Combining Statistics and AI in the Optimization of Semiconductors for Solar Cells

JÖRG RISIUS
GÜNTER SEIDELMANN
Knowledge Based Systems
Hahn-Meitner-Institut Berlin GmbH
Glienicker Str. 100
14109 Berlin, Germany

Abstract

In the framework of a research project in photovoltaics, a flexible method of parameter optimization is developed. Target of the optimization is the quality of semiconductor materials for solar cells. The quality depends on the parameter values chosen for the semiconductor production process. The optimization method is based on the combined application of statistics and artificial intelligence. Experiment design is used to collect and analyze experimental data from the process in order to acquire knowledge about the relationship between parameter values and semiconductor quality. Classifiers built by machine learning algorithms help to determine semiconductor quality by the inspection of special signals obtainable from the running process. A final on-line hillclimbing search for optimal parameter values is guided by both the classifier and the knowledge about process behaviour derived from previous experiments.

1 Introduction

The combination of methods from statistics and artificial intelligence is distinguishing the SOLEIL-O approach[1] to the optimization of semiconductors for solar cells. Thin semiconductor films are produced by a physical-chemical process known as plasma deposition. The purpose of the project is to develop a general method by which optimal parameter values for the control of plasma deposition can be found for arbitrary semiconductor materials. With respect to their application in solar cells the films produced should possess optoelectronic properties allowing for a maximum energy conversion rate.

The SOLEIL-O optimization method requires basically two series of experiments, referred to as off-line and on-line experiments. By the series of off-line experiments, data pertaining to the problem of quality determination for a semiconductor material are collected. Knowledge gained from an analysis of those data is the basis for the second series of experiments. By these experiments an on-line search for optimal deposition parameter values is conducted.

In particular, three aspects of the approach taken in SOLEIL-O have to be emphasized:

- Acquisition of knowledge about the influence of process parameters on semiconductor quality is based on statistical methods from *experiment design*
- *Machine learning* is applied to build a component for judging the film quality by classifying signals taken from the running process
- An *iterated hillclimbing search* for optimal control parameter values is performed on-line, guided by the knowledge about process behaviour and the quality ratings from the classification component

The remaining sections present a closer look at the building blocks of the optimization method. First, the physical background and the experimental environment have to be explained as far as needed. Section 3 gives an overview of the SOLEIL-O optimization method. Sections 4 and 5 deal with the two series of experiments in greater detail. The final section summarizes the essential aspects of the optimization method.

[1] SOLEIL-O is the computer science subpart of a research project in photovoltaics

2 Physical background and experimental environment

Central to the SOLEIL-O project is a plant for a sophisticated physical-chemical process called plasma deposition. Roughly explained, from a plasma of semiconductor molecules a film of a final thickness between 2 and 3μ is slowly depositing on a substrate at a deposition rate of about $1\mu/h$. All statements in this paper, which like the preceding one give details on the plasma deposition process, are referring to the production of amorphous hydrogenated silicon (a-Si:H). Other materials may require modified processing conditions, the optimization method, however, can be applied to all of them.

The plasma deposition process is controlled by several parameters. Six parameters considered to be most important have been included in the optimization. They are listed here to give the reader an idea of what is involved in controlling the process: temperature of substrate (T), power of RF source (P), pressure in reaction chamber (p), silane gas flow (F_{SiH_4}), helium gas flow (F_{He}), and electrode gap spacing (d). Their actual values during the deposition critically influence the quality of the semiconductor film.

Film quality which is the target of optimization, can be determined by two fundamentally different kinds of measurements:

- For *ex-situ* measurements of optoelectronic properties films must be removed from the process chamber
- *In-situ* measurements are obtainable directly from the process chamber while the deposition is in progress

In-situ measurements are based on the TRMC method (*Time Resolved Microwave Conductivity*) [NK92]. TRMC signals obtained by that method somehow reflect the number, mobility and lifetime of charge carriers induced in the film by a short laser pulse of about $20ns$ (nanoseconds). One single TRMC signal consists of 2000 values measured during a $200ns$ interval including one pulse of the laser gun. The generation of charge carriers caused by light is the basic mechanism by which solar cells work, so from the shape of a TRMC signal one can draw conclusions on the film quality.

It must be remarked that because TRMC signals are subject to various factors of noise, they have to be appropriately preprocessed before they can be used for an evaluation of film quality.

3 The SOLEIL-O Approach to Optimization

The most important aspect of the SOLEIL-O project is the fact that the optimization of semiconductors can be carried out on-line or *in-situ*. This is possible because by the TRMC signals information on the quality of the material being produced is available from the running plasma deposition process. Depending on what is being detected from an on-line analysis of the TRMC signals, control of the deposition process can be altered appropriately in order to improve on quality.

Algorithmically, the in-situ optimization of SOLEIL-O is organized as a hillclimbing search. However, for being able to guide that search, two basic problems have to be addressed:

- There must be a well-founded understanding of the principle effects the different process parameters have on semiconductor quality
- TRMC signals must be automatically analyzed for judging the quality of the semiconductor material produced

The solution to both problems adopted in SOLEIL-O is to conduct a series of carefully chosen off-line experiments for knowledge acquisition first. The in-situ optimization is then performed by another sequence of experiments. The way how experiment design, machine learning and hillclimbing are working together in these two series of experiments is described in the next two sections.

4 Acquisition of Basic Knowledge

4.1 Experiment Design

In general, for any semiconductor material only few is known about the depency of its optoelectronic properties on the control parameters of the plasma deposition process. One source of information is the basic theoretical and empirical knowledge provided by the physicists participating in the project. But, although they must be considered as domain experts, they cannot offer sufficient quantitative models for the relations between control parameter values, TRMC signals, and film quality.

Hence, additional knowledge has to be acquired. This can be done by performing and evaluating a number of deposition experiments. However, due to limitations in cost and time, that number must be kept reasonably small. Contrary to that, the parameter space for plasma deposition is very large. A systematic and efficient exploration of the parameter space is possible, if methods from experiment design [MGH89] are applied. They are employed in SOLEIL-O for planning, executing and statistically evaluating the series of off-line experiments. For example, the effects of the six control parameters (T), (P), (p), (F_{SiH_4}), (F_{He}), and (d) mentioned above can be studied using a full factorial design at two levels, which requires $2^6 = 64$ deposition experiments.

All films produced in the designed experiments are measured ex-situ in order to provide data on their optoelectronic properties (band gap, activation energy, defect density). These data are statistically analyzed to compute the effects of single control parameters (*factors* in design terminology), to discover factor interactions, and to derive coarse linear models of the relationship between optoelectronic properties and parameter values. If necessary, suitable refinement of the results (e.g. by quadratic response surfaces) can be achieved by designing and evaluating additional experiments.

4.2 Machine Learning

During the on-line experiments for in-situ optimization, TRMC signals from the running deposition must be automatically analyzed to judge the quality of a semiconductor film. The most difficult problem with the analysis of TRMC signals is the lack of a precise quantitative model relating TRMC signals to film quality. In SOLEIL-O, the analysis is therefore done by a classification component which is built using a machine learning approach. By machine learning algorithms, the unknown relation between signals and quality is derived as close as possible from a set of selected examples. Each example consists of an observation (preprocessed signal) and its classification (quality). Examples are collected by performing appropriate additional off-line experiments.

In the course of the project it turned out, however, that for certain technical reasons (not explained here, see [RS94]), TRMC signals do not allow for the determination of quality in absolute terms. What can be done instead is discover the trend of quality development during one deposition by considering the sequence of TRMC signals. Therefore, the approach for building classifiers had to be based on differences between consecutive TRMC signals. By that it became possible to detect improvement, stability or degradation of quality during a deposition experiment.

Classifiers for TRMC signals have been built using both a tailored version of the induction method ID3 [Sei93] and artificial neural networks using backpropagation. Misclassification rates of less than 15% have been achieved with either paradigm, what is considered acceptable. However, the decision trees generated by the ID3 algorithm have proved to be preferable, because by their inspection it is possible to understand how classifications are arrived at. This property did support the process of finding the most relevant attributes to extract in the preprocessing of raw TRMC signals.

5 Hillclimbing to Better Materials

An iterated hillclimbing search for optimal values of the control parameters is performed by running another series of deposition experiments. In each experiment of this series, the development of quality is permanently evaluated on-line by the classification component. The search strategy is relying on both these evaluations and the knowledge about process behaviour provided by the analysis of the off-line experiments.

Each iteration of hillclimbing (each deposition experiment) is started with parameter values known to produce a semiconductor film of good quality. Search proceeds during the deposition experiments by modifying parameter values such that based on the knowledge available an improvement of quality seems likely. Besides certain time constraints defined by the deposition process and the TRMC method, the fact that by the automatic classification of TRMC signals quality cannot be rated in absolute terms has strong influence on the search strategy.

The strategy can be roughly sketched as follows: Whenever the classifier rates the quality to be constant (i.e. it does not detect any trends in the sequence of TRMC signals), the next parameter modification is tried. Once the quality has been reported to be increasing, parameter values are left unchanged until the quality has stabilized at a higher level. However, if a degradation of quality is detected, parameters are immediately reset to their preceding values, such that the quality can reach the previous, higher level before the next search step is made. On-line experiments are continued until further improvements seem impossible. All search steps and their effects are recorded as an additional source of information.

6 Conclusions

Semiconductor optimization in SOLEIL-O is achieved by two series of experiments. One is executed for learning about the problem domain using experiment design and machine learning for data acquisition and analysis. The other one is used to perform an on-line hillclimbing search for optimal processing conditions. At present, on-line evaluation of quality is not possible in absolute terms. This problem has been accounted for in the generation of classifiers and the design of the search strategy.

Currently, after preliminary studies and a major reconstruction of the deposition plant have been finished, a complete optimization procedure for a-Si:H is being run.

References

[MGH89] R.L. Mason, R.F. Gunst, and J.L. Hess. *Statistical Design and Analysis of Experiments with Applications to Engineering and Science*. Wiley, 1989.

[NK92] H.C. Neitzert and M. Kunst. Microwave detected transient photoconductivity measurements during plasma deposition of intrinsic hydrogenated amorphous silicon. *Applied Physics A, Solids and Surfaces*, A 55:378–386, 1992.

[RS94] J. Risius and G. Seidelmann. Optimization of semiconductor films for solar energy cells. *PD-Vol.59 Software Systems in Engineering, ASME Energy-Sources Technology Conference*, pages 181–188, New Orleans, 1994.

[Sei93] G. Seidelmann. Using heuristics to speed up induction on continuous-valued attributes. In *Machine Learning: ECML-93*, pages 390–395. Springer, 1993.