Adv. Radio Sci., 10, 153–158, 2012
www.adv-radio-sci.net/10/153/2012/
doi:10.5194/ars-10-153-2012

Advances in
Radio Science

The influence of different PAST-based subspace trackers on DaPT parameter estimation

M. Lechtenberg and J. Götze

Information Processing Lab, TU Dortmund University, Otto-Hahn-Str. 4, 44227 Dortmund, Germany

Correspondence to: M. Lechtenberg (matthias.lechtenberg@tu-dortmund.de)

Abstract. In the context of parameter estimation, subspace-based methods like ESPRIT have become common. They require a subspace separation e.g. based on eigenvalue/-vector decomposition. In time-varying environments, this can be done by subspace trackers. One class of these is based on the PAST algorithm. Our non-linear parameter estimation algorithm DaPT builds on-top of the ESPRIT algorithm. Evaluation of the different variants of the PAST algorithm shows which variant of the PAST algorithm is worthwhile in the context of frequency estimation.

1 Introduction

Parameter estimation is a common task in signal processing. It is used in DoA-applications (*direction of arrival*) like radar and sonar or in frequency estimation applications like radio transmission or mass spectrometry – just to name a few. We are researching parameter estimation in the context of power quality applications, i.e. finding the signal components next to the fundamental frequency to extrapolate the transmission system's state and fault-conditions.

Next to transformation-based parameter estimation via Fourier methods like the DFT (*discrete Fourier transformation*) and STFT (*short-time Fourier transformation*) or wavelets, there are parameter estimators based on subspace analysis. Estimators based on the separation of subspaces have advantages that justify the higher computational complexity: They have a significantly higher resolution not bounded by the sampling frequency and the window length and do not suffer from *spectral leakage*.

Common subspace-based parameter estimators are MUSIC (*multiple signal classification*, Schmidt, 1986) and ESPRIT (*estimation parameters via rotational invariance techniques*, Paulraj et al., 1985). MUSIC identifies the frequency components excluded from the noise subspace, ESPRIT identifies the components from the signal subspace. For both algorithms, noise and signal subspace have to be separated. A wide spread way to do this is to perform a decomposition into eigenvectors and separate those in noise and signal. Apart from the expensive eigenvalue decomposition, in a time-varying environment, this can also be done by efficient subspace trackers.

A subspace tracker can be circumscribed EV (*eigenvector*) updating, but this is not exactly true. The vectors generated by subspace trackers (in general) form one possible basis of the signal's space. This basis does not have to be the EV of the correlation matrix, nor does it have to be orthogonal. The properties of the basis depend on the subspace tracker used. YAST (*yet another subspace tracker*, Badeau et al., 2008) and PROTEUS (*plane rotation-based EVD-updating scheme*, Champagne and Liu, 1998) generate an orthonormal basis, whereas PAST (*projection approximation subspace tracking*, Yang, 1995b) in its original form provides a basis that only approximates orthonormality over time.

Depending on the application, an orthonormal basis is not needed. To verify that our application does not require an orthonormal basis, we compared the influence of different subspace trackers all based on the mentioned PAST algorithm. Variants of the original PAST algorithm introduce forced orthonormality (OPAST, *orthonormal PAST*, Abed-Meraim et al., 2000) and additional steps for numerical stability (OPAST-PGS, *PAST with pairwise Gram-Schmidt-orthogonalization*, Bartelmaos et al., 2005). We also included an unrolled version of PAST reducing computational complexity (PASTd, *PAST with deflation*, Yang, 1995a) and – as reference – pure eigenvalue decomposition (EVD).

Section two describes the environment in whose context the subspace trackers are compared. Section three presents the different expansion stages of the subspace tracker. The simulation environment is depicted section four and the simulation results are presented in five.

Published by Copernicus Publications on behalf of the URSI Landesausschuss in der Bundesrepublik Deutschland e.V.

2 Application environment

Modern energy transmission systems can be characterized by two main features. On the one hand, they were designed and constructed for a more static system state with constant 75 centralized generation – on the other hand, nowadays, they are stressed by decentralized *green* generation (e.g. solar, wind, biomass), economic interests (e.g. stock exchange) and cross-border transmission. Decades ago, transmission management was a question of several minutes and could easily 80 be done by humans tweaking a few parameters (like the excitation of a power plant's machine). Today's transmission grid management depends on a lot more factors and is sometimes a question of seconds. To overcome these new challenges, it is desirable to have a system of measurement units being 85 able to characterize the system state without the need of human interpretation.

The signal model describing a power transmission is composed of *superposed sinusoids*:

$$x(n) = \sum_{i=1}^{p} c_i e^{2\pi j \cdot f_i \frac{n}{f_s}} + w_{\mathrm{awgn}}(n) \tag{1}$$

$$\boldsymbol{x}(n) = \sum_{i=1}^{p} c_i e^{2\pi j \cdot f_i \frac{\boldsymbol{n}}{f_s}} + \boldsymbol{w}_{\mathrm{awgn}}(n) \tag{2}$$

$$\boldsymbol{x}(n) = \mathbf{A}(\boldsymbol{n}, \boldsymbol{f})\boldsymbol{c}(n) + \boldsymbol{w}_{\mathrm{awgn}}(n) \tag{3}$$

(scalar model (1) with rank p, complex amplitude $c_i = a_i e^{j\varphi_i}$ and noise $w_{\mathrm{awgn}}(n)$; vector model (2) with window length N and vector $\boldsymbol{n} = [n \ldots n+N-1]$; matrix-vector model (3) with $N \times p$ matrix $\mathbf{A} = [\boldsymbol{a}(f_1) \ldots \boldsymbol{a}(f_p)]$, where vectors $\boldsymbol{a}(f) = \left[e^{2\pi j \cdot f \frac{n}{f_s}} \ldots e^{2\pi j \cdot f \frac{n+N-1}{f_s}} \right]^T$). The dominating sinusoid will naturally be the fundamental frequency (50 Hz/60 Hz, depending on the region). Next to this sinusoid, there may be interim sinusoids. These originate in oscillations from capacitances and inductances characterizing the transmission lines and routing devices triggered by changes (intended or by fault) of the net configuration. They may also be the result of not correctly synchronized/well-controlled areas of the net (*interarea modes*).

2.1 Database assisted parameter estimation

We have designed a database-assisted parameter estimation method (DaPT, Lechtenberg and Götze, 2011). This algorithm rates incoming parameter estimations over time and provides a noise-reduced parameter estimation with reliable rank estimation.

This database has entries for each signal component containing frequency f_i and (complex) amplitude c_i. Additional fields hold

1. *Frequency drift*: this field tracks the change of the frequency estimations against the database field for threshold-based frequency field updates.
2. *Phase difference*: it tracks the change of the phase gradient as second indicator for a misestimated frequency. The phase is assumed to be constant, so a linearly changing phase estimation (based on the database's frequency) indicates a mismatch.
3. *Severity of frequency drift/phase difference*: indicators for the prior mentioned threshold-based updating rules.
4. *Quality of entry*: this is the actual rating. It rates how often a frequency was part of the last estimation recursions. Based on that, components can be marked reliable (i.e. *present*), accounted for rank estimation and output.

In every recursion, the incoming frequency estimation is mapped to the most suitable database entry. Ideally, both values match neglecting the measurement noise. For entries that have been successfully mapped, the quality is increased, for the other decreased. For the drift/difference properties, the estimation is compared to the entry's features and exponentially weighted added to the field. Severity is updated like the quality mentioned before. After the drift fields have been updated, the drift's/difference's severity is checked whether the frequency component needs to be updated. Depending on the quality, the relevant (or reliable/present) entries are selected for output and counted for rank estimation.

2.2 Estimating parameters via rotational invariance techniques

Remembering the signal model in Eq. (3), the signal can also be described by subspaces like in

$$\boldsymbol{x}(n) = \mathbf{A}(\boldsymbol{n}, \boldsymbol{f})\boldsymbol{c}(n) + \boldsymbol{w}_{\mathrm{awgn}}(n) = \mathbf{S}(n) + \sigma^2_{\mathrm{awgn}}\mathbf{I} \ . \tag{4}$$

Within that equation, $\mathbf{S}(n)$ describes the signal subspace, which can be written as a product of the eigenvectors $\mathbf{W}(n)$ and the eigenvalues (on a diagonal matrix) $\mathbf{D}(n)$:

$$\mathbf{S}(n) = \mathbf{W}(n)\mathbf{D}(n)\mathbf{W}(n)^H \ . \tag{5}$$

Estimating frequencies (more general: *parameters*) via rotational invariance relates to the eigenvectors of the autocorrelation matrix of the input samples. In terms of subspaces described by eigenvectors, one time-step can be described by a multiplication with a diagonal matrix $\boldsymbol{\Phi}$ with diagonal entries $e^{2\pi j \cdot \frac{f_i}{f_s}}$:

$$\mathbf{S}(n+1) = \mathbf{W}(n)\mathbf{D}(n)\boldsymbol{\Phi}\mathbf{W}(n)^H \ . \tag{6}$$

In other words, the multiplication with $\boldsymbol{\Phi}$ is a rotation in the complex plane or a phase-shift in the time domain. This is visualized by Fig. 1.

The key idea of Esprit is to estimate the rotation $\boldsymbol{\Phi}$. Due to noise, $\boldsymbol{\Phi}$ will have off-diagonal elements. So an EVD is performed on $\boldsymbol{\Phi}$. The frequencies can be extracted from these

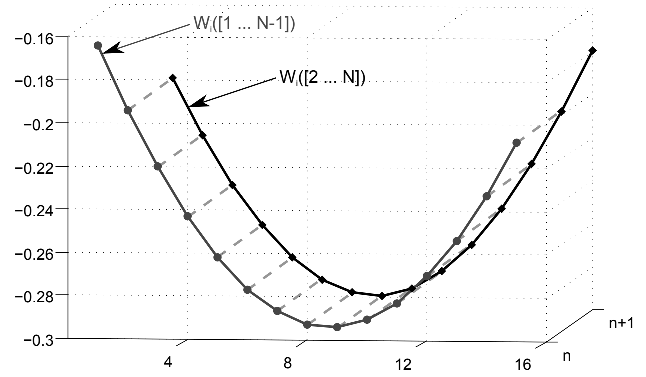

Fig. 1. Rotation of exemplary eigenvector from time-step n to $n+1$.

eigenvalues by calculating their angle. These frequencies are the input of DaPT.

Phase and amplitude estimation is obtained by a simple LS approach referring the signal model Eq. (3). Since time index and frequency are known, the matrix $\mathbf{A}$ can be built. Together with the samples the complex amplitude can be estimated.

3 Subspace tracking

The autocorrelation matrix of the input samples is defined as $\mathbf{R}_{xx} = E\left[\boldsymbol{x}\boldsymbol{x}^H\right]$. The expectation value is a statistical expression that cannot be calculated for a stream of samples, but it can be estimated e.g. via exponential filtering:

$$\mathbf{R}_{xx} \approx \tilde{\mathbf{R}}_{xx}(n) = \beta\tilde{\mathbf{R}}_{xx}(n-1) + \boldsymbol{x}\boldsymbol{x}^H \tag{7}$$

(β is called *forgetting factor*). An ordinary EVD can be applied to this approximated autocorrelation matrix to generate the eigenvectors for ESPRIT. Depending on the algorithm used, this costs $\mathcal{O}(N^2)$ to $\mathcal{O}(N^3)$, which is expensive considering that in most cases $N \gg p$. Subspace trackers realize the extraction of eigenvectors with computational costs in the range of $\mathcal{O}(p^2)$ by an recursive concept of calculating projections and deviations to a previous calculation instead of complete decompositions.

This paper will focus on one branch of subspace trackers, i.e. PAST and derivations of PAST. This is in order to be able to benchmark the influence of features like orthonormality of the subspace basis on the parameter estimation without having deviations due to different calculation approaches.

3.1 PAST

The *projection approximation subspace tracking* concept (Yang, 1995b) is based on a non-constraining cost function:

$$J(\mathbf{W}) = E\|\boldsymbol{x} - \mathbf{W}\mathbf{W}^H\boldsymbol{x}\|^2 = \mathrm{tr}(\mathbf{R}_{xx}) - \mathrm{tr}\left(\mathbf{W}^H\mathbf{R}_{xx}\mathbf{W}\right) + \mathrm{tr}\left(\mathbf{W}^H\mathbf{R}_{xx}\mathbf{W}\cdot\mathbf{W}^H\mathbf{W}\right) \tag{8}$$

(tr(...) is the trace operator). This cost function has some important properties: it is minimized *if and only if* $\mathbf{W}$ describes a basis of the signal subspace; there are no more constraints regarding $\mathbf{W}$, especially not orthonormality. In consequence $J(\mathbf{W})$ is minimized for every basis of the signal subspace, not only the basis consisting of the eigenvectors of the autocorrelation matrix.

To find the minimum of the cost function J, a *straightforward* recursive gradient-descend method is applied (step size μ). The Gradient of J is written

$$\bigtriangledown J = \left[-2\mathbf{R}_{xx} + \mathbf{R}_{xx}\mathbf{W}\mathbf{W}^H + \mathbf{W}\mathbf{W}^H\mathbf{R}_{xx}\right]\mathbf{W} \tag{9}$$

(here, $\mathbf{W}$ is short for $\mathbf{W}(n-1)$). The resulting updating equation for a $\mathbf{W}$ approaching the minimum is given as follows:

$$\mathbf{W}(n) = \mathbf{W}(n-1) - \mu\cdot\bigtriangledown J(n-1) \tag{10}$$

By rearranging this equation with the projection $\boldsymbol{y} = \mathbf{W}^H\boldsymbol{x}$ and the approximation $\mathbf{W}^H\mathbf{W} \approx \mathbf{I}$, this becomes

$$\mathbf{W}(n) = \mathbf{W}(n-1) - \mu[\boldsymbol{x}(n) - \mathbf{W}(n-1)\boldsymbol{y}(n)]\boldsymbol{y}^H(n)\ . \tag{11}$$

With this knowledge (Eq. 11) and the approximation concept of the autocorrelation matrix (Eq. 7) the cost function can be reformulated as an exponentially weighted LS (*least squares*) problem

$$J'(\mathbf{W}(n)) = \sum_{i=1}^{n}\beta^{n-i}\|\boldsymbol{x}(i) - \mathbf{W}(n)\boldsymbol{y}(i)\|^2. \tag{12}$$

This is also known from adaptive filtering. The minimization of this cost function is expressed by the following equations

$$\mathbf{W}(n) = \mathbf{R}_{xy}(n)\mathbf{R}_{yy}^{-1}(n) \tag{13}$$

$$\mathbf{R}_{xy}(n) = \beta\mathbf{R}_{xy}(n-1) + \boldsymbol{x}\boldsymbol{y}^H \tag{14}$$

$$\mathbf{R}_{yy}(n) = \beta\mathbf{R}_{yy}(n-1) + \boldsymbol{y}\boldsymbol{y}^H\ . \tag{15}$$

The PAST algorithm by Yang is based on the RLS (*recursive LS*) algorithm described by the Eq. (13) and the inversion of $\mathbf{R}_{xy}$. Whereas the $\mathbf{W}$ in the first definition of J (Eq. 8) was not constrained to hold orthogonal columns but produced an orthogonal solution, the reformulated J' (Eq. 12) only approaches an orthogonal solution while convergence depends on SNR and β.

This section shall only give the idea of PAST. An implementation can be found in other publications like Yang (1995b) or Abed-Meraim et al. (2000). The computational complexity is $3Np + \mathcal{O}(p^2)$, see Table 1.

3.2 OPAST

As mentioned before, the original PAST algorithm does not produce orthogonal results in every run. The OPAST (*orthogonal PAST*) by Abed-Meraim et al. (2000) is based on

Table 1. Comparison of the algorithms computational complexities (N window length and p rank, note $N \gg p$).

EVD	$\mathcal{O}(N^2)$
PAST	$3Np + \mathcal{O}(p^2)$
OPAST	$4Np + \mathcal{O}(p^2)$
OPAST-PGS	$5Np + \mathcal{O}(p^2)$
PASTd	$4Np + \mathcal{O}(p^2)$

the same cost function, but introduces an additional orthogonalizing step. This step is informally written like $\mathbf{W} = \mathbf{W}(\mathbf{W}^H\mathbf{W})^{-0.5}$, with its brackets including the exponent called *inverse square root*.

This orthogonalizing step ensures an orthogonal basis (represented by $\mathbf{W}$) in every time-step. This additional calculation is incorporated in the updating algorithm of PAST. The implementation can be found in the corresponding publication. The computational complexity is $4Np + \mathcal{O}(p^2)$, see Table 1.

3.3 OPAST-PGS

Due to Bartelmaos et al. (2005), a problem of OPAST is its sensitivity to numerical rounding errors. These are originated in the inverse square root concept. The authors propose a *pairwise Gram-Schmidt orthogonalization* (OPAST-PGS) to oppose this problem.

Remembering that the recursion is indexed by n and the rank is p, two indices are defined: $k_n = \mathrm{mod}(n, p)$ and $k_{n+1} = \mathrm{mod}(n+1, p)$. With these indices, the k_{n+1}-vector of $\mathbf{W}$ is made orthogonal to the k_n-vector, before it is normalized:

$$\boldsymbol{w}_{k_{n+1}}(n) = \boldsymbol{w}_{k_{n+1}}(n) - \boldsymbol{w}_{k_n}(n)\boldsymbol{w}_{k_n}^H(n)\boldsymbol{w}_{k_{n+1}}(n) \qquad (16)$$

$$\boldsymbol{w}_{k_{n+1}}(n) = \frac{\boldsymbol{w}_{k_{n+1}}(n)}{\|\boldsymbol{w}_{k_{n+1}}(n)\|} \,. \qquad (17)$$

The computational complexity is $5Np + \mathcal{O}(p^2)$ (Table 1).

3.4 PASTd

Another derivation of the PAST algorithm was proposed by Yang (1995a). This approach focuses on a further reduction of computational costs by unrolling the PAST algorithm. This *deflation* describes the sequential estimation of the principal components (or: basis-vectors) beginning with the most dominant. The most dominant is defined by the maximal (pseudo-)eigenvalue corresponding to the principal component.

The process is the same as for the original PAST, except that it is done for each vector of $\mathbf{W}$ separately. So, PAST is performed for the most dominant vector, then a projection of the current data onto this vector is removed from the data vector. The current vector is now excluded and the process will be repeated for the new dominant vector. This method results in less matrix computations resulting in a computational complexity of $4Np + \mathcal{O}(p)$. Note that it is $\mathcal{O}(p)$ and not $\mathcal{O}(p^2)$.

4 Simulation environment

The signal will always embody a 50 Hz-component (sloppy for *fundamental system frequency ideally being 50 Hz or 60 Hz, depending on the region*), slightly varying over time, and spontaneous, temporally present components with random parameters.

The simulation is Monte-Carlo based and performs 400 iterations, each having 20 sections. The following results are consistent and bias-free.

4.1 Signal generator

The signal generator is based on a concept of random sections. A default section has a length of 5 s. The actual sections will have a length that is a small, random integer multiple of the default length. Each section varies the parameters (frequency, amplitude) of the fundamental component a little. For each section, a small random number of additional components is added and/or removed. In consequence, the additional components will endure a small random number of sections. If an additional component endures more than one section, its parameters are also slightly changed randomly.

The noise applied to the generated signal is based on the AWGN (*additional white Gaussian noise*) model and is leveled 20 dB.

4.2 Exemplary estimation

Figure 2a–b shows a typical estimation of such a signal. When viewing the total simulation time, no details of the estimation are observable. To get an insight into these algorithms, two cut-outs of typical estimations will be discussed.

Since the database concept is non-linear and relies on its own statistics, especially the first estimations of a component have a high influence. Outliers from pure incoming frequency estimations can have high impact at first. Figure 3 shows this for – in this case – faulty initial estimations of PAST(d). When the database fields become statistically reliable, the output values become consistent again.

To visualize the principle of the phase difference based frequency updating, Fig. 4 gives a cut-out of a phase estimation. The solid line is not constant but linearly in-/decreasing (with steps). The gradient of this line is a hint on the deviation between the true frequency and the current estimation. At the positions where steps occur, this gradient has been used to update the estimation.

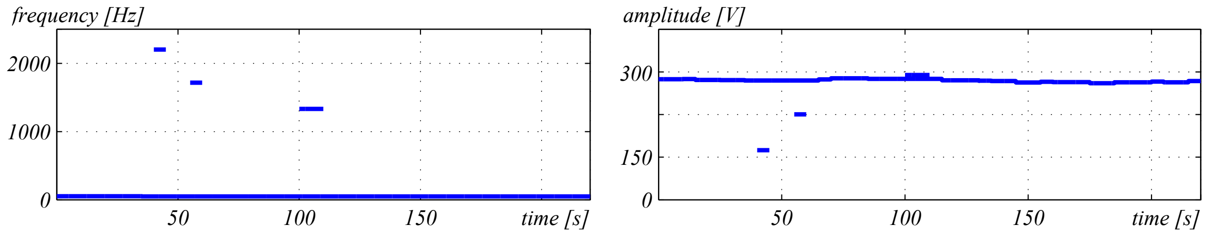

Fig. 2. A typical estimation picked from one of the simulations.

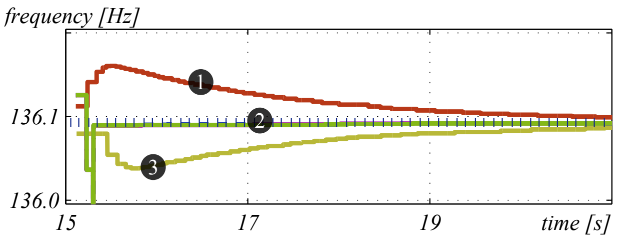

Fig. 3. A cut-out of a typical frequency estimation. (1) and (3) show aberrations due to erroneous parameter estimation and how they converge when estimations become reliable (here from PASTd and PAST); (2) points to the ideal dotted line and the EVD/OPAST-(PGS) estimation.

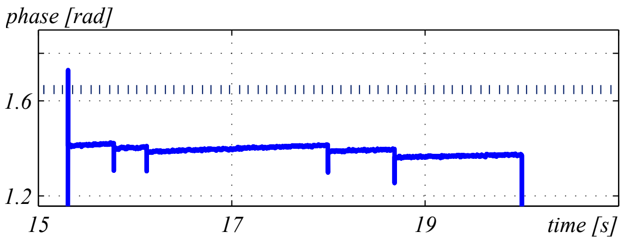

Fig. 4. A cut-out of a typical phase estimation. In the solid line, linearly increasing trends can be seen and the steps where DaPT reaches the threshold for phase difference severity and uses the gradient to correct the frequency estimation.

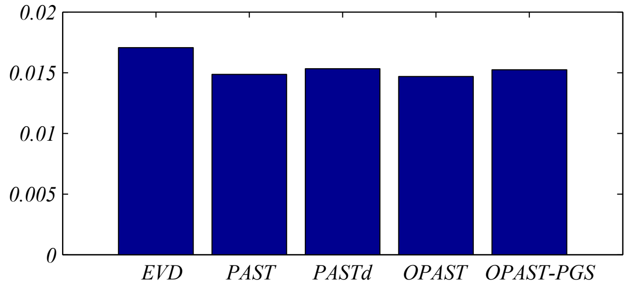

Fig. 5. Deviation of frequency estimation against true values, summed and averaged.

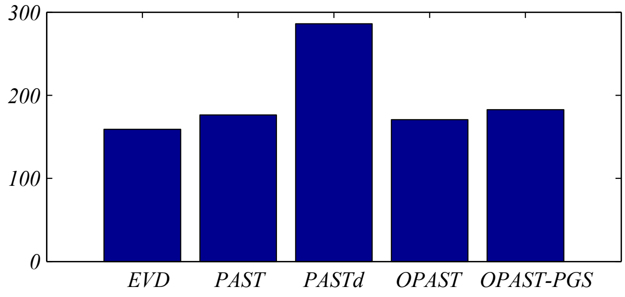

Fig. 6. Number of frequency components not found by the estimation per algorithm, summed and averaged.

5 Simulation results

To summarize the results of the Monte-Carlo simulation, three cases will be presented. For the first, we measured the deviation of the frequency estimation and the theoretical values from the signal generator. We summed them for all contemporaneous components and averaged the sums over the simulation time and all simulations. Figure 5 presents the results.

First of all, it emerges that the differences between the different algorithms are small. If there are findings at all, then EVD has slight disadvantages and (O)PAST has slight advantages. EVD might suffer from the position of the exponential filtering. In EVD the autocorrelation matrix itself is exponentially filtered, which is indirect since not $\mathbf{R}_{xx}$ but the EV are desired. The other algorithms perform that step directly on the basis of the subspace.

Also, we counted the number of frequency components that were not found by the estimator. Reasons may be (from the signal's perspecitve) a two weak amplitude or (from the statistical perspective) too high variance in the estimations. The numbers are per second and include the mandatory non-hits before DaPT regards them as present (resulting in an offset in the figure). In Fig. 6, the most significant fact is, that PASTd cannot keep up with the others. Also, EVD leaves a better mark. Both facts might be justified by the handling of estimations with high variance. A high variance in the estimations may confuse the sequential processing of PASTd. The joint cost of the serial cost functions might be higher than the cost of the (matrix-) PAST.

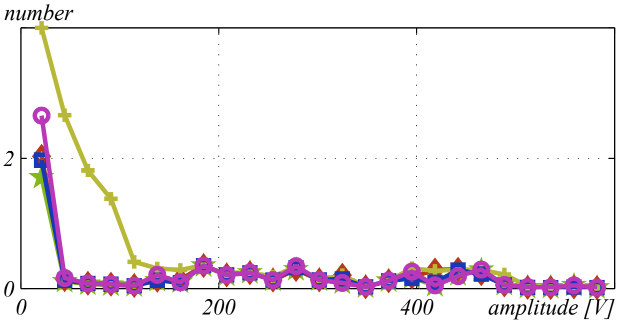

Fig. 7. Summed and averaged not-found components plotted against the amplitude: ★ EVD, ◇ PAST, □ OPAST, ○ OPAST-PGS, + PASTd.

High variances (indicating less reliable estimations) are most likely for low amplitudes (in relation to noise and dominant components). The last case surveys the unfound components against the amplitude, as (relative) amplitude is a good indicator for less reliable estimations. Figure 7 approves the previous results. PASTd is especially unsuitable for low amplitudes which may hint high variances in estimations. All the other perform equally well with slight advantages for EVD and (O)PAST. OPAST-PGS does not seem to bring advantages.

6 Conclusions

We have simulated the performance of different subspace trackers, based on the PAST theory, to evaluate the impact of the properties of the subspace basis presented to the ESPRIT algorithm. The estimated components supply the input of our DaPT non-linear parameter estimation post-processing. The subspace's basis of the signal generated from the (estimated) autocorrelation matrix of the input samples is not unique. The subspace can be described by various basis', some of them orthogonal, some even orthonormal and one of them being the signal eigenvectors of the autocorrelation matrix.

The simulations have shown, that it is not necessary for the subsequent steps (ESPRIT, DaPT) to be provided with exact eigenvectors (EVD). The basis presented to ESPRIT does not even have to be orthogonal (OPAST(-PGS)). The classic PAST algorithm, which does not guarantee an orthonormal basis, performs as well as the variations providing an orthonormal basis and as EVD providing actual eigenvectors.

A closer look reveals, that the sequential PASTd performs worse due to the lack of a joint cost function. The orthonormal OPAST is marginally better than PAST, because the approximation $\mathbf{W}^H\mathbf{W} \approx \mathbf{I}$ is more accurate for really orthonormal columns of $\mathbf{W}$. The algorithm with additional pairwise Gram-Schmidt orthogonalization does not improve the results.

In future work, the phase- and amplitude estimation shall be investigated further, i.e. other algorithms as LS relating the signal model. Also, cooperative estimations of regionally displaced measurements shall be investigated.

Acknowledgements. This work was supported by DFG (German Research Foundation).

References

Abed-Meraim, K., Chkeif, A., and Hua, Y.: Fast orthonormal PAST algorithm, IEEE Signal Proc. Let., 7, 60–62, doi:10.1109/97.823526, 2000.

Badeau, R., Richard, G., and David, B.: Fast and Stable YAST Algorithm for Principal and Minor Subspace Tracking, IEEE T. Signal Proces., 56, 3437–3446, doi:10.1109/TSP.2008.925924, 2008.

Bartelmaos, S., Abed-Meraim, K., and Attallah, S.: Mobile localization using subspace tracking, in: 2005 Asia-Pacific Conference on Communications, 1009–1013, doi:10.1109/APCC.2005.1554216, 2005.

Champagne, B. and Liu, Q.-G.: Plane rotation-based EVD updating schemes for efficient subspace tracking, IEEE T. Signal Proces., 46, 1886–1900, doi:10.1109/78.700961, 1998.

Lechtenberg, M. and Götze, J.: Database Assisted Frequency Estimation, in: 4th IEEE International Conference on Computer Science and Information Technology, 2011.

Paulraj, A., Roy, R., and Kailath, T.: Estimation Of Signal Parameters Via Rotational Invariance Techniques, in: IEEE Conference on Circuits, Systems and Computers, 83–89, 1985.

Schmidt, R.: Multiple emitter location and signal parameter estimation, IEEE T. Antenn. Propag., 34, 276–280, 1986.

Yang, B.: An extension of the PASTd algorithm to both rank and subspace tracking, 9, 179–182, doi:10.1109/97.410547, 1995a.

Yang, B.: Projection approximation subspace tracking, IEEE T. Signal Proces., 43, 95–107, doi:10.1109/78.365290, 1995b.