MATHEMATICAL FRAMEWORKS OF STATISTICAL DESIGN OF
THE MEASUREMENT SYSTEMS

A.A. Platonov
Institute of Electronic Systems, Warsaw University of Technology,
Nowowiejska 15/19, 00-665, Poland

Abstract: In the paper, there are discussed the backgrounds of unified approach to the analysis and statistical synthesis of optimal (best in the given conditions) analogue-digital measurement systems (MS), working in a presence of noises, random disturbances and observation errors. Methods of derivation of equations and algorithms for optimal estimation of measurands and for adaptive control the measurement experiment are presented. Apart from optimal software, the obtained results determine optimal structure, characteristics, the way of functioning and locked interaction of the main elements of analogue and digital parts of the MS. The approach enables effective mathematical support of the MS design at the most responsible, initial phase of design. It may also serve as a basis for development of more efficient, mathematically argued computer – aided methods of the MS design.

Keywords: optimal synthesis of MS, Bayesian approach, adaptive measurements

1. INTRODUCTION

Fast extension of requirements to new generations of analogue-digital MS causes growing complications in the MS design. Additional problems have appeared with inclusion of powerful microprocessors as an integral part of MS. Apart from estimation of measurands (measured parameters or characteristics of the model of a process at the sensor input), they solve additional complex tasks such as auto-calibration, control of the measurement process, etc. New properties of modern MS – multifunctionality, adaptation, flexibility can be realised also only in the software. Its design needs a developed theory. According to a common point of view [1-3], there is no general mathematical approach to the analysis of MSs with microprocessors, capable to support their design at the system level of decisions.

As the results of works [4-7] and [8] shows, promising preconditions for solution of this problem appear with adaptation and transfer of the methods of optimal statistical synthesis [9-13] into the metrology. The mathematical apparatus of statistical synthesis is a verified efficient tool of analytical support in design of different classes of complex analogue-digital radio and automatic systems. Its application allows the designers to determine optimal versions of systems at the most difficult and responsible initial stage of design, decisive for characteristics of the final product. This means:

1. The derived expressions or equations for optimal estimates of measurands not only enable a design of corresponding software, but establish the sequence of operations to be carried out during data acquisition and processing for these estimates were formed in most accurate and the fastest way.

2. In turn [9-12], this sequence determines a set of technical elements needed for realisation of these operations, as well as the main characteristics and the way of connection of these elements. Thus, the results of calculations allow us to find optimal structure and characteristics of both analogue and digital parts of the MS, i.e. the theoretically best version of MS before the engineering stage of design.

3. Solution of optimal synthesis problem puts the basis for the optimal decomposition of the initial design task into the simpler tasks for design of the main elements and subsystems of MS.

The paper presents new models and changes, which are necessary for the conventional methods of the statistical synthesis [9-12] and corresponding mathematical apparatus could be applied to the MS design.

Concerning a great delay in appearance of similar to analytical synthesis methods in metrology: we think that one of the main hindering factors was a lack of relatively simple and sufficiently adequate mathematical model of the sensor (or analog part of the measuring chain). Let us explain this.

- In conditions of normal functioning, saturation of signal receivers in radio-systems is a rare event. However, saturation of sensors is frequent, always taken into account source of temporary or full losses of information. For this reason, classic statistical synthesis may successfully employ the linear models of physical receivers. Models of the sensors, however, should take into account their working ranges and possible saturation, and must be non-linear. This radically complicates calculations and makes impossible direct application of known tools of statistical synthesis to the MS design.
Model of the sensor should take into account noises, drifts, setting and calibration errors, which may substantially worsen a final quality of measurements. For radio systems, these factors are less crucial.

2. ROLE OF SENSORS STATISTICAL FITTING IN STATISTICAL DESIGN OF MS

Universal, acceptably adequate and simple non-linear model of the sensor has been introduced in [14] and successfully used for optimal synthesis of different classes of MS ([4-7] and other works). The model is presented by the relationship (the case of discrete time \( n = 1, 2, \ldots \) is considered):

\[
\hat{Y}_n = \begin{cases} C_n (Y_n - B_n) + \xi_n, & \left| Y_n - B_n \right| \leq D / C_n \\ D \text{sgn}(Y_n - B_n) + \xi_n, & \left| Y_n - B_n \right| > D / C_n \end{cases}
\]  

(1)

where \( Y_n \) is a signal at the sensor output, \( Y_n \) describes the input analysed process, \( B_n, C_n \) describe the shift of zero point and sensitivity of the sensor respectively, and \([-D, D]\) are the boundaries of its output range. The internal noise of the sensor \( \xi_n \) (white, Gaussian, with the variance \( \sigma_n^2 \)) describes fluctuations in a sensing element, noises in secondary converters and amplifiers, etc.

For the statistical synthesis of MS could be performed, model (1) should be linearised. That can be done in the way most appropriate to the specifics of measurements by means of introduced in [14] (see also [4-7]) condition of statistical fitting of the sensors with possible values of the analysed processes \( Y_n \).

In different situations it has a different form [4-7,14], but common physical sense.

Definition. We say that saturation of the sensor (1) on the interval \([1, N]\) will be practically excluded, and the sensor will be statistically fitted with the analysed input process if for each \( n \in [1, N] \) parameters \( B_n, C_n \) of the sensor have the admissible values, guaranteeing that the probability of the output signal saturation for each \( n \in [1, N] \) will not exceed a given (confidence) value \( \mu, 0 < \mu << 1 \).

In this case, the probability of sensor saturation during the whole interval \([1, N]\) will not exceed the value:

\[
P^\text{sat} = 1 - (1 - \mu)^N = \mu N + O(\mu N^2)
\]  

(2)

Illustrative example: Let the process \( Y_n \) be quantified with a frequency \( F_s = 10 \text{ kHz} \). The probability of its saturation during one hour of observation \((T = 3600 \text{ s})\) by means of the sensor fitted with the process on the confidence level \( \mu = 10^{-10} \), will be not greater of \( P^\text{sat} = \mu F_s T \sim 10^{-8} \).

Claim 1. (For the case of possible adaptation of a sensor to the conditions of experiment [4]):

1. For the sensor parameters \( B_n, C_n \) were statistically fitted with the process \( Y_n \) at the level \( \mu, (0 < \mu << 1) \), they should belong to the set \( U_n \) of admissible values (in general case depending on previous observations: \( U_n = U_n (\hat{Y}_n^{m-1}) \)), determined by the relationship:

\[
\hat{U}_n (\hat{Y}_n^{m-1}) = \left\{(B_n, C_n): \Pr \left[ \left| Y_n - B_n \right| > \frac{D}{C_n} \hat{Y}_n^{m-1} \right] = 1 - \int_{-\frac{B_n + D/C_n}{C_n}}^{\frac{B_n + D/C_n}{C_n}} \right. p (Y_n) \hat{Y}_n^{m-1} dY_n < \mu \right\}
\]  

(3)

where \( \hat{Y}_n^{m-1} = (\hat{Y}_1, \ldots, \hat{Y}_{m-1}) \) is a data sequence at the sensor output; \( p (Y_n) \hat{Y}_n^{m-1} \) is the posterior probability density function (PDF) of the input process values; \( k \) is a time delay in the feedback (if any exist).

2. For each fitted sensor with parameters \( B_n, C_n \in U_n \) for each \( n \) the non-linear model (2) can be replaced by practically equivalent, strictly linear model:

\[
\hat{Y}_n = C_n (Y_n - B_n) + \xi_n
\]  

(4)

3. The results of processing of realisations \( \hat{Y}_n \) observed using fitted observer (1), and its linearised analogue (4) will be identical with a probability not less than \( 1-\mu \). For the Gaussian input processes, realisations of the signal at the sensor output preserve, almost always, the Gaussian properties.

Remark 1. Despite of strictly linear form of model (4), a nonlinearity of the sensor will be taken into account in calculations in the form of limitation (3) on admissible values of the sensor parameters.

Remark 2. Designers and users of MSs take always into account the range of future applications of the system. For this could be done, the prior PDF of possible values of the analysed process should be known. If this distribution is unknown, then the expected mean and variance, at least the majoring boundaries of the range of possible values of the process should be determined. On their basis, designers and users select the sensors with most appropriate parameters.

Remark 3. Bayesian methods have the important for applications advantage over non-Bayesian ones: optimal Bayesian estimates and controls remain optimal for the short intervals of observation, whilst the non-Bayesian ones are optimal only asymptotically, for a long time of observation \( n \to \infty \). The latter has a great meaning for applications, because one of the goals of optimal synthesis is a design of MSs for the fastest and accurate measurements.

This brief discussion shows that just the Bayesian approach is most adequate for the aims of the MS design. The listed above and other reasons [16] explain our increased attention to this approach.
3. **MATHEMATICAL BACKGROUNDS OF THE MS STATISTICAL SYNTHESIS**

Restoration of the form of analytical criterion of the measurement quality on the basis of prior knowledge is the main task of the first stage of optimal MS synthesis (see [9-12], Sect.3.2 in [8]).

3.1. **Analytical measure of quality and optimal design of measurement systems**

According to the main principles of statistical synthesis and estimation theory [9-12,17,18], quality criterion is to be constructed initially in the most general form, using **probabilistic models** of the processes and elements participating in the measurement experiment (elements of metasystem [3,19]).

These models have the form of usual or conditional **PDFs** of corresponding values and possess following, very useful for the **MS** design, properties:

1. They can be built by standard methods on the basis of “usual” mathematical models of corresponding processes or objects. “Usual” mathematical models [1,3,9-13,20] may have the form of mappings of the sets, algebraic, differential or difference (deterministic, quasi-deterministic or stochastic) equations, response functions, spectrums, etc. These models are constructed on the basis of physics, mathematics, theoretic fundamentals of technical disciplines, as well as experimentally.

2. There exist well known, not complex standard methods of decomposition or restoring of probabilistic models of the higher level using models of the lower levels. This property allows us to build general probabilistic models of measurement experiment and MS using analytically, heuristically or empirically constructed probabilistic models of their main components. They are: a) analysed process or object, b) sensor or analogue part of the measuring chain, c) controlling or adjusting elements, d) software – the way of data processing, obtaining and utilisation of the measurement information.

**Remark 4.** Decomposition of probabilistic models works as a peculiar “zoom” which can be applied to the detailed consideration of each chosen part of the measuring chain.

For the analytical measure a quality could be formed, designer should have following prior information:

1. Probabilistic model \( p(Y_i^n|\theta) \) of the analysed process or object \((n=1,2,...)\). Here \( Y_i^n=\{Y_i^1,...,Y_i^n\} \) are the samples of the process, \( \theta=(\theta^1,...,\theta^n)^T \) is the vector of unknown parameters, possibly changing with time.

2. **Measurands** - all or some number of the vector \( \theta \) components. Sometimes measurands are given in the form of definite functions \( f_\theta = f_\theta(\theta) \) of this vector, with a given domain \( \Theta \subset \mathbb{R}^d \).

3. **Probabilistic model** \( p(Y_i^n|\theta_0) \) of the sensor, where \( Y_i^n=\{Y_i^1,...,Y_i^n\} \) is the sequence of samples of the process at its input, \( Y_i^n \) is the realisation of a signal at its output (this model has not been used earlier in the statistical synthesis, and its appearance is caused by the specificity of the MS design).

4. Probabilistic model \( p(\hat{\theta}|Y_i^n) \) of the blocks, which form the estimates of measurands (usually called the randomised estimates [17]). Below we assume that estimates are formed deterministically, as definite functions \( \hat{\theta} = \hat{\theta}(Y_i^n) \) of observations \( Y_i^n \). The **PDF** \( p(\hat{\theta}|Y_i^n) = \delta(\hat{\theta} - \hat{\theta}(Y_i^n)) \) The domain of the estimates \( \hat{\theta} \in \Theta \) is assumed to be given (in general case, domains \( \Theta \subset \mathbb{E} \) and \( \Theta \subset \mathbb{R}^d \) are not identical).

5. The prior distribution \( p(\theta) \) of unknown parameters \( \theta \), \( \theta \in \Theta \).

6. Usually one-dimensional loss function \( p(\hat{\theta},\theta)=p(\hat{\theta} - \theta) \) or \( p(f_\theta, f_{\theta^n})=p(f_\theta - f_{\theta^n}) \) - the non-negative even function, minimal in the point \( \hat{\theta} = \theta \). The loss function can be treated as a model of comparator with a scale determined by the form of dependence \( p(\theta) \).

Calculations performed using these models and general probabilistic identities, gives following expression for the **full** probabilistic model \( p(\hat{\theta}|Y_i^n) \) of the measurement process:

\[
p(\hat{\theta}|Y_i^n) = p(\hat{\theta}) \int \delta(\hat{\theta} - \hat{\theta}(Y_i^n)) p(Y_i^n|\hat{\theta}) \, d\hat{\theta} \quad \text{(5)}
\]

belonging to the highest level in hierarchy of models. It determines a dependence of statistics of the estimates of measurands, formed by the designed MS, on probabilistic models of the analysed process, sensor, and the way of data processing. Using (5), criterion of quality can be determined as the averaged loss function \( p(\hat{\theta}|Y_i^n) \). As a result, we obtain the adapted to the aims of MS synthesis Bayesian risk of estimates \( \theta_0|Y_i^n \) of the measurand \( \theta \) :

\[
R_n = \int ... \int p(\hat{\theta}_n - \hat{\theta}_n) p(\hat{\theta}_n,\hat{\theta}_n) \, d\hat{\theta}_n \quad \text{(6)}
\]

In formula (6), the shortened notations: \( d\hat{\theta} = d\hat{\theta^1}...d\hat{\theta^n} \), \( dY_n = dY_n \times ... \times dY_{n^n} \) are used.

Next, following from general theory [9-12,17] assertion is of a great importance for the MS synthesis: **Claim 2.** Under weak additional constraints, for each even loss function and symmetrical posterior PDF \( p(\hat{\theta}|Y_i^n) \), optimal (minimising risk (6)) estimates of measurands \( \hat{\theta}_n|Y_i^n \) are equal to the conditional (posterior) averages:
\[
\hat{e}_n^\text{opt} = E(\hat{e}_n | \tilde{Y}_i^n) = \int \hat{e} \cdot p(\hat{e}, Y_i^n) \, d\hat{e}
\]  

The posterior PDF \( p(\hat{e}_n | \tilde{Y}_i^n) \) can be determined similarly to (5) using known Bayes’ formula:

\[
p(\theta_n | \tilde{Y}_i^n) = \frac{1}{p(\tilde{Y}_i^n)} p(\hat{e}) \int p(\hat{e}_n | \tilde{Y}_i^n) \, p(Y_i^n | \hat{e}) \, dY_i^n
\]

The latter relationships solve, in the most general and universal form, the task of optimal statistical synthesis of MSs. They are always valid in Gaussian approximation, widely used in theoretical researches, and in many other situations. Formulas (6)-(8) can be used also as the most convenient first approximation in the iterative form of the MS quasi-optimal synthesis.

3.2. The probabilistic models constructing

Many probabilistic models of the analysed processes and objects can be found experimentally or heuristically, utilising the particularities of the problem and selecting most suitable solutions. Apart from this, there exists a group of standard, general and widely used methods for construction of the probabilistic models on the basis of “usual” mathematical stochastic models. They are:

a) Static quasi-random models of the processes [9-12]. These models are presented by linear over \( e \) expressions of the regression type:

\[
Y_n = \sum_{i=0}^{L} \theta_i X_i^n + U_n + \nu_n = \hat{e}^T X_n + U_n + \nu_n
\]

and sufficiently well describe a large number of physical and technical objects and processes. Vector \( X_n = \{X_0^n,...,X_L^n\} \) in (10) describes a deterministic, known type of evolution with time of unknown parameters \( e \). The variable \( U_n \) describes the controlled actuation of the process or object. Process \( Y_n \) is observed using sensor (1).

b). Dynamic Markovian or auto-regressive models in the form of stochastic equations [9,12,15]:

\[
\theta_i^n = \sum_{j=0}^{L} \rho_{ij} \theta_j^{(n-1)} + \sum_{j=1}^{L} \beta_{ij} U_{ij}^{(n-1)} + \eta_{ij}^{(n-1)} \quad (i=1,...,L) \quad \text{or} \quad \hat{e}_n = \hat{\nu}_n + \hat{\nu} U_{n-1} + \hat{\nu}_n
\]

where matrices \( \hat{\nu}, \hat{\nu} \) are completely or partially known. Both elements of the vector \( \hat{e} \) and matrices \( \hat{\nu}, \hat{\nu} \) can be measurands of the process. Restoration of the models can be carried out using connections between (11) and Fokker-Plank-Kolmogorov equations [9,12,15] or by direct calculations.

c) In many cases it is known that the analysed process has Gaussian or close to Gaussian properties. Then, its probabilistic model can be built on the basis of measured or calculated mean values and covariance function. In stationary case, the spectral density function can be applied.

Many details of transitions from physical-mathematics models to the probabilistic ones can be found in [4-7,9-15,18]. The volume of the paper does not allow the wider discussion. We will present here only the probabilistic model of the fitted sensor (1) constructed on the basis of equation (4):

\[
p(\tilde{Y}_i^n | Y_i^n) = p(\tilde{Y}_i^n) \int_{Y_i^n - c_s(Y_i^n - B_i)} = (2\pi \sigma_\xi^2)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma_\xi^2} \sum_{i=1}^{L} \left[ \tilde{Y}_i^n - C_s(Y_i^n - B_i) \right]^2 \right\}
\]

Recently, intensive investigations have been carried out in the field of the probabilistic models construction both in signal theory, automatics and mathematics.

Remark 5

For optimal estimates, mean risk (6) is minimal. Then, value \( R_m^{\text{min}} \) refers to the theoretically highest, “potential” accuracy of measurements carried out by “ideal” MS. For this reason, it can be used as a reference standard for analytical and numerical comparison of different technical solutions used for simplification of theoretically optimal version of the MS, or for its fitting with specification.

This puts the basis for development of the formalised methods of optimal or sub-optimal decomposition of the initial design task into the particular ones.

4. EXAMPLE: SYNTHESIS OF OPTIMAL ADAPTIVE STABILISER OF VOLTAGE

We assume the source of the constant voltage \( V_0 \) is not ideal and the real voltage \( V_n \) at its output may deviate in a random way, but not too fast, from \( V_0 \). For not too long time intervals, this voltage can be presented as \( V_n = V_0 + \gamma \cdot n \), where \( \gamma \) is random, normally distributed rate of drift. This voltage is measured in a presence of additive harmonic disturbance \( A \sin(2\pi fn + \Phi) \) with a known frequency and unknown random amplitude and phase and of broadband (white) Gaussian noise \( \nu_n \) with the variance \( \sigma^2 \). Measurements are performed using voltmeter with adaptively adjusted characteristic (1).
The measured voltage is corrected using voltage $U_n=U_n(Y_{n-m}^{opt})$ generated with technically conditioned delay $m$. The task is to find optimal structure of stabiliser and characteristics of subsystems needed for a synthesis of its initial, theoretically best version. Analytical criterion of quality - minimal mean square deviation of the stabilised voltage $V_n$ from the required reference value $V_0^*$: $R_n=E[(V_n-V_0^*)^2]=\min$.

In the given conditions, signal at the voltmeter input has the form:

$$Y_n=V_n=V_0^*+\gamma n+A\sin(2\pi fn+\phi)+U_n+\nu_n=\theta^T X_n+U_n+V_0^*+\nu_n$$  \hspace{1cm} (12)

where $\theta=(\gamma,A\cos, A\sin)^T$ describes the transformed vector of unknown parameters of the input process, considered as normally distributed values with zero mean $\theta_n^*=[0,0,0]^T$ and covariance matrix $P_n=\text{diag}(\sigma^2_\theta,\sigma^2_\theta,\sigma^2_\theta)$; $X_n=(n, \sin 2\pi fn, \cos 2\pi m)^T$ is deterministic vector-process describing a form of the voltage and disturbance time-evolution.

Restoration of corresponding probabilistic models, risk, next a solution of the extremum task (in the way similar to the used in [3-6]), results in a complete group of the relationships for:

1. Optimal correcting and stabilised voltages:

$$U_n^{opt}=-\hat{\theta}_{n-m}^T X_n; \quad \bar{V}_n^{opt}=V_0^*+(\hat{\theta}_n^* - \hat{\theta}_{n-m}^*)^T X_n + \nu_n$$  \hspace{1cm} (13)

2. Optimal rules for the sensor adaptive adjusting:

$$B_n=V_0^*+ (\hat{\theta}_{n-1}^* - \hat{\theta}_{n-m}^*)^T X_n; \quad C_n=D/\alpha_q\sigma^2_\theta+X_n^T P_n X_n$$  \hspace{1cm} (14)

3. Recursion for optimal estimates of unknown parameters of the input voltage:

$$\hat{\theta}_n^* = \hat{\theta}_{n-1}^* + K_n \bar{Y}_n; \quad \text{where} \quad K_n = C_n P_n^{-1} X_n [\sigma^2_\theta + C_n^2 (\sigma^2_\theta + X_n^T P_n X_n)]^{-1}$$  \hspace{1cm} (15)

4. Minimal stabilisation errors – MSE of deviations:

$$R_n = \sigma^2_\theta + X_n^T P_n X_n$$  \hspace{1cm} (16)

5. Recursion for the covariance measurement error matrix $P_n=E[(\hat{\theta}_n^* - \hat{\theta}_n )^T ]$:

$$P_n = P_{n-1} - K_n [\sigma^2_\theta + C_n^2 (\sigma^2_\theta + X_n^T P_{n-1} X_n)] K_n^T$$  \hspace{1cm} (17)

Initial conditions for the recursions are as follows: $\hat{\theta}_0^* = \nu_0^*; \quad P_0 = \text{diag}(\sigma^2_\theta, \sigma^2_\theta, \sigma^2_\theta); \quad B_n=V_0^*+(\hat{\theta}_{n-1}^* - \hat{\theta}_n )^T X_n$ and $U_n^{opt}=-\hat{\theta}_n^* X_n$ for $1 \leq n \leq m$.

![Figure 1. Block-diagram of theoretically optimal adaptive analogue-digital stabiliser of the voltage.](image)

Formulas (13)-(17) determine, step by step, the types, characteristics and sequence of optimal operations to be performed synchronically both in the analogue and digital part of the system to guarantee the most accurate stabilisation. This means that (13)-(17), apart from optimal software, determine theoretically best structure and characteristics of main subsystems of the stabiliser that solves a task of optimal synthesis. Block-diagram of theoretically optimal, “ideal” stabiliser is presented in Fig.1. The potentially achievable accuracy of stabilisation is determined by value (16).

The obtained results can be used for optimisation, systematisation and fastening of the design works. Simultaneously, they ensure the highest or close to the highest quality of the final product of design.
The considered example shows a possibility to use the approach not only for the MSs synthesis, but also for synthesis of more complex classes of embedded systems.

CONCLUSIONS

Model (1) of the non-linear observer and fitting condition (3) enable an efficient adaptation of the statistical synthesis, widely used in radio-technique, to the measurement systems design. This enables complex optimisation of their software and hardware -of observing, data processing and controlling parts of the designed systems including their interaction. The latter ensures full utilisation of the resources envisaged in the initial specification, and theoretically highest quality of the system work.

The approach may help the designers to make, at the initial stage of design, the argued analytically choice of the best in the given conditions version of MS. This can be used for next optimal or sub-optimal decomposition of a design task, as well as for optimisation of the program and schedule of the works. The approach can be used for development of mathematically based computer tools for support and optimisation of the measurement, control, and signal processing systems design.

The analysis shows, that Bayesian approach to synthesis of the analogous-digital MS has a number of advantages over non-Bayesian ones, especially for the short-time and dynamic measurements.

REFERENCES


ANATOLIY A. PLATONOVD, Assoc.-Prof., PhD, Institute of Electronic Systems, Warsaw University of Technology, Nowowiejska 15/19, 00-665, Warsaw, Poland, Phone int. (48-22)-660-79-52, Fax int. (48-22)-825-23-00, E-mail: plat@ise.pw.edu.pl.