure.

The main result of the present work can now be formulated and proven.

Theorem 2. *Let us consider a linear signal* $f_n = \theta_1 n + \theta_0$, *where* $\theta_1 \neq 0$ *and a noise that is a linear combination of harmonics at* $n = 0, 1 \ldots, N-1$:

$$e_n = \sum_{\ell=1}^{r} \tau_\ell \cos(2\pi n\omega_\ell + \varphi_\ell), \tag{13}$$

where $\tau_\ell \neq 0$, $\omega_\ell \neq \omega_p$ *at* $\ell \neq p$ *and* $0 < \omega_\ell < 1/2$.

Let us assume that $x_n = f_n + \delta e_n$, *where* δ *is the formal parameter of perturbation and taking odd* N *and* L *equal to* $(N+1)/2$, *apply the variant of the SSA method described in Introduction to the series* x_n, $n = 0, \ldots, N-1$ *to extract the signal with* $d = 2$.

If we denote the reconstruction result of the series $\{x_n\}_{n=0}^{N-1}$ *using this variant of the SSA method as* $f_0(\delta), \ldots, f_{N-1}(\delta)$, *then for any* $\delta \in \mathbb{R}$ *at* $N \to \infty$

$$\max_{0 \le n < N} |f_n(\delta) - f_n| = O(N^{-1}).$$

Proof. Since, as was already mentioned, $\|\mathcal{S}\mathbf{A}\|_{\max} \le \|\mathbf{A}\|_{\max}$, then, at $r = 1$, the result immediately follows from Propositions 1 and 2.

Let us now proceed from $r = 1$ to an arbitrary r. First, according to ([3], Lemma 3.1), $\|\mathbf{E}\| \sim CN$ not only at $r = 1$ but also in the case when the noise has the form (13). Further, since

$$\left\|\mathbf{H}(\mathbf{E}_1 + \mathbf{E}_2)^{\mathrm{T}}\right\|_{\max} \le \left\|\mathbf{H}\mathbf{E}_1^{\mathrm{T}}\right\|_{\max} + \left\|\mathbf{H}\mathbf{E}_2^{\mathrm{T}}\right\|_{\max}$$

and $\|\mathbf{S}_0(\mathbf{E}_1 + \mathbf{E}_2)\| \le \|\mathbf{S}_0\mathbf{E}_1\| + \|\mathbf{S}_0\mathbf{E}_2\|$, then the assertions of Lemmas 1 and 3 (and as a result of Proposition 1) remain true for the noise (13) as well. In the same way

$$\left\|\mathbf{P}_0^{\perp}(\mathbf{E}_1 + \mathbf{E}_2)\right\|_{\max} \le \left\|\mathbf{P}_0^{\perp}\mathbf{E}_1\right\|_{\max} + \left\|\mathbf{P}_0^{\perp}\mathbf{E}_2\right\|_{\max}$$

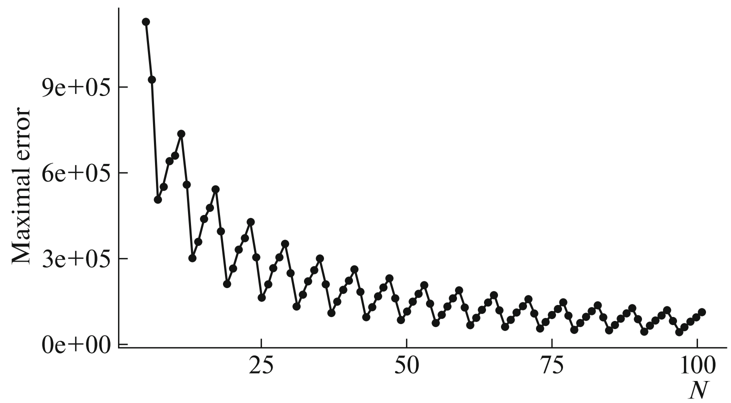

Fig. 1. Maximal reconstruction errors depending on the length of the series N for $x_n = 2n + 1 + \cos(2\pi\omega + \varphi)$, where $\omega = 1/3$, $\varphi = 0$.

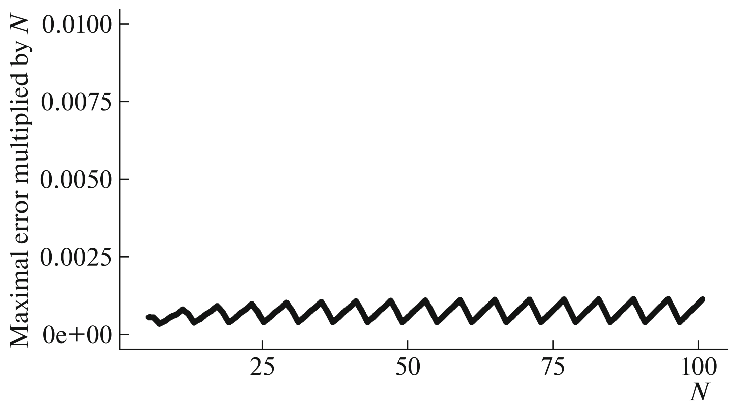

Fig. 2. Maximal reconstruction errors, multiplied by N, depending on the length of the series N for $x_n = 2n + 1 + \cos(2\pi\omega + \varphi)$, where $\omega = 1/3$, $\varphi = 0$.

implies that the assertions of Lemma 2 and Proposition 2 are also true for this noise. The theorem is proven.

Note 2. The condition $L = K$ is technical and is used only for proving Proposition 1. General considerations and computation experiments allow for assuming that the result of the theorem will be preserved for any polynomial signal and window length $L \sim \alpha N$ at $N \to \infty$ if $\alpha \in (0, 1)$. However, this is not yet proved is such generality.

It is necessary to note that a similar line of reasoning was carried out in [6], where a growing exponential signal and a sinusoidal noise were studied. In this case, it turned out that $\|F_N(\delta) - F_N\|_{\max} \nrightarrow 0$ at $N \to \infty$, which is in sharp contrast to the case of the linear signal under consideration.

APPENDIX. THE RESULTS OF NUMERIC EXPERIMENTS

As an example, let us consider the series

$$x_n = 2n + 1 + \cos(2\pi\omega + \varphi), \quad \text{where} \quad \omega = 1/3, \quad \varphi = 0,$$

when $n = 0, \ldots, N - 1$ with $N = 5(1)101$. We use window length $L = \lfloor N/2 \rfloor$.

Calculation results are shown in Figs. 1 and 2. It can be seen in Fig. 1 that the maximal in absolute value reconstruction errors of the series really tend to zero with growth of N. At the same time, Fig. 2 shows that maximal errors become bounded after multiplication of the terms of the series in Fig. 1 by N. This confirms the result of Theorem 2.

ACKNOWLEDGMENTS

We thank both anonymous reviewers whose comments have undoubtedly contributed to the improvement of our work.

FUNDING

The study was supported by the Russian Foundation for Basic Research, project no. 20-01-00067.

REFERENCES

1. N. Golyandina, V. Nekrutkin, and A. Zhigljavsky, *Analysis of Time Series Structure. SSA and Related Techniques* (Chapman & Hall/CRC, Boca Raton, Fla., 2001), in Ser.: *Monographs on Statistics and Applied Probability*, Vol. 90.
2. N. Golyandina and A. Zhigljavsky, *Singular Spectrum Analysis for Time Series*, 2nd ed. (Springer-Verlag, 2020), in Ser.: *Springer Briefs in Statistics*.
3. V. Nekrutkin, "Perturbation expansions of signal subspaces for long signals," Stat. Its Interface **3**, 297–319 (2010).
4. G. H. Golub and Ch. F. Van Loan, *Matrix Computations*, 4th ed. (Johns Hopkins Univ. Press, Baltimore, Md., 2013).
5. T. Kato, *Perturbation Theory for Linear Operators* (Springer-Verlag, Berlin, 1966; Mir, Moscow, 1972).
6. E. Ivanova and V. Nekrutkin, "Two asymptotic approaches for the exponential signal and harmonic noise in Singular Spectrum Analysis," Stat. Its Interface **12**, 49–59 (2019).

Translated by K. Gumerov